BIM - Based process mining

BIM based process mining

Enabling knowledge reassurance and fact-based problem discovery within the AECFM industry

Graduation thesis for Construction Management and Engineering Msc by:

S. (Stijn) van Schaijk

2016

TNO

TU/e

Colophon

Title: Building Information Model (BIM) based process mining

Subtitle: Enabling knowledge reassurance and fact-based problem discovery within the Architecture, Engineering, Construction and Facility Management Industry

Author: S. (Stijn) van Schaijk
University student number: 0856091
E-mail: stijnvanschaijk@gmail.com
University: Eindhoven University of Technology
Graduation program: Master Construction Management and Engineering

Graduation committee:
Prof.dr.ir. B. (Bauke) de Vries
University supervisor (chairman graduation committee)
b.d.vries@tue.nl

Dr.dipl.-ing J. (Jakob) Beetz
University supervisor
j.beetz@tue.nl

Ir. L.A.H.M. (Léon) van Berlo
TNO External supervisor
Leon.vanberlo@tno.nl

Thesis defence Date: 01-03-2016

This thesis contains several pictures and URLs which are intended for digital usage. The digital version of this report can be downloaded via the link or QR-code below:
https://www.dropbox.com/s/ww1htr6kahuv7l7/Master%20Thesis%20Stijn%20van%20Schaijk%20Digital%20Version.pdf?dl=0

Preface

I am very proud to present this thesis as a result of my graduation research. The study is done in collaboration with the Eindhoven University of Technology and the Netherlands organization for applied scientific research TNO. I look back at an interesting time with hard work. I spoke a lot of inspiring people and want to thank all who has helped me during the research, among which all who helped me with conducting the case studies within this research.

Special thanks go to my university supervisor Jakob Beetz with whom I had a lot of interesting meetings and he really helped me defining the roadmap of this study. Infinite endless everlasting thanks go to my TNO supervisor Léon van Berlo who really guided me throughout the whole research. His advice, visions, network, good sense for coffee and humour really brought my research to a higher level.

Lastly I want to thank my family and my girlfriend for supporting me during my master track Construction Management and Engineering at the university. For that I am very thankful.

I hope you will enjoy reading this thesis as much as I enjoyed conducting the research.

Stijn van Schaijk

Eindhoven, February 2016

Table of contents

Summary

Research among practitioners from the construction industry in the Netherlands has indicated that failures in planning, in specific planning deviation, cause failure costs while executing the project. A realistic planning is suggested as a solution for this problem. The planning is based on how previous projects are executed and how this process lasted, or the company thinks it was lasted. To gain insight into how these processes actually proceeded and where bottlenecks occur projects are being monitored by the construction company. Interviews with practitioners revealed that these monitoring systems mainly depend on human written notes. It was noticed that despite the importance, the current practices of monitoring systems are still-time-consuming, costly and prone to errors. Thereby monitoring is rarely be done and the data is nearly used. This results in feeling based improving instead of fact based problem solving within construction companies.

Regarding the requirement to monitor construction processes and the fact that the currently used observation tools are inefficient and inaccurate the need for research regarding this problem is grounded. Within the research domain of process mining a possible solution for this problem can be found. Process mining provides an approach to gain insight and improve processes in a lot application domains. The goal of process mining is to gain event data, extract process-related information stored in event logs and discover a process model. Process mining is all about exploiting event data in a meaningful way in order to provide insights, identify bottlenecks, anticipate problems, record policy violations, recommend countermeasures, and streamline processes.

During the 'data explosion' from last decades the capabilities of information systems expanded rapidly. As a result the digital universe and the physical universe are becoming more and more aligned. The growth of the technological possibilities with Radio Frequency Identification (RFID), Global Positioning System (GPS), Intelligent Imaging Camera systems, sensor networks and Internet of Things will stimulate further alignment of the digital and physical universe. Even in the so called 'old fashioned' construction industry these new technologies are gradually implemented. Specifically the expansion of the use of Building Information Modelling (BIM) enables the industry to combine those new technologies within the design databases. This data explosion development makes it possible to record physical events during construction projects as 'event logs' and analyse them digitally.

Given (a) the interest in monitoring, observation of construction projects and the interest in construction process models, (b) the limited quality of current monitoring and observation tools and the unrealistic hand-made models, and (c) the possibilities to autonomous create, store, and extract event data with new technologies and BIM, it seems legit to study the value of relating event data to process models. Therefore this research targets on shortening construction projects by discovering bottlenecks with help of process mining and BIM.

In order to reach the target of the research state-of-the-art technologies are studied and a workflow is proposed. The Plan-Capture-Analyse-Reuse workflow enables continuously learning loops for contractors. It is seen that BIM based planning tools can be used to make as-planned models including planned tasks linked to building elements. Several monitoring technologies are available to capture the progress and planning deviation on construction sites. Within the case study executed within this research drone images are used to generate

as-built point clouds models which are compared to as-planned BIM models. The deviations of the as-planned model and as-built process where measured, but no tool which can generate event logs from those models was available. Event logs are necessary for the analysing phase. By use of the Eventlog service, a software tool which is made within this study, the as-planned model is translated into event logs. The as-built data is merged with those logs resulting in as-built event logs. The event log data format enables process mining analytics. It is noticed that with process mining tools bottlenecks from the event logs can be discovered. The case study which is executed and described during this research proved that it is possible to gain insight in bottlenecks and planning deviation. By use of the Plan-Capture-Analyse-Reuse workflow and several BIM and process mining tools, continuously learning loops for construction companies where realized. As part of this workflow the Planning consult software tool is developed. By use of this tool it is proven that event logs from previous projects can be reused in order to advice construction planners and identify risks in early phases of new to build construction projects. Thereby it serves as a basis for continuously learning and enables construction project time decrease.

Current state-of-the-art technologies regarding planning and monitoring systems for construction projects enable to autonomously measure progress. It is recommended to contractors to implement the concept of *process-oriented data warehouse*, a database containing information about relevant events happened in the company and his construction projects. As-planned BIM models can serve as a solid bases for such warehouses. BIM tools are able to make as-planned models which are suitable for project scheduling. At this moment as-planned BIM is mainly used on single projects and mainly for visualisation purposes. However companies are not aware of the value of those models and applications and reusability of the accompanied data. With process mining analytics in mind a contractor finds an additional purpose for his as-planned models and records its requirements in the BIM protocol. The applicability of BIM based process databases in practice is an interesting topic for further research.

Samenvatting

Onderzoek in de bouw sector in Nederland heeft aangetoond dat fouten in de planning, vooral afwijkingen in de planning, faalkosten veroorzaken tijdens bouwprojecten. Een meer realistische planning zou de oplossing moeten bieden voor dit probleem. Planningen voor bouwprojecten worden gemaakt met kennis en inzichten die zijn ontstaan gedurende eerder uitgevoerde projecten. Om inzicht te krijgen in het verloop van bouwprojecten worden ze gemonitord door bouwbedrijven. Interviews met experts uit de praktijk hebben aangetoond dat de systemen om te monitoren vooral zijn gebaseerd op hand geschreven notities. Ondanks dat men beseft dat monitoren belangrijk is zijn de huidige observatie systemen duur, tijd rovend, en foutgevoelig. Daarom wordt het monitoren zelden goed uitgevoerd en wordt de verzamelde data weinig gebruikt. Dit resulteert in het verbeteren van bouwprocessen op basis van gevoel, in plaats van op feiten gegenereerd uit accurate data.

Gezien de noodzaak van het monitoren van bouwprocessen en het feit dat huidige observatie tools inefficiënt en inaccuraat zijn, is onderzoek naar dit probleem verantwoord. Het wetenschappelijk onderzoek domein process mining biedt een mogelijke oplossing voor dit probleem. Process mining biedt methodes om inzicht te krijgen in processen en ze daadwerkelijk te verbeteren. Het doel van process mining is om 'event logs' te verzamelen, proces gerelateerde informatie eruit te halen en proces modellen te ontdekken. Process mining probeert de proces modellen te analyseren zodat nieuwe inzichten ontstaan, knelpunten ontdekt kunnen worden, op problemen geanticipeerd kan worden, afwijkingen ontdekt kunnen worden, oplossingen aangedragen kunnen worden, en processen uiteindelijk gestroomlijnd kunnen worden.

Gedurende de 'data-explosie' van de laatste decennia zijn de mogelijkheden van informatie systemen flink uitgebreid. Een resultaat hiervan is dat de digitale wereld steeds meer met de fysieke wereld is verbonden. Het groeiende aantal mogelijkheden met Radio Frequency Identification (RFID), Global Positioning System (GPS), Intelligente camera systemen, sensor netwerken, en het internet der dingen stimuleert de verbondenheid tussen fysiek en digitaal nog meer. Zelfs in de zogenaamde traditionele bouw sector worden deze technieken geleidelijk aan geïmplementeerd. Specifiek het groeiende gebruik van Bouw Informatie Modellen (BIM) zorgt ervoor dat het mogelijk is om deze nieuwe technieken te combineren met de ontwerp database van bouwprojecten. Deze data explosie maakt het mogelijk om fysieke events tijdens bouwprojecten te monitoren en op te slaan als event logs, om ze vervolgens digitaal te analyseren.

Gegeven (a) de interesse in het monitoren en observeren van bouw projecten en de interesse in proces modellen, (b) de beperkte kwaliteit van huidige monitoring tools en de onrealistische hand gemaakte proces modellen, en (c) de mogelijkheden om autonoom event data te meten en op te slaan met nieuwe technologieën en BIM, geeft genoeg aanleiding om onderzoek te doen naar het combineren van deze technieken. Daarom heeft dit onderzoek als doel om bouw projecten te verkorten door het ontdekken van knelpunten in projecten met behulp van process mining en BIM.

Om het doel van dit onderzoek te behalen zijn state-of-the-art technologieën bestudeerd en een werkmethode is voorgesteld. De Plan-Capture-Analyse-Reuse werkmethode maakt het mogelijk voor aannemers om continu te leren van projecten. Gedurende het onderzoek is er

gezien dat BIM gebaseerde planning tools gebruikt kunnen worden om 'zoals geplande'-modellen te maken waarbij geplande taken gekoppeld zijn aan bouw elementen. Verschillende monitoring technieken zijn beschikbaar om het proces vast te leggen en planning afwijking op bouwplaatsen te registeren. Met behulp van een case studie tijdens dit onderzoek zijn drones gebruikt om foto's te maken van bouwplaatsen. Deze foto's zijn vervolgens omgezet in 3D punten wolk modellen, de 'zoals gebouwd-'modellen. Deze punten wolken zijn vervolgens vergeleken met de zoals gepland- modellen. De verschillen tussen het geplande en het gebouwde zijn daarmee gemeten. Door gebruik te maken van de Eventlog service, een software tool gemaakt tijdens dit onderzoek, zijn er event logs gemaakt van het zoals gepland-model. Vervolgens is de daadwerkelijke gebouwde data toegevoegd aan de event logs. Deze event logs zijn dan geschikt voor process mining analyses. De case study heeft aangetoond dat het mogelijk is om knelpunten en planning afwijkingen te ontdekken met BIM en process mining. Met behulp van de Plan-Capture-Analyse-Reuse workflow en verschillende BIM en process mining tools, is het continue leren van projecten gerealiseerd. Als onderdeel van deze workflow is de Planning consult software tool ontwikkeld. Met behulp van deze tool is aangetoond dat event logs van eerder uitgevoerde projecten hergebruikt kunnen worden om advies te kunnen geven aan project planners en managers van bouwbedrijven. Planning- en risico advies kan automatisch gegenereerd worden in vroege stadia van nieuw te bouwen projecten. Daarmee biedt dit onderzoek een basis voor het continu leren en daarmee het verkorten van bouwtijden.

Het is aangetoond dat huidige state-of-the-art technieken het mogelijk maken om autonoom bouw projecten te monitoren en deze data op te slaan en te hergebruiken. Als gevolg hiervan is het aanbevolen aan bouwbedrijven om het concept van proces georiënteerde data warenhuizen te implementeren. Een database waarin alle relevante evenementen staan die zijn gebeurd binnen het bedrijf of binnen bouwprojecten. Geplande BIM modellen kunnen dienen als een basis voor dit soort databases. BIM software tools zijn geschikt om zoals geplande- BIM modellen te maken, maar op dit moment worden de modellen alleen projectmatig gebruikt voor visualisatie doeleinden. Bouwbedrijven zijn zich er niet bewust van dat deze modellen waardevolle informatie bevatten welke geanalyseerd en hergebruikt kan worden. Met de toepassingen van BIM en process mining, zoals beschreven in dit onderzoek, kan een bouwbedrijf extra doelen hebben voor de zoals-geplande modellen. De toepasbaarheid en meerwaarde van BIM gebaseerde proces databases in de praktijk is daarom een interessant onderwerp voor vervolgstudie.

1   Reading guide

The research is divided in a theoretical- and a practical part. As can be seen in Figure 1 the report follows this structure. The report starts first with elaborating the problem definitions and research questions, thereafter a literature review (theoretical) is depicted followed by a description of the development of several tools and case studies (practical). The report ends with conclusions and recommendations for further work.

The classification of the chapters is elaborated and the target audience is addressed in this chapter.
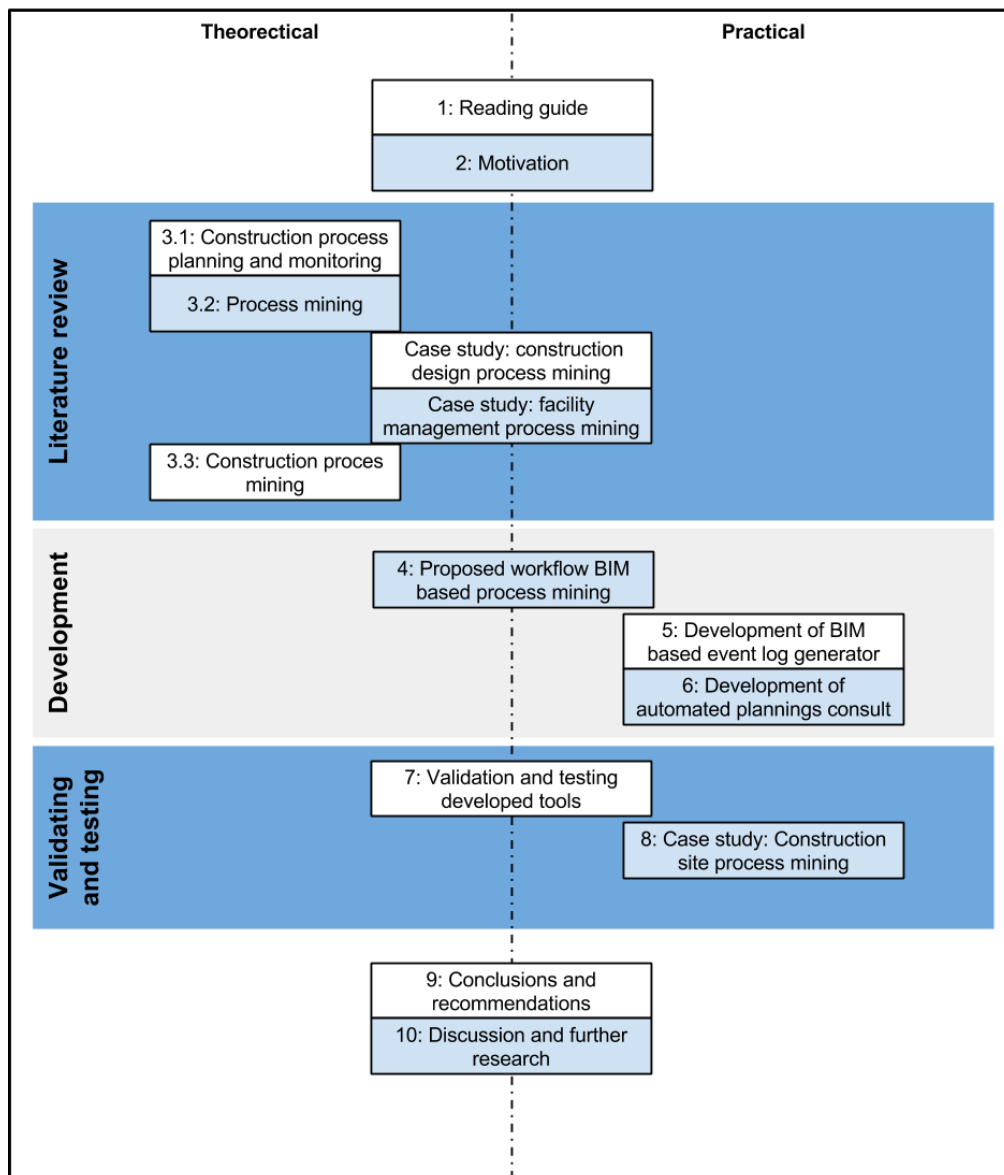
1.1   Chapter classification

Figure 1 Chapter classification

Chapter 1 describes the reading guide, the chapter classification and target group are addressed as well in this part.

Chapter 2 elaborates the motivation of the research. In addition the target of the study is explained. Moreover the research questions and methods are grounded in this chapter.

Chapter 3 gives background information and expounds previous research about the topics in the research. It concludes with research gabs and elaborates the motivation of the study even more.

Chapter 4 elaborates a proposed workflow which gives methods to reach the target of this study. In addition gabs in current technologies are identified.

Chapter 5, 6 and 7 describes the development and testing of the tools which fill the technology gabs as identified in chapter 4.

Chapter 8 describes the developed workflow. In addition the developed tools are tested in practice by use of a case study. Moreover this chapter analysis the applicability of the tools and workflows in practice.

Chapter 9 gives conclusions and recommendations for applicability of the proposed workflow and technologies.

Chapter 10 finally elaborates discussions and possibilities for further research into the field of construction process mining.

1.2   Target group

The report elaborates this research as detailed as possible but it can be used in different manners depending on the reader's interests. The report is written for several target groups: (1)Researchers or students in the fields of process mining, Building Information Modelling (BIM) and construction process management, (2) Practitioners in the construction industry such as contractors managers, project managers, project planners and BIM managers, (3) Software developers focussed on the construction industry, (4) Consultants in the field of Business Process improvement, and (5) innovators or people who follow the development of state-of-the-art technologies and accompanying possibilities.

Process mining researchers

Researchers within the process mining field will be interested in the applicability of process mining in the construction industry. They use this report to find opportunities for process mining in this specific sector. They also can have interest in the data sets of event logs which are used in this research and are available for scientific purposes (see Appendix 14).

BIM researchers

Researchers within the field of BIM can use this report as inspiration and find new applicability's and possibilities of BIM usage. In addition the data sets which are used in the study contain several BIM files and are available for further research (see Appendix 14).

Process management researchers

Researchers in the field of construction process management can be interested in the technologies and tools which are used in this research. Those give state-of-the-art methods of analysing and managing of construction companies and projects. Process management

researchers can use the proposed workflows and technologies within the report to study the applicability and usefulness in projects.

Practitioners

Practitioners within the construction field such as contractors managers, project managers, project planners or BIM managers will be most interested in the practical benefits and opportunities of the developed tools. They can use this report to see alternative opportunities and workflows to manage their projects and enable continuously learning loops.

Software developers

Software developers can use the developed tools as inspiration for developing new applications within their own software. Moreover the source code of the developed tools is publically available via GitHub and therefore it gives a basis for further development.

Consultants

Consultants in the field of business process improvement can use this research to find new opportunities within the construction sector. Within this research alternative ways of process improvement methods are presented which could be of interest for consultants in this field.

Innovators

Innovators, or people who just follow technological developments, will be attracted to the case studies presented in the report, where a combination of different technologies is applied and the state-of-the-art regarding construction planning and monitoring is experienced. In addition they can use this report to see data science applications and opportunities in the construction sector.

2   Motivation

Construction projects take too long, occasionally finish on time and cause a lot of nuisance. In general a lot of people benefit when the execution of construction projects are finished as soon as possible. First the residents in the neighbourhood of the construction site have nuisance of the project and therefore they will be pleased when a project is completed. Second the construction company has less costs if execution processes will be shorter. Because construction site costs a large amount of money every day. In addition also the client will be happy if his project is finished as soon as possible.

The construction company is responsible for the realization of the project. The company will make a planning and this will be followed during execution. Research among practitioners from the construction industry in the Netherlands has indicated that failures in planning, in specific planning deviation, cause failure costs while executing the project (Bouwkennis, 2013). A realistic planning is suggested as a solution for this problem.

A planning is made with knowledge of the construction process. Knowledge is a result of experience from the planner and the construction company. Experience is gained by doing and practicing the profession. The planning is based on how previous projects are executed and how this process lasted, or the company thinks it was lasted. To gain insight into how this processes actually proceeded and where bottlenecks occur projects are being monitored by the construction company. However Quirijnen and van Schaijk (2013) indicated that employees of construction companies have a feeling that some specific process are struggling but that even if people are aware of inefficiency they are still not improving the right problems because they do not know exactly what can be improved. This results in feeling based improving instead of fact based problem solving.

Currently a lot of construction companies trying to optimize their processes with help of Building Information Modelling (BIM) and LEAN approaches. Companies claim to be 'LEAN. Where they try to optimize processes in order to increase customer satisfaction. One of the fundamentals of LEAN is knowing what you are doing and have insight into your own processes (Sayer & Williams, 2013). Business Process Management (BPM) and Process Intelligence (PI) approaches are used to monitor and control LEAN process flows.

Interviews with practitioners conducted by Quirijnen & van Schaijk, (2013) revealed that these monitoring systems mainly depend on human written notes. It was noticed that despite the importance, the current practices of monitoring systems are still-time-consuming, costly and prone to errors. Thereby monitoring is rarely be done and the data is nearly used. Nevertheless, research in monitoring performance is not mature enough to enable control related studies (Yang, Park, Vela, & Golparvar-Fard, 2015).

Regarding the requirement to monitor construction processes and the fact that the currently used observation tools are inefficient and inaccurate the need for research regarding this problem is grounded. Within the research domain of process mining a possible solution for this problem can be found. Process mining provides an approach to gain insight and improve processes in a lot application domains. The goal of process mining is to gain event data, extract process-related information and discover a process model. Most organizations detect process problems based on fiction rather than facts. Van der Aalst (2011) describes process mining as

an "emerging discipline providing comprehensive sets of tools to provide fact-based insights and to support process improvements" (p. 7). In comparison with other data mining techniques, like for example Business Process Management (BPM) and Business Intelligence (BI), process mining provides a full understanding of as-is to end-to-end processes. BI focus on dashboards and reporting rather than clear process insights. BPM heavily relies on the experts which ware modelling the to-be process and not help organizations to understand the actual as-is processes. (van der Aalst, 2011)

During the 'data explosion' from last decades the capabilities of information systems expanded rapidly. As a result the digital universe and the physical universe are becoming more and more aligned. The growth of the technological possibilities with RFID (Radio Frequency Identification), GPS (Global Positioning System), Intelligent Imaging Camera systems, and sensor networks will stimulate further alignment of the digital and physical universe. Even in the so called old fashioned construction industry these new technologies are slowly implemented. Specifically the expansion of the use of BIM enables the industry to combine those technologies within the design databases. (i.e. (Meadati, Irizarry, & Amin, 2010) (Lu, Huang, & Li, 2011) (Bügler, Ogunmakin, Teizer, Vela, & Borrmann, 2014)) This data explosion development makes it possible to record physical events as 'event logs' and analyse them digitally. Events may range for example from buying a bread, step in a train, or arrival of a truck and installation of a prefab element. Process mining is all about exploiting event data in a meaningful way in order to provide insights, identify bottlenecks, anticipate problems, record policy violations, recommend countermeasures, and streamline processes. (van der Aalst, 2011)

Process mining is impossible without proper event logs. Sometimes it can be a challenge to store and extract the event logs from a range of data sources such as, databases, flat files, message logs, transaction logs, ERP systems, and document management systems. In the construction industry all of these data sources are used, but every company has his own data structure and software systems. In general these sources are not used to extract event logs, but are used to run daily activities. Thereby companies do not realize that they are creating valuable data. As mentioned earlier the monitoring and observation data are unstructured and thereby not suitable to easily generate event logs. In contrast with these old data structures a BIM process requires a structured data protocol. Therefore project teams and companies agree about how to develop a structured BIM in order to store, add, extract and re-use information. But in these protocols hardly any agreements are made about process information. Not at all about storing event logs in BIM. This can be explained by the fact that it is labour-intensive to add process data to these models, and that people do not realize the value of this data. It is hard to estimate how much insight the event logs will generate, how processes can be improved with this information and what kind of value this has. Therefore it is reasonable that people do not take the effort to store and track monitoring data in a consistent way. But with the technological developments autonomous generation of event logs will be more easy. In addition it would be possible to structure the data autonomously and connect it with BIM. So that one can get insight into the actual process of the elements designed in the model.

Little research is done in studying the potential of process mining with BIM in construction projects. A small study is executed by Terlouw and Mulder (2014) who explored the potential

of process mining with the ISO standard for communication VISI. They used the events extracted from VISI archives in order to gain insight into communication processes in civil projects. They gained insights within three projects from a social interaction perspective and concluded that the potential of this data analysing approach in the construction industry is high. When looking at other industries several process mining studies have been done from a product perspective. This means that it describes everything that happens with a specific product in order to gain insight into the process bottlenecks from that specific product. For example Paszkiewicz (2015) conducted a case study in order to analyse inventory processes. The study revealed that the actual process is deviating from the process which was assumed. Also the distribution of work was not very efficient. As a result the process was changed and seemed more efficient after two iterative improvement rounds. This approach can be useful when describing a process from a building element perspective. As proved in the study by Quirijnen & van Schaijk, (2013) there is probably a lot of inefficiency in the process of building elements. But how much this is and how these processes occur in real time is very time intensive to measure by hand. Therefore this study focusses on measuring and analysing the processes of building elements with new technologies.

Given (a) the interest in monitoring, observation of construction projects and the interest in construction process models, (b) the limited quality of current monitoring and observation tools and the unrealistic hand-made models, and (c) the possibilities to autonomous create, store, and extract event data with new technologies and BIM, it seems legit to study the value of relating event data to process models. Therefore this research will explore the possibilities of shortening construction project with help of process mining and BIM. Moreover this study consists of two parts, a theoretical and a practical part. Within the theoretical part the requirements to enable process mining with BIM will be described and the practical part tests with use of a case study the potential of these new technologies.

2.1 Problem definition

2.1.1 Problem analysis and research objectives

The problem where this study is focussing on is that construction processes too often deviate from planning and thereby take too long. Even though practitioners know it went wrong in previous projects, still they do not improve this processes or adjust their planning formats. Theoretically it should be possible to identify failures and improve them in future projects so that a process will succeed in a most efficient way. Practically this is very difficult to realize because a lot of factors influence the construction time. The failure identification and improvement of construction processes is studied with process mining techniques and BIM. The objective of this research is split into two parts, a theoretical and a practical target.

The theoretical objective is:

"To realize a shorter construction time for the execution phase of construction projects"

The practical target is:

"To determine bottlenecks and planning deviations (-and their causes) with process mining techniques and Building Information Models within the execution phase of construction projects by using autonomously gained observation data."

Based on the theoretical- and the practical target a hypothesis for this study is determined:

"If bottlenecks and planning deviations can be identified with autonomously gained data, process mining and BIM, the planning can be made more realistic and construction projects can be shortened."

2.1.2 Problem definition

Based on the objectives two problem definitions are defined. Problem definition 1 is the theoretical approach of the subject. Problem definition 2 is defined to realize the practical target of this study.

Problem definition 1:

"Can execution processes from construction projects be shortened based on insights gained with BIM and process mining techniques?"

Problem definition 2:

"Can bottlenecks and planning deviations (and their causes) be identified with BIM and process mining techniques (which are) employed in early construction process stages?"

2.2 Research question(s)

In order to achieve the objectives of this research and to solve the problem stated above several research questions are specified.

Questions that belong to problem definition 1:

Research question 1:
- How can processes of construction projects be characterized?

To shorten construction projects it is useful to understand how these processes are structured, characterized and executed in practice.

Research question 2:
- What kind of data can be used during the analysis for improving construction processes with process mining?

During construction projects a lot of data is generated, it is useful to understand which data has to be described in event logs such that it is valuable for improving execution processes.

Research question 3:
- How can process mining methods be used for analysing automatically harvested data to identify bottlenecks in construction projects?

There are different ways of analysing the gathered data, therefore it is useful to understand how the algorithms work and which ones are suitable for analysing construction projects.

Research question 4:
- How can process mining methods be used for identifying planning deviations in construction projects?

There are different ways of analysing the gathered data, therefore it is useful to understand how the algorithms work and which ones are suitable for analysing construction projects.

Research question 5:
- How can process data be linked or related to elements in BIM?

From the answers on the questions above it is known what kind of data should be collected and which algorithms could be used. This question answers how this data can be related to construction elements which are stored in BIM.

Questions that belong to problem definition 2:

Research question 6:
- What are the requirements for a toolbox which can be used to address the challenges posed in problem definition 2?

To solve problem definition 2 several tools could be used in practice. Therefore it is useful to understand what requirements the tools should have to solve the problem. This toolbox should be primarily focusing on contractors' managers or project leaders but could also be useful for other participants in the process. An example of elements of this toolbox could be: BIM software, Track and trace software and hardware, Data analysing software, Camera systems, etc. How these elements complement each other should be answered in this question.

2.3   Research approach

This study is divided in two parts, a theoretical and a practical part. The first problem definition[1], belongs to the theoretical part which is addressed along with the sub questions by knowledge gained with desk research. More specific, research question five is addressed by the employment of open building data models such as the Industry Foundation Classes (IFC). In addition it is described and prototypically demonstrated how process data can be stored and extracted in IFC using data processing tools such as the BIMserver.org platform.

The practical part is problem definition 2[2]. To answer this question and the sub questions and validate the proposed workflow an in depth case study is conducted. This approach allows to narrow down a very broad field of research into one easily researchable topic (Martyn, 2015).

2.4   Expected results

The results of this research are twofold. First a theoretical scientific approach is given which describes if construction processes can be shortened with use of BIM and process mining. In addition big data- and process mining methods will be explored for applicability in the construction industry especially with use of BIM databases. Process mining is rarely applied in the field of construction, therefore this study will contribute to the knowledge of this research field in the applicability in the construction industry. This combination of approaches has not been done to date. As a result of this thesis a new application area of BIM is introduced.

Second this study will explore the state-of-the-art of monitoring techniques. Therefore it is useful for companies, especially contractors, which are increasingly interested in process monitoring and process models. This study will be describe the potential of this new developments. Thereby contractors can determine if they are willing to invest in data analysis capacity in software, hardware and in human resources. Also, educational institutions will be able to determine if these kind of skills are useful to their students.

---

[1] "Can execution processes from construction projects be shortened based on insights gained with BIM and process mining techniques?"

[2] "Can bottlenecks and planning deviations (and their causes) be identified with BIM and process mining techniques (which are) employed in early construction process stages?"

3 Literature review

In order to answer research questions one until five a literature study is conducted. The literature review consists of three parts. Starting with (1) a clear understanding of the past research into construction process planning and monitoring. Moreover the first part identifies the global technological state-of-the-art of progress monitoring and information systems in construction. Thereafter it contributes with an overview (2) of the past research into process mining in construction and other sectors where process mining has proven its value. In addition this part attentions on identifying bottlenecks and planning deviation methods with process mining algorithms since this is the focus of the study. The last part (3) is about the identified research gabs and the application of process mining within the construction industry.

Since the focus of this research is on the execution phase of construction projects it is reasonable that the scope of this review is also on the execution phase. Although to understand the complexity of construction processes and the potential of using process mining throughout the whole building life cycle some concepts are described also in other phases (like design and maintain stages).

3.1 Construction process planning and monitoring

In this section the characteristics of the construction sector are presented. The purpose of this section is not to give a complete and detailed report, but by use of previous research it pictures shortly the complexity of processes within construction projects. In addition the difficulty of manage project information and involved stakeholders which exists in the lifecycle of a building is described. In order to manage those projects IT systems are used. The trends in usage of those systems are elaborated. More specific focus on using IT systems in order to manage project schedules are elaborated.

3.1.1 The complexity of construction processes

Over time lot of studies have been done to construction processes. A study performed in the fifties by Cox and Goodman (1956) already concluded that the complexity of construction operations regarding housing projects demands a subsequent problem solving capability by the involved stakeholders. Winch (1987) stated that construction projects are amongst the most complex of all activities. Gidado (1996) supplemented this view by indicating that "there is continuous demands for speed in construction, cost and quality control, safety in the work place and avoidance of disputes, together with technological advances, economic liberalization and globalization, environmental issues and fragmentation of the construction industry"(p.3). This results in a rapid increase in the complexity of construction processes. Studies state that firms in the construction industry are blamed for inefficiency of operations (A. Cox & Thompson, 1997). Moreover, research has concluded that characteristics of the construction process seems to favour short term productivity while hampering innovation, technical development and learning (Dubois & Gadde, 2001). A common view among most authors seems the construction processes are complex and this complexity doesn't stimulate process improvement.

As a solution to manage this common complexity studies state that construction has to adopt techniques that have improved performance in other industries, for example just-in-time delivery (Pheng & Hui, 1999), LEAN management (Aziz & Hafez, 2013), supply chain management (Vrijhoef & Koskela, 2000) and industrialization of manufacturing processes (Gann, 1996). However Weick (1976) (as cited by Dubois & Gadde, 2001) suggest that the

construction industry is described as a 'loosely coupled system' with complex particularities. Therefore they state that management techniques which improve performance in other industries are not readily transferable to the construction sector because of its particularities.

One of the particularities in construction is the amount of actors involved in each project. Cox & Thompson (1997) argued that because the composition of actors is continually changing among projects it is difficult to make use of experience gained in previous projects. They concluded that this causes repeatedly the same problems and cost inefficiencies because 'learning curves have to be climbed each time'.

It can be concluded that construction processes became more complex over time. As a result practitioners in planning are facing a lot of difficulties and projects repeatedly run over time. Complicated individual parts within construction projects may cause problems for the whole project. In order to help them managing the project information systems including soft- and hardware are more common used to monitor the projects' progress. But what is exactly a construction process? Which phases does it exist of and how are soft- and hardware tools used to plan and monitor this processes? Those questions are answered in next paragraphs.

3.1.1.1 A construction process description

This section describes per phase the complexity in terms of stakeholders, information exchange and monitoring. The phases as described are non-exhaustive and give a short identification of a construction process.

The complexity of a construction project already starts with the notification that the phases can differ from one project to another (Zwikael, 2009), depending on the contract, stakeholders and project type. However a common structure can be established (Picard, Paszkiewicz, Strykowski, Wojciechowski, & Wojciech, 2014): starting with the preparation or plan phase, thereafter a construction phase, followed by an operation and maintenance phase. Finally the building can be demolished or recycled but since this is out of the scope of the study this phase will not be described. Each phase could contain several sub phases which are elaborated more.

Plan phase

Each project is started with a preparation or plan phase where a construction permit must be obtained. A project initiator (i.e. developer, real-estate agent, private client) acquires a plot of land and hires (usually) architects to design the blueprint of the building. Construction projects involve a lot of organizations, however mainly public agencies have to deliver agreements, certificates, and permits. In order to sketch the complexity Picard et al. (2014) listed a set of departments which often are consulted concerning construction permits: Area planner, Heritage planner, Industrial Lands planner, City Drug Policy, Coordinator, Social planning, Housing Centre, Building Processing Centre, Fire and Rescue Services, Environmental Protection, Licenses and Inspections, Legal Services, Police, Park Board. Each department has its own requirements regarding the constructions design. Validating the design with this large amount of requirements is difficult to manage. Often a project manager is hired by the initiator to lead and monitor the process to gain a construction permit. A growing amount of project teams use information systems to manage those complex design processes (JBKnowledge, 2015), more on this in chapter 3.1.2. When the permit is approved by the local public agency the preparation phase is over and evolves to the construction phase.

Construction phase

The goal of the construction phase is to realize the building according to the design which is approved with a permit during the preparation phase. In the construction phase a contractor gets involved. He is responsible for realizing the building, in order to do so a lot of activities are delegated to subcontractors. So during the construction phase a lot of specialists with various competences are involved to build the buildings. To identify the complexity and variety of parties which should be managed: Electricians, Plumbers, Painters, Bricklayers, Roof coverer, Glazier, Tiler, etcetera. Those specialist are usually contracted with short term project contracts. It happens that even in middle- or small sized projects a lot of contracts are signed by the main contractor. As a result the management of all parties, processes and activities in the construction phase is a complex task. The construction phase ends when the building is ready and a handover form is signed off. This procedure is involved with a series of inspections and quality checks.

Operate and maintenance phase

During the operation and maintenance phase the building is used for daily activities. The building should function proper according to the functions for which it was designed and constructed. In order to assure the quality operation and maintenance activities have to be performed like: Cleaning, Elevator maintenance, Electrical replacements, Painting, etcetera. Like in the construction phase these tasks are often outsourced to specialized organizations. As a consequence a wide spectrum of contracts have to be arranged. Optionally a property is transferred of ownership. Before selling potential buyers are interested in information about the property. In order to assess the value of the property a property condition assessment (PCA) is done (Straub, 2009). Several aspects are taken into account such as: building location and site, structural elements, interior elements, roofing, plumbing systems, electrical systems, safety systems, etcetera. An important part for value estimating concerns the costs of immediate and necessary future repairs to building parts (Picard et al., 2014). A PCA provides precise information about the state of a property, therefor it is often a basis for negotiations and financial planning. In addition also the financial archives of the operate and maintenance phase is key information for potential buyers that need to value the profitability of a property. Therefore several documents related to financial aspects have to be prepared. In case of large buildings the management of such documents and involved organizations within this phase is a complex task. As well as capturing, updating, and resolving customer issues is a crucial and demanding process.

Problem identification in construction processes

From previous chapters it is clear that the construction industry can be characterized with large fragmentation. In this regard, (Caldas, Soibelman, & Gasser, 2005) notice that in such a dynamic environment the exchange of information between two organization information systems and sources is crucial for efficient project management. However there is a growing adoption of tech solutions (JBKnowledge, 2015) in order manage all those information exchanges. More specific BIM is gaining more momentum to manage the requirements of all parties in order to make a good design.

3.1.2 BIM and IT systems adoption in construction

In order to effectively manage construction projects and stakeholders throughout their life cycle information technology (IT) systems in the construction industry are implemented more

and more (Jung & Gibson, 1999) IT budgets are gradually growing, more IT staff is hired and research and development departments within companies are focusing on IT applications more and more (JBKnowledge, 2015). Nowadays most firms within the construction industry are using IT systems as management tools and a growing amount of information is captured within those systems (Ma & Lu, 2010). JBKnowledge (2015) surveyed the usage of IT systems. Most are used to run daily activities such as accounting, estimating, biding, project scheduling and management, client relationship management, file storage and collaboration, conferencing and communications, on company level and on project level.

Traditionally 2D CAD drawings and documents where used in order to express the projects design. Now intelligent building databases within BIM environments are created in order to manage the essential building design and project data throughout the buildings life cycle. Growing computer performances contributed to growth of BIM platforms and applications (JBKnowledge, 2015). Over last year's BIM technology has been developed rapidly and a broad variety of programs used by different disciplines such as architects and structural engineers makes collaborations difficult. Varying file formats lose reliability as they are exchanged across platforms and software. To combat this inefficiency the Industry Foundation Class (IFC) is developed. Nowadays IFC tends to be mature and has widely been used.

3.1.2.1 IFC

Industry Foundation Classes (IFC) is a data schema for representing buildings and associated activities for designing, constructing, and maintaining them (BuildingSmart, 2016a). Simply explained IFC is a set of agreements which describes how building elements (such as a door, or a wall) can be stored digitally. The IFC specification is developed and maintained by buildingSmart International as a 'Data Standard'. The IFC open and neutral data standard enables a wide range of possibilities and is well suited for the construction domain (BuildingSmart, 2016a). IFC is designed to accommodate many different configurations or levels of detail, therewith it supports interoperability across software applications and industry domains. IFC can serve as a basis to describe, exchange and share building project data. Building components are expressed as objects with attributes (Tah, Carr, & Howes, 1999). IFC allows most of the required object definitions for building elements. Moreover an advantage of IFC is that missing attributes can be easily added if needed.

The specifications of IFC are determined within a schema. This schema describes a collection of entities (or classes), attributes, and relationships between entities. An IFC model is a population of the schema, following the patterns, templates and constraints as stipulated by the schema. (Liebich, 2009) Such a model describes in practice for example a project data model or a building information model. All implementations of IFC are called Model View Definition (MVD). A MVD defines a subset of the IFC schema that is needed for particular use (BuildingSmart, 2016a).

Although much of the initial focus of IFC has been on describing the physical components of buildings, the IFC schema also includes non-physical project data such as documents, costs, organisational aspects, schedules, and construction resources (T. M. Froese, 1999). Several studies have indicated that IFC can be used for process related intensions. For example several early studies claimed that information for planning and cost estimation can represented usefully with IFC (T. Froese, Fischer, & Ggrober, 1999; T. M. Froese, 1999). Zhang, Zhang, Hu, and Lu, (2008) used IFC to represent building element information and to simulate

construction processes. Fu, Aouad, Lee, Mashall-Ponting, and Wu, (2006) described the nD model using IFC to describe planning related information such as task and costs.

Process related properties in IFC

In order to answer research question five[3] it is studied how process data can be linked to elements in IFC. Processes within the IFC scheme can be independently described and represented in project models (T. M. Froese, 1999). Schedule activities can be associated with the work processes. There are two major entities in IFC to support construction processes: IfcProcess and IfcTask. IfcProcess represents an action taking place in building construction. Processes can have predecessors and successors of the process defined by IfcRelSequence. In addition processes can have resources assigned to it by IfcRelAssignsToProcess. A task is captures in IfcTask and is described as an identifiable unit of work carried out independently of any other units of work in a construction project (BuildingSmart, 2016b). Information such as construction work methods, schedule dates, and duration are defined at IfcTask level. The association between processes and the objects upon they operate is described in an objectified relationship entity named IfcRelAssignToProcess.

Specific process related properties which are used in this study, TaskName and TaskID, can be found in IfcTask[4]. The specific task is connected to schedule information by use of IfcRelAssignsTasks. Time-related information such as task start time and task end time can be found in IfcScheduleTimeControl[5]. This counts for IFC 2x3, with the development of IFC 2x4 this changed in IfcTaskTime (van Merriënboer, Dijkmans, Klerks, van Berlo, & Ree, 2013). However since 2x4 is still not adopted in every software 2x3 is adopted in this study.

3.1.3 Planning, capturing and monitoring on construction sites

Schedules can be used to effectively support the execution of the construction project. In addition, time schedules, resource schemes, or other documents for a feasible construction process control can be developed. In practice those documents are generated in early stages of the project. However, on-site works are characterized by permanent changes and modifications in project conditions (Mikulakova, König, Tauscher, & Beucke, 2010).

As stated by (Saidi, Lytle, & Stone, 2003) efficient data collection, a timely data analysis and a communication of the results in a well interpreted way are major concerns for construction companies. This section explores related work regarding construction process planning and monitoring reviewing the traditional methods and the state-of-the-art regarding automated planning generation with BIM.

3.1.3.1 Traditional planning methods

Construction project planners contribute to a construction organisation by ensuring that estimating and tendering are based on good understanding of the methods, time and space which are necessary to carry out the tasks for each project taken into account the corresponding risks (Graham M. Winch & Kelsey, 2005).

The project plan can be developed with a plethora of techniques and methods (Kenley & Seppänen, 2009). Harris and Ioannou, (1998) listed a long selection of methods, which can be

---

[3] "How can process data be linked or related to elements in BIM?"
[4] http://www.buildingsmart-tech.org/ifc/IFC2x3/TC1/html/ifcprocessextension/diagrams/diagram_0003.htm
[5] http://www.buildingsmart-tech.org/ifc/IFC2x3/TC1/html/ifcprocessextension/diagrams/diagram_0002.htm

categorized in two main categories *activity-based and location-based methodologies* (Kenley & Seppänen, 2009).

Activity based scheduling is the current dominant scheduling technique in the construction industry and was first developed in the 1950s but origins in early work of Taylor and Gannt in early 1900s. They introduced Gantt charts (see Figure 2) independently. Kelley and Walker, (1959) first coined the term *critical path method* (CPM) which is a method which has a deterministic structure. CPM provides a graphical view of the projects, predicts the time required to finish the project and shows which activities are critical in order to maintain the schedule and which are not. CPM can be implemented in Gantt charts. CPM describes activities and events as a network. CPM is developed for complex but fairly routine projects with minimal uncertainty in the completion times. However for less routine projects CPM is not suitable. Therefor an another activity based method is more suitable, for example is *Program Evaluation and Review Technique* (PERT) which has a probabilistic structure (as can be seen in Figure 3) While CPM is easy to understand and use, time variations are not considered. PERT is a method that allows for randomness in completion times. PERT is useful since it provides expected completion times of projects, probabilities of completions of events, critical path activities that impact the completion time, activities which may have slack and can lend resources to critical path activities. However the use of PERT can be limited as little experience is available about durations of activities.
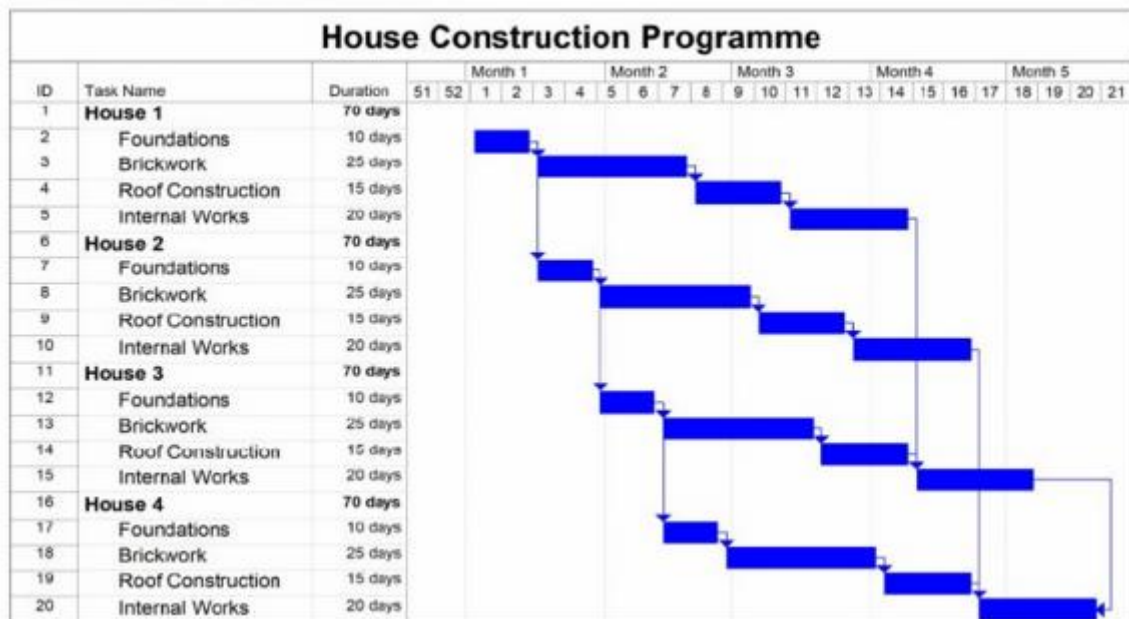

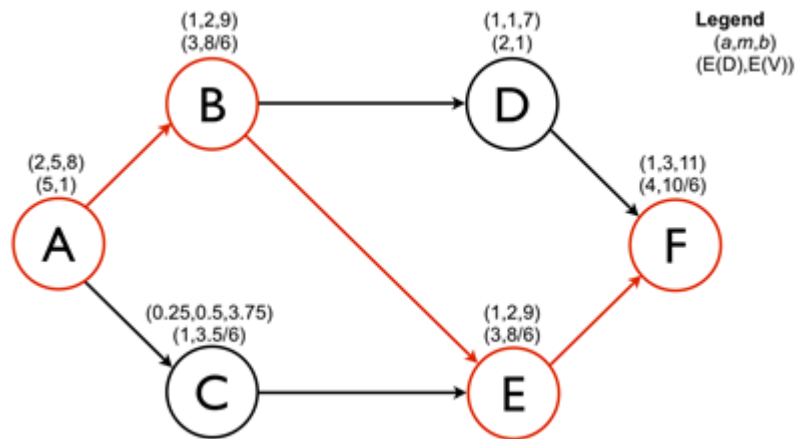**Figure 2 Example of a Gantt chart**

Figure 3 Example of a PERT network

Common to both CPM and PERT is the underlying logical structure of the model where discrete activities are joined by logical relationships. Each activity is part of a activity model and maintains a relationship with its predecessors and successors. Activity based planning methods are considered elegant and powerful and has enabled improving project performance within the construction industry (Kenley & Seppänen, 2009).

Location based scheduling is an methodology based on tracking the continuity of crews working on tasks. Those techniques are based on methods developed in early 1900 and used since the 1929 on projects like the Empire State Building (Kenley & Seppänen, 2009). But until now has found limited support in commercial construction, however modern methods have shown that CPM and location based scheduling can be combined within modern software tools (such as Vico[6]) which allows powerful scheduling methods which thus expected to gain popularity. The location based methods such as line of balance and flowline method (see are concerned with movement of resources trough locations or places. In comparison to activity based methods it doesn't focus on repetition, but concentrates on tasks. Possibly this techniques should be called task-based methods, however this is indicated as confusingly with the term activity-based. Therefore location-based is adopted (Kenley & Seppänen, 2009). The major benefits of this method are that it provides production rates and duration information in form of easily interpreted graphics format and t allows a smooth and efficient flow of resources.
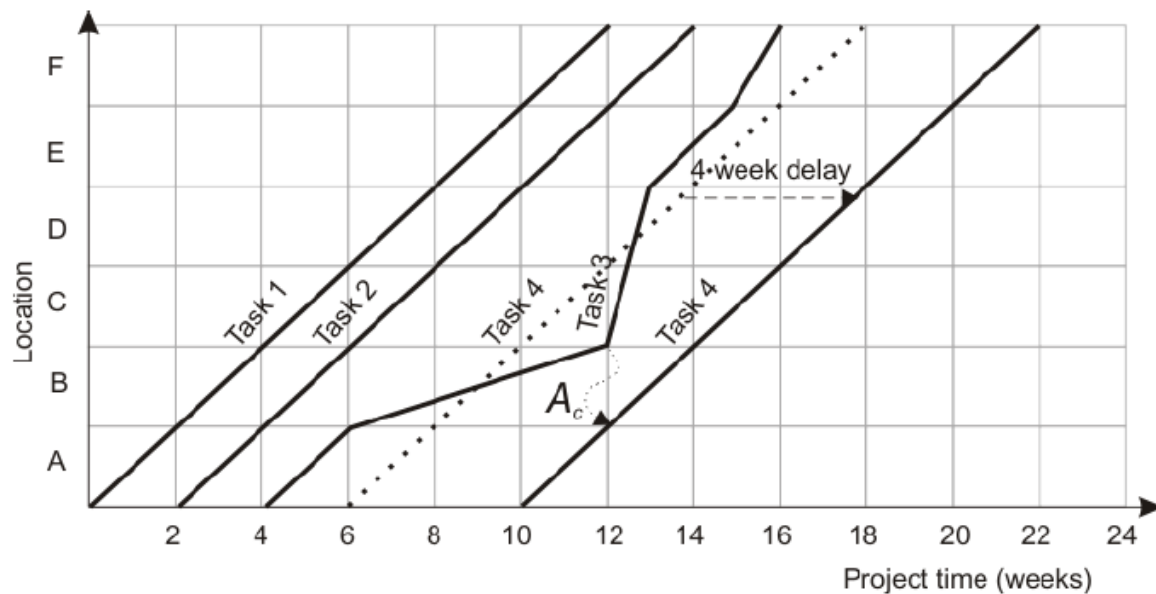
---

[6] http://www.vicosoftware.com/

Figure 4 Flowline of four planned balanced tasks, showing effects of delay (Kenley & Seppänen, 2009).

3.1.3.2 More advanced and automated planning methods

Traditional methods have proven its value regarding construction process management. However methods such as CPM or PERT support the documentation of construction schedules but not their generation. The schedules created with those methods are produced manually and are thereby error prone (Mikulakova et al., 2010). Since the adoption of IT systems is growing new methodologies have been developed on top of existing planning methods. The aim of those procedures is to improve accuracy and quality of schedules based on predefined algorithms. (Huhnt & Enge, 2006) This chapter elaborates shortly current usages of automated planning methods.

BIM planning

The as-planned models, often referred to as 4D models (Wang, Zhang, Chau, & Anson, 2004), help to increase the understanding, coordination and communication of construction processes (Mikulakova et al., 2010). By linking building components from the BIM to construction processes an as-planned model is created. However the linkage can be time consuming, as-planned models can demonstrate the entire construction process in a bright way and show potential conflicts in early stages (Wang et al., 2004). Collier and Fischer, (1995) applied CAD geometry and project schedules in order to generate 4D movies. Since then several applications and tools are built in order to connect BIM to schedules and create as-planned models (such as Autodesk Navisworks[7] or Synchro[8]) In addition the usage of those planning tools is established in the construction industry (Kassem, Dawood, & Chavada, 2015). Studies have extended as-planned models with incorporating other aspects as cost assessment and resource management.

Planning simulation

Linking BIM data and process information is an important step to support the automated generation of schedules. Computer simulation is not widely been used in the industry due to low generality and difficulties of development. Although several studies have been done in automating planning and production generation.

Chang and Jeong, (2015) managed to produce a construction operation simulation that can predict productivity at the operational level. They used literature and interviews in order to gain productivity data and used this as input. However they didn't evaluate the as-planned simulation with the as-built results.

As mentioned by (Mikulakova et al., 2010) experiences from previous projects which can provide guidance not accessible for companies or project participants. As identified by (Mikulakova et al., 2010) structured storing of process data enables the reuse of planning information. In order to do so they automated the generation of schedules based on experiences from previous projects. By use of a reasoning system construction alternatives can be evaluated automatically when processes changes or deviate.

3.1.3.3 Monitoring

As indicated by Quirijnen and van Schaijk (2013) and surveyed by JBKnowledge (2015), see Figure 5, monitoring on the construction site is mainly a manual process. This makes the data inaccurate and not easy to analyse and reuse. However there is a growing adoption of tech solutions which make it possible to easily collect and store data.
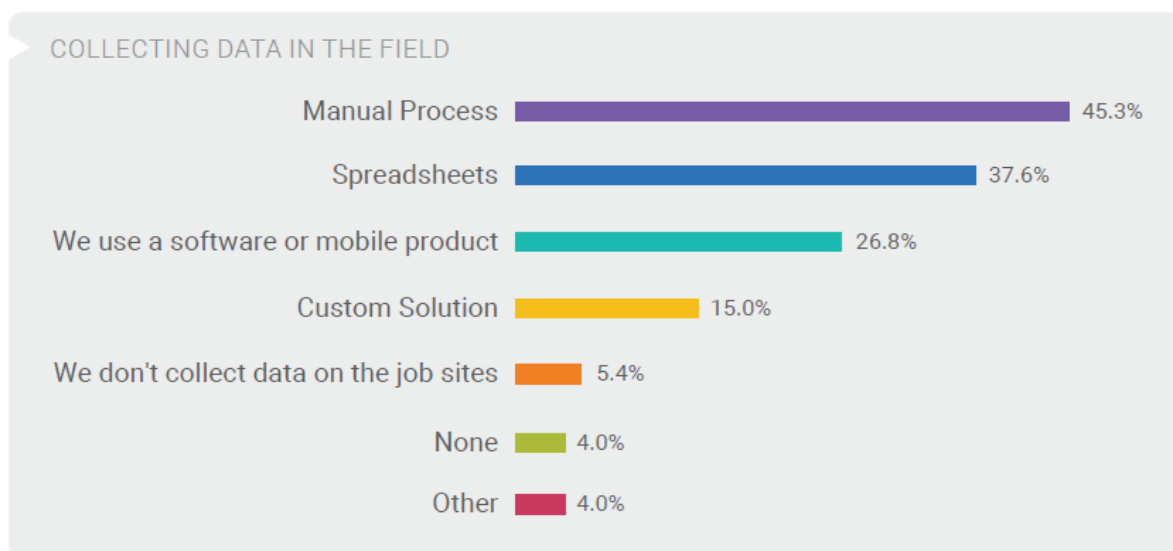
Figure 5 Data collection on construction sites as surveyed by JBKnowledge (2015)

A comprehensive overview of the current state-of-the-art in construction monitoring is given by Kopsida, Brilakis, and Vela (2015). They interpreted the technologies in terms of "(a) data acquisitions, as in technologies that are used for capturing the as-built scenes,(b) information retrieval, as in extracting the information needed from the as-built data,(c) progress estimation, which includes the comparison between the as-built and as-planned model in order to define the progress status and (d) visualisation of the results." (Kopsida et al., 2015, p1). Techniques such as Augmented Reality, RFID, Laser Scannners, Vision Static Images, and Vision Based Reconstruction are reviewed. As is concluded within the study the choice for best suitable monitoring method varies in each situation. The review nicely gives a recommendation of the most appropriate technologies to use base on the activity type, built environment and the needs of the inspection. The overview of monitoring technologies and there suitable applicability can be seen in Figure 6.

Mobile Augmented Reality (AR) systems are indicated as suitable in most situations because they are cheap and easy to use in every environment. However those systems require additional time and cost for installation and maintenance regarding Geospots, use of a WiFi network, or they do not perform in real time. Model bases AR algorithms are developed that could be used for comparison between the as-planned and as-built model, but their performance in real team operation on construction sites is not explored yet.

Laser scanning, image processing and computer vision techniques were frequently applied in research in order to automate progress estimation. Laser scanning is promising for as-built data acquisition due to its high accuracy. However it is expensive and labour intensive to obtain information about every room or element on a project. An alternative image processing method which uses images to create point clouds with structure from motion techniques can be used to identify progress deviations. However those point clouds do not capture interior elements and their tasks so progress measurement for interior tasks cannot be performed. So indoor progress monitoring provides only low level of automation. The user has to manually perform comparison between as-planned and as-built models.

	Mobile AR	Stationary AR	RFID	Laser Scanners	Vision Static Image	Vision Based Reconstruction	Ideal Case
Utility	multiple occasions	multiple occasions	multiple occasions	only spatial data	limited applications	only spatial data	multiple occasions
Time Efficiency	time spent on manual navigation of BIM	time spent on manual registration within registered view	instant information retrieval	time needed for scans	time spent on assigning manual information	time spent on the reconstruction	automatic information retrieval and assignment
Accuracy	AR registration, errors, subjective evaluation of progress	accurate AR registration, subjective evaluation of progress	subjective evaluation of progress	accurate	accurate for simple tasks	variable results, spatial data	accurate
Level of Automation	automated document management & data acquisition, no data analysis	automated document management & data acquisition, no data analysis	automated document management & data acquisition, no data analysis	automated document management & automated data analysis	partially automated data acquisition, automated data analysis	partially automated data acquisition, automated data analysis	fully automated
Required Preparation	minimal set-up required	set the equipment (<1h)	installation and maintenance of tags (>1h)	set the equipment (<1h)	minimal set-up required	minimal set-up required	minimum
Training Requirements	none	none	none	trained personnel for using the scanner	none	trained personnel for the reconstruction	none
Cost	consumer hardware	tracking cameras (£10000)	cost of installation & maintenance	laser scanner (£30000)	consumer hardware	consumer hardware	operates on commercial hardware
Mobility	handheld equipment	large equipment on tripod	handheld equipment	large & heavy equipment	handheld equipment	handheld equipment	handheld equipment

Figure 6 Performance of solutions for progress monitoring inspection made by Kopsida et al., (2015).

3.1.4 Research gab in construction process planning and monitoring

Construction projects are characterized as fragmented and as a result learning curves have to be climbed repeatedly. Construction projects traditionally have to work with manual processes and traditional communication techniques such as phone calls, faxes and e-mails (Dave, Kubler, Främling, & Koskela, 2015). This problem has been discussed over time broadly but the issue still is unsolved. Nowadays IT systems are more and more adopted in all phases of construction projects. Construction process related information is sometimes available within systems, however unstructured data formats containing tasks without any deeper information besides time. Since BIM is adopted more and it is possible to connect planning information to building elements in BIM. The process related information becomes accessible true databases and is interconnected with actual building data. Automatic monitoring technologies exists which can be applied to capture progress. At the same time methods are developed to automatically generate schedules from an event based database. However combining monitoring data within event based databases is seldom been studied. Luci and Shen, (2015) suggest that the creation of intelligent 4D databases should improve managing construction sites and that this should be studied more. In order to achieve the target of this study the possibility of using

event based databases in order to shorten construction project with autonomous gained process data is studied.

3.2 Process mining

Process mining provides an approach to gain insight and improve processes in a lot application domains. The goal of process mining is to gain event data, extract process-related information and discover a process model. Most organizations detect process problems based on fiction rather than facts. Van der Aalst (2011) describes process mining as an "emerging discipline providing comprehensive sets of tools to provide fact-based insights and to support process improvements" (p. 7). In comparison with other data mining techniques, like for example Business Process Management (BPM) and Business Intelligence (BI), process mining provides a full understanding of as-is to end-to-end processes. BI focus on dashboards and reporting rather than clear process insights. BPM heavily relies on the experts which are modelling the to-be process and not help organizations to understand the actual as-is processes. Applications of process mining can be found in various economic sectors and industries like healthcare, governments, banking and insurance, educational instances, retail, transportation, cloud computing, capital goods industry. (van der Aalst, 2011)

The idea of process mining is not new. The roots of the research field can be found a half century ago. (Nerode, 1958) already presented an approach to synthesize finite-state machines from traces, (Petri, 1962) introduced the first modelling language capturing concurrency and (Gold, 1967) first explored different notions of learnability. While data mining gained more attention during the nineties little attention was given to process related mining. Since the first survey on process mining in 2003 (van der Aalst et al., 2003) a lot of progress in the research field has been made. Several techniques have been developed and various tools have come into existence. A comprehensive overview of the state-of-the-art in process mining is given in the book of Wil van der Aalst "Process Mining Discovery, Conformance and Enhancement of Business Processes".

Process mining is all about exploiting event logs in a meaningful way in order to provide insights, identify bottlenecks, anticipate problems, record policy violations, recommend countermeasures, and streamline processes. (van der Aalst, 2011) With process mining operational questions can be answered, examples are:

- What *really happened in the past?*
- *Why did it happen?*
- *What is likely to happen in the future?*
- *When and why do organizations and people deviate?*
- *How to control a process better?*
- *How to redesign a process to improve its performance?*

Event logs enable three types of process mining (van der Aalst, 2011). The first one is *discovery* where an event log is taken as input and a process model (i.e. a Pertri net or BPMN model) is generated as output. In addition it is also possible to discover resource-related models, such as social networks, when the log contains information about resources.

The second type of mining is *conformance.* This method compares reality with an existing process model using event logs of the process. In addition conformance checking can be used to detect locate and explain deviations, and to measure the value of these deviations.

The third type of process mining is *enhancement.* This type focus on extending or improving existing processes using event logs. Where conformance checking measures alignment, enhancement focusses on repairing, changing or extending the model. A model can be extended by adding performance data. Doing this enables one to show bottlenecks, service levels, throughput times, and frequencies.

Process mining applications are found and determined over the last years. Still new applications are being developed. However (van der Aalst, 2011) defines the following perspectives regarding applications of process mining:

- *Control-flow perspective:* Focuses on the control-flow and ordering of activities. The goal of this perspective is to characterize all possible paths and visualize them in Petri nets or other notations like EPC, BPMN or UML Ads.
- *Organizational perspective:* Focuses on resources included in the log. For example actors or machines which are involved in the process. The goal is to gain insight in the structure of organizations by discovering for example the social network.
- *Case perspective:* focuses on properties of cases by characterizing its path in the process, or the originators working on it. Also its relation with other cases or corresponding data elements can be revealed.
- *Time perspective:* focuses on timing and frequency of events. Timestamps enable to discover bottlenecks, measure service levels, monitor utilization of resources or predict remaining process time of running cases.

Note: Those perspective are somehow overlapping and non-exhaustive, although they provide insight in the aspects of process mining and its aims.

This chapter starts with describing the 'oil' of process mining, event logs. Thereafter the focus points, bottlenecks and planning deviation, of this study are elaborated. However this study focusses on bottleneck and planning deviation , interesting applications of organizational mining within the construction industry are discovered during this research. Therefore the background of those mining techniques is elaborated as well.

3.2.1 Event logs

The starting point of process mining is typically a raw data source. A raw data source may be every file, for example an Excel spreadsheet, transaction log, or a database. Often the necessary data is scattered over different files or data sources. For cross-organizational mining those sources may be distributed over multiple organisations. The raw data source has to be converted into event logs in order to be suitable for process mining analytics. The availability of high quality event logs is essential in order to enable process mining (van der Aalst, 2011).

An important characteristic of process discovery algorithms is its *representational bias* (van der Aalst, 2011)*.* This is the type of process model which can be discovered with an algorithm. This can for example be a Petri net, a BMPN model, or a Causal net. The representational bias determines the potentials of a discovered model. Some types of representation are more searchable or analysable then others for example.

In order to answer research question two[9] this chapter describes how an event log should look like to enable construction process mining from a 'building element perspective'. First event logs in general is described shortly, thereafter the requirements regarding event logs to enable process mining in construction are described. Finally an indication of the results is given.

3.2.1.1 Structure

Using Figure 7 as a reference, the following is assumed regarding event logs (van der Aalst, 2011):

- A process consists of cases.

- A case consists of events such that each event related to precisely one case.

- Events within a case are ordered.

- Events can have attributes. Examples of attributes are activity, time, costs, and resource.

---

[9] "What kind of data can be used during the analysis for improving construction processes with process mining?"
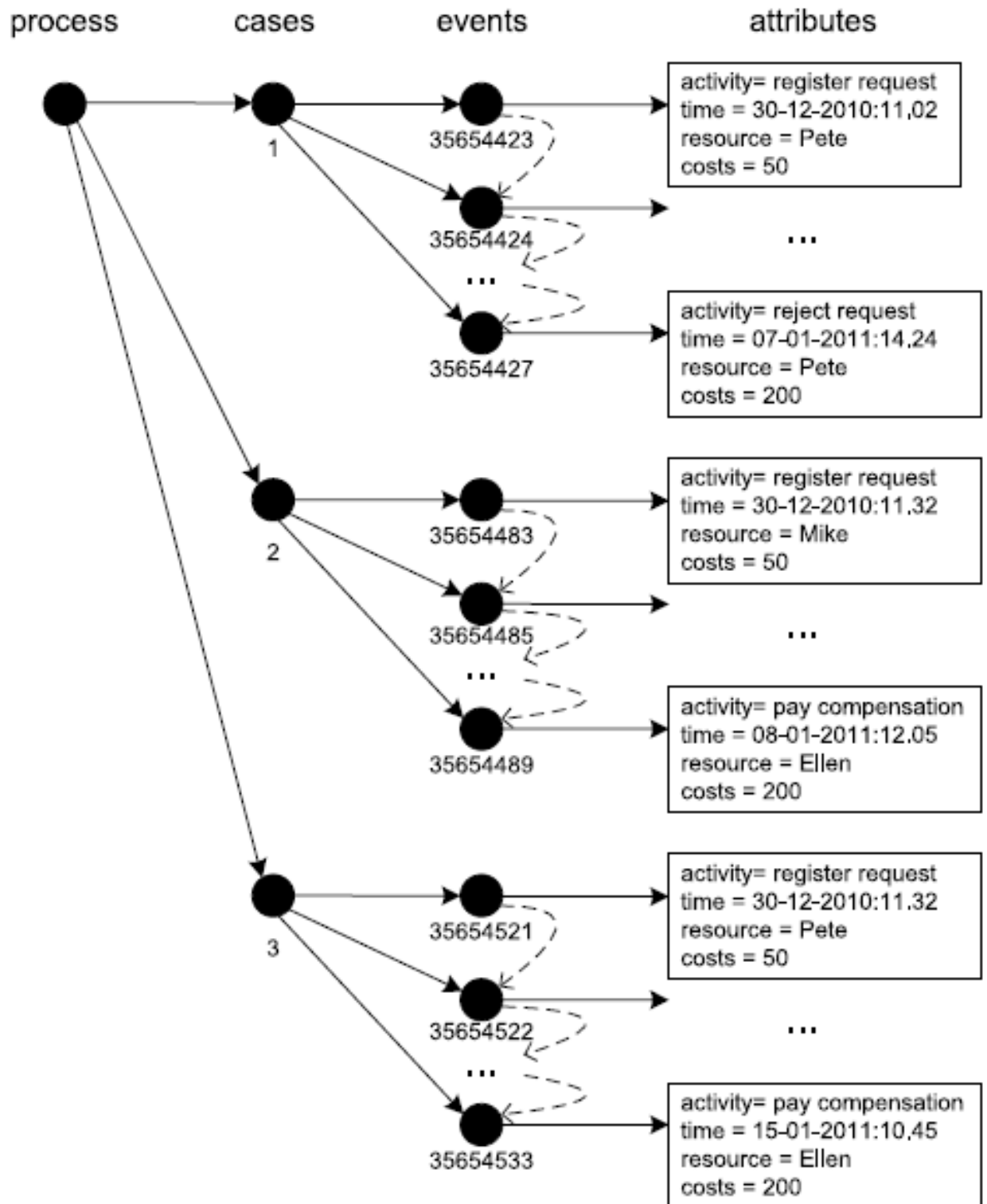
Figure 7 Principle structure of event logs (van der Aalst, 2011)**.**

In order to describe the requirements for process mining in construction a case is used. In this case the process of two prefab elements are described, from arrival on the construction site until installation. The log example as described in Figure 8 is the preferred result.

Process	Case ID (GUID)	Event ID	Attributes			
Prefab elements on construction site			Timestamp	Activity	Resource	...
	1TyTkZ9yv5EeHtNFxzmA9S	1	22-06-2015: 11.36	Arrived	Transporter	
	1TyTkZ9yv5EeHtNFxzmA9S	2	22-06-2015: 15.00	Stored	Contractor	
	1TyTkZ9yv5EeHtNFxzmA9S	3	24-06-2015: 10.00	Moved	Contractor	
	1TyTkZ9yv5EeHtNFxzmA9S	4	30-06-2015: 15.00	start installation	Sub-Contractor	
	1TyTkZ9yv5EeHtNFxzmA9S	5	30-06-2015: 16.00	end installation	Sub-Contractor	
	1JWxxtHnDEd8ND2YJ76EPx	1	22-06-2015: 12.00	Arrived	Transporter	
	1JWxxtHnDEd8ND2YJ76EPx	2	22-06-2015: 15.00	Stored	Contractor	
	1JWxxtHnDEd8ND2YJ76EPx	3	24-06-2015: 14.00	Moved	Contractor	
	1JWxxtHnDEd8ND2YJ76EPx	4	30-06-2015: 15.20	start installation	Sub-Contractor	
	1JWxxtHnDEd8ND2YJ76EPx	5	30-06-2015: 17.00	end installation	Sub-Contractor	
	1JWxxtHnDEd8ND2YJ76EPx	6	01-07-2015: 09.00	Defect recognized	Contractor	

Figure 8 Example of an event log generated from an as-planned BIM

The bare minimum to enable process mining requires a Case ID and an Activity log:

1. Case ID: A case identifier, also called process instance ID, is necessary to distinguish different executions of the same process. What precisely the case ID is depends on the domain of the process. For example in a hospital this would be the patient ID.
   In the case of construction process mining with BIM the case ID would be the Global Unique Identifier (GUID) of the building element when one is interested in this specific process. When someone is interested in the generic process of a project he could use the unique building ID as a case ID. In our case the GUIDs of the prefab elements are:
   a. *1TyTkZ9yv5EeHtNFxzmA9S*
   b. *1JWxxtHnDEd8ND2YJ76EPx*

2. Activity: There should be names for different process steps or status changes that were performed in the process. If you have only one entry (one row) for each process instance, then your data is not detailed enough.
   The data needs to be on the transactional level (you should have access to the history of each case) and should not be aggregated to the case level.
   In the case of construction process mining of a building element there should be a history of this element. In our case the activities are: 'Arrived at construction site', 'stored on construction site', 'Moved on construction site' start installation on construction site', 'end installation on construction site'. With one element something is wrong, so there is a 'defect recognized'.

To analyse performance related properties additional attributes are useful in the event log like:

3. Timestamp: At least one timestamp is needed to bring events in the right order. This time stamp is also needed to identify delays between activities and bottleneck identification.

> *In our case the timestamp is described with the following syntax: DD-MM-YYYY:*
> *HH.MM, where D = Day, M = Month, Y=year, H=hour, M=minute.*
> *This gives the following example: 22-06-2015: 11.36*

4. Resources: For example the person or the company who executed the activity

   In our case this could be the transporter, contractor, or sub-contractor for example.

5. … More attributes can be added…

3.2.1.2 Data formats

However the structure and information within the event logs may be the same, event logs can be described with different data formats. Detailed information about those formats can be read in (W.M.P. van der Aalst, 2011, Chapter 4), but the most occurring ones are named in this paragraph.

The Mining eXtensible Markup Language (MXML) standard was used for a long time and adopted by process mining tools. Using an XML-based syntax MXML stored event logs as described in the previous paragraph. This approach worked quite well in practice, however the ad-hoc extensions of the standard revealed shortcomings (van der Aalst, 2011). This resulted in the development of the eXtensible Event Stream (XES) which the successor of MXML. The XES format is less restrictive and truly extensible. The format is supported by major process mining tools such as PrOM, Disco, XE-Same and Open XES. At www.xes-standard.org detailed information about the standard can be gathered.

However XES seems to be the standard adopted in the process mining field this file format can be difficult to understand for freshman (such as the researcher of this study) in this research field. Therefore it is sometimes more understandable and practical to store event logs within excel sheets or CSV files. Within this study this approach seemed suitable since it is focused on workflows in small scales instead of in depth analytics on large scale. When larger scale analytics will be done using more comprehensive file formats may be necessary.

3.2.1.3 Conclusion event logs

In conclusion event logs are a suitable data format for improving construction processes with help of process mining analytics. The log should be structured as described in this chapter. While analysing specific processes of construction elements throughout projects and companies the case ID could be the unique identifier of the element (GUID). When one is more interested in the general process of a project the unique building identifier (buildingGUID) could be used as the case ID.

3.2.2 Bottlenecks

The presence of timestamps in event logs enables the discovery of bottlenecks, the analysis of service levels, the monitoring of resources utilization and the predication of remaining processing times of running cases (van der Aalst, 2011). Some timestamps may differ in pattern. It is critical that within one log the timestamps have the same patterns. By use of the timestamps process mining algorithms can compute statistics such as mean, standard deviation, minimum and maximum durations. In addition waiting times between events can

be derived easily. Various kinds of performance related information can be extracted from process models including time such as:

- *Visualization of waiting and service times.* Activities with a large variation in time could be highlighted in the process model.

- *Bottleneck detection and analysis.* The durations attached to each place can be used to find and analyse bottlenecks. Traces where most time is spend can be highlighted. Cases which spend a long time on one place can be further investigated. Delayed cases can be find and analysed separately in order to find the cause for the delay.

- *Flow time and SLA analysis.* Flow computations can be done with use of the timestamps. Average flow times between activities can be shown for example. An application in Service Level Agreement (SLA) is found in flowtimes. It could be that there are contractual obligations to solve problems between a specific time. This could easily be computed.

- *Analysis of frequencies and utilization.* Times and frequencies can be used to show routing probabilities in the process model.

In conclusion timestamps are essential for finding bottlenecks in processes. Process mining tools use the timestamps to calculate statistics. Bottlenecks have to be defined per project. The process models can be searched and filtered for specific durations which results in finding repeatable events with to long durations. This answers research question three[10].

3.2.3 Planning deviation

Most process mining applications are done afterwards. After processes are being executed they are analysed to see how they can be improved. However more techniques are being developed to enable real time process mining. An example is the detection of non-conformance at the moment the deviation actually takes place. This seems applicable to the construction industry, where often questions occur like '*are we on schedule?'* Conformance checking, or conformance analysis, aims to detect inconsistencies between a process model and its corresponding execution log (Rozinat & van der Aalst, 2008). Questions like '*Do the model/or planning and the event log conform to each other?'* may be answered. Conformance checking measures the alignment between model and reality. Algorithms are developed in order to quantify and diagnose deviations. A sample of a conformance checking algorithm is given in (Rozinat & van der Aalst, 2008). They developed a tool called Conformance Checker which is part of the ProM framework[11]. The tools uses an existing process model represented in a Petri net. As follows executed event logs are mapped onto the Petri net in order to visualize deviations.

Conformance checking assumes the presence of a prescriptive process model (Rozinat & van der Aalst, 2008). Since construction projects use several planning methods (see chapter 3.1.3) those can be used as a starting point. Using the schedule which is made within construction projects, use measurement tools, and check on conformance with process mining algorithms such as presented in (Rozinat & van der Aalst, 2008) enables to measure planning deviations.

---

[10] "How can process mining methods be used for analysing automatically harvested data to identify bottlenecks in construction projects?"

[11] http://www.promtools.org/

Thereby an theoretical answer towards research question four[12] is found. The applicability is tested in practice, moreover in chapter 8 Case study construction site process mining.

3.2.4 Organizational mining

Mining from an organizational perspective is enabled when a *resource* attribute is present in event logs (Song & van der Aalst, 2008; van der Aalst, Reijers, & Song, 2005). A resource can for example be an actor, company, role, department, team or machine performing the event. As described by van der Aalst (2011) using this information enables techniques to learn more about people, machines, organizational structures, work distribution and work patterns. Three application groups organizational mining are given by van der Aalst (2011): Social network analysis, Organizational structure analysis and Resource behaviour analysis.

Social network

The nodes in a social network correspond to the aggregate resource. Using Figure 9 as a reference the arcs indicate a relationship between such entities. Both arcs and nodes may have weights which indicate the *'importance'* of the relationship. (Importance can be interpret differently depending per process or network.) Several examples of social networks are elaborated by van der Aalst et al. (2005). With such analytics it is possible to identify groups of entities which are strongly interconnected. Also finding 'the spider in the web' meaning important actors within an organization is thinkable. Other algorithms indicate the degree of connectedness between actors within an organization or process. In addition it is possible to count the number of times work is handed from one actor to another.
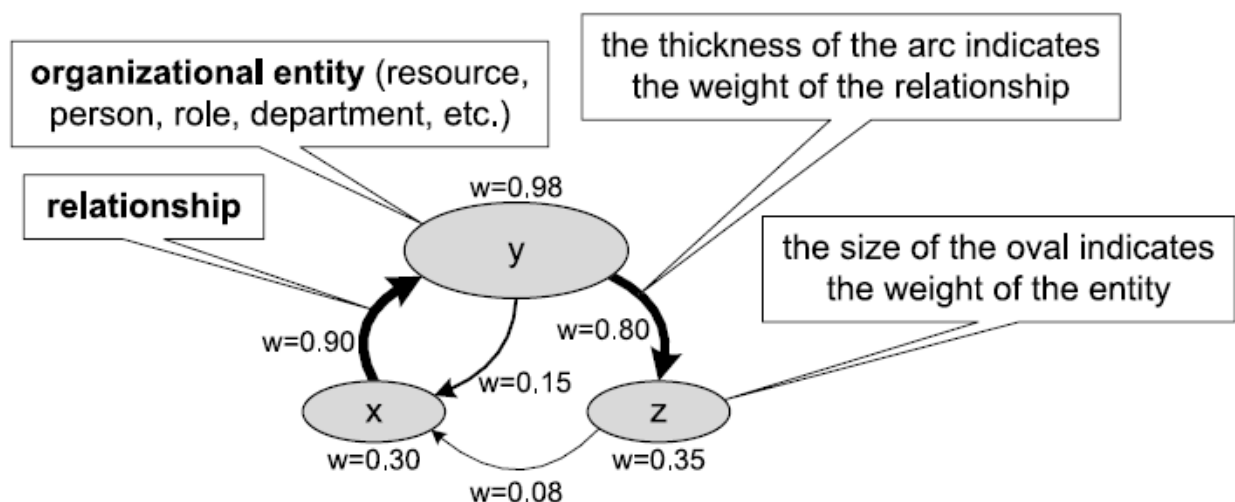

Figure 9 A social network consists of nodes representing organizational entities and arcs representing relationships. Both nodes and arcs can have weights indicated by "w = . . ." and the size of the shape. (van der Aalst, 2011)

Organizational structures

Resources can be characterized by a profile, using this clustering techniques can be used to discover similar resources. After clustering the resources into groups, these can be aligned to activities. Based on those activities roles can be discovered. Using Figure 10, three roles can be determined in this example namely Expert, Manager and Assistant. If one is interested in

---

[12] "How can process mining methods be used for identifying planning deviations in construction projects?"

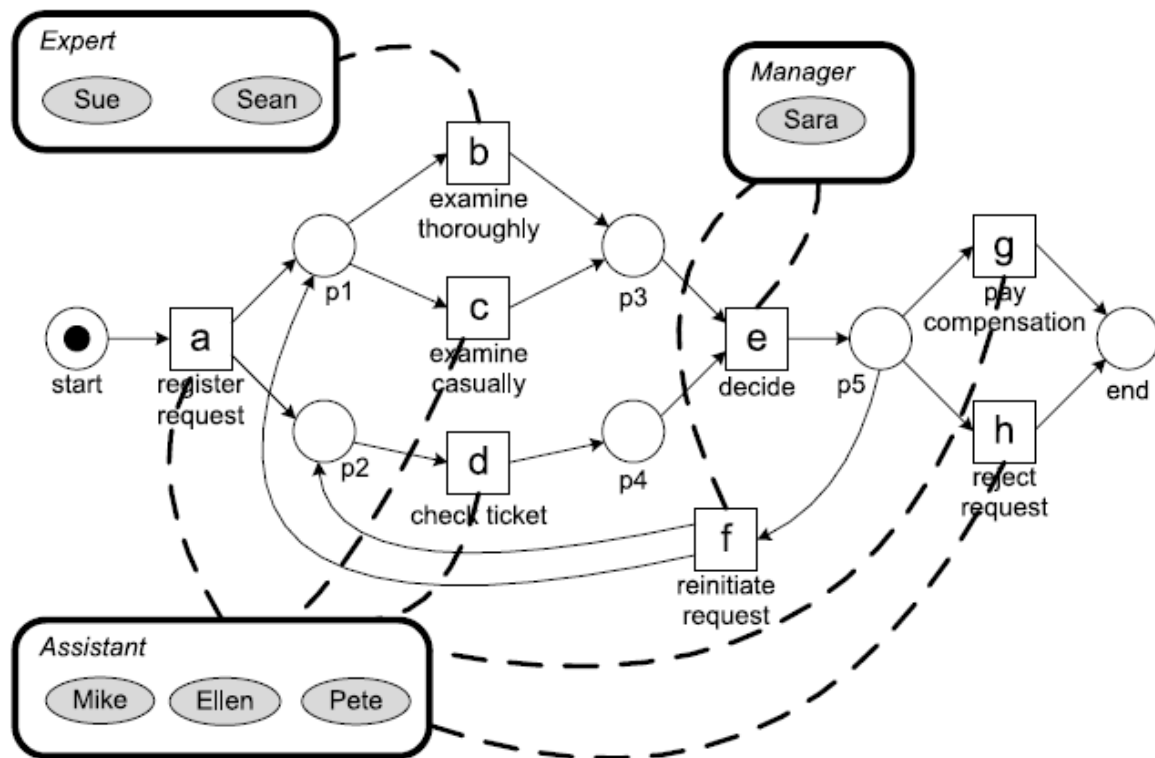details and more information about organizational mining, this can be found in (Song & van der Aalst, 2008).

Figure 10 Concept of an organizational model discovered from event logs

Resource behaviour

Since events in the event log contain resources performance measures extracted from the log can be projected onto the models (van der Aalst, 2011). Frequencies can be visualised onto activities, entities or resources. Also the utilization and response time of resources can be shown.

Organizations may prefer analytics on an aggregate level instead of analysing individuals. Therefor a log can be anonymized, In addition when privacy issues play a role this is preferable. More examples and elaboration of resource behaviour based process mining can be read in (van der Aalst, 2010).

3.2.5 Experimental case studies

Cleary the research field of process mining is relatively new and not been applied often in the construction industry. Process mining uses databases of existing IT systems to gain major insights in processes. Currently the construction industry adapts IT systems within all phases more and more. Which gives the possibility to use process mining techniques to gain insight in the processes of construction projects. As part of this study two experimental case studies are conducted in order to figure out the potential of process mining in the construction industry. Firstly a study is done within the design phase of a construction project. Thereafter a study is done within the operational phase of a building. In this chapter the abstracts and conclusions about both studies are elaborated. When one wants more in depth information about those studies. The paper about construction design process mining can be found in the

appendix in chapter 12. And the paper of process mining with facility management data can be found in the appendix in chapter 13.

3.2.5.1 Construction design process mining

Techniques that improved performance in other industries are more often adopted in the construction industry such as Systems Engineering (SE). Interviews with SE experts conducted within this study have indicated that SE IT systems supports with managing their project and helps to prove that the clients specifications are realized. However it is not known if those systems support an efficient process. Practioners indicated that a lot of people are involved in such projects who all work in parts of the IT systems but nobody has a clear overview of the total process. This study explores the possibilities of discovering parts of the design process with process mining techniques. By use of a case study at a large civil project the potential of process mining within the design phase of construction projects is discovered. It took some time to prepare the data but valuable analytics could be done with process mining techniques. Real bottlenecks are found, process variants are discovered, social networks are exposed and improvements in the design process can be made. Due to this experimental study the process engineers realized that the organization doesn't control the field of information- and data modelling enough. It is concluded that it is valuable to use process mining to give continuously feedback loops to the project managers at contractors. Process mining gives unique visualisations which enables refreshing insights. In addition process mining analytics did realize that some IT systems are used to store information, but those systems don't automatically facilitate an efficient process.

3.2.5.2 Process mining with facility management data

This experimental study explored if facility management data is suitable for analysing processes around building elements with process mining techniques. By use of a case study at a hospital the potential of combining process mining with maintenance data is discovered. It can be concluded that with some data transformation maintenance data is suitable for process mining. Moreover the facility managers were surprised by the visualisation techniques and they gained clear insight in the error handling process. As a result they discovered problematic building elements and odd processes. In addition as a result of this analysis the facility managers were surprised about the amount of money which was spend to short (and unrealistic) jobs and they are going to monitor those errors for next months to figure out how this maybe can save them money. A notifiable quote was mentioned by one of the facility managers, she said "We can probably save more money with investing in data analytics, than with firing our own people".

The facility managers where definitely interested in using those kind of analytics more in the future. In order to enable process mining on a larger scale it would be useful that facility management systems adapt an 'export to event log' function. Also naming of elements should be consistent in order to make analytics more easy and reliable.

This study just explored the topic of facility management based process mining and has proven some useful applications. More (case-) studies should be elaborated to indicate the potential of this topic. In addition it would be useful to study the potential of integrating different data sources of other phases, for example the design -or construction- phase, in building elements lifecycle in order to gain insight in the process on a longer time span.

3.3 Conclusion construction process mining

There is a clear gap in the literature regarding construction monitoring spanning across the project lifecycle, especially related to construction site related process information. However nowadays IT systems are more and more adopted in all phases of construction projects.

At the same time the research field process mining as part of the data science knowledge field is gaining more interest. In addition algorithms are developed to enable fact based problem discovery within processes in IT systems. As part of this research experimental studies are done regarding process mining in the plan- and operate phase of buildings. As a result of those experimental studies process mining within the construction industry is assessed potentially valuable. More specific it is indicated that it can be useful to study the possibility of the integration of event log databases from different phases of construction projects. Meaning that event logs from a design phase, construction phase and operate phase could be merged in order to give unique insight in the lifecycle of a construction object. However, in contradiction to the plan and operate phase, event logs from the construction phase are difficult to find or generate. In addition clearly little research is done regarding data science using BIM with process information. This study fills the gab by exploring the potential of BIM based process mining within the construction phase. More specific it provides workflows to autonomously gain process related data on construction sites, create event logs, identify bottlenecks and planning deviation, and finally gives options of reusing data in order to learn.

4 Proposed workflow BIM based process mining

In order to reach the target of this study[13], solve problem definition two[14], answer research question six[15] and to fill the literature gab identified in the previous chapters, a workflow is suggested. The target of the workflow is to enable a continuous optimization of project schedules by eliminating bottlenecks and planning deviations and thereby shorten construction projects. More specific the workflow uses BIM tools to manage the execution phase of construction projects and reveal planning deviations. Subsequently process mining tools are used to identify bottlenecks over single and multiple projects. The proposed workflow is elaborated in this chapter. The workflow based on principles of the Plan-Do-Check-Act (PDCA), or Deming Cycle (Bushell, 1992) which is often used within LEAN working philosophies. The PDCA cycle enables a continuous improvement while eliminating waste in the value stream (Sayer & Williams, 2013). This cycle is supplemented within the workflow with specific technologies in every step. Using Figure 11 as a reference the workflow consists of four phases starting with (1) the *plan* phase, (2) the do and *capture* phase, (3) the check and *analyse* phase, and (4) the act and *reuse* phase. Every phase is described by the desired input and the desired output. Different tools can be suitable for generating the desired output within a phase. Therefore it is possible that what happens exactly within a phase can differ per project. This chapter does not elaborate in detail the technological possibilities within a phase. However when one wants for example technical options of planning or capturing these are given in the literature study. Although some of the technologies which can be used in phases are tested in practice with a case study (see chapter 8).
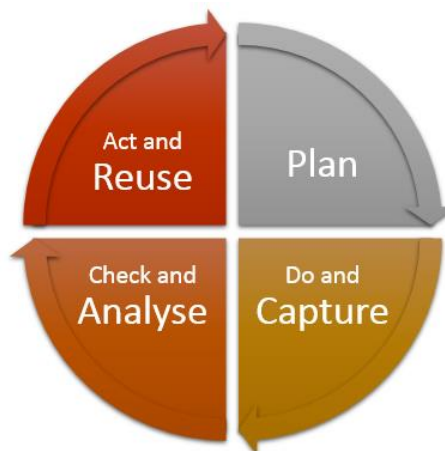

Figure 11 Proposed workflow inspired on PDCA cycle

4.1 Plan

In order to identify planning deviation as described in the target of this study a plan has to be made first. Within a planning phase domain experts define the steps necessary to achieve a

---

[13] "To realize a shorter construction time for the execution phase of construction projects."

[14] "Can bottlenecks and planning deviations (and their causes) be identified with BIM and process mining techniques (which are) employed in early construction process stages?"

[15] "What are the requirements for a toolbox which can be used to address the challenges posed in problem definition 2?"

desired result (Sayer & Williams, 2013). Traditional planning methods (as described in chapter 3.1.3.1) define processes and put them in a time perspective, which results in a schedule. BIM allows to connect process information to building information which creates more context and a more understandable schedule. Several tools are available to connect or make schedules with BIM. Since this workflow proposes a target to reuse all the information it is useful to use scheduling tools which adapt the open standard IFC in importing and exporting functionality. This model can be used in several tools and therefore be reused. The proposed output of the plan phase should be an as-planned IFC model. A project planner aims at creating a schedule which is as reliable possible. In order to realize this target some input is needed: (1) as-designed IFC model. (2) Schedule tasks based on specific ratios and indicators which are available in different formats. Some schedulers have experience and have ratios as part of their internalized knowledge and rules of thumbs. Some task ratios can be described in a ratio database. Independent of the form, some ratio input is used to determine time aspects of processes within schedules. In addition each process may have an (3) error sensitivity, or a risk identification. This describes the probability of failing for a particular task in the process. In conclusion the following input and output are desired within the plan phase:

Input:

- As-designed IFC

- Schedule ratios

- Risk identification

Output:

- As-planned IFC

4.2 Capture

Collecting data about the as-built process is done in the capture phase. After the plan phase the *do step* is defined in the PDCA circle. Carrying out the plan on the construction site (Sayer & Williams, 2013) . When the planned construction project is executed it is useful to capture what is happening during this process. Therefore the do step is supplemented with capturing techniques. Capturing construction projects can be done in several ways. As described in chapter 3.1.3.3 technology may be used to autonomously gain process related data in order to monitor the progress. However, in order to be suitable for analysing and reusing the data it is necessary to think about a data format which is suitable for this kind of purposes. As indicated in chapter 3.2 event logs can be used to store process related information, as well as building related information, and are perfectly suitable for analysing and reusing. Therefore the proposed workflow generates as-built event logs as output during the capture phase.

Input:

- As-planned IFC

Output:

- As-built event logs

4.3 Analyse

The third step in the PDCA cycle is called *Check*. Within this step the results of the do phase are examined. The proposed workflow aims on analysing the captured data. Since event logs

are generated during the capture phase those can be analysed in several existing process mining tools. Analysis can be done on several subjects as described in chapter 3.2. The output of the analysis can be in several manners. In the context of this thesis, the focus lies on analysing bottlenecks and planning deviation, therefore the output should be useful for project managers and schedulers.

Input:

- As-built event logs

Output:

- Document with analytics in bottlenecks and planning deviation

4.4 Reuse

The PDCA cycle ends with an *act* step. Which implements failures from the check phase in order to improve the next project. The proposed workflow orates for reusing the data from previous projects within new projects to avoid failures. So what is happened during the execution phase can be used as input for the next project. Based on this experience and knowledge, projects schedule ratios can be determined, automated planning consultation can be given and risk identification can be advised. This output can be used for planning new projects and avoiding failures which happened before.

Input:

- As-built event logs
- As-designed IFC

Output:

- Schedule ratios
- Risk identification

4.5 Conclusion and discussion

In order to reach the target of this study a workflow is proposed which enables a continuously learning loop for better construction schedules where failures are identified over multiple projects. Several technological tools have to be used in phases within this workflow. As identified in the literature study for some phases the requirements for input and output are met by existing tools. Using Figure 12 as a reference currently several tools exist to achieve the result of the plan phase, namely create as-planned IFC models by use of BIM planning tools. In addition monitoring technologies exist to capture as-built information. Event log analysing algorithms exists in current process mining tools. However the proposed workflow misses some tools in order to function. There is need for a tool which translates an as-planned or as-built IFC into event logs. Therefore a BIM based event log generator is made, the development of the tool is described in chapter 5. Also for reusing the as-built event logs in order to identify risk in new projects there is no tool accessible, therefore the Planning consult service is made (see chapter 6).
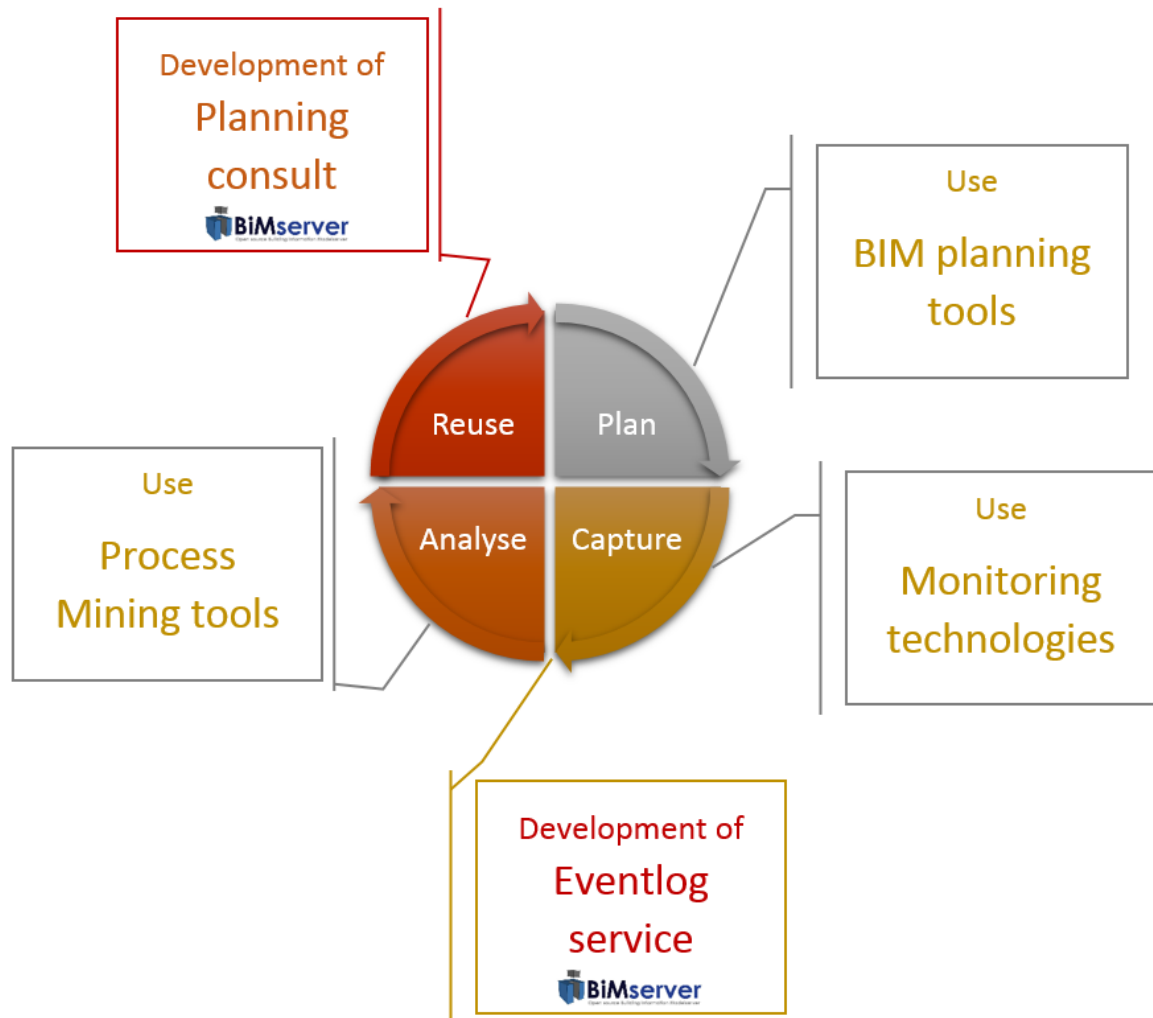
Figure 12 Need for tools which translate BIM files in event logs and enable them to for reusing purposes.

5 Development of BIM based event log generator

In order to find bottlenecks in the analyse phase, by use the information gathered in the plan and capture phase, event logs have to be available. Literature study revealed that event logs contain at least case ID's, event ID's, and a time stamp. Since as-planned IFC models can be made those models can be used because they have all required information in order to generate an event log. The information is in the IFC model, but it is not in the right data structure to enable process mining analytics. No tool currently exist which can generate event logs from an as-planned IFC. Therefore a BIM based event log generator is made. Called 'Eventlog service'. Using BIMserver[16] a new tool is created which has as input an as-planned IFC and gives an event log as output. Using GUID's as case ID, and task names as event ID. This event log can be used in event log based applications, like 'the planning consult' (see chapter 6) Disco[17] or MyInvenio[18]. The Eventlog service is schematic elaborated in Figure 13. This chapter elaborates first the desired output, thereafter the requirements of the input IFC model, and the proposed workflow of the Eventlog service within BIMserver is described. Subsequently followed by an output description. Lastly the tool is discussed. The tool is tested and validated in chapter 7. The tutorial of the tool can be seen on YouTube[19] as well.
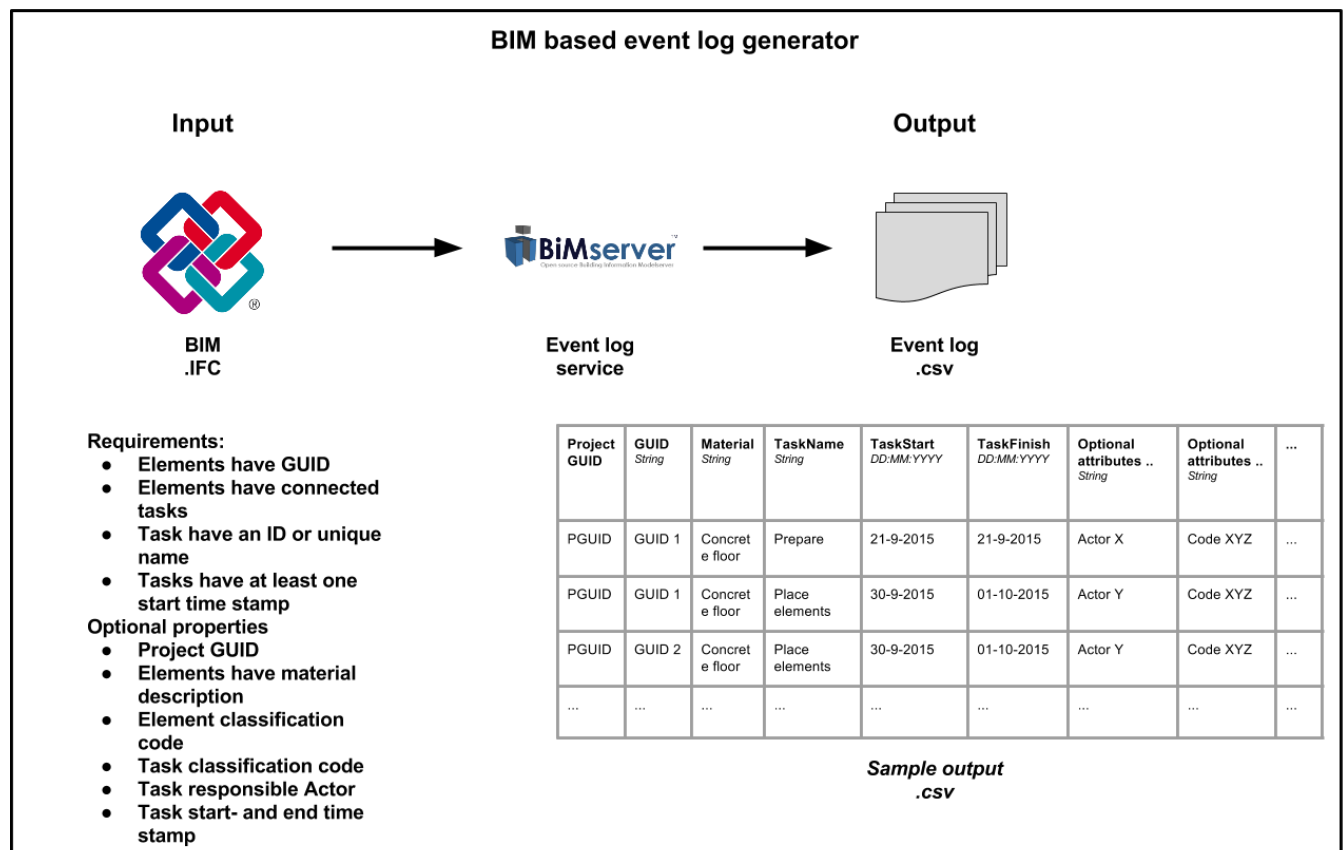

Figure 13 Schematic description of the Eventlog service

---

[16] http://bimserver.org/
[17] https://fluxicon.com/disco/
[18] https://www.my-invenio.com/
[19] BIMserver Eventlog tutorial: https://www.youtube.com/watch?v=LOJsHvGq-KE

5.1 Desired output

In order to describe the requirements for process mining in construction a case is used. In this case the process of two prefab elements is described, from arrival on the construction site until installation. The following log example is the preferred result:

Process	Case ID (GUID)	Event ID	Attributes			
Prefab elements on construction site			Timestamp	Activity	Resource	...
	1TyTkZ9yv5EeHtNFxzmA9S	1	22-06-2015: 11.36	Arrived	Transporter	
	1TyTkZ9yv5EeHtNFxzmA9S	2	22-06-2015: 15.00	Stored	Contractor	
	1TyTkZ9yv5EeHtNFxzmA9S	3	24-06-2015: 10.00	Moved	Contractor	
	1TyTkZ9yv5EeHtNFxzmA9S	4	30-06-2015: 15.00	start installation	Sub-Contractor	
	1TyTkZ9yv5EeHtNFxzmA9S	5	30-06-2015: 16.00	end installation	Sub-Contractor	
	1JWxxtHnDEd8ND2YJ76EPx	1	22-06-2015: 12.00	Arrived	Transporter	
	1JWxxtHnDEd8ND2YJ76EPx	2	22-06-2015: 15.00	Stored	Contractor	
	1JWxxtHnDEd8ND2YJ76EPx	3	24-06-2015: 14.00	Moved	Contractor	
	1JWxxtHnDEd8ND2YJ76EPx	4	30-06-2015: 15.20	start installation	Sub-Contractor	
	1JWxxtHnDEd8ND2YJ76EPx	5	30-06-2015: 17.00	end installation	Sub-Contractor	
	1JWxxtHnDEd8ND2YJ76EPx	6	01-07-2015: 09.00	Defect recognized	Contractor	

Figure 14 Example of an event log generated from an as-planned BIM

The bare minimum to enable process mining requires a Case ID and an Activity log:

1. Case ID: A case identifier, also called process instance ID, is necessary to distinguish different executions of the same process. What precisely the case ID is depends on the domain of the process. For example in a hospital this would be the patient ID.
   In the case of construction process mining with BIM the case ID would be the GUID of the building element. In this sample case the GUIDs of the prefab elements are:
   a. *1TyTkZ9yv5EeHtNFxzmA9S*
   b. *1JWxxtHnDEd8ND2YJ76EPx*

2. Activity: There should be names for different process steps or status changes that were performed in the process. If you have only one entry (one row) for each process instance, then your data is not detailed enough.
   The data needs to be on the transactional level (you should have access to the history of each case) and should not be aggregated to the case level.
   In the case of construction process mining of a building element there should be a history of this element. In our case the activities are: 'Arrived *at construction site',* 'stored *on construction site', '*Moved *on construction site'* start installation *on construction site', '*end installation *on construction site'. With one element something is wrong, so there is a '*defect recognized'.*

To analyse performance related properties additional attributes are useful in the event log like:
3. Timestamp: At least one timestamp is needed to bring events in the right order. This time stamp is also needed to identify delays between activities and bottleneck identification.
   In our case the timestamp is described with the following syntax: DD-MM-YYYY: HH.MM, where D = Day, M = Month, Y=year, H=hour, M=minute.
   This gives the following example: 22-06-2015: 11.36

4. Resources: For example the person or the company who executed the activity
   In our case this could be the transporter, contractor, or sub-contractor for example.

5. … More attributes can be added…

The desired output file should contain all the described requirements. Like described in chapter 3.2.1 the file format can be in several forms like .csv, .xes , MXML, as long as it keeps the structure as described in Figure 7 at page 40.

5.2   Input requirements

In order to be able to generate BIM based event logs with BIMserver the input file should be an IFC model, containing at least the following information:

- Elements have a GUID

  o *Can be used as Case ID.*

- Elements have assigned tasks or activities

- Tasks have an ID or name

  o *Can be used as Event ID.*

- Tasks have at least one start time stamp

Optional several properties can be useful while analysing within process mining tools:

- Project GUID

- Elements have a material description

- Elements have a classification code

- Tasks have a classification code

- Responsible actors are connected to tasks

- Start- and end time stamp

- … More attributes can be added …

In conclusion the file should have two types of properties (1) general properties which describe the building element, and (2) process related properties which describe as-planned and/or as-built processes. Those properties can be found within the IFC scheme as described in chapter 3.1.2.1.

5.3   Eventlog service in BIMserver

Since BIMserver is open source it enables one to make their own additions to the platform. Extra functionality can be added in several ways, to generate the Eventlog service an 'internal service'[20] is made. The algorithm will not be explained in detail but can be found in Appendix 15.

---

[20] Source code of the Eventlog service can be found at: https://github.com/opensourceBIM/BIMserver

The Eventlog service works as follows (see also video tutorial[21]):

1. One starts up his own server. [22]

2. One adds the 'Eventlog' service to a project.

3. One edits the Eventlog service by giving the parameters where 'Material' and 'Classification' are located in the IFC model.

Note: During the validation phase of the Eventlog service it is noticed that the parameter 'Material description' and 'Element classification code' can be on different places in the IFC model or can be named different each project. Therefore the Eventlog service can be adjusted per model so that the material and classification can be found in each project.

4. One checks in an IFC model which meets all the requirements.

Note: During this study all files where checked manually for all the requirements (as described in section 5.2). It is useful to have a file analyser tool which does this automatically. When not all the requirements are met the output event log will not be sound.

5. Automatically the IFC model is analysed by the Eventlog service and an event log (.csv file) is generated and added as external data to the server. This file can be downloaded.

5.4 Output

The output of the Eventlog service is a .csv file containing the general properties of the IFC model in the columns: BuildingGUID, GUID, IfcClass, Nl-sfb, Material, Resource. And containing process related properties in the columns: Task, TaskName, TaskStart, TaskFinish. As can be seen in Figure 15. Every row contains an event with a specific task as described in the process properties, which is connected to a building element which is described via the general properties.

Figure 15 Output Eventlog service containing general properties and process related properties

Note: the resource which is exported here is the name of the building element, this can be confusing since resources refer normally to supportive elements connected to the event.

Since the source code of the Eventlog service is open source one can edit the desired properties for specific mining analytics. For example this prototype uses only as-planned tasks,

---

[21] BIMserver Eventlog tutorial: https://www.youtube.com/watch?v=LOJsHvGq-KE
[22] More info on how to start your own BIMserver at: https://github.com/opensourceBIM/BIMserver/wiki

but if one also has as-built tasks, actors, or specific resources within the IFC model it can be added to the Eventlog service.

5.5   Discussion Eventlog service

The Eventlog service is made because there was no known tool which can make event logs out of BIM files. This tool can now make .csv files structured as event logs from one IFC project which meets the requirements. However process mining will be most interesting by analysing over multiple projects. Therefore it is useful to create one event log from multiple IFC models. This is not possible with the Eventlog service, the IFC models of single projects have to be converted to event logs, which can then be merged manually in one event log database. For this study this is suitable. When one wants to do process mining analytics over a large amount of IFC models it would be useful to add extra functionality to the Eventlog service to merge all event logs into one.

6 Development of automated Planning consult

During the literature study it was discovered that event logs can be analysed and reused in several ways (chapter 3.2). Event logs can for example be used to predict situations and identify process risks. Since construction companies are continuously struggling with planning deviation and have problems quantifying risks it is useful to study if an event log based approach can support with this struggling. Therefore a tool is developed, called 'Planning consult' which re-uses event logs to make an automated planning and identify risks in the process steps. The planning and risk identification can be used as input for new construction projects and should avoid failures and thereby shorten construction projects which is the target of this study.

The Planning consult is an additional service to the open source BIMserver. An schematic representation of the Planning consult is visualised in Figure 16 Schematic description of the Planning consult service. This chapter describes the development of the Planning consult by starting with the desired output, thereafter the input requirements are elaborated. Followed by a description of the proposed workflow of the tool within BIMserver. Subsequently followed by the output file. Finally a discussion about the tool is described. The tool is validated in chapter 7. The tutorial on how to use the planning consult can be seen on YouTube as well[23].
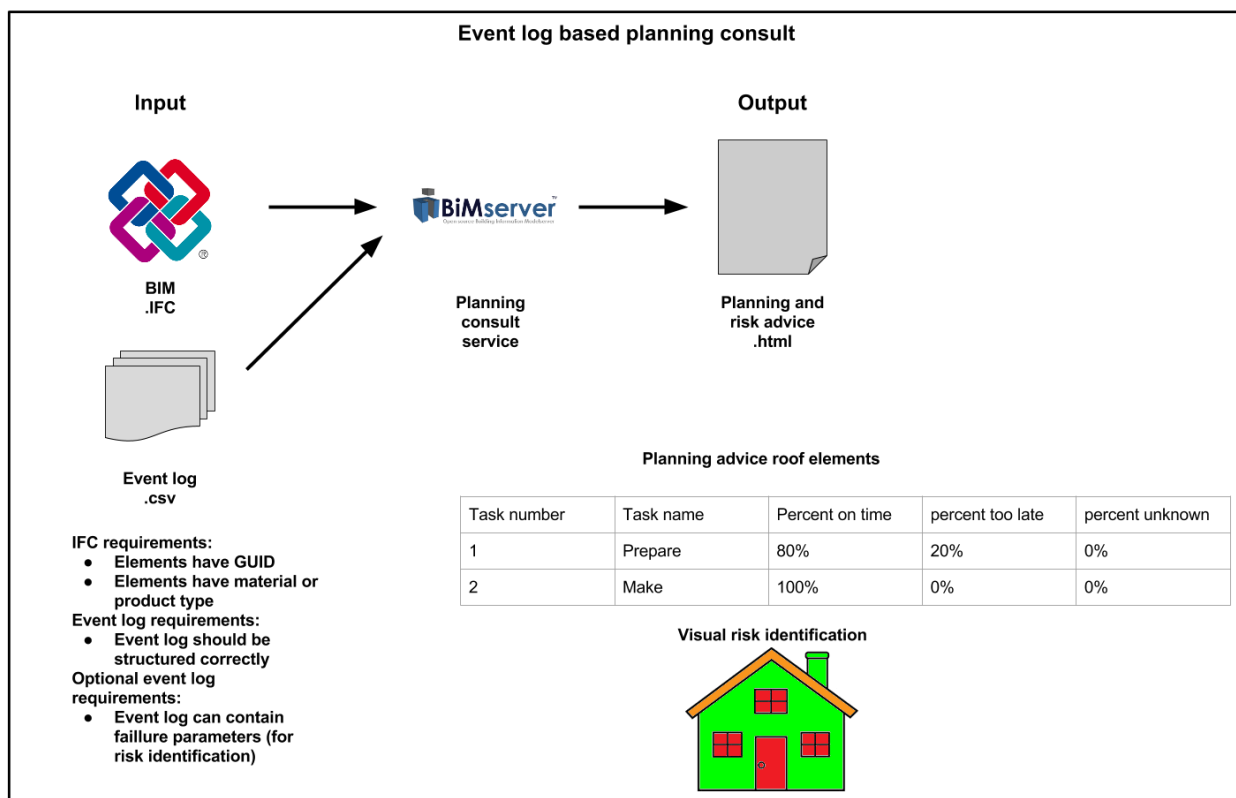

Figure 16 Schematic description of the Planning consult service

6.1 Desired output

Nowadays it is difficult for construction companies to identify process related risks within construction projects. Risk identification is done in all project phases. However risk

_____

[23] Tutorial BIMserver Planning consult: https://www.youtube.com/watch?v=nJTh_0Xmra0

identification by a contractor is usually done in early phases after a design phase. The design is usually analysed manually and risk identification is feeling based and results in a risk overview document. The Planning consult should identify risks per building element from a design BIM of a project. It is assumed that a BIM is available within all project phases from preparation until construction. So this BIM can be used to identify the risks regarding construction planning. The output of the Planning consult should advice which process steps per element have the highest risks. In addition the risks should be made visual. So the output should be twofold. One is a document file which describes the process steps of each element based on the event log. If failure parameters are available in the event log also a risk identification per step can be given in the document. Second output is a visual representation of the risk per element. Elements with a high failure risk are coloured red. Elements with a small failure risk are coloured orange and elements which didn't fail according to the event log are coloured green.

6.2 Input requirements

The Planning consult needs two input files. First an IFC model has to be uploaded which should be analysed by the Planning consult. This IFC model is developed in early phases and can be an architectural design for example. Second an event log is used to give planning advice and risk identification about the elements in the IFC model.

6.2.1 IFC requirements

In order to be able to analyse the construction model with BIMserver the input file should be an IFC model, containing at least the following information:

- Elements have a GUID.

- Elements have a material or product type.

Optional requirements:

- Elements have a classification code.

In conclusion the IFC model should have general parameters as described in section 5.4.

6.2.2 Event log requirements

The event log which can be used by the planning consult should be structured as the event log as generated by the Eventlog service (chapter 5.4).Which means it should contain general- and process related parameters of building elements. The file should be a .csv format. The exact same rows have to be available, containing: BuildingGUID,GUID,ifcClass,Nl-sfb,Material,Task,Resource,TaskName,TaskStart,TaskFinish,Faillure data . Each row should exists, however it may be empty. This will not crash the Planning consult but will give less detailed advice.

6.3 Planning consult in BIMserver

Since BIMserver is open source it enables one to make their own additions to the platform. Extra functionality can be added in several ways, to generate the Planning consult an 'internal service' is made.[24] The algorithm will not be explained in detail but can be found in Appendix 16.

---

[24] Source code of the Planning consult service can be found at: https://github.com/opensourceBIM/BIMserver

The Planning consult works as follows (the workflow tutorial can also be seen on YouTube[25].):

1. One starts his own server.

2. One adds the Planning Consult HTML service.

3. One adds the Planning Consult Visualization service.

4. One adjusts the Planning consult by giving the parameters where 'Material' and 'Classification' are located in the IFC model.

Note: During the validation phase of the Planning consult it is noticed that the parameter 'Material description' and 'Element classification code' can be on different places in the IFC model or can be named different each project. (Depends on used software or modelling template.) Therefore the Planning consult can be adjusted per file so that the material and classification can be found in each project.

5. One uploads an IFC model which meets all the requirements.

Note: During this study all files where checked manually for all the requirements. It is useful to have a file analyser tool which does this automatically. Keep in mind that the output of the Planning consult will not be sound when not all the requirements are met.

6. Automatically the IFC model is analysed by the Planning consult service and a .html file is generated and added as external data. This file can be downloaded. Also a risk visualisation option is added within the 3D view.

6.4 Output

As mentioned the output is twofold, a Planning consult document and a risk visualisation. The document is a .html file containing the following information:

- List of elements which are recognized in the IFC and in the event log containing:
  - Amount of elements found in IFC.
  - Amount of elements found in event log.
  - Process variants with tasks connected to the elements.
  - Risk identification for each task by telling the percentage:
    - Finished task on time.
    - Finished task too late.
    - Finished task unknown.
- List of elements which are recognized in the IFC but are not recognized in the event log.

The risk visualisation is a 3D view of the IFC with elements which are coloured as follows:

- Elements final task finished on time in more than 99%: Green (No risk)

- Elements final task finished on time in 50 to 99%: Orange (Low risk)

- Elements final task finished on time in less than 50%: Red (High risk)

---

[25] https://www.dropbox.com/s/b1x982yx5oyzrn6/Planning%20consult%20tutorial.mp4?dl=0

6.5   Discussion Planning consult

The Planning consult will give an advice on the steps to take for building elements. However the advice can only be as detailed as the event log used as input. So if a high detailed event log comes in, a more detailed advice can be given. Moreover the Planning consult gives advice on consecutive tasks on an element level. It is useful to study the possibility of giving planning advice on a project level, in order to consult the sequence of building elements. A new tool can be built or an extension on the Planning consult can be made.

7   Validation and testing of the developed tools

It is recognized that validation of the developed tools is important. Since the tools are prototypically made it is not achievable to officially validate the tools. When value of the tools is recognized in practice they can be further developed into everyday applications and then be validated scientifically. However to present the developed tools and to test the proposed workflow a fictive case study is done. The following case is proposed:

An imaginary small construction company is specialized in making houses. In 2015 they built four houses and in the coming period they will make one more. The four projects from 2015 are planned and monitored. All projects where designed in BIM and an IFC model was available. The IFC model was used to plan their project. So for each project an as-planned IFC is available. The as-built data is captured and stored in excel sheets. The company wonders if they can use their data in order to analyse the process and identify bottlenecks and risks. In addition they want to have a risk analysis of the project which will be made in the coming period.

This research will answer the questions of the construction company by use of the following workflow:

1. Convert as-planned IFC of all built projects in event logs with the Eventlog service.
2. Merge as-built data with event logs.
3. Merge event logs within one event log database.
4. Analyse event log database for bottlenecks and risks.
5. Use Planning consult service to do risk analyses for the new project

The company has built four projects which were identical in design (see Figure 17), but differ in schedule. So there are four variants of the project. This is elaborated shortly.

The design contains the following elements:

- 4 Roof elements

- 4 Floor elements (2 different types)

- 4 Wall elements
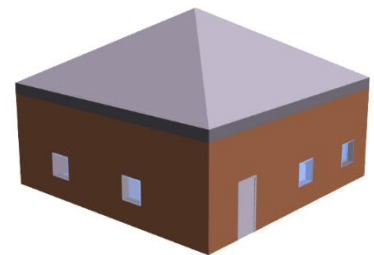
- 1 Door element

- 9 Window elements

Figure 17 Sample model used in test case

Project 1
Project 1 has the schedule as indicated in Figure 18. For each element a preparation task and an make task is assumed. During the project it was captured that the windows where delayed.

Note: This schedule is simplified and not realistic and only meant to test the event log generator.

ID	Name	Duration	Start	Finish
ST00020	Prepare flooring	1d	14:00 22-12-2015	14:00 23-12-2015
ST00030	Lay floors	2d	14:00 23-12-2015	14:00 25-12-2015
ST00060	Prepare windows	1d	14:00 25-12-2015	14:00 28-12-2015
ST00070	Make windows	2d	14:00 28-12-2015	14:00 30-12-2015
ST00080	Prepare doors	1d	14:00 30-12-2015	14:00 31-12-2015
ST00090	Make doors	1d	14:00 31-12-2015	14:00 1-1-2016
ST00040	Prepare walls	1d	14:00 1-1-2016	14:00 4-1-2016
ST00050	Make walls	2d	14:00 4-1-2016	14:00 6-1-2016
ST00100	Prepare flooring	2d	14:00 6-1-2016	14:00 8-1-2016
ST00110	Lay floors	1d	14:00 8-1-2016	14:00 11-1-2016
ST00120	Prepare roof	2d	14:00 11-1-2016	14:00 13-1-2016
ST00130	Make roof	3d	14:00 13-1-2016	14:00 18-1-2016

Figure 18 Schedule sample project 1

Project 2
Project 2 has the same design but the schedule differs in sequence, the doors where first prepared and made thereafter the windows are prepared and made. During the execution everything was built according to schedule.

Project 3
As project 1 and 2 the design of project 3 is the same. The schedule differs in sequence from previous projects. First the doors and windows are prepared. Thereafter both types of elements are made. Within this project the doors where made too late.

Project 4
The design is the same as previous projects. The schedule is the same in sequence. However preparing the windows takes longer than in previous projects. This results in a longer time span of the project. The windows where delivered too late.

7.1 Validation of the Eventlog service

The Eventlog service is tested and validated with the as-planned IFC models from projects 1,2,3,4 from the fictive company. The as-planned IFC is generated with Synchro software.

Each element is connected to two tasks according to Figure 18.

Since the model contains 22 elements and each element has two connected tasks the total amount of events within this one project is 44. After uploading the as-planned IFC of project 1 to the BIMserver with Eventlog service added it is seen that the event log contains indeed 44 events (See Figure 19). This means that every connected task is exported and translated into a single event per building element. In addition the tasks start and task finish time are according to the planning. The object GUIDS and Building GUIDs are available in the event log. The material information is in there as well. Only the classification code of the elements are missing in the event log. However the model didn't contain any. This attribute is useful while analysing large models. Since this models are small this will not block the workflow. So all attributes which are necessary for process mining are available.

The event log is imported in two different process mining tools, first one is Disco, second one is MyInvenio. Both tools generate the expected process map from the .csv file. (See Figure 20 and Figure 21) It can be seen that the event log which is generated by the Eventlog service contains all the expected data and thereby is useful to analyse statistics and bottlenecks. In order to show the workflow and to answer the questions of the construction company the event logs of all projects have to be merged. This is done in the next chapter.

	Project GUID	GUID	ifcClass	NI-	Material	Task	Resource	TaskName	TaskStart	TaskFinish
2	2w_AZO4MP4mfMVfWoTXGi9	0D$8_5ZnTF3QuiIzF4VwJI	IfcSlab		Floor: Kanaalplaatvloer	ST00020	Floor:Kanaalplaatvloer:305021_(#161)	Prepare flooring	22-2-2016	23-2-2016
3	2w_AZO4MP4mfMVfWoTXGi9	0qE4Fx1zbFyBpddG9XUJ$o	IfcSlab		Floor: Kanaalplaatvloer	ST00020	Floor:Kanaalplaatvloer:305078_(#316)	Prepare flooring	22-2-2016	23-2-2016
4	2w_AZO4MP4mfMVfWoTXGi9	0D$8_5ZnTF3QuiIzF4VwJI	IfcSlab		Floor: Kanaalplaatvloer	ST00020	Floor:Kanaalplaatvloer:305021_(#161)	Prepare flooring	22-2-2016	23-2-2016
5	2w_AZO4MP4mfMVfWoTXGi9	0qE4Fx1zbFyBpddG9XUJ$o	IfcSlab		Floor: Kanaalplaatvloer	ST00020	Floor:Kanaalplaatvloer:305078_(#316)	Prepare flooring	22-2-2016	23-2-2016
6	2w_AZO4MP4mfMVfWoTXGi9	0D$8_5ZnTF3QuiIzF4VwJI	IfcSlab		Floor: Kanaalplaatvloer	ST00030	Floor:Kanaalplaatvloer:305021_(#161)	Lay floors	23-2-2016	25-2-2016
7	2w_AZO4MP4mfMVfWoTXGi9	0qE4Fx1zbFyBpddG9XUJ$o	IfcSlab		Floor: Kanaalplaatvloer	ST00030	Floor:Kanaalplaatvloer:305078_(#316)	Lay floors	23-2-2016	25-2-2016
8	2w_AZO4MP4mfMVfWoTXGi9	0D$8_5ZnTF3QuiIzF4VwJI	IfcSlab		Floor: Kanaalplaatvloer	ST00030	Floor:Kanaalplaatvloer:305021_(#161)	Lay floors	23-2-2016	25-2-2016
9	2w_AZO4MP4mfMVfWoTXGi9	0qE4Fx1zbFyBpddG9XUJ$o	IfcSlab		Floor: Kanaalplaatvloer	ST00030	Floor:Kanaalplaatvloer:305078_(#316)	Lay floors	23-2-2016	25-2-2016
10	2w_AZO4MP4mfMVfWoTXGi9	39RDxZvOfBUuaZaaoBuyR7	IfcWindow		M_Fixed: Stelkozijn	ST00060	M_Fixed:Stelkozijn:306749_(#2026)	Prepare windows	25-2-2016	28-2-2016
11	2w_AZO4MP4mfMVfWoTXGi9	1$a1JbSNj04POHGM3JLkOP	IfcWindow		M_Fixed: Stelkozijn	ST00060	M_Fixed:Stelkozijn:306475_(#1731)	Prepare windows	25-2-2016	28-2-2016
12	2w_AZO4MP4mfMVfWoTXGi9	2uqVQT3mXCZh0sDm3qLEqe	IfcWindow		M_Fixed: Stelkozijn	ST00060	M_Fixed:Stelkozijn:306398_(#1567)	Prepare windows	25-2-2016	28-2-2016
13	2w_AZO4MP4mfMVfWoTXGi9	36AkN3VHjF6PGoSIXI2cRb	IfcWindow		M_Fixed: Stelkozijn	ST00060	M_Fixed:Stelkozijn:306547_(#1790)	Prepare windows	25-2-2016	28-2-2016
14	2w_AZO4MP4mfMVfWoTXGi9	3wUkOmcAr7LAfanRBD$EnJ	IfcWindow		M_Fixed: Stelkozijn	ST00060	M_Fixed:Stelkozijn:306705_(#1967)	Prepare windows	25-2-2016	28-2-2016
15	2w_AZO4MP4mfMVfWoTXGi9	0P7y2qKjv5oxJ1V4ERi_xD	IfcWindow		M_Fixed: Stelkozijn	ST00060	M_Fixed:Stelkozijn:306663_(#1908)	Prepare windows	25-2-2016	28-2-2016
16	2w_AZO4MP4mfMVfWoTXGi9	2LTPBqLRP6BgkymeOMgvB9	IfcWindow		M_Fixed: Stelkozijn	ST00060	M_Fixed:Stelkozijn:306609_(#1849)	Prepare windows	25-2-2016	28-2-2016
17	2w_AZO4MP4mfMVfWoTXGi9	1aG8Gq7Tv3ORqbB2Hudddw	IfcWindow		M_Fixed: Stelkozijn	ST00060	M_Fixed:Stelkozijn:306827_(#2085)	Prepare windows	25-2-2016	28-2-2016
18	2w_AZO4MP4mfMVfWoTXGi9	1nvE_1Qt58OhPjFFsyHVEy	IfcWindow		M_Fixed: Stelkozijn	ST00060	M_Fixed:Stelkozijn:306437_(#1671)	Prepare windows	25-2-2016	28-2-2016
19	2w_AZO4MP4mfMVfWoTXGi9	39RDxZvOfBUuaZaaoBuyR7	IfcWindow		M_Fixed: Stelkozijn	ST00060	M_Fixed:Stelkozijn:306749_(#2026)	Prepare windows	25-2-2016	28-2-2016
20	2w_AZO4MP4mfMVfWoTXGi9	1$a1JbSNj04POHGM3JLkOP	IfcWindow		M_Fixed: Stelkozijn	ST00060	M_Fixed:Stelkozijn:306475_(#1731)	Prepare windows	25-2-2016	28-2-2016
21	2w_AZO4MP4mfMVfWoTXGi9	2uqVQT3mXCZh0sDm3qLEqe	IfcWindow		M_Fixed: Stelkozijn	ST00060	M_Fixed:Stelkozijn:306398_(#1567)	Prepare windows	25-2-2016	28-2-2016
22	2w_AZO4MP4mfMVfWoTXGi9	36AkN3VHjF6PGoSIXI2cRb	IfcWindow		M_Fixed: Stelkozijn	ST00060	M_Fixed:Stelkozijn:306547_(#1790)	Prepare windows	25-2-2016	28-2-2016
23	2w_AZO4MP4mfMVfWoTXGi9	3wUkOmcAr7LAfanRBD$EnJ	IfcWindow		M_Fixed: Stelkozijn	ST00060	M_Fixed:Stelkozijn:306705_(#1967)	Prepare windows	25-2-2016	28-2-2016
24	2w_AZO4MP4mfMVfWoTXGi9	0P7y2qKjv5oxJ1V4ERi_xD	IfcWindow		M_Fixed: Stelkozijn	ST00060	M_Fixed:Stelkozijn:306663_(#1908)	Prepare windows	25-2-2016	28-2-2016
25	2w_AZO4MP4mfMVfWoTXGi9	2LTPBqLRP6BgkymeOMgvB9	IfcWindow		M_Fixed: Stelkozijn	ST00060	M_Fixed:Stelkozijn:306609_(#1849)	Prepare windows	25-2-2016	28-2-2016
26	2w_AZO4MP4mfMVfWoTXGi9	1aG8Gq7Tv3ORqbB2Hudddw	IfcWindow		M_Fixed: Stelkozijn	ST00060	M_Fixed:Stelkozijn:306827_(#2085)	Prepare windows	25-2-2016	28-2-2016
27	2w_AZO4MP4mfMVfWoTXGi9	1nvE_1Qt58OhPjFFsyHVEy	IfcWindow		M_Fixed: Stelkozijn	ST00060	M_Fixed:Stelkozijn:306437_(#1671)	Prepare windows	25-2-2016	28-2-2016

Sampleproject eventlog incl pro

Figure 19 Eventlog generated from as-planned IFC from sample project 1. (Possibly not every detail of this picture can be read clearly. The event log can be accessed digitally in the dataset, see appendix 14.)
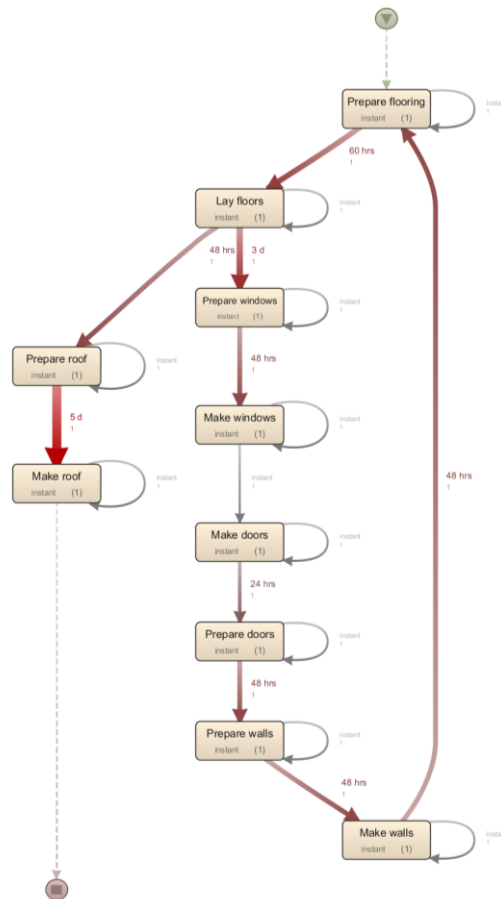
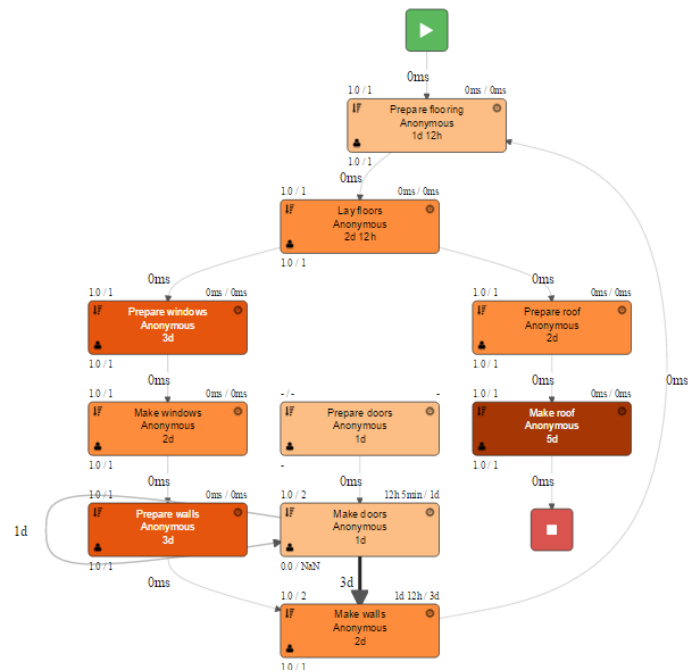**Figure 20 Process map generated with Disco**

Figure 21 Process map generated with MyInvenio

7.1.1   Multiple event log projects combined

In order to identify bottlenecks for the construction company and to show the usage of process mining for bottleneck analytics the event logs of all four projects are merged into one event log (a .csv file). This file contains 176 events including all failure data which is captured by the construction company. The file is imported in process mining tool Disco. This results in a process map on how the projects are made and a variant analysis. The process map indicated that a straight process is undertaken in parts of the process (see Figure 23). But within the windows and door phases event's sequences are changing. As expected there are three different process variants (see Figure 22). Where variant 1 is describing project 1 and 4. Variant 2 is project 2 and variant 3 is project 3. Three projects contained 27 days, one project's duration was 29 days (Reason is the delayed windows in project 4). The bottlenecks of the process are also within making the doors and windows. The longest durations take part in there, as well in making the roof. This can be seen in Figure 23, the darkest red blocks and arrows indicate the most time consuming parts of the processes.

Figure 22 Discovered process variants
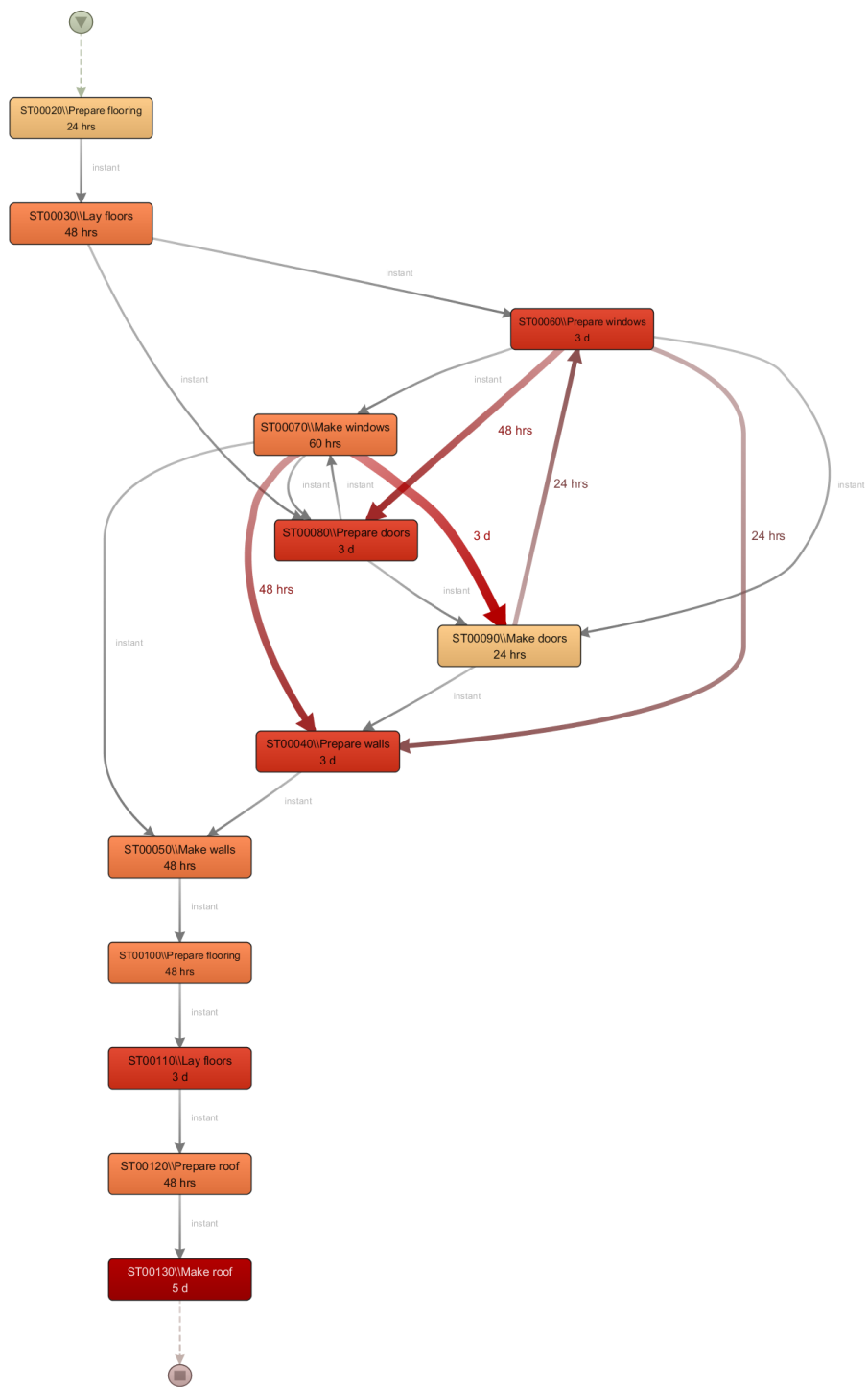
Figure 23 Process map of the combined event log of four projects created with Disco.

7.2   Validation of the Planning consult

The Planning consult is tested and presented with the projects from the fictive construction company. The company is willing to make a new project and wonders which elements have highest risks in process steps and how those elements have to be planned. As described in chapter 6 the Planning consult needs two types of input, an event log and an IFC model. The event log which is used as input in this case is the combined event log from the four projects of the company. This event log can be seen as the process database of the company and contains also failure data. The IFC which is used as input is a new sample project (see Figure 24). This project contains:

- 4 Wall elements, Type: 'Baksteen'.

- 1 Door element, Type: M_Single-Flush: 0915 x 2134 mm.

- 1 Roof element, Type: Basic Roof: Generic – 400mm.

- 1 Window element, Type: 'Stelkozijn'.

- 4 Floor elements, 2 Types: 'Kanaalplaat', 2 Types: 'Breedplaat'.
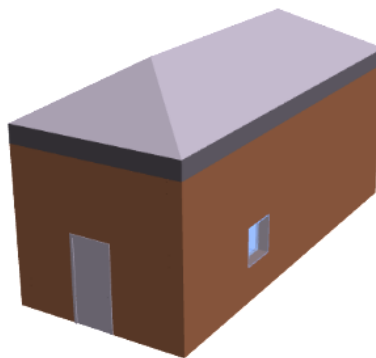
Figure 24 Sample project 5

The output of the Planning consult is visualized in Figure 25 and Figures Figure 26, Figure 27 Figure 28. As can be seen the wall and floor elements are coloured green (low risk), and the window and door elements are coloured orange(middle risk). This is as expected because the data which was captured by the construction company indicated failures by window and door elements. The task steps which are generated by the planning consult are identical to the schedules. As a result this part is presented.
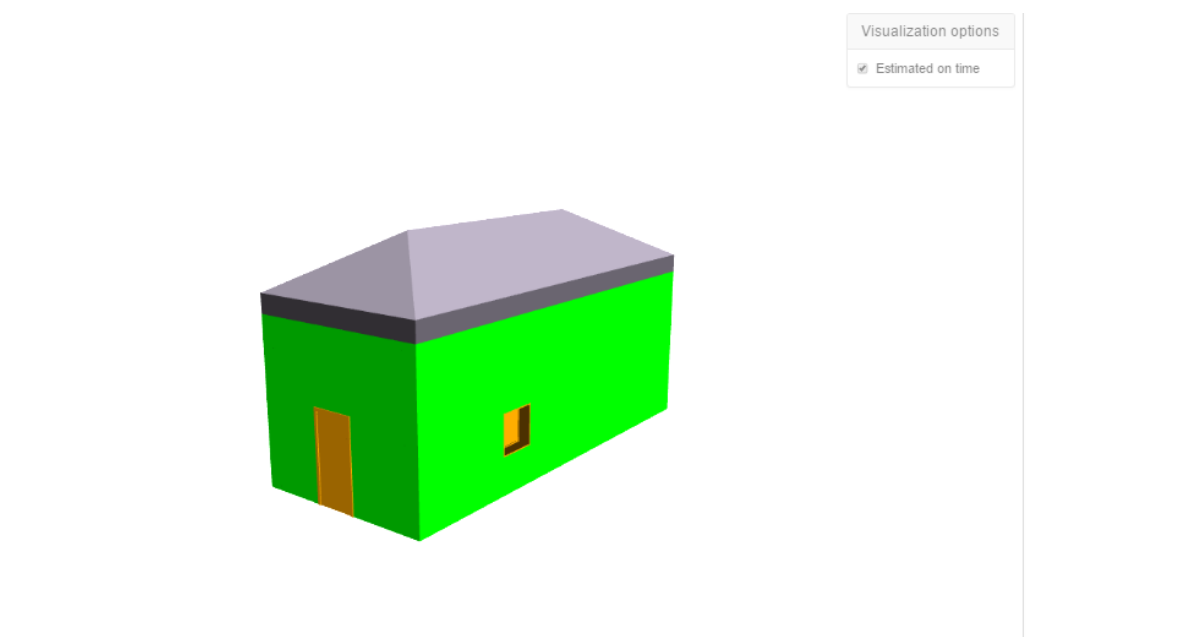
Figure 25 Visualisation output of the planning consult containing risk identification in colour codes. Green means no risk. Orange means middle risk. Red means high risk. Grey means no data available.

The following elements are not recognized in the database:

Product type	Amount of elements in model
Level: 8mm Head	2
Basic Roof: Generic - 400mm	1

Planning Consult Results

M_Single-Flush: 0915 x 2134mm

In the model are found 1 elements of this type of product.

Planning advice

Amount of elements in the database: 4

In the planningsdatabase the following process variants are found

Variant 1: 100 % of the elements from this product type are planned according to these steps:

Task number	Task name	Percent on time	Percent too late	Percent unknown
1	ST00080 Prepare doors	100%	0%	0%
2	ST00090 Make doors	75%	25%	0%

Figure 26 (1) Planning advice and risk identification HTML output of the planning consult.

Kanaalplaat

In the model are found 2 elements of this type of product.

Planning advice

Amount of elements in the database: 8

In the planningsdatabase the following process variants are found

Variant 1: 100 % of the elements from this product type are planned according to these steps:

Task number	Task name	Percent on time	Percent too late	Percent unknown
1	ST00020 Prepare flooring	100%	0%	0%
2	ST00030 Lay floors	100%	0%	0%

Baksteen

In the model are found 4 elements of this type of product.

Planning advice

Amount of elements in the database: 16

In the planningsdatabase the following process variants are found

Variant 1: 100 % of the elements from this product type are planned according to these steps:

Task number	Task name	Percent on time	Percent too late	Percent unknown
1	ST00040 Prepare walls	100%	0%	0%
2	ST00050 Make walls	100%	0%	0%

Figure 27 (2) Planning advice and risk identification HTML output of the planning consult.

Stelkozijn

In the model are found 1 elements of this type of product.

Planning advice

Amount of elements in the database: 36

In the planningsdatabase the following process variants are found

Variant 1: 100 % of the elements from this product type are planned according to these steps:

Task number	Task name	Percent on time	Percent too late	Percent unknown
1	ST00060 Prepare windows	61%	38%	1%
2	ST00070 Make windows	61%	38%	1%

Figure 28 (3) Planning advice and risk identification HTML output of the planning consult.

8 Case study construction site process mining

In order to answer research question six[26] and test the proposed toolbox. The proposed workflow ,as stated in chapter 4, is tested and presented by use of a case study. The target of the workflow is to support a continuous optimization of project schedules for construction companies by eliminating bottlenecks and planning deviations. Thereby it gives an achievable solution on problem definition 2. In addition this case study explores the boundaries of the current state-of-the-art regarding progress monitoring on construction sites. The target within this case study is to capture as-built process data, compare it to as-planned BIM, measure deviations, find bottlenecks and store this information in data formats which can be used as a reference in future projects. This study focusses on analytics on a project level, however it keeps in mind the applicability of BIM based process mining on a company- and industry level. Which means that every part of the workflow should be applicable on small- and large scale. The project is introduced shortly thereafter the phases of the proposed workflow are described in detail. The used tools are elaborated and the results of each phase are discussed. A final conclusion about the case study and the workflow is given at last. The data used in this project is available for scientific purposes. More info about the dataset can be found in appendix 14.

8.1 The project

The case study's project is called 'Schependomlaan'[27]. The project contains 10 apartments and is located in Nijmegen, The Netherlands. The building is developed and built by Hendriks Bouw en Ontwikkeling[28]. The execution phase started in February 2015 and at December 2015 the building was finished. Originally the building was designed in 2D, however those flat plans where converted into a high quality BIM model by ROOT[29]. This resulted in an as-designed IFC model which is used as a reference during the whole project.

Figure 29 as-designed model Schependomlaan

8.2 Method

The proposed workflow of chapter 4 is used which consists of four phases. Plan, capture, analyse, reuse. Within every phase specific tools are used (as can be seen in Figure 30). The

---

[26] "What are the requirements for a toolbox which can be used to address the challenges posed in problem definition 2?"

[27] http://www.schependomlaan.nl/

[28] http://www.hendriksbouwenontwikkeling.nl/

[29] http://www.root-bv.nl/

tools are chosen with in mind that the final target, reusing the information in order to shorten construction projects, could be achieved. During the plan phase Synchro software is chosen since it can import and export IFC models. As-built capturing is done by use of the BIMserver tool of RAAMAC which compares as-planned IFC models and as-built point clouds. This tool is chosen since drone images where available for the project which could be used to generate the point clouds. In order to create event logs the Eventlog service is used and analytics are done with Disco and MyInvenio since those tools had academic licenses available. Reusing the information is done by use the Planning consult service.

Figure 30 Used tools within this case study.

8.3   As-planned model generation

The first phase is planning the execution of the project. The input which is used by the project planner is threefold, (1) as-designed IFC, (2) Schedule ratios, (3) Risk identification. The output which is generated in this phase is an as-planned IFC. The dataflow can be seen in Figure 31.
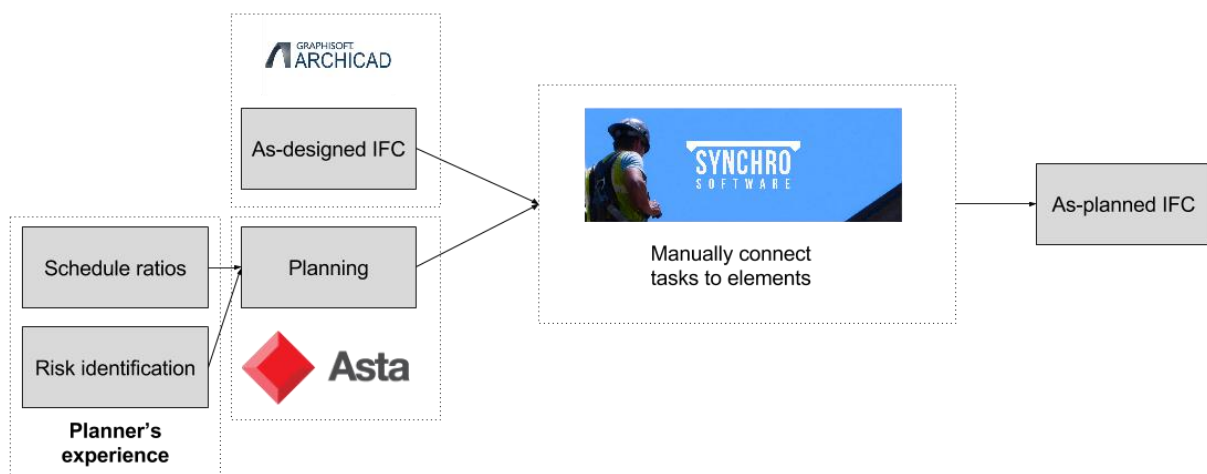
Figure 31 Schematic dataflow within the plan phase of the case study.

An as-planned IFC is generated with use of Synchro[30] software. The as-designed IFC is imported in the tool. The original planning is made with Asta Powerproject[31]. From Asta the planning is exported to a .XML file which is imported in Synchro. Every object in the as-designed IFC is then connected to the tasks belonging to it. This resulted in an as-planned model. During the project this model is used to explain and visualize the sequence of the construction schedule to practitioners. Synchro enables one to export to IFC, which resulted in as-planned IFC models (as can be seen in Figure 32).
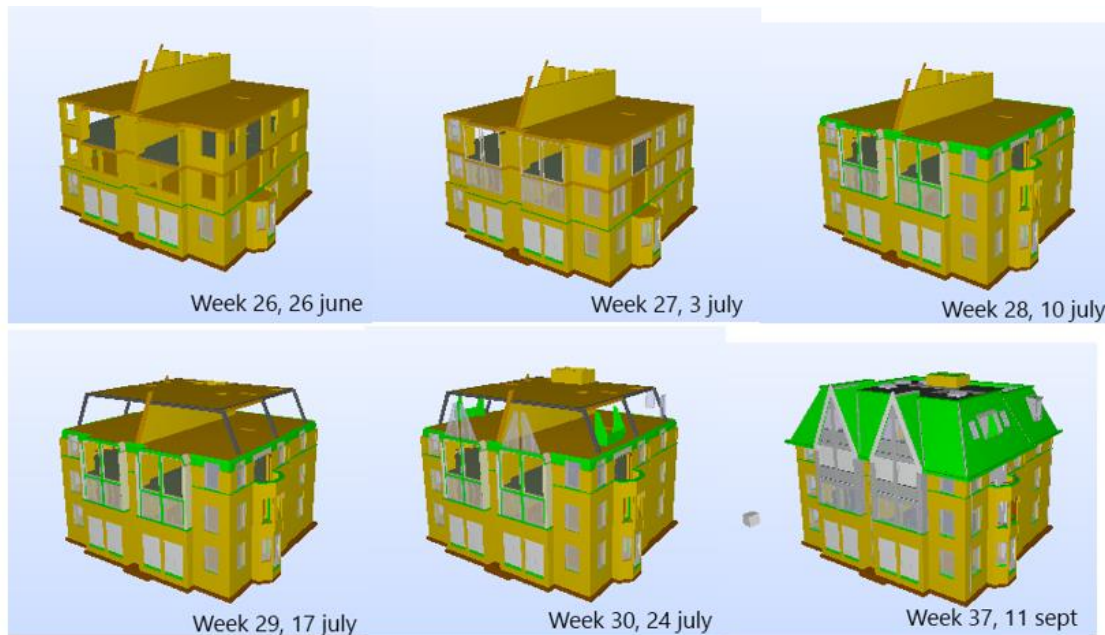
Figure 32 Exported as-planned IFC models.

8.3.1 Discussion as-planned model generation

The as-planned models where made with a schedule which was made in the beginning of the project. During the project this schedule was partly outdated. This resulted in data output which was partly not reusable. Another point of discussion is the IFC exporting function of Synchro which has some errors. For example the object tree within IFC is changed by creating the as-planned IFC model. For this study this was not problematic since the tool which reuses the IFCs has been newly developed, and is adjusted to the incorrect IFC object tree. However when a company wants to adapt this workflow on larger scale and will be able to reuse the as-planned IFC models it is necessary that the IFC model contains its original structure. More details about the structure of the as-planned IFC can be found in the dataset Appendix 14.

8.4 As-built capturing

During this project drone movies[32] where made by a future inhabitant. Those movies where originally intended for marketing purposes. In order to capture the as-built process on the construction site those movies were also suitable. As indicated in chapter 3.1.3.3, several

---

[30] https://synchroltd.com/

[31] http://www.powerproject.nl/oplossingen/asta-powerproject/

[32] The drone movies can be seen on the YouTube page of the drone pilot:
https://www.youtube.com/channel/UCHsKVCxqYaZn39JNZH7v9LQ

studies have proven that photos of an object can be combined in order to generate an point cloud model with structure from motion techniques. During the project several drone flights have been executed. Images of five of those flights are used within this case study. Resulting in point clouds of the days: 26 June, 3 July, 10 July, 17 July, and 24 July. So every Friday in week 26 until week 30 the progress is monitored. With help of the 'Real-Time and Automated Monitoring and Control' (RAAMAC) Lab from the University of Illinois at Urbana-Champaign the as-built situation is captured. The schematic workflow of the capturing phase is visualised in Figure 33.
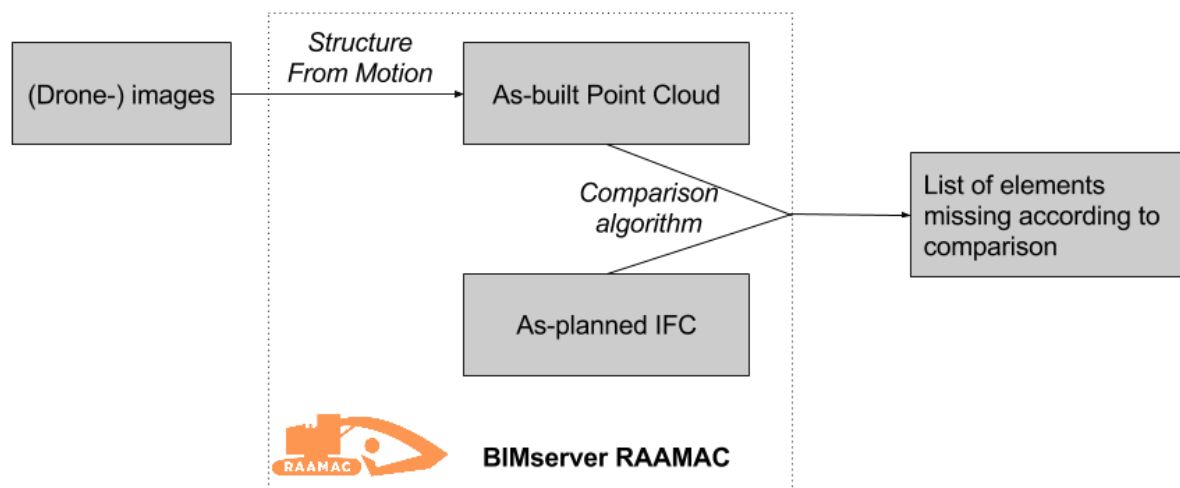
Figure 33 Schematic dataflow within the capture phase of the case study.

Details about the work of RAAMAC related to IFC and point cloud comparison can be found in (Dimitrov & Golparvar-Fard, 2014; Golparvar-Fard, Pena-Mora, & Savarese, 2009; Yang et al., 2015). However it is roughly explained next. RAAAMAC developed the D4AR tool for automating construction progress monitoring data collection processing and communication (Golparvar-Fard et al., 2009). They developed their tool with the vision that regular construction site images should feed the tool in order to track progress. They assumed that on almost every construction site images are taken and collected into a database.

8.4.1 Point cloud creation with Structure from Motion

Since Structure from Motion (SfM) methodologies are applied in the case study within this research this is elaborated more in this chapter. The procedure of figuring out correspondences between images and how they are interrelated to another in a 3D coordinate systems is commonly called Structure from Motion (Golparvar-Fard et al., 2009). SfM targets to reconstruct an unknown 3D scene and calculates unknown camera positions and orientations form a set of feature similarities among a set of images. Research (Snavely et al, Bown and Lowe) has shown that bundling the images to obtain a 3D reconstruction is robust regarding changes in images, resolution, time, focal length variability and illumination changes. These techniques are originally applied for image based walkthroughs and virtual tours (Snavely, Seitz, & Szilski, 2006). However (Golparvar-Fard et al., 2009) have successfully demonstrated a SfM technique applied to "geospatially photographs that are capturing dynamic construction scene over the time span of its construction" (Golparvar-Fard et al., 2009, p138).

Reconstructing the 3D scene is based on finding feature points in each image. SIFT keypoint detector is used (Lowe, 2004). Pairing those points creates tracks which matches keypoints across multiple images. Once those points are detected over the dataset it is necessary to detect which features are paired in imaged.

In order to be able to create a robust 3D model the application needs accurate information about the camera like relative location and orientation, focal length and distortion of the lenses. In order to be able to reconstruct the model over the as-planned model also the absolute location of the camera is needed. However this computation is not done with GPS or wireless location trackers, (Golparvar-Fard et al., 2009) uses computer vision techniques to compute this information.

The detected keypoints are combined with the camera parameters which results that each keypoint consists of a 3D location and a colour which is determined out of the average from all observed images. A mapping is made of all the keypoints resulting in a 3D point cloud, in addition the cameras observing those points are mapped as well.

8.4.2 As-built scene representation

The Structure from Motion procedure calculates the relative camera locations (Golparvar-Fard et al., 2009). However the final step is to align the reconstructed point cloud with the as-planned IFC. Therefore locations of both models is aligned with use of "the closed-form solution of absolute orientation using unit quaternions"(Horn, 1987). If models are not aligned correctly due to procedure fails control points can be added to the models in order to manually rectify them. Once the registration of the IFC model and the reconstructed scene is done only the images with respect to the as-planned model are taken into account. The as-built scene representation including the as-planned IFC can be seen in Figure 34.
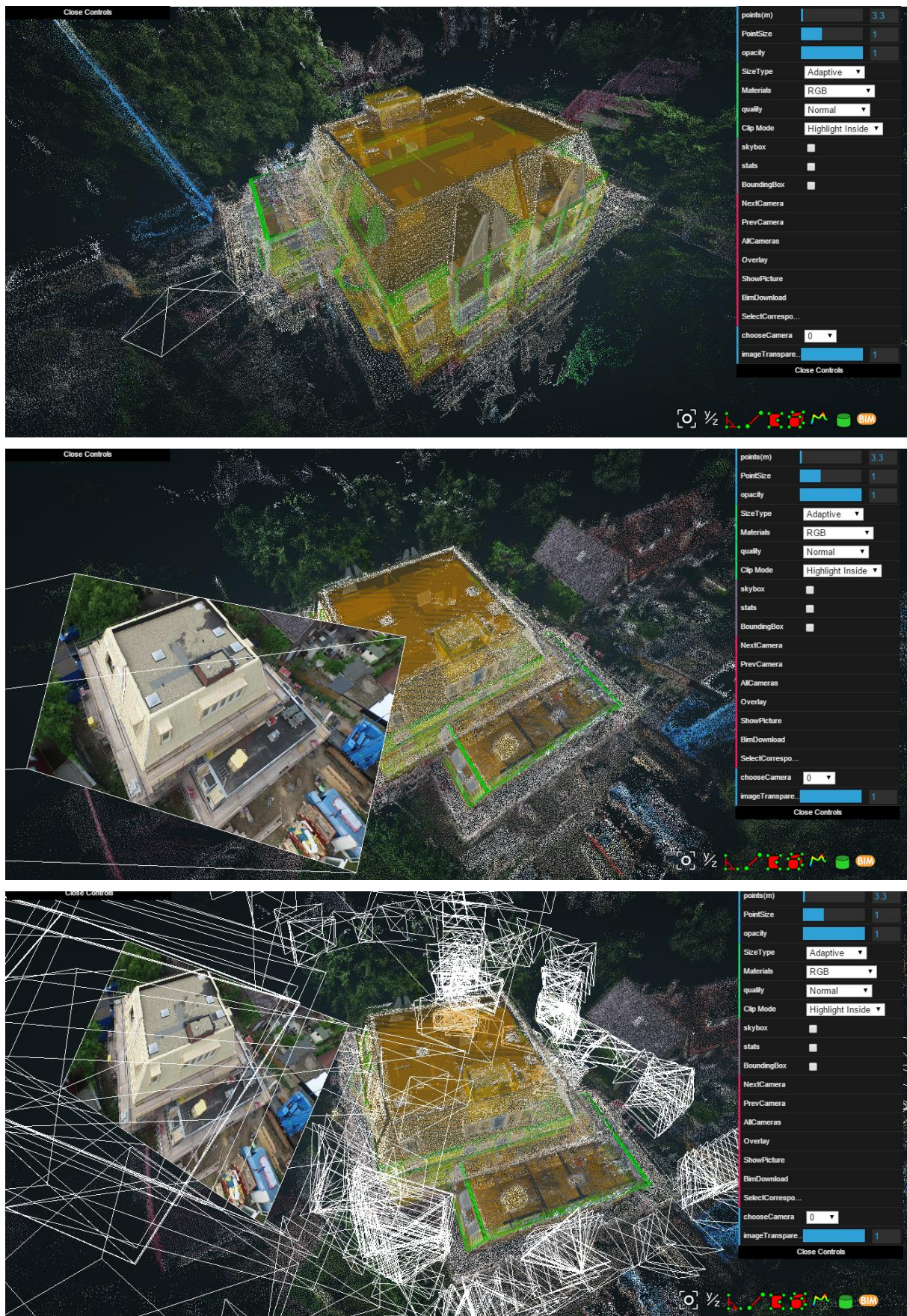
Figure 34 As-built point cloud versus as-planned IFC comparison.

8.4.3 Comparison as-built point cloud and as-planned IFC

Using the integrated scene with the as-built point cloud and the as-planned IFC D4AR reasons about the occupancy and visibility of both models. Together with this assessment a supervised machine learning method is applied in order to observe if an element is built or not (Yang et al., 2015). As a result elements are indicated built (coloured green) or not built (coloured red).

For this study weekly comparisons where made between the as-built point clouds and the as-planned IFC models. Every week a list with elements is created containing missing elements according to the drone scan. Moreover this means that every element in the list has been too late that day according to schedule. An example of the created list is shown in Figure 35.

	A	B	C	D
1	Elements missing according to comparison point cloud week 26 and as planned model week 26			
2	Name	GUID	GlobalID (SynchroResourceProperty)	GlobalID (IFC)
3	dakisolatie_(#797154)	b10bb459-349b-48b6-ad0f-1bd9d238ae50	2UXSKqZJ97I9trCo$wjoG6	2n2xHPD9j8jgqF6zdIEAvG
4	dakisolatie_(#797314)	94c0da4c-a0ba-4032-ab92-40f8b0a2ca89	2F7VeKs_LAaRr4HPesM$ZW	2KmDfCeBf0CgkIGFYmeig9
5	merk F-R_(#187290)	614be6e4-6ecf-4307-b9a2-6773ee9c94d3	335BWviar9ZQ5ecfCfZenu	1XI_RaRiz31xcYPtFkd9JJ
6	merk F-R_(#189383)	c08c1017-4d1e-4cef-8c34-ed1e057c3424	3OvN5YRun1k8pQauRQQC1S	30Z10NJHvCxumqxHu5V3Ga
7	merk F-R_(#77756)	0fed9314-8108-4875-9ff9-031eef9e9de5	2NlJ56MP9BBxSRulUZYZuW	0FxPCKWGX8TP$v0nxIdftb
8	merk F-R_(#81462)	bb9ed45b-7aa7-45d6-8217-55c9fb22505b	0TsCZPJtz5tALncV2598El	2xdjHRUgT5re8NLSdx8b1R
9	merk N_(#132846)	361b167b-f72d-4095-bfe8-e8503e3492fb	1QKgjjdDP3J9Q8EaAxh21h	0s6nPxzor0bR$ew50_D9Bx
10	merk A2_(#128666)	359377c6-52d5-4b87-8954-c8538edf5fb7	1AgoIy9ePFOBt24JNbOQtA	0ratV6KjLBXubKo5EEtr_t
11	merk N_(#131772)	bc267f21-edaa-4f59-9135-741a1b841bf5	2Sw92ltcj6seLSG_J5epq2	2y9dyXxQfFMP4rT1eRX1lr
12				

Figure 35 Missing elements according to comparison point cloud week 26 and as-planned IFC model week 26.

- Week 26 contained 9 missing elements including windows and insulation parts.

- Week 27 contained 50 missing elements including windows and additional parts belonging to the same windows.

- Week 28 contained 89 missing elements including windows, additional parts, insulation, brick facades, and prefab elements.

- Week 29 contained 18 missing elements including windows, insulation and brick facades.

- Week 30 contained 4 missing elements all being windows.

The project manager indicated that the results are reliable. He mentioned that a problem with the windows in early phases resulted in a delay of process related to those elements. However this was updated in a new schedule.

8.4.4 Discussion as-built capturing

Only five days are captured with the drone scan. So only on five moments the progress could be captured. It took quite some time to generate point clouds of the images, as a result more than five moments of analysing was not feasible within this study. However it is proven that measuring planning deviation throughout the project with such a workflow is possible with the current state of technology. The process of generating point clouds and comparing them with IFC models and generating lists of missing elements can possibly be fully automated. As a result it should be possible to scan construction sites on a more regular basis and the level of detail of the monitoring data would be a lot higher. In conclusion automating this process is recommendable for future research.

The schedule which is used was already partly outdated. More specific it was known that the window elements where too late because of earlier delay. If those workflows are used in practice it should be necessary that the as-planned model is from the latest version.

Planning deviation is measured with use of technology, this addresses a part of problem definition 2. However this study only indicates the planning deviation, it does not study the causes of the deviation. But with the use of the drone videos causes could be detected. One could identify deviation on a certain moment, find this moment in its context on the videos and search for causes. The development and usage of such a system would be recommendable for future research.

8.5 As-planned and as-built analytics

8.5.1 Event log generation

In order to analyse the project with process mining tools an event log is necessary. Using Figure 36 as a reference this workflow is explained.

With help of the Eventlog service in BIMserver (see chapter 5) the as-planned IFC is converted into an as-planned event log. Thereafter the results of the drone scan are implemented in the event log which results in an as-built event log. This log is suitable for analytics. The as-planned event log of the case study contains 7166 events, a fragment is visualised in Figure 37.

In order to create an as-built event log the results of the drone scan have to be implemented in the as-planned event log. Both files are merged manually using excel. The as-built indication results in an extra row in the event log named 'Event on time executed or not'. As can be seen in Figure 38 this is a value 'On time' or 'Too late'. 194 events of the total log are indicated as 'Too late'.
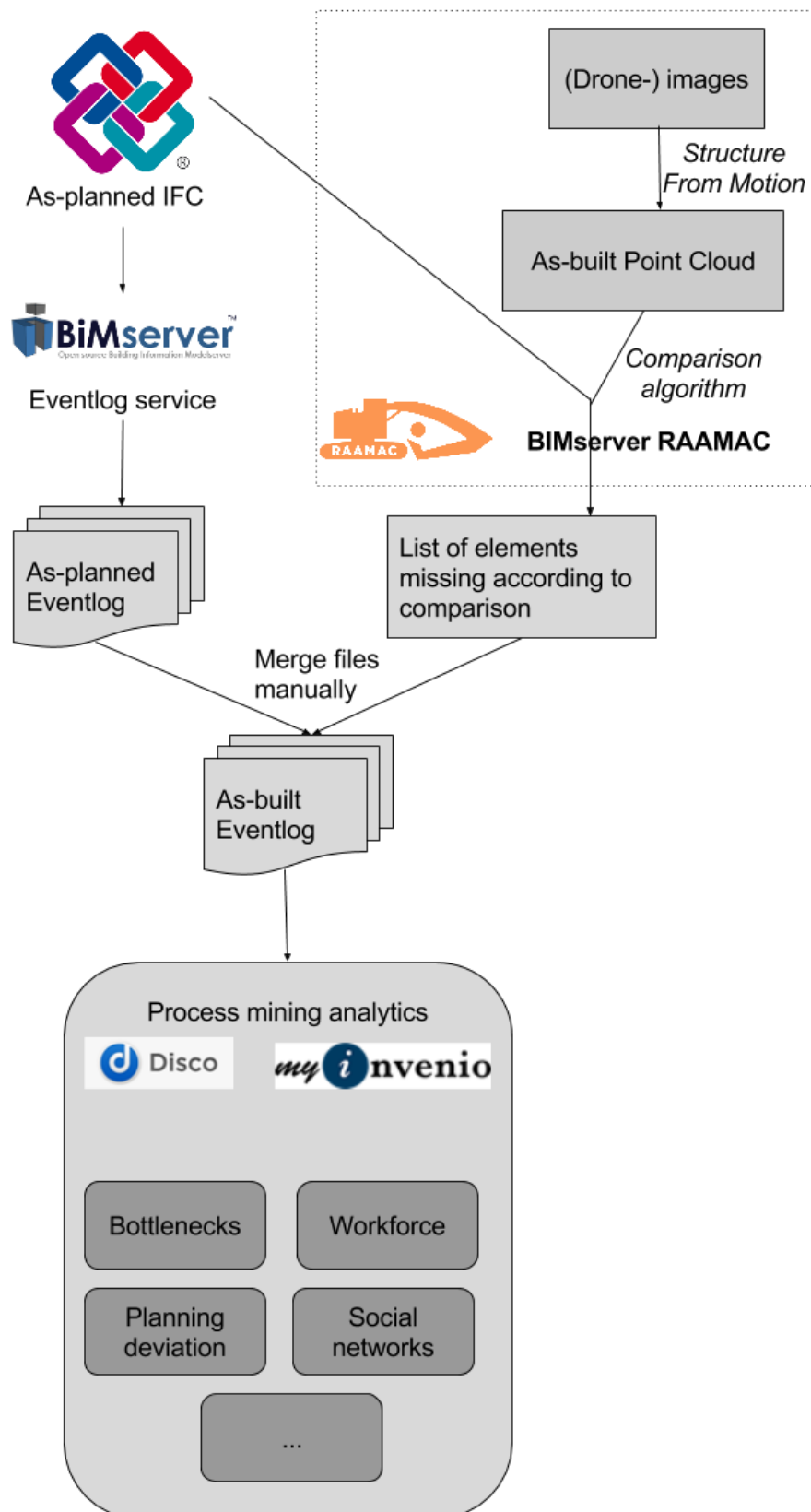
Figure 36 Dataflow in order to enable event log analytics.

	A	B	C	D	E	F	G	H	I	J	K
1	BuildingGUID	GUID	ifcClass	NI-sfb	Material	Task	Resource	TaskName	TaskStart	TaskFinish	
2	2CKEPCy9T9zQMKorZoDrni	3iHy1LOhfAtfLt80$LZBqn	Slab	1320_vloeren op grondslag - construct	IFC_betonvloer_ihw_300mm	ST00060	liftputvloer_(#5715)	Stellen bruggetjes	26-2-2015	27-2-2015	
3	2CKEPCy9T9zQMKorZoDrni	3iHy1LOhfAtfLt80$LZBqn	Slab	1320_vloeren op grondslag - construct	IFC_betonvloer_ihw_300mm	ST00070	liftputvloer_(#5715)	Nuts invoeren	2-3-2015	4-3-2015	
4	2CKEPCy9T9zQMKorZoDrni	3iHy1LOhfAtfLt80$LZBqn	Slab	1320_vloeren op grondslag - construct	IFC_betonvloer_ihw_300mm	ST00080	liftputvloer_(#5715)	Werkvloer liftput	3-3-2015	3-3-2015	
5	2CKEPCy9T9zQMKorZoDrni	3iHy1LOhfAtfLt80$LZBqn	Slab	1320_vloeren op grondslag - construct	IFC_betonvloer_ihw_300mm	ST00100	liftputvloer_(#5715)	Bekisten liftvloer	4-3-2015	4-3-2015	
6	2CKEPCy9T9zQMKorZoDrni	3iHy1LOhfAtfLt80$LZBqn	Slab	1320_vloeren op grondslag - construct	IFC_betonvloer_ihw_300mm	ST00090	liftputvloer_(#5715)	Wapening liftvloer	4-3-2015	4-3-2015	
7	2CKEPCy9T9zQMKorZoDrni	3iHy1LOhfAtfLt80$LZBqn	Slab	1320_vloeren op grondslag - construct	IFC_betonvloer_ihw_300mm	ST00110	liftputvloer_(#5715)	Stort liftvloer	5-3-2015	5-3-2015	
8	2CKEPCy9T9zQMKorZoDrni	2GTwJrGhDAzvl7WZJUCeZE	Wall	2220_binnenwanden - constructief	IFC_kalkzandsteen_214mm	ST00130	liftwand_(#29901)	Bekisten liftwanden	6-3-2015	6-3-2015	
9	2CKEPCy9T9zQMKorZoDrni	2BvlWOGA18JuJGl8ryIXFV	Beam	1610_funderingsconstructies - voeten	IFC_betonopstort_370x965	ST00130	liftputwand_(#5543)	Bekisten liftwanden	6-3-2015	6-3-2015	
10	2CKEPCy9T9zQMKorZoDrni	0ZOrCFqKrF_Ox7WOTGyQmt	Beam	1610_funderingsconstructies - voeten	IFC_betonopstort_220x965	ST00130	liftputwand_(#5377)	Bekisten liftwanden	6-3-2015	6-3-2015	
11	2CKEPCy9T9zQMKorZoDrni	1naZcReNz7PvhQAmj6Jq3l	Beam	1610_funderingsconstructies - voeten	IFC_betonopstort_220x345	ST00130	liftputwand_(#4278)	Bekisten liftwanden	6-3-2015	6-3-2015	

Figure 37 Fragment of the as-planned event log from the case study project

	A	B	C	D	E	F	G	H	I	J	K
1	BuildingGUID	GUID	IfcClass	NI-sfb	Material	TaskID	Resource	TaskName	TaskStart	TaskFinish	Element on time or no
2	2CKEPCy9T9zQMKorZoD	3iHy1LOhfAtfLt80$LZBqn	IfcSlab	1320_vloeren op grondslag - constructief	IFC_betonvloer_ih	ST00060	liftputvloer_(#5715)	Stellen bruggetjes	26-2-2015	27-2-2015	On time
3	2CKEPCy9T9zQMKorZoD	3iHy1LOhfAtfLt80$LZBqn	IfcSlab	1320_vloeren op grondslag - constructief	IFC_betonvloer_ih	ST00070	liftputvloer_(#5715)	Nuts invoeren	2-3-2015	4-3-2015	On time
4	2CKEPCy9T9zQMKorZoD	3iHy1LOhfAtfLt80$LZBqn	IfcSlab	1320_vloeren op grondslag - constructief	IFC_betonvloer_ih	ST00080	liftputvloer_(#5715)	Werkvloer liftput	3-3-2015	3-3-2015	On time
5	2CKEPCy9T9zQMKorZoD	3iHy1LOhfAtfLt80$LZBqn	IfcSlab	1320_vloeren op grondslag - constructief	IFC_betonvloer_ih	ST00100	liftputvloer_(#5715)	Bekisten liftvloer	4-3-2015	4-3-2015	On time
6	2CKEPCy9T9zQMKorZoD	3iHy1LOhfAtfLt80$LZBqn	IfcSlab	1320_vloeren op grondslag - constructief	IFC_betonvloer_ih	ST00090	liftputvloer_(#5715)	Wapening liftvloer	4-3-2015	4-3-2015	On time
7	2CKEPCy9T9zQMKorZoD	3iHy1LOhfAtfLt80$LZBqn	IfcSlab	1320_vloeren op grondslag - constructief	IFC_betonvloer_ih	ST00110	liftputvloer_(#5715)	Stort liftvloer	5-3-2015	5-3-2015	On time
8	2CKEPCy9T9zQMKorZoD	0n3uFweoDDvekPpYzTttt4	IfcWall	2220_binnenwanden - constructief	IFC_kalkzandsteer	ST00130	binnenwand_(#44016)	Bekisten liftwanden	6-3-2015	6-3-2015	On time
9	2CKEPCy9T9zQMKorZoD	2rxEfS5c1BRBMAjC9Weumf	IfcBeam	1610_funderingsconstructies - voeten en bal	IFC_betonopstort_	ST00130	liftwand_(#1315)	Bekisten liftwanden	6-3-2015	6-3-2015	On time
10	2CKEPCy9T9zQMKorZoD	2Y$eh0ohrDEepO4_LPdoEr	IfcBeam	1610_funderingsconstructies - voeten en bal	IFC_betonopstort_	ST00130	liftwand_(#1963)	Bekisten liftwanden	6-3-2015	6-3-2015	On time
11	2CKEPCy9T9zQMKorZoD	1naZcReNz7PvhQAmj6Jq3l	IfcBeam	1610_funderingsconstructies - voeten en bal	IFC_betonopstort_	ST00130	liftwand_(#4278)	Bekisten liftwanden	6-3-2015	6-3-2015	On time

Figure 38 Fragment of the as-built event log from the case study project. This log differs from the log in figure 37 with the parameter 'Element on time or not'

8.5.1.1 Discussion event log generation

In order to gain insight in the actual building process the as-built event log is generated. However this log only describes if planned tasks where executed on time or not. It doesn't describe activities which were executed but where not planned. As a result probably some events are not taken into account. Since this is an experimental study focussing on the total workflow this is accepted. However it could be interesting for further research to implement unplanned tasks within the as-built event log and do planned versus unplanned task analytics.

Moreover the events are indicated with an 'on time' or 'too late' attribute. This doesn't describe at which moment an event is exactly executed. Due to the few drone scans this was not possible to realize. However it is interesting to study the possibilities to add exact execution times of events. Probably other measuring technologies should be used in order to optimize process monitoring.

Not every event could be monitored with the drone scan. Several events describe supporting steps which are used to finally finish the building element. With the drone scan only the last steps in a process could be verified for sure. However when an element was not recognized it was assumed that supporting steps were also too late. This result in a questionable event log. However as mentioned earlier this study proposes a workflow which must be optimized in some details. This is for further research.

8.5.2 Event log analytics

As indicated in chapter 3.2 several analytics can be done with the event logs. This study intends to focus on bottlenecks and planning deviation. Those are elaborated below. However interesting analytics such as social network mining or workforce mining was also possible. The results of those analytics are also elaborated shortly. The analytics are done using the process mining tools Disco and MyInvenio. Other tools are available for analysing the event logs,

however both of these tools provide academic licences to students hence they were accessible..

8.5.2.1 Total process bottlenecks

As can be seen in the structure of the event log a BuildingGUID is included (see Figure 37). Using this as a case ID gives a process map of the complete building process as can be seen in figure (process map buildingGUID). The process map clearly indicates with red colour which tasks take the longest time, those are task ST01120 (21 days) and ST01130 (20 days) both are named "Plaatsen scharnierkap incl. goot en platte daken". The project manager validated this result and said this long duration is caused by the chosen design and building method. The roofs where highly detailed which causes the long duration. The project manager mentioned this data and duration could be interesting to see before projects are starting. It could identify different durations between different building methods and designs. Moreover the applicability of automated duration calculation can be read in chapter 8.6. In addition the process map in Figure 39 indicates waiting times in between tasks. The project manager mentioned those times are interesting to study because this times are waste and cost money. The construction company is continuously applying the LEAN philosophy which states that waiting times should be minimized. It is concluded by the project manager that process mining enables unique insight in those durations and that it makes him more aware of the feelings he already had. In addition he mentioned that it was easier to communicate those results and to use them in conversations with higher managers as well as craftsmen on the construction site. This is not done within this study and would be interesting for further research. The applicability of total project's process analytics is in case of just one project not that interesting, since there is already a planning which can be analysed without event logs. However the visualisation of the process map easily enables one to find bottlenecks. But it is recognized that total process analytics is more interesting when several comparable project's event logs are merged into one database which can be analysed. As suggested by van der Aalst, (2011) this database is one where companies can create a *process-oriented data warehouse* containing information about relevant events happened in the company.
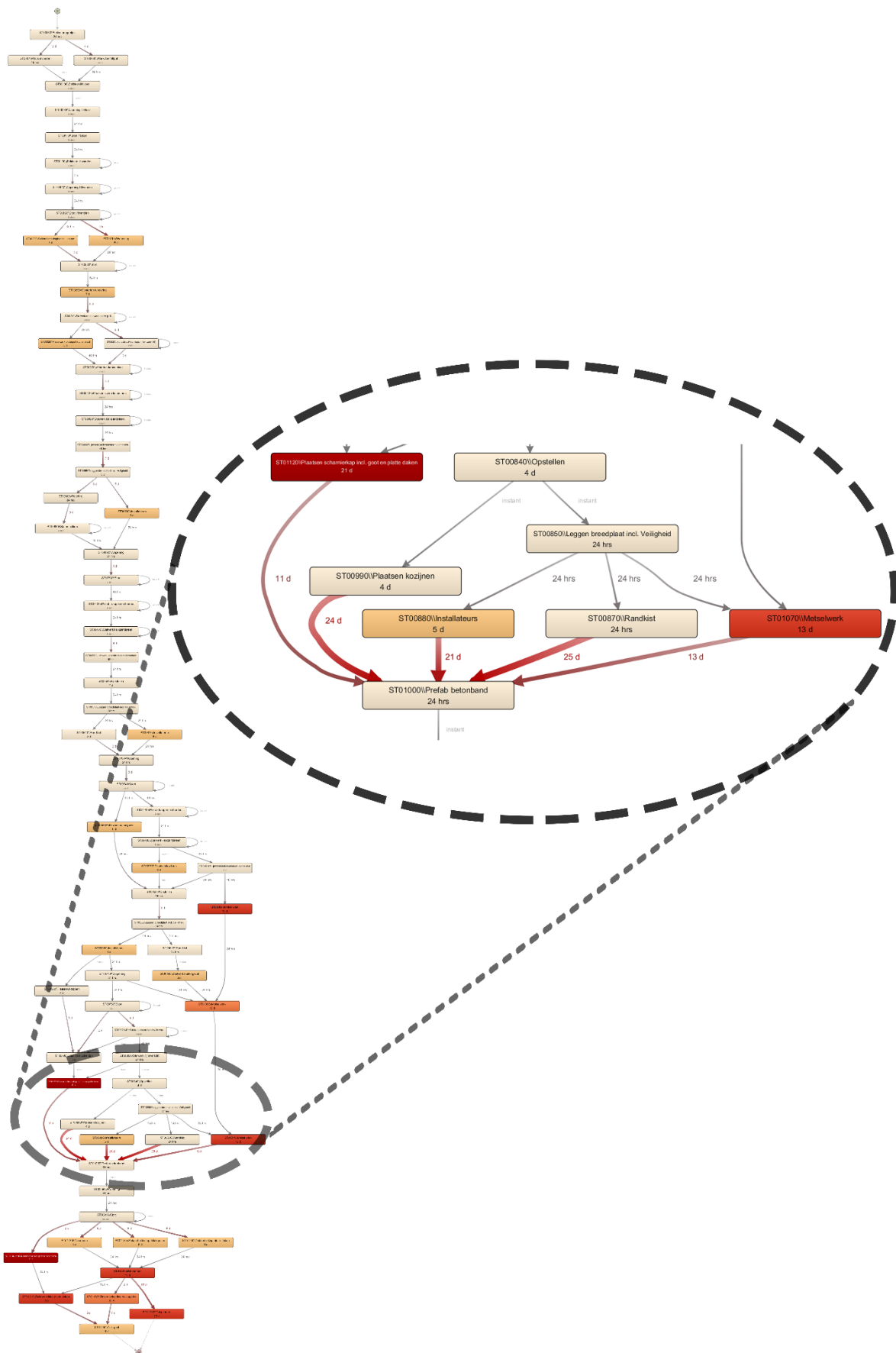
Figure 39 Process map of complete case study using the buildingGUID as case ID. (The details are probably not readable but the red visuals indicate long durations. This image is just to indicate the possibilities, not to explain the details.)

8.5.2.2 Element process bottlenecks

As can be seen in the structure of the event log a GUID of each building element is included (see Figure 37). When using this GUID as a case ID in process mining analytics process maps about specific building element types can be discovered. So not focussed on the total building process, but on separate building elements. With this approach process variants of elements can be discovered within projects and over several projects. Questions which can be answered are for example: "What is the process of concrete floors?" and "Which process variants of concrete floors do exists?". Examples of process maps and statistics of concrete floors, foundations are given.

Foundations

Using Disco a filter is made for all events which contain the attribute "Strookfundering" (which is Dutch for foundation). This resulted in process map as shown in Figure 40. 31 cases where found. All elements passed the same process steps with the same sequence. The mean duration was 13 days. 100% of the elements which passed this process where on time. This results are validated by the project manager. He also mentioned that some more process steps where actually executed but those were not included in the initial planning and therefore not translated into the event log.
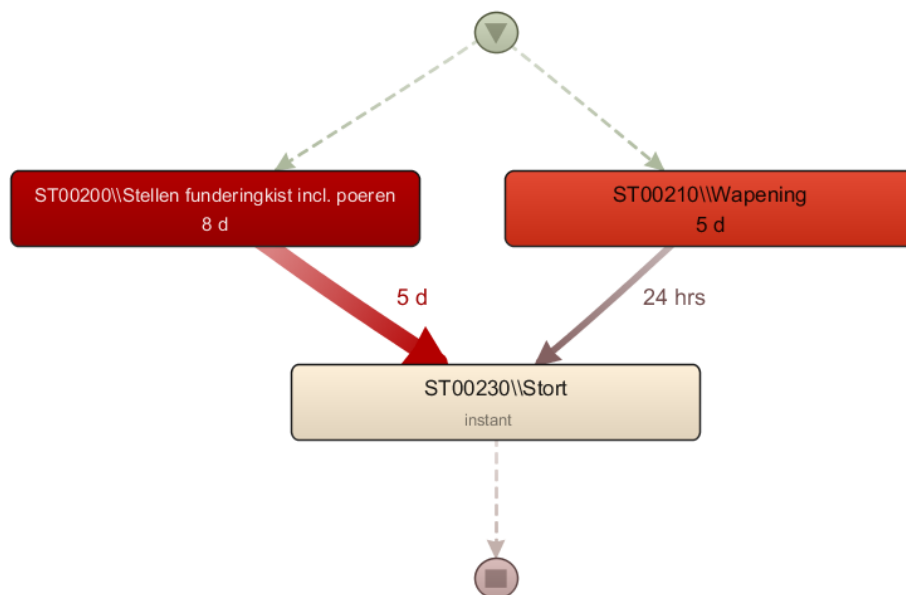
Figure 40 Process map of creating the foundation.

Concrete floors

Using Disco a filter is made for all events which contain the attribute "Breedplaatvloer" (which is Dutch for concrete floor). This resulted in process map as shown in Figure 41. Six cases where found. Four process variants where discovered. This was indicated as strange by the project manager, because every element should pass the same process. He discovered that some elements where not connected to the planning correctly. As a result it is seen that process mining can be used for as-planned model validation. It can be checked if elements with the same attributes have the same tasks. However probably also other methods can be used for this. This can be interesting for further research. Interesting for this research is that is proven

that process variants can be discovered easily with process mining analytics. The events containing the most time is in every variant the same, which is placing all the installations on the floor (task: "Installateurs"). The project manager indicated that if this task can be shorten the total process of laying concrete floors will be shortened most efficient.
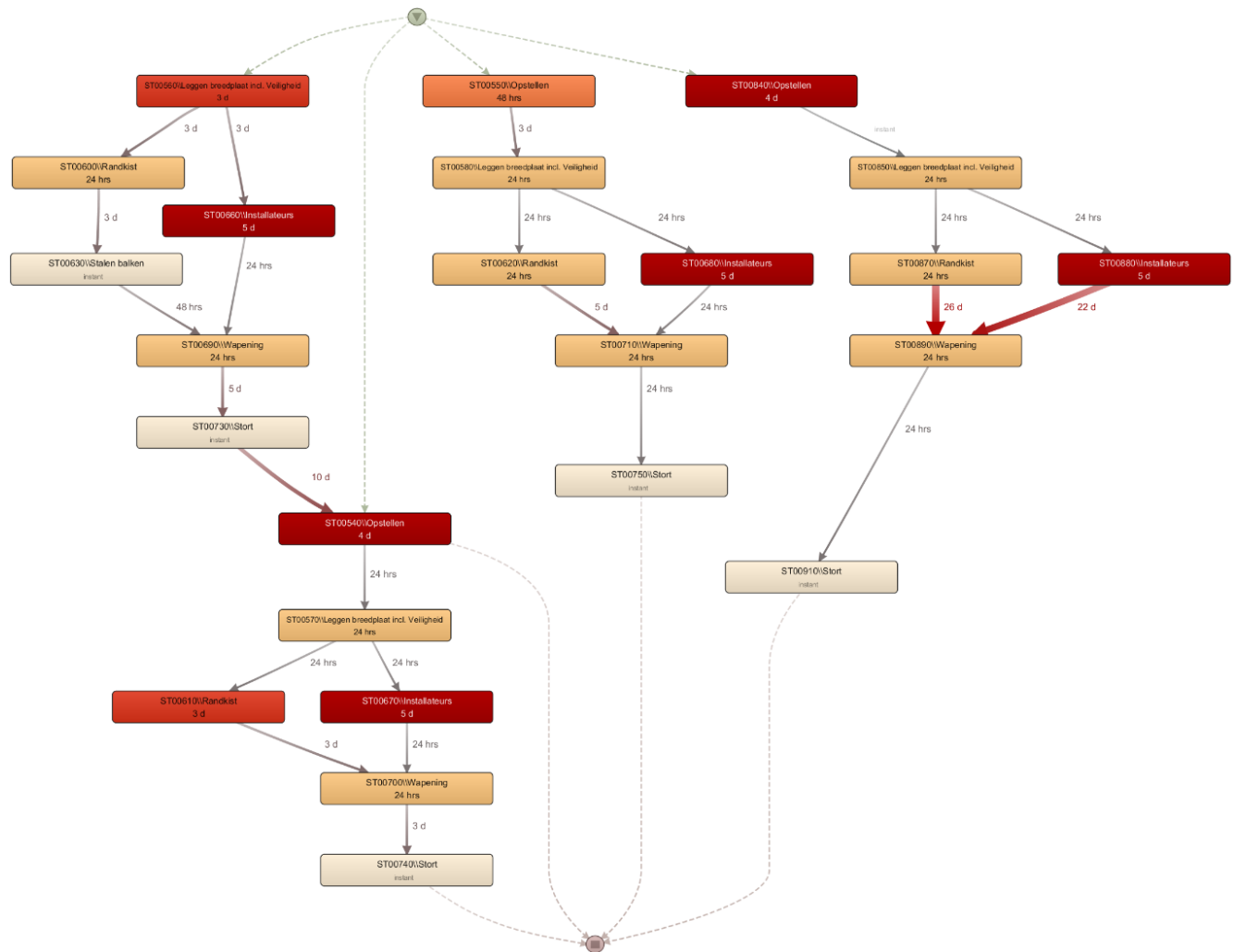
Figure 41 Process map of making a concrete floor. The red arrows and blocks indicate the longest durations.

8.5.2.3 Organizational mining

As described in section 3.2.4 process mining enables organizational network analytics. Meaning it is possible to discover social networks within processes, determining organizational structures and analyse individual behaviour. A resource attribute must be added to the events in the log to enable such analytics. In order to explore the use of those analytics within construction companies the event log of the case study is analysed with a social analytic tool[33]. However no resources where added to the event log initially. In addition there was no time to ask all the involved actors to accordance, therefore fake actors are added to the event log. So the results from this analytics cannot be seen as legitimate, but the results are discussed with the project manager in order to determine the applicability in future projects. As can be seen in

---

[33] https://www.my-invenio.com/

Figure 42 the event log of the case is supplemented with the attributes *company* and *actor.* This resources are added as realistic as possible and are validated as realistic by the project manager of Schependomlaan.

Figure 42 Resources 'company' and 'actor', see green rectangle, added to the event log of the project.

Activity analytics

Activity maps can easily be generated with the event log (see Figure 43).The tasks which are executed during the project are related to the corresponding actor. The project manager indicated that this is valuable on company level to find out which partners or actors are hired for specific tasks. Current methodologies for finding partners are based on experience and feelings about those experiences. When a project manager searches for companies which can execute a specific task during a project he talks to the relation manager who gives advice based on his experience. Using event logs containing resource attributes will enable a method where partners can be found and selected based on data. It will be possible to judge a partner on facts and select the best one for the job based on previous projects.

Social network discovery

Social networks can be discovery easily from the event log (see Figure 44). Actors who follow each other in activities are interconnected. The more they are related the stronger the connection, so the bigger the line. As a result it is noticed that 'Timmerman 1' (Carpenter) from company 'Hoofdaannemer' (main contractor) is the most influential within the project. He has the most communication with the most parties. The project manager from case study Schependomlaan indicates that it is nicely visualized, but he is more interested in these kind of studies from the office perspective. He likes to know who is communicating with each other in office related work like engineering and project preparation. In order to enable this kind of analytics event logs have to be extracted from project management systems. As described in chapter 3.2.5.1 ,and elaborated more in appendix 0, this is possible and would be interesting in further research.
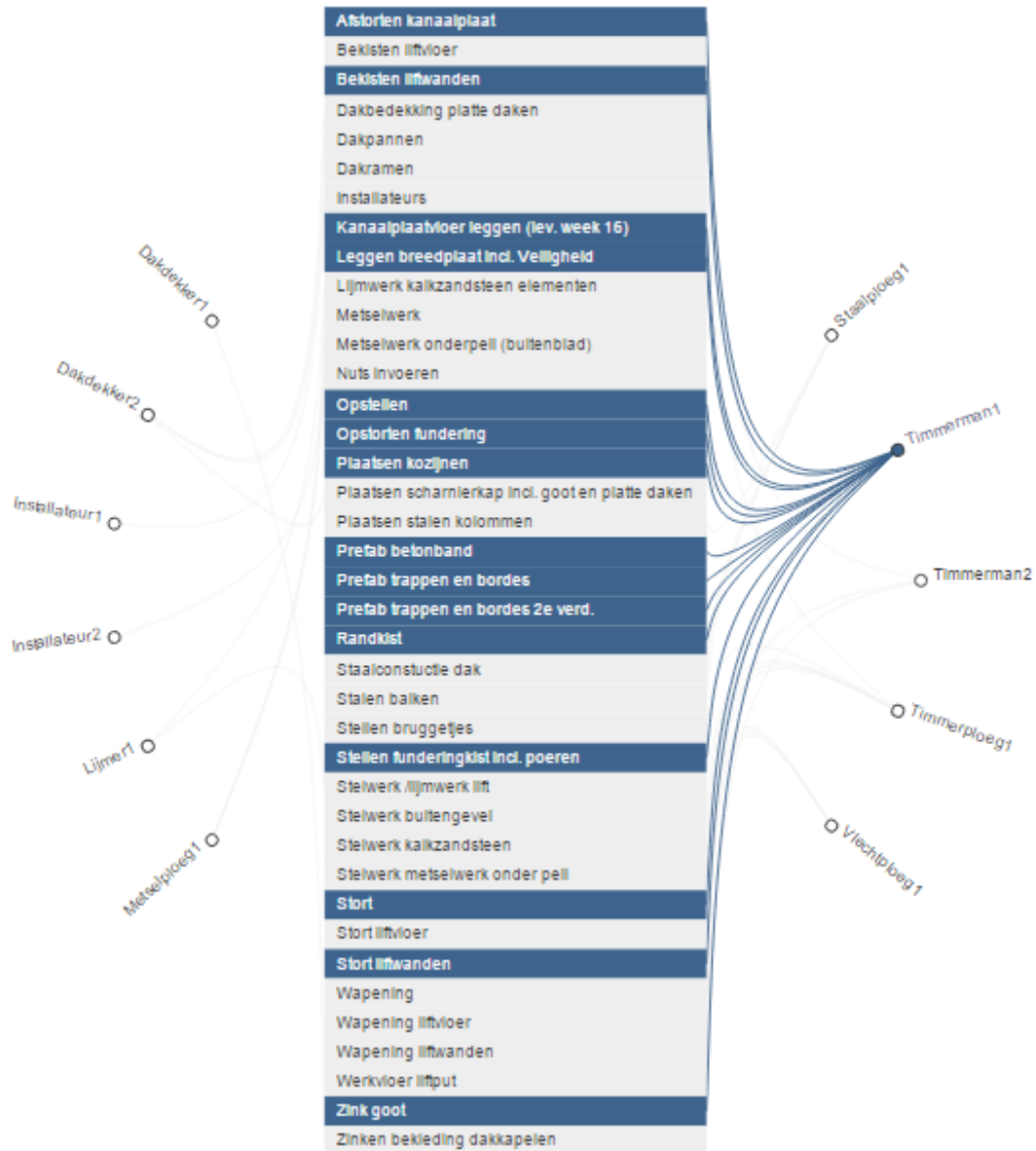
Figure 43 Activity map of the actor 'Timmerman' (Dutch for carpenter).
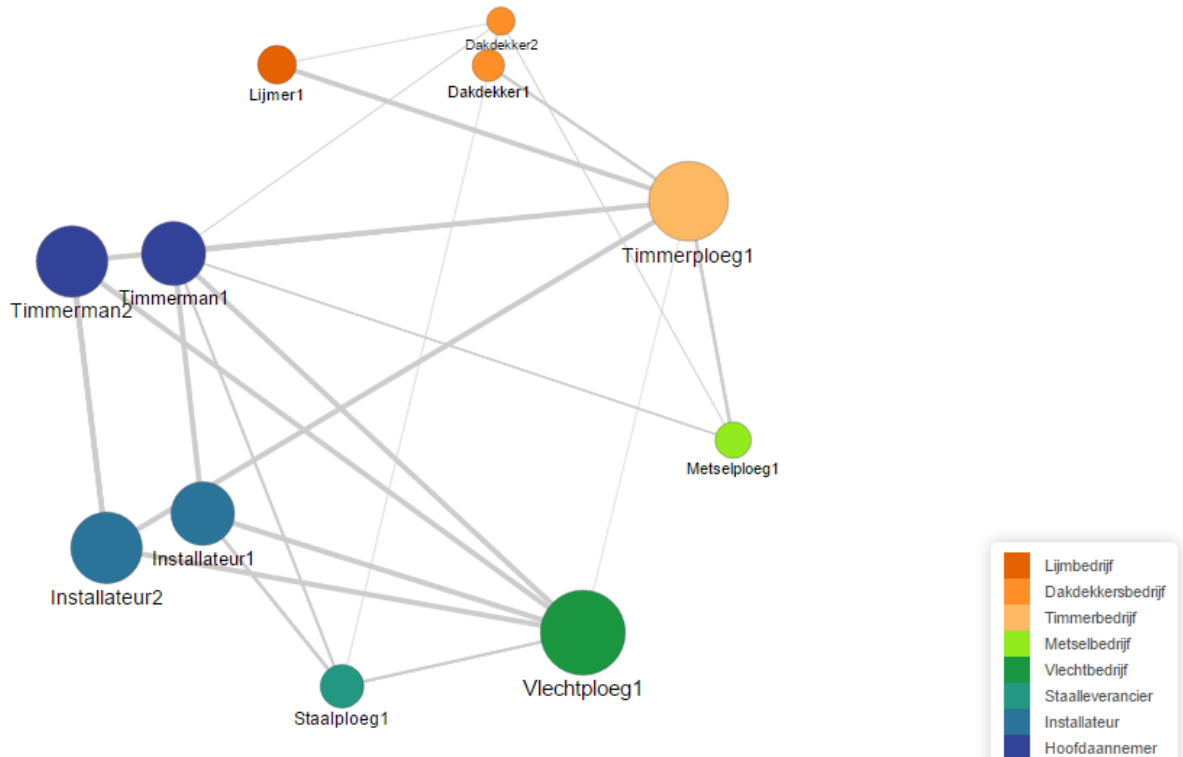
Figure 44 Social network mined from the event log of the project.

8.5.3 Discussion analytics

Since the data in the event log is questionable the results of this analytics cannot be rated with a real value. But the workflow is proven as possible and further research could determine real benefits in terms of efficiency and money.

In order to apply task related analytics on company level it is useful to agree about task or activity names. To compare similar tasks they should have the same unique name or number. The application of schedule naming standards is not been studied extensively during this study and is interesting for further research.

8.6 Reuse of process data

In order to actually help the project planner to avoid failures and short construction projects learning from failures should be enabled. Therefore reusing data is important to study. The last step of the proposed workflow, as described in chapter 4, is reusing the data from previous projects. This data is stored in event logs and contains (1) general properties about building elements and (2) as-planned process related properties about those elements. As-built properties are added as well. In order to reuse the data a Planning consult service is made for BIMserver (chapter 6). The Eventlog which is generated with this case study is suitable for identifying risks and giving some planning consultation. The workflow is shown to the project manager and several valuable applications are seen. First the Planning consult outcome is shown, thereafter the applications are described and a discussion is placed in the last paragraph of this chapter.

8.6.1 Planning consult

Reusing the event log of case study Schependomlaan is possible with the Planning consult. A new as-designed IFC model can be uploaded to BIMserver. The model will be searched for

objects with materials which are used in earlier projects stored in the event log. A planning consult including risk identification will be generated for all objects which are previously built.

In order to show the potential the sample project of created in chapter 7 is used as input (see Figure 45). The output is twofold, a planning consultation and a risk identification. First an 3D visualisation is shown with elements containing a risk identification by colour (see Figure 46). Windows and walls are indicated with low risk profiles in Orange, floor elements have no risk and are indicated green. Second parts of the output is the planning consult which is shown partly in the screenshots in Figure 47. The complete output file can be seen in Appendix 17.
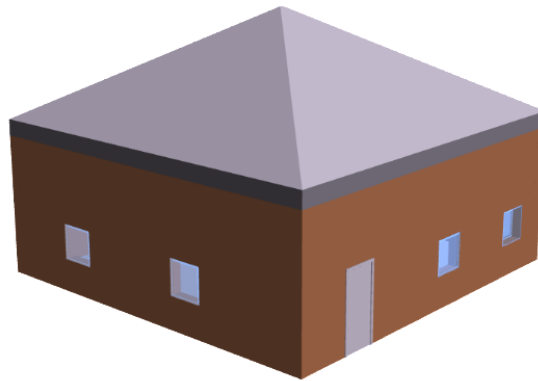
Figure 45 Sample project
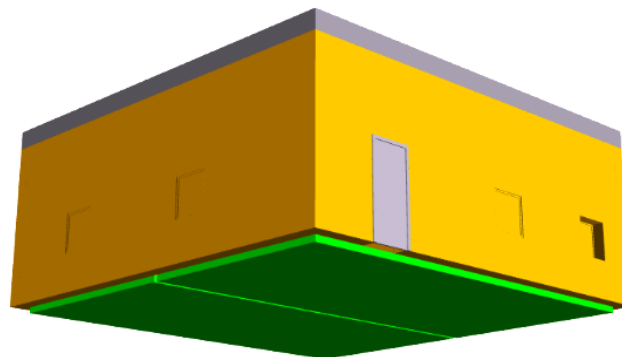
Figure 46 Visual risk identification generated by the Planning consult.
Risk is determined with the amount of tasks finished on time in more than 99%: Green (No risk, like floors). Amount of tasks finished on time in 50% to 99%: Orange (Low risk, like walls and windows).Amount of tasks finished on time in less than 50%: Red (High risk, no elements in this model). Element which aren't built before have a grey colour (like roof and door).

Kanaalplaat

In the model are found 2 elements of this type of product.

Planning advice

Amount of elements in the database: 46

In the planningsdatabase the following process variants are found

Variant 1: 100 % of the elements from this product type are planned according to these steps:

Task number	Task name	Percent on time	Percent too late	Percent unknown
1	ST00350 Kanaalplaatvloer leggen (lev. week 16)	100%	0%	0%

Stelkozijn

In the model are found 9 elements of this type of product.

Planning advice

Amount of elements in the database: 135

In the planningsdatabase the following process variants are found

Variant 1: 28 % of the elements from this product type are planned according to these steps:

Task number	Task name	Percent on time	Percent too late	Percent unknown
1	ST00970 Plaatsen kozijnen	84%	15%	1%

Variant 2: 72 % of the elements from this product type are planned according to these steps:

Task number	Task name	Percent on time	Percent too late	Percent unknown
1	ST00940 Stelwerk buitengevel	85%	14%	1%
2	ST00970 Plaatsen kozijnen	84%	15%	1%

Figure 47 Planning consult including process variants and risk identification per step.

8.6.2 Applications of the Planning consult

In consultation with the project manager this is seen as an interesting concept which enables fact based improvements. However the prototype needs more development to be actually useful, but the concept is interesting according to the project manager. Nowadays the companies are continuously trying to improve processes. However this is feeling based. The concept such as a Planning consult gives clear identification of processes around elements which are most inefficient. In addition the project manager told that each new project is analysed on risks which results in a document called 'Risk Matrix Overview' (RMO). The document is made during the calculating phase. Elements with high risks are named in this document and control measurements are described. In addition often extra budget is reserved in order to tackle the risk. However the RMO is based on experience and feeling of the

calculators. The risk identification function in the Planning consult could be useful to apply in the RMO. Definitely more research should be done in order to prove the actual value, however another potential application of event logs is discovered.

8.6.3 Conclusion and discussion reuse

The usage of the Planning consult is tested with an event log from just one project, this is a very small scale. It would be better to study with more projects combined in one event log database. This will result in a better risk identification.

Probably more, better and advanced reusing methods can be done. The proposed reusing method is just one option. However the proposed workflow may inspire some to study this topic even more and find more appropriate ways of reusing data within the construction industry.

8.7 Conclusion and discussion case study

This case study gives an example of tools which can supplement each other to improve process planning, capturing, analysing and reusing. However this study still needed to align all tools and data formats. Since this approach serves to demonstrate the system's functionality no definitions, specifications, ontologies, rules, knowledge bases or standards for process definitions were considered. Further research may increase the quality of incorporating such terms which should make it possible to streamline this whole workflow in order to enable automated process planning, capturing, analysing and reusing.

This study used drone images to capture deviation which are not suitable for recognizing elements inside buildings, therefore other technologies can be used (as mentioned in chapter 3.1.3.3). In addition causes of deviations can be found in the camera images and enables one to do fact based problem identifying. Moreover it is noticed that the data in form of an event log can be easily analysed for bottlenecks on a project level, and merged in an event log database which is useful on a company level.

Within this study event logs are stored in .csv files. However when large amounts of event log are generated and combined more comprehensive databases will be necessary. In addition .csv files where useful during this study, but it would recommendable to study the XES data format for more comprehensive and interchangeable process mining applications within the construction industry. In addition an open source java library for reading, storing, and writing XES is available (see: http://www.xes-standard.org). Since BIMserver is a java based application aligning the XES library within BIMserver can be useful to study.

9 Conclusions and recommendations

9.1 State-of-the-art

This research targets on shortening construction projects by identifying bottlenecks and planning deviation. In order to reach this target state-of-the-art technologies are studied and a workflow is proposed. It is seen in chapter 3.1.3.2 that BIM based planning tools can be used to make useful models including planned tasks linked to building elements. Several monitoring technologies are available to measure the progress and planning deviation. As described in chapter 3.1.3.3 depending on the preferred measurement a technology can be suitable. The deviations of the as-planned model and as-built process can be measured, but no tool which can generate event logs from those models was available. By use of the Eventlog service, a tool which is made within this study, the as-planned model is translated into event logs. The as-built data is merged with those logs resulting in as-built event logs. The event log data format enables process mining analytics. As elaborated in chapter 3.2 several analytics can done. It is noticed in chapter 3.2.2 that with process mining tools bottlenecks from event logs can be discovered. The case study which is executed and described in chapter 8 proved, by use of the Plan-Capture-Analyse-Reuse workflow and several BIM and process mining tools, that it is possible to gain insight in bottlenecks and planning deviation. Thereby the practical target within this research ("Determining bottlenecks and planning deviations with process mining techniques and Building Information Models within the execution phase of construction projects by use of autonomously gained observation data.") is achieved. The following can be concluded and recommended:

Conclusion 1

Current state-of-the-art technologies regarding planning and monitoring systems for construction projects enable to autonomously measure progress. However those technologies are not perfectly aligned. Planning tools are focussed on making good schedules. Monitoring systems are focussed on capturing processes. But schedules or as-planned models are not directly reusable for monitoring purposes. However it is proven that such technologies can be part of a construction process and deliver valuable information.

Recommendation 1a

This first recommendation is addressed to managers at contractors: Think about ways and tools in order to enable structured data collection within every process within the company. Data forms a basis for fact based improvement. However data extraction should be driven by questions rather than the availability of lots of data. Organisations should have a clear target before starting process mining data preparation. Without a good goal it can be a time wasting project. Technologies and workflows used in this study can be applied in practice, but other options may be more suitable depending on the company and projects. Combine as-planned BIM and as-built event logs in order to gain product specific information. Event logs is a data structure which enables useful analytics, but storage of this data may be difficult. A possible solution as indicated in (van der Aalst, 2011) is one where companies can create a *process-oriented data warehouse* containing information about relevant events happened in the company. Possibilities for companies in the AEC industry such as contractors can be studied as well.

Recommendation 1b

This recommendation is addressed to software vendors: Align planning and monitoring tools and focus on structured process planning, process measurement, process data collection, process data storage and process data reuse, and not just on one of those aspects. This will enable a more complete service for companies in the construction sector. The tools developed within this study may act as inspiration and starting point.

Conclusion 2
BIM tools are able to make as-planned models which are suitable for project scheduling. At this moment as-planned BIM is mainly used on single projects and mainly for visualisation purposes. However companies are not aware of the value of those models and applications and reusability of the accompanied data. With process mining analytics in mind a contractor maybe finds a new purpose for his as-planned models and records its requirements in the BIM protocol.

Recommendation 2a
This recommendation is addressed to scientific researchers, project planners or BIM managers at contractors: Study the possibilities and requirements for BIM based process databases. As stated within this research parameters with GUIDs, classification codes and material names can be used to describe events of specific building elements. As experienced in this study it is required to adopt as-planned BIM preparation and agree on task names or task identification numbers which can describe events. However how does this have consequence for agreements in BIM protocols? And how should the requirements be checked during model checking?

9.2 Potentials for process mining in the construction industry
The technologies and methods proposed in this research can be implemented in current processes of construction companies. They will be useful and with the vision to reuse all captured data it will form the basis for knowledge reassurance and fact based improving. Based on the case study executed within this research it can be said that the insights gained with the proposed methods have positive influence on small projects. But is it noticed that the influence will be higher when more and bigger projects will be studied. The knowledge database which is captured will be more interesting when more complex projects occur. As a result it is questionable if the theoretical objective of this study ("To realize a shorter construction time for the execution phase of construction projects".) will be realized, this depends on the implementation of the proposed philosophy, workflow and accompanying tools by practitioners within the construction industry. As concluded in the case study only a part of the total construction process is captured, namely only the part which is executed by the main contractor. Processes executed by subcontractors or other stakeholder where not clearly captured. The literature study clearly stated that a large amount of stakeholders are involved in the construction process and this causes nuisance in the project. Also within plan- and maintain- phase a lot of stakeholders are involved. The experimental case studies as described in chapter 3.2.5 concluded that in those phases already IT systems are used which have data which can result in useful insights with process mining analytics. This understanding has led to conclusions wherefrom recommendations can be done:

Conclusion 3

Due to the fragmentation, large projects, and complex processes, companies within the construction sector not always have clear sight on their processes. A lot of those processes are managed with IT systems. Especially in the plan- and maintain phase of buildings several tools are already applied. As a result of the experimental case studies within this research it can be concluded that some of these tools do not have functionality to give clear insight in overall processes. However the data which is in those tools can be extracted and analysed with process mining algorithms.

Recommendation 3a
This recommendation is targeted to data scientist, process mining enthusiasts, or companies which offer business intelligence services: Do a market analysis for big data services and potentials in the construction industry. In addition perform more case studies to prove the value of those services. Interviews with domain experts revealed that several companies are interested in such services. Since data is available and existing algorithms are suitable such services are directly applicable, however companies have to be convinced of the power of such analytics.

10 Discussion and further research

A workflow is proposed in order to enable continuously learning loops. This workflow is probably one of many methods in order to reach the higher target of shortening construction projects. However the workflow and used tools are proven as useful within this project. But it is good to emphasize that it is not about which technology will be used, but more about how the philosophy of continuously learning will be integrated in the company culture. Technology is just an instrument to achieve the target. Still the usefulness of process mining within the construction industry looks valuable. Since this study has applied it on just one project and it has already benefits, then it is logical to say that process mining on a larger scale will be useful even more. However without data those analytics will not be enabled. So capturing of processes is an important part. But does every contractor then has to implement advanced capturing technologies? It is seen that technology will be implemented more and more in order to manage every day processes. When a contractor choses new technology he can think about the applicability of data reuse and if it can be part of a continuous data driven learning culture within the company. When complex projects occur where the overall picture is difficult to manage, then process mining applications will be most useful. When companies say they manage everything and measuring progress is not necessary this can be questionable. When you think about construction companies as top athletes, which are measuring every detail of their performance in order to improve and become the best, than process mining will be suitable to identify improvement points within details for construction companies.

This study evokes for capturing construction processes. The implementation of monitoring systems, such as intelligent camera systems, may result in privacy concerns of the people who are getting monitored. This study did not take into account the privacy aspect. However process improvement should never be meant to blacken individual persons. It is good to know that anonymizing the data is always possible and so it should not be a big deal, but that is an assumption and may be an topic to research.

The tools proposed in this study are far from perfect since some of them are prototypical made for this study (like Eventlog generator and Planning consult). However the researcher indicated that a combination of technologies and tools may enable new learning methods. This may inspire researchers and companies to study the combination of process mining and BIM in order to improve construction projects. Since both research fields are gaining more and more momentum they can supplement each other even more in the future. Researchers can start with exploring the dataset which is published along with this study, see appendix 14.

From the literature study it is stated that complexity in construction projects and stagnating learning curves is mainly caused by changing stakeholders. During the study to process mining in the design phase it was indicated that social network mining was valuable. However during the case study it was not possible to integrate actor information. It is recommendable to further study the possibility to gain insight in stakeholder involvement in construction projects by use of process mining.

BIM based process planning may be implemented at contractors, but may also be integrated at subcontractors and suppliers. If this happens it will be possible to arrange supply chain alignment by combining those data logs. As a result process mining will not only be possible on project- and company level but also on an industry level will be enabled (the conceptual

idea can be seen in Figure 48). This application of BIM based process planning can be interesting to study.

As a result of demands of clients and laws on private quality control, as-built BIM models are soon likely to be required in all construction projects (JBKnowledge, 2015) An event log approach may add the process of how the as-built BIM has come into existence. The applicability of an BIM based event log approach may be subject for studies.

Process mining is just one part of the data science research field. More interesting data science areas are gaining momentum. Since data is generated more and more within the construction industry it seems logical to study potentials of Computational Intelligence Approaches. This is described as a family of the process discovery research field (van der Aalst, 2011). Examples of techniques are colony optimization, genetic programming, genetic algorithms, simulated annealing, reinforcement learning, machine learning, neural networks, fuzzy sets, rough sets, and swarm intelligence. Possibilities of these comprehensive techniques within the construction industry will be interesting to study.

Conceptual scheme and ideas of BIM based Process mining applications on Project -, Company, and Industry level
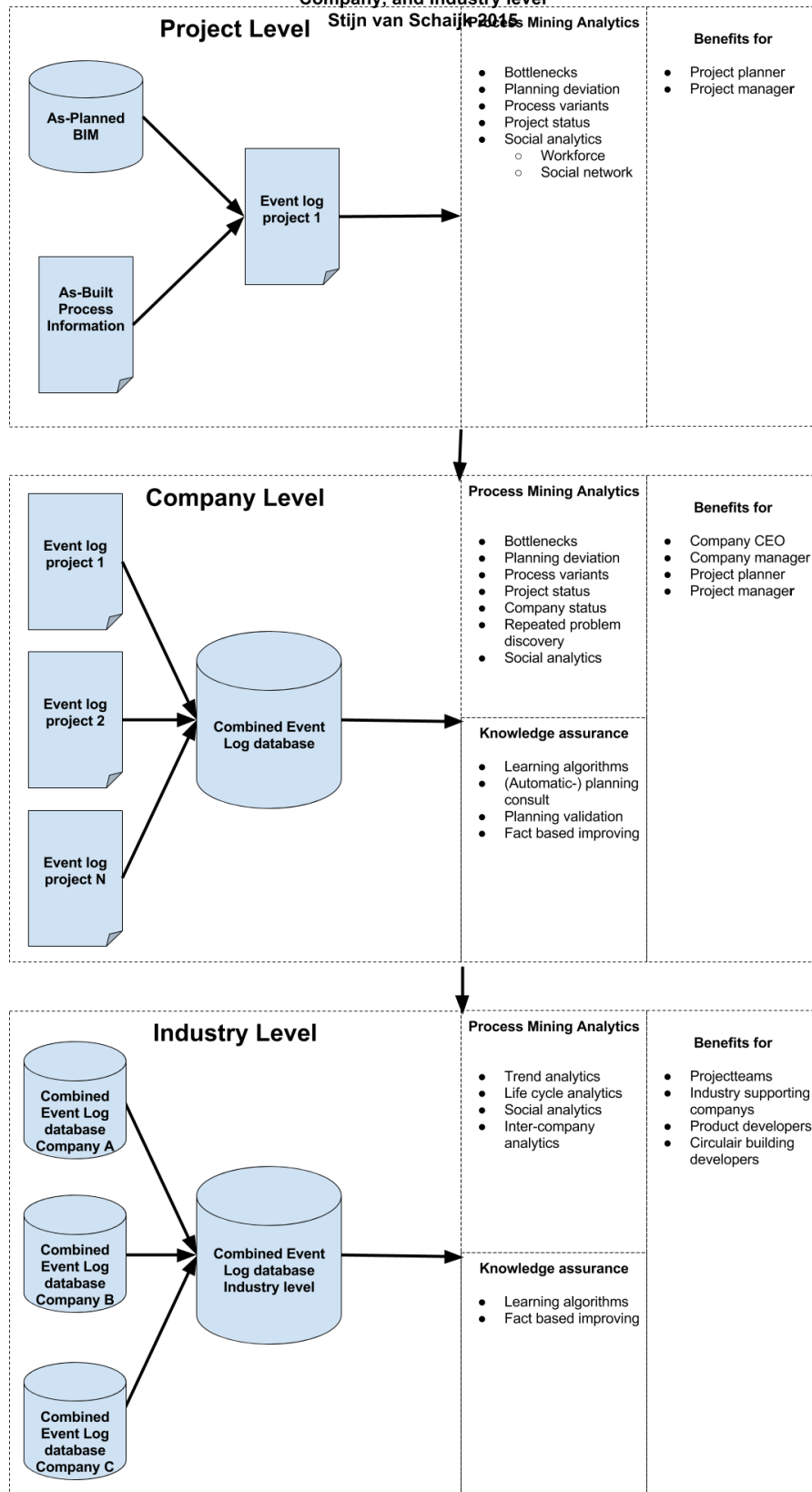
Project Level Stijn van Schaijle 2015

Figure 48 Conceptual ideas of process mining applications on project-, company- and industry level.

11 References

Aziz, R. F., & Hafez, S. M. (2013). Applying lean thinking in construction and performance improvement. *Alexandria Engineering Journal*, *52*(4), 679–695. http://doi.org/10.1016/j.aej.2013.04.008

Beerda, E. (2010). ProRail stuk goedkoper uit. *Cobouw*. Retrieved from http://www.cobouw.nl/artikel/865546-prorail-stuk-goedkoper-uit

Bouwkennis. (2013). Faalkosten Uiting, oorzaken, preventie remedies.

Bügler, M., Ogunmakin, G., Teizer, J., Vela, P. a, & Borrmann, A. (2014). A Comprehensive Methodology for Vision-Based Progress and Activity Estimation of Excavation Processes for Productivity Assessment. *Eg-Ice*, 1–10.

BuildingSmart. (2016a). IFC Industry Foundation Classes. Retrieved January 12, 2016, from http://www.ifcwiki.org/index.php/Main_Page

BuildingSmart. (2016b). IfcTask. Retrieved January 28, 2016, from http://www.buildingsmart-tech.org/ifc/IFC2x3/TC1/html/ifcprocessextension/lexical/ifctask.htm

Bushell, S. (1992). Implementing Plan, Do, Check and Act. *The Journal for Quality and Participation*, *15*(5), 58. Retrieved from http://search.proquest.com/docview/219169994/abstract

Caldas, C. H., Soibelman, L., & Gasser, L. (2005). Methodology for the Integration of Project Documents in Model-Based Information Systems. *Journal of Computing in Civil Engineering*, *19*(1), 25–33. http://doi.org/10.1061/(ASCE)0887-3801(2005)19:1(25)

Chang, S., & Jeong, W. (2015). BIM-Integrated Construction Operation Simulation for Reliable Production Management. In *CIB W78* (pp. 89–97). Eindhoven.

Collier, E., & Fischer, M. (1995). *Four-dimensional modeling in design and construction.*

Cox, A., & Thompson, I. (1997). "Fit for purpose" contractual relations: determining a theoretical framework for construction projects. *European Journal of Purchasing & Supply Management*, *3*(3), 127–135. http://doi.org/10.1016/S0969-7012(97)00005-1

Cox, R., & Goodman, C. S. (1956). Marketing of Housebuilding Materials. *The Journal of Marketing*, *21*(1), 36–61. http://doi.org/10.2307/1247698

Dave, B., Kubler, S., Främling, K., & Koskela, L. (2015). Opportunities for enhanced lean construction management using Internet of Things standards. *Automation in Construction*, *61*, 86–97. http://doi.org/10.1016/j.autcon.2015.10.009

Dimitrov, A., & Golparvar-Fard, M. (2014). Vision-based material recognition for automated monitoring of construction progress and generating building information modeling from unordered site image collections. *Advanced Engineering Informatics*, *28*(1), 37–49. http://doi.org/10.1016/j.aei.2013.11.002

Dubois, A., & Gadde, L. E. (2001). The construction industry as a loosely coupled system–implications for productivity and innovativity. *Paper for the 17th IMP Conference, Oslo, Norway*, (September), 9–11. http://doi.org/10.1080/01446190210163543

Froese, T., Fischer, M., & Ggrober, F. (1999). INDUSTRY FOUNDATION CLASSES FOR PROJECT MANAGEMENT - A TRIAL IMPLEMENTATION. *Journal of Information Technology in*

Construction (ITCON), *4*(November), 17–36.

Froese, T. M. (1999). *INDUSTRY FOUNDATION CLASS MODELING FOR ESTIMATING AND SCHEDULING*. *Durability of Building Materials*.

Fu, C., Aouad, G., Lee, A., Mashall-Ponting, A., & Wu, S. (2006). IFC model viewer to support nD model application. *Automation in Construction*, *15*(2), 178–185. http://doi.org/10.1016/j.autcon.2005.04.002

Gann, D. M. (1996). Construction as a manufacturing process? Similarities and differences between industrialized housing and car production in Japan. *Construction Management and Economics*, *14*(5), 437–450. http://doi.org/10.1080/014461996373304

Gidado, K. I. (1996). Project complexity: The focal point of construction production planning. *Construction Management and Economics*, *14*(3), 213–225. http://doi.org/10.1080/014461996373476

Gold, E. M. (1967). Language identification in the limit. *Information and Control*. http://doi.org/10.1016/S0019-9958(67)91165-5

Golparvar-Fard, M., Pena-Mora, F., & Savarese, S. (2009). D4AR - a 4 dimensional augmented reality model for automation construction progress monitoring data collection, processing and communication. *Journal of Information Technology in Construction*, *14*(June), 129–153.

Harris, R. B., & Ioannou, P. G. (1998). Scheduling Projects with Repeating Activities. *Journal of Construction Engineering and Management*, *124*(4), 269. http://doi.org/10.1061/(ASCE)0733-9364(1998)124:4(269)

Horn, B. K. P. (1987). Closed-form solution of absolute orientation using unit quaternions. *Journal of the Optical Society of America A*, *4*(4), 629. http://doi.org/10.1364/JOSAA.4.000629

Houdt, S. Van Den, & Vracken, J. L. M. (2013). Rolling out Systems Engineering in the Dutch Civil Construction Industry, 1–9.

Huhnt, W., & Enge, F. (2006). Can algorithms support the specification of construction schedules. … *Journal of Information Technology in Construction*, *11*(April), 547–564. Retrieved from http://itcon.org/cgi-bin/works/Search&noform=1&max=100&search=year:2012&title=Volume/Show?2006_39

Incose. (2012). *Systems Engineering Handbook* (3rd ed.). San Diego: Incose.

JBKnowledge. (2015). *The 4th Annual Construction Technology Report*. *Construction Technology Surveys*.

Jung, Y., & Gibson, G. E. (1999). Planning for Computer Integrated Construction. *Journal of Computing in Civil Engineering*. http://doi.org/10.1061/(ASCE)0887-3801(1999)13:4(217)

Kassem, M., Dawood, N., & Chavada, R. (2015). Construction workspace management within an Industry Foundation Class-Compliant 4D tool. *Automation in Construction*, *52*, 42–58. http://doi.org/10.1016/j.autcon.2015.02.008

Kelley, J. E., & Walker, M. R. (1959). Critical-Path Planning and Scheduling. In *Papers presented at the December 1-3, 1959, eastern joint IRE-AIEE-ACM computer conference on - IRE-AIEE-ACM '59 (Eastern)* (Vol. 32, pp. 160–173). http://doi.org/10.1145/1460299.1460318

Kenley, R., & Seppänen, O. (2009). Location-based management of construction projects: Part of a new typology for project scheduling methodologies. *Proceedings - Winter Simulation Conference*, 2563–2570. http://doi.org/10.1109/WSC.2009.5429669

Kopsida, M., Brilakis, I., & Vela, P. A. (2015). A Review of Automated Construction Progress Monitoring and Inspection Methods. In *CIB W78* (pp. 421–431). Eindhoven.

Liebich, T. (2009). IFC 2x Edition 3.

Lowe, D. G. (2004). Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, *60*(2), 91–110. http://doi.org/10.1023/B:VISI.0000029664.99615.94

Lu, W., Huang, G. Q., & Li, H. (2011). Scenarios for applying RFID technology in construction project management. *Automation in Construction*, *20*(2), 101–106. http://doi.org/10.1016/j.autcon.2010.09.007

Luci, G., & Shen, Z. (2015). A Critical Review of Recent Progress on Construction Workspace Planning Using BIM. In *CIB-W78* (pp. 506 – 517). Eindhoven.

Ma, Z., & Lu, N. (2010). Approach to representing information resources in construction firms by using IFC based on object-oriented method. *Proceedings of the International Conference on Computing in Civil and Building Engineering (ICCCBE)*.

Martyn, S. (2015). Case Study Research Design - How to conduct a Case Study. Retrieved June 30, 2015, from https://explorable.com/case-study-research-design

Meadati, P., Irizarry, J., & Amin, K. a. (2010). BIM and RFID Integration: A Pilot Study. *Advancing and Integrating Construction Education, Research and Practice*, 570–578.

Mikulakova, E., König, M., Tauscher, E., & Beucke, K. (2010). Knowledge-based schedule generation and evaluation. *Advanced Engineering Informatics*, *24*(4), 389–403. http://doi.org/10.1016/j.aei.2010.06.010

Nerode, A. (1958). Linear automaton transformations. *Proceedings of the American Mathematical Society*, *9*(4), 541–541. http://doi.org/10.1090/S0002-9939-1958-0135681-9

Paszkiewicz, Z. (2015). Case Study: Process Mining for Analyzing Inventory Processes — Flux Capacitor. Retrieved June 30, 2015, from http://fluxicon.com/blog/2014/01/case-study-process-mining-for-analyzing-inventory-processes/

Petri, C. A. (1962). Kommunikation mit Automaten. *Fakultät Für Mathematik Und Physik*.

Pheng, L. S., & Hui, M. S. (1999). The application of JIT philosophy to construction: a case study in site layout. *Construction Management and Economics*, *17*(5), 657–668. http://doi.org/10.1080/014461999371268

Picard, W., Paszkiewicz, Z., Strykowski, S., Wojciechowski, R., & Wojciech, C. (2014). *Advanced SOA Tools and Applications* (Vol. 499). http://doi.org/10.1007/978-3-642-38957-3

Quirijnen, R., & van Schaijk, S. (2013). *Meten = Weten*. Avans Hogeschool, Tilburg.

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

Saidi, K. S., Lytle, A. M., & Stone, W. C. (2003). Report of the NIST Workshop on Data Exchange Standards at the Construction Job Site. In *Proceedings of the 20th International Symposium on Automation and Robotics in Construction* (pp. 617–622).

Sayer, N. J., & Williams, B. (2013). *Lean voor Dummies* (second edi). Amsterdam: Pearson Benelux bv.

Snavely, N., Seitz, S. M., & Szilski, R. (2006). Photo Tourism: Exploring Photo Collections in 3D. *Proceedings of ACM Transactions on Graphics*, *Vol. 25*(No.3), 835–846. Retrieved from http://phototour.cs.washington.edu/Photo_Tourism.pdf

Song, M., & van der Aalst, W. M. P. (2008). Towards comprehensive support for organizational mining. *Decision Support Systems*, *46*(1), 300–317. http://doi.org/10.1016/j.dss.2008.07.002

Straub, A. (2009). Dutch standard for condition assessment of buildings. *Structural Survey*, *27*(1), 23–35. http://doi.org/10.1108/02630800910941665

Tah, J. H. M., Carr, V., & Howes, R. (1999). Information modelling for case-based construction planning of highway bridge projects. *Advances in Engineering Software*, *30*, 495–509. http://doi.org/10.1016/S0965-9978(98)00128-8

Terlouw, L., & Mulder, H. (2014). Schatgraven in bouwprojecten. *Informatie*, *1408*, 32 – 29.

U.S. Department of Energy. (2008). *SENSITIVE Managing Design and Construction Using Systems Engineering for Use with*. *DOE O 413.3A*. Washingtoc, D.C.

van der Aalst, W. M. P. (2010). Business Process Simulation Revisited. *Enterprise and Organizational Modeling and Simulation, 6th International Workshop, EOMAS 2010, Held at CAiSE 2010, Hammamet, Tunisia, June 7-8, 2010. Selected Papers*, 1–14. http://doi.org/10.1007/978-3-642-15723-3_1

van der Aalst, W. M. P. (2011). *Process Mining: Discovery, Conformance and Enhancement of Business Processes*. *Media* (Vol. 136). http://doi.org/10.1007/978-3-642-19345-3

van der Aalst, W. M. P., Reijers, H. a., & Song, M. (2005). Discovering social networks from event logs. *Computer Supported Cooperative Work*, *14*(6), 549–593. http://doi.org/10.1007/s10606-005-9005-9

van der Aalst, W. M. P., Van Dongen, B. F., Herbst, J., Maruster, L., Schimm, G., & Weijters, a. J. M. M. (2003). Workflow mining: A survey of issues and approaches. *Data and Knowledge Engineering*, *47*(2), 237–267. http://doi.org/10.1016/S0169-023X(03)00066-1

van Merriënboer, S., Dijkmans, T., Klerks, S., van Berlo, L., & Ree, D. De. (2013). BIM bouwlogistiek rekenmodel. *TNO-Rapport*.

Vrijhoef, R., & Koskela, L. (2000). The four roles of supply chain management in construction. *European Journal of Purchasing & Supply Management*, *6*(3-4), 169–178. http://doi.org/10.1016/S0969-7012(00)00013-7

Wang, H ., Zhang, J ., Chau, K ., & Anson, M. (2004). 4D dynamic management for

construction planning and resource utilization. *Automation in Construction*, *13*(5), 575–589. http://doi.org/10.1016/j.autcon.2004.04.003

Winch, G. M. (1987). The construction firm and the construction process: the allocation of resources to the construction project. In *C.I.B. W-65*. London.

Winch, G. M., & Kelsey, J. (2005). What do construction project planners do? *International Journal of Project Management*, *23*, 141–149. http://doi.org/10.1016/j.ijproman.2004.06.002

Yang, J., Park, M.-W., Vela, P. a., & Golparvar-Fard, M. (2015). Construction performance monitoring via still images, time-lapse photos, and video streams: Now, tomorrow, and the future. *Advanced Engineering Informatics*. http://doi.org/10.1016/j.aei.2015.01.011

Zhang, J., Zhang, Y., Hu, Z., & Lu, M. (2008). Construction Management Utilizing 4D CAD and Operations Simulation Methodologies. *Tsinghua Science and Technology*, *13*(SUPPL. 1), 241–247. http://doi.org/10.1016/S1007-0214(08)70156-1

Zwikael, O. (2009). Critical Planning Process in Construction Projects. *Construction Innovation*, *9*(4), 372–387. http://doi.org/10.1108/14714170910995921

12 Appendix Case study: Construction design process mining

The paper is published on the next page, but can also be downloaded at:
http://www.slideshare.net/StijnvanSchaijk/case-study-construction-design-process-mining

Case study: Construction design process mining

Exploring the applications of process mining techniques within the construction design phase.

08-01-2016, Eindhoven University of Technology

Stijn van Schaijk, s.v.schaijk@student.tue.nl

Abstract

Techniques that improved performance in other industries are more often adopted in the construction industry such as Systems Engineering (SE). Interviews with SE experts conducted within this study have indicated that SE IT systems supports with managing their project and helps to prove that the clients specifications are realized. However it is not known if those systems support an efficient process. Practitioners indicated that a lot of people are involved in such projects who all work in parts of the IT systems but nobody has a clear overview of the total process. This study explores the possibilities of discovering parts of the design process with process mining techniques. By use of a case study at a large civil project the potential of process mining within the design phase of construction projects will be discovered. It took some time to prepare the data but valuable analytics could be done with process mining techniques. Real bottlenecks are found, process variants are discovered, social networks are exposed and improvements in the design process can be made. Due to this experimental study the process engineers realized that the organization doesn't control the field of information- and data modelling enough. It is concluded that it is valuable to use process mining to give continuously feedback loops to the project managers at contractors. Process mining gives unique visualisations which enables refreshing insights. In addition process mining analytics did realize that some IT systems are used to store information, but those systems don't automatically facilitate an efficient process.

Keywords: Process mining, construction design, case study, systems engineering, data.

1. Introduction

It is known that construction projects are amongst the most complex of all activities (G.M. Winch, 1987). In addition (Gidado, 1996) indicated that "there is continuous demands for speed in construction, cost and quality control, safety in the work place and avoidance of disputes, together with technological advances, economic liberalization and globalization, environmental issues and fragmentation of the construction industry"(p.3). This results in a rapid increase of complexity of construction processes in all phases of a project. Designing complex buildings demands a subsequent problem solving capability by the involved stakeholders. In order to manage the clients demands and construction specifications within the design phase information systems (IT systems) are more often used (JBKnowledge, 2015). In addition techniques that improved performance in other industries are adopted such as Systems Engineering (SE) (Houdt & Vracken, 2013; U.S. Department of Energy, 2008). Interviews with SE experts conducted within this study have indicated that SE IT systems supports with managing their project and helps to prove that the clients specifications are realized. However it is not known if those systems support an efficient process. Practitioners

indicated that a lot of people are involved in such projects who all work in parts of the IT systems but nobody has a clear overview of the total process. Questions occur like: How are specifications generated and approved? How long does it take to approve an specification? Who is involved in most process steps? How often changes this responsibility? Is the process of approving specifications always the same? If not, which kind of specifications have an efficient process and which ones don't? The research field of process mining may bring an answer to those questions. Process mining provides an approach to gain insight and improve processes in a lot application domains. The goal of process mining is to gain event data, extract process-related information and discover a process model.

The outline of this paper is as follows, first the research goals and proposed method will be elaborated, than some background about SE and process mining will be given. Thereafter a case study is described and finally the conclusion and discussion will be elaborated.

2. Research

Research goals

This study explores the possibilities of discovering parts of the design process with process mining techniques. By use of a case study at a large civil project the potential of process mining within the design phase of construction projects will be discovered.

Method

Because little research is done to process mining within SE design phases this study has an experimental purpose. By use of a case study the possibilities and usefulness will be explored. A civil project at a large contractor is chosen as a case because a lot of data was available in the SE IT system. In addition the process engineers of the contractor are curious about the applicability of process mining and where enthusiastic to help with this study. The first research step is to explore if the data is directly useful for process mining, in other words: can event logs directly be exported from the SE tools database? If yes, analysis can be done. If no, how should the data be transformed into event logs? And how much effort does this take? If it is possible to prepare the data then specific analysis can be done and questions of the process engineers can be answered.

3. Background

Systems engineering

The ultimate goal of Systems Engineering (SE) as defined by Houdt and Vracken (2013) is "to create a successful system as a whole solution to some complex problem" (p.2). The client of large construction projects increasingly demands more transparency in order to safeguard his functional requirements (Houdt & Vracken, 2013). Both clients and contractors are adopting SE as a method to facilitate the realization of construction projects. Previous research has proven that SE can increase the overall performance of complex engineering projects, reducing the overall schedule, and increasing product quality (Beerda, 2010;Incose, 2012). However as indicated by (Houdt & Vracken, 2013) wrong usage of SE can result ineffective processes and hard implementation. This opinion is verified by the process engineers interviewed within this research. Therefore they supported this research in order to find out how requirements or specifications are validated and verified using SE databases.

Process mining and Event logs

This paragraph describes how an event log should look like to enable construction design process mining from a 'SE specification perspective'. Event logs in general will be described shortly, in addition the requirements regarding event logs to enable process mining in systems engineering will be described.

Using figure 1 as a reference, the following is assumed regarding event logs (van der Aalst, 2011):
- A process consists of cases.
- A case consists of events such that each event related to precisely one case.
- Events within a case are ordered.
- Events can have attributes. Examples of attributes are activity, time, costs, and resource.

Figure 1 Structure of event logs

The bare minimum to enable process mining requires a Case ID and an Activity log:
1. Case ID: A case identifier, also called process instance ID, is necessary to distinguish different executions of the same process. What precisely the case ID is depends on the domain of the process. For example in a hospital this would be the patient ID.
   In the case of construction process mining with systems engineering the case ID would be the unique name or number of the specification
2. Activity/Event ID: There should be names for different process steps or status changes that were performed in the process. If you have only one entry (one row) for each process instance, then your data is not detailed enough.
   The data needs to be on the transactional level (you should have access to the history of each case) and should not be aggregated to the case level.

In the case of SE process mining there should be a history of the specifications.

To analyse performance related properties or additional attributes are useful in the event log like:
3. Timestamp: At least one timestamp is needed to bring events in the right order. This time stamp is also needed to identify delays between activities and bottleneck identification. It is useful if an event has an start- and an end timestamp
4. Resources: For example the person or the company who executed the activity.
5. More attributes can be added, this will be useful to gain extra information while doing the analytics.

4. Case study

The case study is done at a large contractor in the Netherlands. They use a systems engineering tool (Relatics*)* to manage their complex projects. The case involves a large civil road project. The project is financed by the province and a lot of stakeholders (such as municipalities) are involved at the client side of the project. Each stakeholder has its own demands. During the initiation phase the demands where registered as specifications in the systems engineering tool. Every specification has to be implemented into the design. During the project every specification is continuously monitored and at the end every specification should be checked and validated. This ensures the quality of the project. The process engineers of the contractor wonder how this process happens, if it is efficient and who has the most prominent role in it. The systems engineering tool doesn't provide such analytics, so a data log is exported and analysed with process mining tools. This study explores if it is possible to discover process models from this log.

Process steps

The process engineers have a process template which describes the relations between process steps. Several steps can be done during the process of a specification. The process of a specification can be very complicated and can have a lot of variants. The steps which can be taken are described in table 1.

Table 1 Process steps description

Dutch process step name	English translation process step
Aanmaken/Importeren Eis	Create/Import Specification
Koppelen Eis aan Bovenliggende/Onderliggende Eis	Connect Specification to Upper/Lower Specification
Toewijzen Eistekst	Assign Specification description
Toewijzen controleverantwoordelijkheid	Assign control responsibility
Toewijzen Eis aan Projectstadium	Assign Specification to Project phase
Toewijzen Eis aan Persoon	Assign Specification to Person
Koppelen Eis aan Werkpakketactiviteit 1	Connect Specification to Work package activity 1
Koppelen Eis aan Werkpakketactiviteit 2	Connect Specification to Work package activity 2
Koppelen Eis aan Werkpakketactiviteit 3	Connect Specification to Work package activity 3
Aanmaken werk packet (WPA) verificatie	Create work package (WPA) verification
Toewijzen werk pakket (WPA) Verificatie aan Persoon	Assign work package (WPA) verification to Person
Toevoegen Bewijsdocument	Add document of prove
Plaatsen opmerking bij WPA Verificatie	Add note to WPA verification

Explore usefulness data

The raw data which was exported from the SE tool was distributed over 3 .csv files. Those files had to be merged using macro's in excel. It was a complicated task and took a few days to prepare the data. The process engineers succeeded in creating a structured event log with the data structure as visualised in table 2.

Table 2 Sample of event log created from datalog SE tool

EIS_ID	ACTOR	TIMESTAMP	ACTIVITY
Eis1698	persoon1	6-2-2014 08:26	Aanmaken/Importeren Eis
Eis1698	persoon1	6-2-2014 08:31	Koppelen Eis aan Bovenliggende/Onderliggende Eis
Eis1698	persoon1	6-2-2014 08:47	Toevoegen Eistekst
…	…	…	…

Note that the actors are anonymized.

This format is an event log and therefore suitable for process mining. The event log contains processes of 3002 specifications, 139.018 events happened over a time period starting at 06-02-2014 and ending at 19-06-2015. 65 people were involved in the project. The log is made while the project was half way so not every specifications' process was finished. Therefore a lot of logs are incomplete. During the analysis the researchers have taken this into account.

Analysis

The process mining tools Disco[34] and MyInvenio[35] are used to analyse the event log. The first assignment is to study the complete process happened over time within the SE tool. The whole process map is very complicated and can be seen in Appendix 1. The simplified version, where the most occurring process steps are included, can be seen in Figure 2. As a result of this figure the process engineers concluded there are two main process flows. First (1) which can be described as creating a specification and assigning a responsible person to it. And (2) the verification and validation process of the specification. From the fact that two main processes occur can be concluded that the connection between those main processes are very variable and not have a structured workflow. This is seen by the process engineers as ineffective.

---

[34] https://fluxicon.com/disco/
[35] https://www.my-invenio.com/
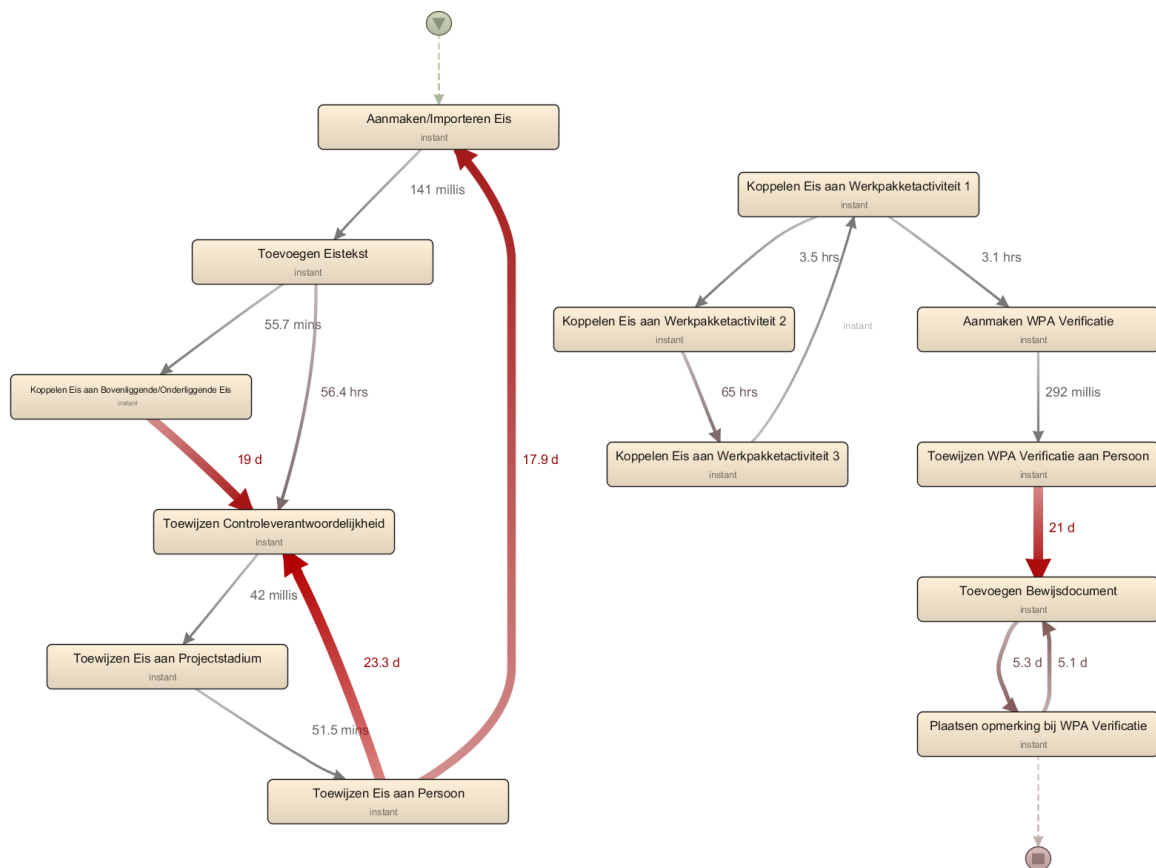
Figure 2 Processes within the SE tool filtered on most occuring process steps

Second question is to analyse the mean durations between creating an specification and validation the specification. In order to answer the question the data is filtered by the following query: Include only the cases which have a process step 'Aanmaken/Importeren eis (Create/import specification)' and 'Toevoegen Bewijsdocument (Add document of prove)'. This resulted in a dataset of 45% of the total specifications. The process map with the most occurring steps can be seen in Figure 3. This map is identified as a regular process by the process engineers. However the complete process map can be seen in appendix 2 and more complicated. It is concluded by the process engineers that several cases went as expected. But a lot of cases do not. 941 different process variants were discovered with a mean duration of 50.7 weeks, the process engineers identified this as a reliable result. The fastest process was done within 12 process steps and a duration of 92 days and 23 hours. The slowest process was done within 784 process steps and a duration of 1 year and 124 days. The case with the most steps contained 1903 process steps and had a duration of 1 year and 48 days. It is seen that a lot of differences are identified within the process of a specification. The process engineers concluded that in a lot cases the process is not happening as expected and that there is a lot of inefficiency. In addition they are willing to find the reasons of this variance and help the project team to improve the processes.
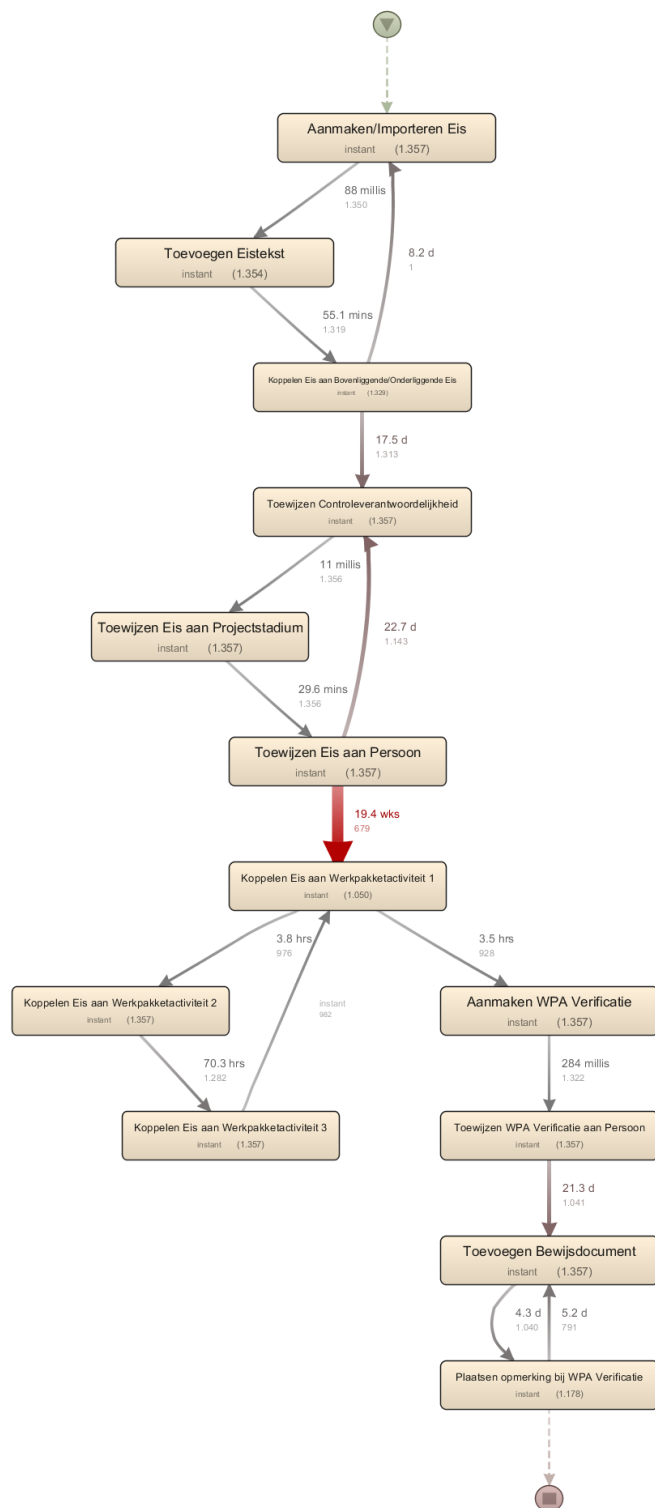
Figure 3 Most occurring process steps of generating until validating a specification

The third question which was of interest was: Who is involved in most process steps? It seemed that person 1 (38%) and person 2 (21%) have the most influence on the project. As can be seen in Figure 4. According to the process engineers this is strange because the projects has

65 people involved which all should contribute to the process. Such a large differences should not occur, as a result they concluded that not everyone is using the tools correctly.
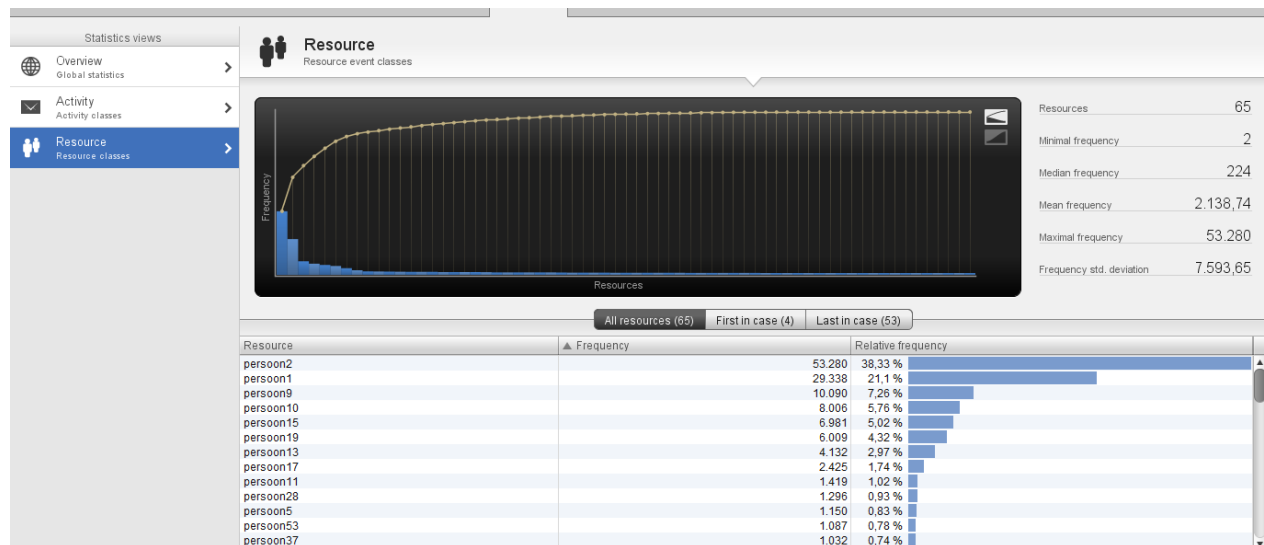
Figure 4 Statistics of persons involved in process steps

Fourth question was to analyse how often the responsibility of specifications changes from one person to another. This is important because failures can occur when responsibilities are changing during the project. The answer on this question is gained by Excel analytics and resulted in the histogram showed in Figure 5. In most situations the specification had just 1 person responsible during the process. But several cases the responsibility changed 2, 3, 4, 5, 6,7 or even 8 times. In the most extreme situation the responsibility changed 11 times. The reason of those changes were not directly found. This is for future research.
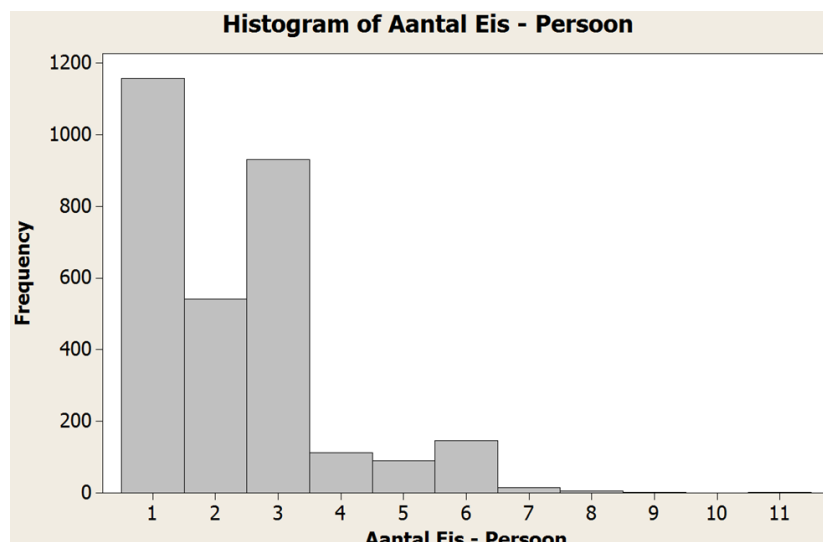
Figure 5 Histogram describing the amount of responsibility changes

As can be seen in Figure 6 and Figure 7 also analytics could be done about which people do similar tasks within the project. The process engineers identified 'knowledge-clusters' of people who do similar tasks, they see applications to connect people with the same responsibilities. In addition they discovered some people where very dominant in the usage of the tool during the project. They mentioned that in fact SE should trigger a process where everyone could participate, however the data proves that this was not the fact during this

project. However no hard conclusions can be given since this analytics was just on one project. Therefore more studies regarding social network mining in SE have to be done. Although it can be concluded that social network mining gives unique visualisations which enables refreshing insights.
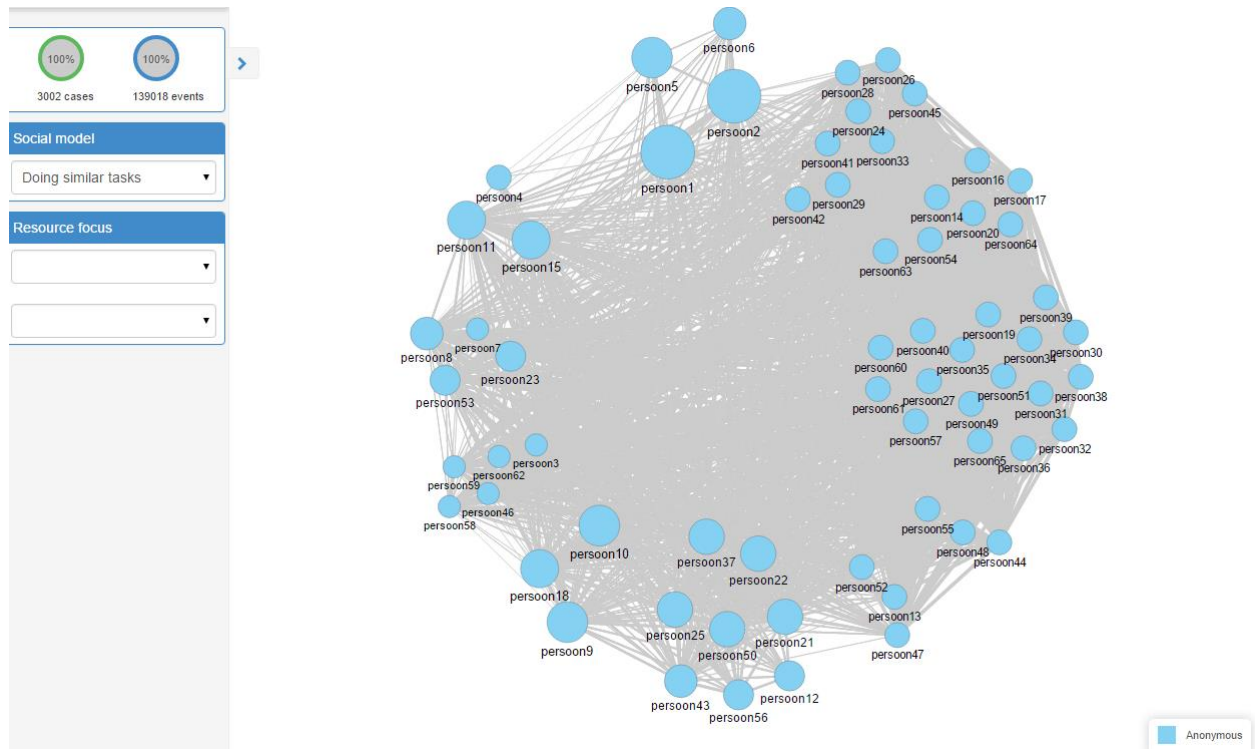
Figure 6 Network of people doing similar process steps
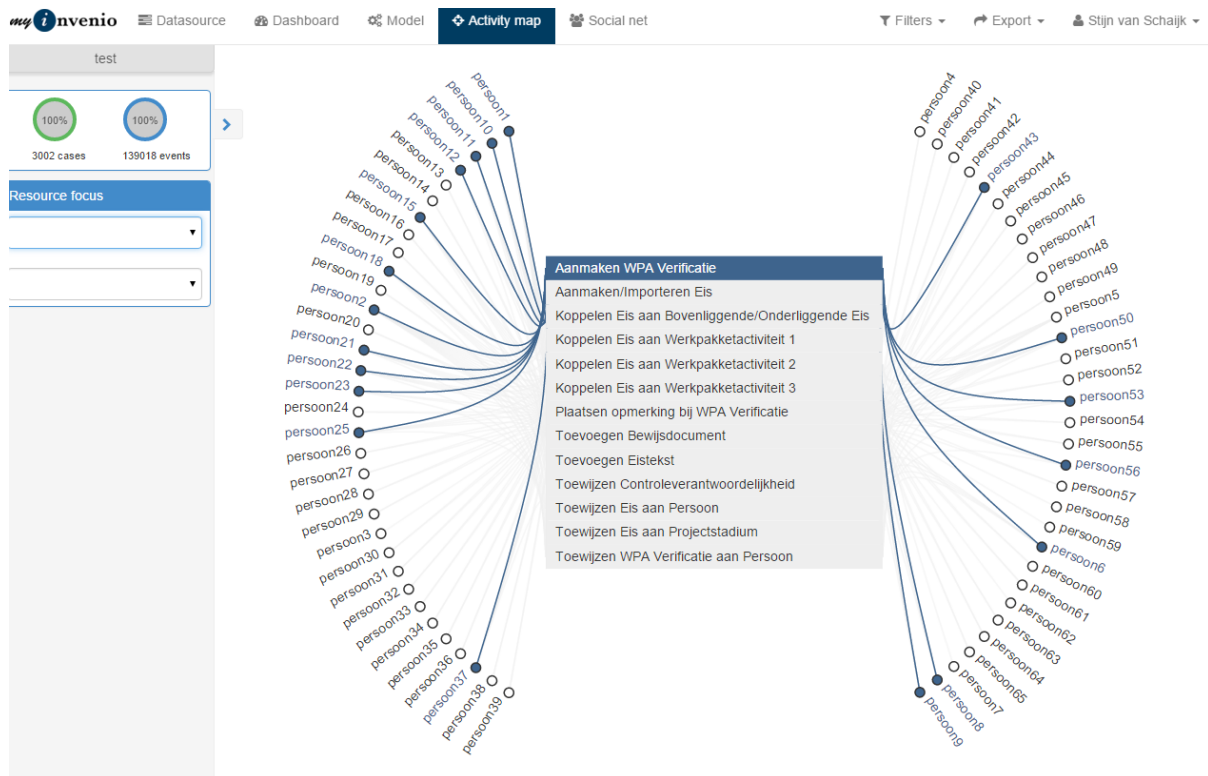
Figure 7 Example of people who are involved in specific process steps

5. Conclusion and discussion

This explorative study proved that it is possible to use Systems Engineering data to gain insight in the design process of complex projects. It took some time to prepare the data but valuable analytics could be done with process mining techniques. Real bottlenecks are found, process variants are discovered, social networks are exposed and improvements in the design process can be made. However the results look promising, they are generated out of data without considering the practical context of the results. Therefore the researcher and process engineers planned to do more case studies in order to monitor the influence of process mining in practice. However no hard conclusions can be given since this analytics was just on one project, it is concluded that it is valuable to use process mining to give continuously feedback loops to the project managers. Process mining gives unique visualisations which enables refreshing insights. In addition the social analytics gave insights which couldn't be generated before.

Due to this experimental study the process engineers realized that the organization doesn't control the field of information- and data modelling enough. People who are involved should have other competences which have to be developed in parts of the organization. The process engineers concluded with the following quote:

"Process mining analytics did us realize that we use our IT systems to store information, but those systems don't automatically facilitate an efficient process. "

6. Acknowledgements

7. References

Beerda, E. (2010). ProRail stuk goedkoper uit. Cobouw. Retrieved from http://www.cobouw.nl/artikel/865546-prorail-stuk-goedkoper-uit

Gidado, K. I. (1996). Project complexity: The focal point of construction production planning. Construction Management and Economics, 14(3), 213–225. http://doi.org/10.1080/014461996373476

Houdt, S. Van Den, & Vracken, J. L. M. (2013). Rolling out Systems Engineering in the Dutch Civil Construction Industry, 1–9.

Incose. (2012). Systems Engineering Handbook (3rd ed.). San Diego: Incose.

JBKnowledge. (2015). The 4th Annual Construction Technology Report. Construction Technology Surveys.

U.S. Department of Energy. (2008). SENSITIVE Managing Design and Construction Using Systems Engineering for Use with. DOE O 413.3A. Washingtoc, D.C.

van der Aalst, W. M. P. (2011). Process Mining: Discovery, Conformance and Enhancement of Business Processes. Media (Vol. 136). http://doi.org/10.1007/978-3-642-19345-3

Winch, G. M. (1987). The construction firm and the construction process: the allocation of resources to the construction project. In C.I.B. W-65. London.

Appendix 1: Process map total process
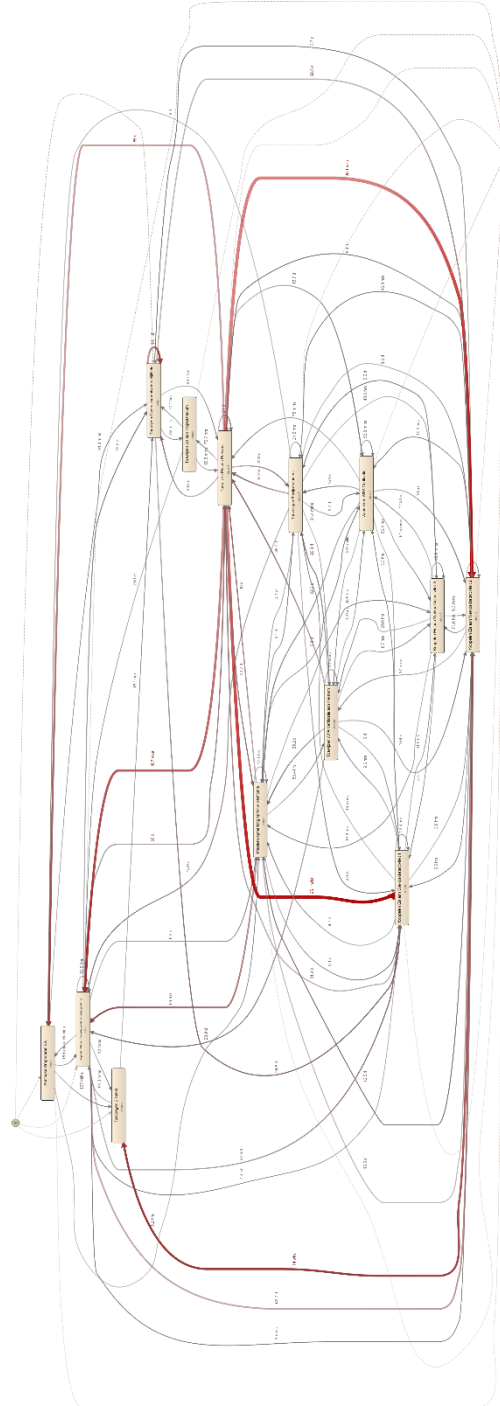
Appendix 2: Process map filter complete cases
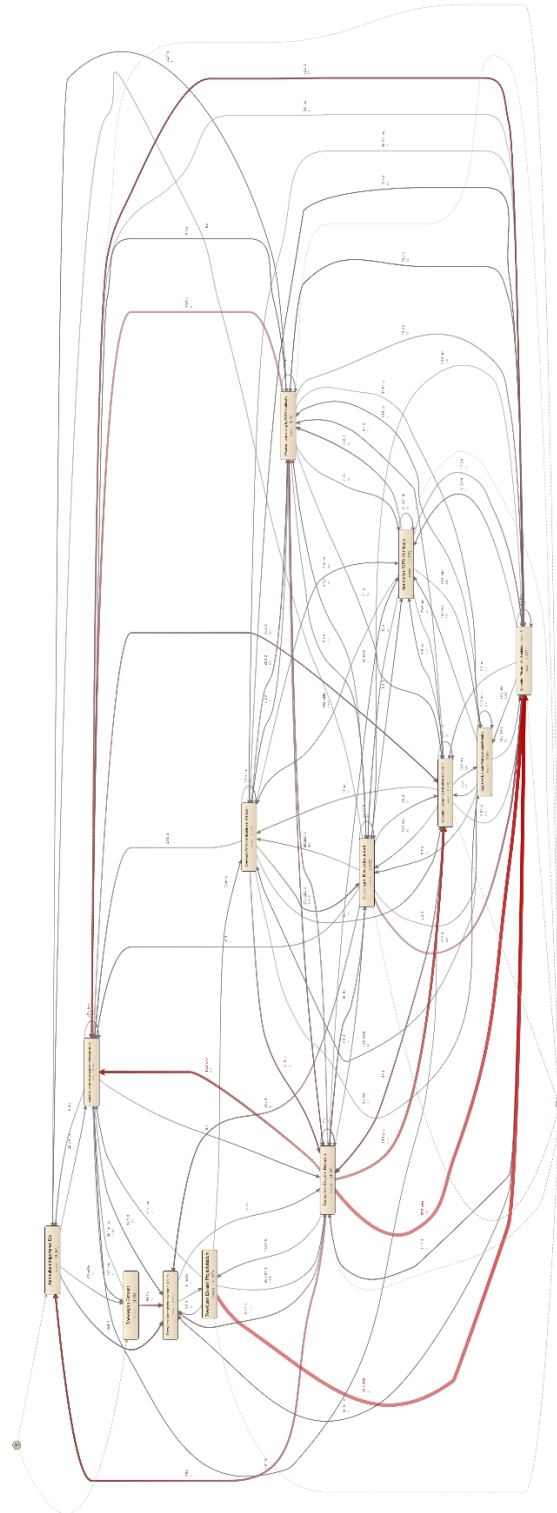
13 Appendix Case study: Process mining with facility management data

The paper is published on the next page, but can be downloaded at: http://www.slideshare.net/StijnvanSchaijk/case-study-process-mining-with-facility-management-data

Case study: Process mining with facility management data

Exploring the applications of process mining techniques with maintenance error data.

06-11-2015, Eindhoven University of Technology
Stijn van Schaijk, s.v.schaijk@student.tue.nl

Abstract

Facility managers, especially ones with a large amount of assets, use facility management systems to capture, store and plan maintenance data. Those systems are meant to plan and monitor current processes. But are hardly able to analyse processes which are already happened. In addition they are not able to compare different process variants of assets over longer time periods. Still this information is useful for a facility manager because he/she wants to learn from things which went wrong in the past. Therefore this study explores if facility management data is suitable for analysing processes around building elements with process mining techniques. By use of a case study at a hospital the potential of combining process mining with maintenance data will be discovered. It can be concluded that with some data transformation maintenance data is suitable for process mining. Moreover the facility managers were surprised by the visualisation techniques and they gained clear insight in the error handling process. As a result they discovered problematic building elements and odd processes.

Keywords: Process mining, facility management, case study, maintenance, data, error handling process.

1. Introduction

What is this type of door? What properties does it have? When is it fabricated? How many times is it broken? How many times does it need a paint job? When is it not fixable and goes it to a destruction place? Is the process of door X the one I expect it to be? And if it deviates, why is that? Does that happen to all my doors? Or to just this one? In conclusion: What happens within a life of a door? These questions occur when a facility manager has to determine his maintaining schedule and budget. In this example it describes a door, but those questions can be about any building element. Currently it is difficult for the facility manager to answer those questions. The answers can be found in data, but often it is not generated, it is not available in a structured way, or when it is, it can be hard to analyse it. Facility managers, especially ones with a large amount of assets, use facility management systems to capture, store and plan maintenance data. Those systems are meant to plan and monitor current processes. But are hardly able to analyse processes which are already happened. In addition they are not able to compare different process variants of assets. Still this information is useful for a facility manager because he/she wants to learn from things which went wrong in the past. The research field of process mining may bring an answer to those questions. Process mining provides an approach to gain insight and improve processes in a lot application domains. The goal of process mining is to gain event data, extract process-related information and discover a process model.

The outline of this paper is as follows, first the research goals and proposed method will be elaborated, than some background about process mining and event logs will be given. Thereafter the case study is described and finally the conclusion and discussion will be elaborated.

2. Research

Research goals

This study explores if facility management data is suitable for analysing processes around building elements with process mining techniques. By use of a case study at a hospital the potential of combining process mining with maintenance data will be discovered.

Method

Because no research is done to process mining with facility management data this study has an experimental purpose. By use of a case study the possibilities and usefulness will be explored. A hospital is chosen as a case because a lot of data is available within the facility management system. First step is to explore if the data is directly useful for process mining, in other words: can event logs directly be exported from the facility management database? If yes, analysis can be done. If no, how should the data be transformed into event logs? And how much effort does this take?

If it is possible to prepare the data then specific analysis can be done. To determine useful cases a brainstorm session will be held with the facility managers. Based on this session the use case with the most potential will be chosen.

3. Background

Process mining and Event logs

This paragraph describes how an event log should look like to enable construction process mining from a 'building element perspective'. Event logs in general will be described shortly, in addition the requirements regarding event logs to enable process mining in construction will be described.

Using figure 1 as a reference, the following is assumed regarding event logs:
- A process consists of cases.
- A case consists of events such that each event related to precisely one case.
- Events within a case are ordered.
- Events can have attributes. Examples of attributes are activity, time, costs, and resource.

Figure 1 Structure of event logs

The bare minimum to enable process mining requires a Case ID and an Activity log:

1. Case ID: A case identifier, also called process instance ID, is necessary to distinguish different executions of the same process. What precisely the case ID is depends on the domain of the process. For example in a hospital this would be the patient ID.

    In the case of construction process mining with BIM the case ID would be the GUID of the building element.

2. Activity/Event ID: There should be names for different process steps or status changes that were performed in the process. If you have only one entry (one row) for each process instance, then your data is not detailed enough.

    The data needs to be on the transactional level (you should have access to the history of each case) and should not be aggregated to the case level.

    In the case of construction process mining of a building element there should be a history of this element.

To analyse performance related properties additional attributes are useful in the event log like:

3. Timestamp: At least one timestamp is needed to bring events in the right order. This time stamp is also needed to identify delays between activities and bottleneck identification.

    Resources: For example the person or the company who executed the activity

    More attributes can be added, this will be useful to gain extra information in while doing the analytics.

4. Case study

The case study has taken place at the Erasmus MC hospital in Rotterdam. This case is chosen because a lot of maintenance data is available in the facility management system (Planon). In addition this case is interesting because the hospital consists of different buildings, with all kinds of different building elements with specific maintenance programs and schedules. Moreover several alerts come in every day about broken doors, elevators and other elements. Currently the hospital is making new buildings and has plans to expand and renovate current buildings. In those plans they want to choose the best types of products. They indicated that they are interested in insight in the maintenance and error handling process of their current building elements in order to choose the best products in the future. Moreover the facility management has a standard process flow to treat errors, they wonder if this process is actually happening in reality. The amount of data, the complexity of the maintenance program and the fact that they are planning to make new buildings makes this case a perfect one to experiment with.

Interviews with the facility managers turn out that a lot of errors occur with building elements in the hospital. A dataset with all errors of the year 2014 is generated from Planon. This one is used in this study.

Explore usefulness data

In order to be useful the maintenance data had to be converted into event logs. Therefor at least a case ID, Event ID and timestamp are required. The data was delivered in .csv format with a lot of information and columns. It occurred that the data was not immediately suitable because in one row several events where captured. Those events had to be converted to single rows. This took half a day to prepare manually. But finally the column 'Number' could be used as Case ID. The column 'StatusCode' could be used as Event ID and the column 'DatumEvent' could be used as timestamp. Rows like 'BuildingCode' and 'InventoryDescription' are used as attributes. The data is tested by importing in the process mining tool Disco. It is concluded that the data is suitable for process mining analytics, but it takes some time to convert the data into event logs.

Analysis

A brainstorm session is held and as a result some questions are determined which can be analysed with process mining. The facility managers are interested in the process of alerts, is it really going as they think it is? They have a scheme which describes the proposed process flow, but they don't know if it is actually happing. In addition they wonder how long the durations in between the steps are and which specific elements have most errors.

Process check

The standard process of error handling as determined by the facility managers can be seen in figure 2. The W numbers describe the process steps and are called StatusCode. The description of this code is in the next column. A normal process starts with W1 'error reported' and ends with W6 'error administrative finished'. In between those steps all kinds of scenarios can occur. It can occur that some steps are skipped. Depending on the kind of error several steps are taken in the process.

The process map of the event log is generated and can be seen in figure 3. The facility managers indicated that this map is like expected. The process steps are according the planned process flow. Also the durations where indicated as realistic. For example the median duration between process step W5 and W6 is calculated for 31 days. The facility managers told that once a month step W6 is taken, so this is close to 31 days. They mentioned that they never had seen this kind of visualisation and that it gives an interesting perspective on their data.

W0	Werkorder ontvangen
W1	Werkorder gemeld
W1b	Werkorder in behandeling genomen
W2	Werkorder offerte aanvraag
W2a	Werkorder offerte akkoord
W2k	Werkorder offerte voor akkoord klant
W2o	Werkorder offerte ontvangen
W3	Werkorder opdracht
w3b	Werkorder in behandeling
w3c	Werkorder On Hold
W3d	Werkorder opnieuw toegewezen
W3m	Werkorder materiaal besteld
W4	Werkorder offerte vervallen
W5	Werkorder technisch gereed
W5d	Werkorder techn. ger. deel factuur
W5v	Werkorder techn. ger. verzamelfactuur
W5z	Werkorder techn. ger. zonder kosten
W6	Werkorder administratief gereed
W7	Werkorder vervallen
W7rvb	Werkorder vervallen na maatregel RvB

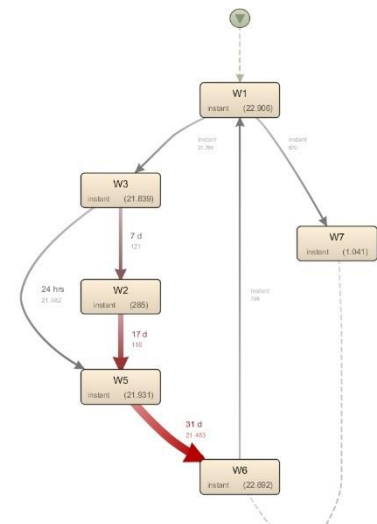**Figure 2 Process steps in error handling at the hospital**

Figure 3 Process map generated with process mining tool Disco

Duration

In the process map it was noticed that between step W3 'Error assigned to subcontractor' and step W5 'Error fixed' the median duration was 24 hours. So within 24 hours the problem was solved by the subcontractor. The most interesting part for the facility managers was to identify the elements which took steps W3 and W5 after each other within a very short time. It was found that in 2414 cases the problem was solved within a median duration of 3 minutes(see figure 4) . Which is ,accordingly to the facility managers, practically impossible. More specific this means that within 3 minutes an error is assigned to a person and he almost immediately solved it, which is only possible if he was on the exact location of the error at the moment of assigning. This notification led to a deeper research in this process. In seemed that for this 3 minute job an average price of 300 euro was charged by the subcontractors (in total €723600,-).

The errors were related to several topics (As can be seen in figure 5). For example 14% of the errors where related to electricity and 11% was related to sanitary and water problems. In addition the product types which faced the most problems could be discovered partly. It seemed that most electrical problems occurred with 'Double electrical revolving doors', 'Personal elevators' and 'Brancard elevators'. But also a lot of product types couldn't be identified because they didn't have a specific description in the data (As can be seen in figure 6). In addition it is seen that some types of the same product, have different descriptions like 'Personenlift' and 'Personen lift' are indicated as separate types, but are in fact the same. This makes it more difficult to identify the problematic product types.

Figure 4 Process steps filtered on cases with W3 and W5

As a result of this analysis the facility managers were surprised about the amount of money which was spend to short (and unrealistic) jobs and they are going to monitor those errors for next months to figure out how this maybe can save them money. A notifiable quote was mentioned by one of the facility managers, she said "We can probably save more money with investing in data analytics, than with firing our own people".

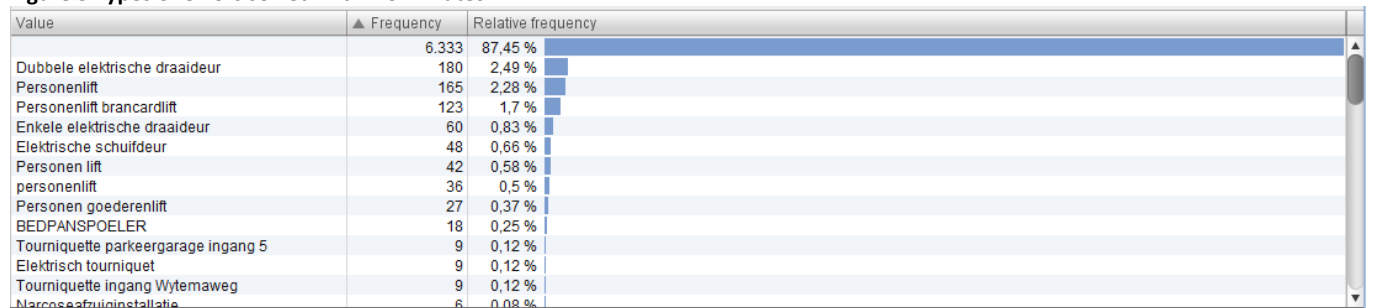| Value	Frequency	▲ Relative frequency	
Opheffen storing elektriciteit	1.041	14,37 %	
Opheffen storing sanitair en water	783	10,81 %	
Opheffen storing verlichting	675	9,32 %	
Opheffen storing elektrische werking deuren	663	9,15 %	
Opheffen storing lift	573	7,91 %	
Opheffen storing luchtbehandeling	507	7 %	
Opheffen temperatuur storing	489	6,75 %	
Opheffen storing waterlekkage	420	5,8 %	
Herstellen schade of defect aan deur of raam	234	3,23 %	
Opheffen storing verwarmingsinstallatie	150	2,07 %	
Opheffen storing waterbehandeling	147	2,03 %	
Opheffen bouwoverlast	144	1,99 %	
Opheffen schade buitenzonwering of rolluik	144	1,99 %	
Opheffen storing buispostinstallatie	117	1,62 %	
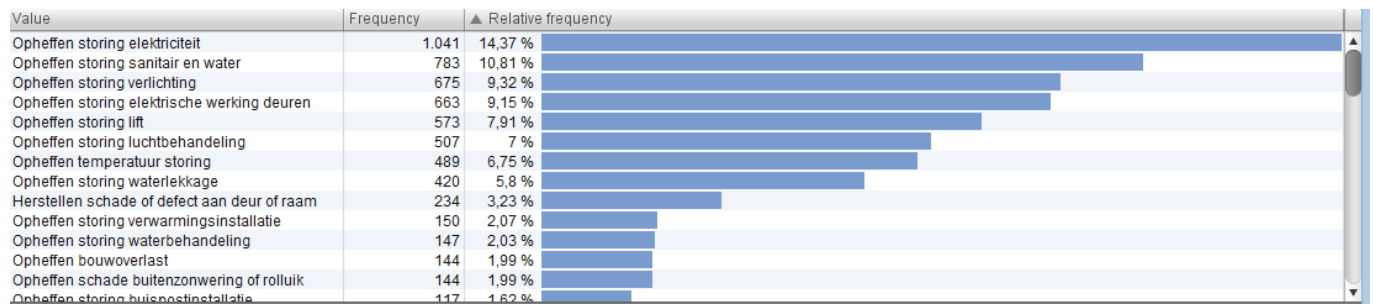
Figure 5 Types of errors solved within 3 minutes

Value	▲ Frequency	Relative frequency	
	6.333	87,45 %	
Dubbele elektrische draaideur	180	2,49 %	
Personenlift	165	2,28 %	
Personenlift brancardlift	123	1,7 %	
Enkele elektrische draaideur	60	0,83 %	
Elektrische schuifdeur	48	0,66 %	
Personen lift	42	0,58 %	
personenlift	36	0,5 %	
Personen goederenlift	27	0,37 %	
BEDPANSPOELER	18	0,25 %	
Tourniquette parkeergarage ingang 5	9	0,12 %	
Elektrisch tourniquet	9	0,12 %	
Tourniquette ingang Wytemaweg	9	0,12 %	
Narcoseafzuiginstallatie	6	0,08 %	

Figure 6 Types of products involved in a lot of errors

5. Conclusion and discussion

This study explored if facility management data is suitable for analysing processes around building elements with process mining techniques. By use of a case study at a hospital the potential of combining process mining with maintenance data is discovered. It can be concluded that with some data transformation maintenance data is suitable for process mining. Moreover the facility managers were surprised by the visualisation techniques and they gained clear insight in the error handling process. As a result they discovered problematic building elements and odd processes. The facility managers where definitely interested in using those kind of analytics more in the future. In order to enable process mining on a larger scale it would be useful that facility management systems adapt an 'export to event log' function. Also naming of elements should be consistent in order to make analytics more easy and reliable.

This study just explored the topic of facility management based process mining and has proven some useful applications. More (case-) studies should be elaborated to indicate the potential of this topic. In addition it would be useful to study the potential of integrating different data sources of other phases, for example the design -or construction- phase, in building elements lifecycle in order to gain insight in the process on a longer time span.

6. Acknowledgements

This study was not possible without the help of the facility managers of Erasmus MC Hospital. Special thanks go to them. Without their open mind, enthusiasm and assistance with data preparation and analysation this paper was not possible.

14 Appendix Dataset

The dataset can be downloaded via:
https://www.dropbox.com/s/a0smmmsx5kopyho/Dataset%20Schependomlaan.zip?dl=0

14.1 Dataset Schependomlaan

This document contains information about the data set regarding the construction project Schependomlaan. All involved parties have given permission to use the data for scientific and academic purposes. The data is gathered during the master thesis project of Stijn van Schaijk. In collaboration with Hendriks Bouw en Ontwikkeling[36], ROOT[37], TNO[38] and RAAMAC[39] the data is collected. General information about the project can be found at the website http://www.schependomlaan.nl/ .

The dataset contains the following elements:

- Design model in .IFC
- Subcontractor models in .IFC and .DWG
  - Flooring
  - Walls
  - Stairs
  - Fencing
  - Steel
  - Roofs
  - Prefab
  - …
- Coordination models in .TBP (Tekla BIMsight Package)
- Schedule/Planning in .pdf and .xml
- As-planned models in .IFC and Synchro file format.
- As-built models in point cloud formats .ASCII and .PLY
- Results comparison as-planned and as-built models in .xlsx
- As-planned Event log in .xlsx and .csv
- As-built Event log in .xlsx and .csv
- Event log with actors in .xlsx and .csv

Some files contain some errors and therefore have to be treated carefully. The description of all files is written next.

14.1.1 Design model

The design model IFC Schependomlaan.ifc is made in Archicad by ROOT bv. They modelled the project by order of Hendriks Bouw en Ontwikkeling. The project was originally drawn in 2D cad formats and was translated into a high quality BIM model. The design model is used as a reference during the engineering phase in collaboration with the subcontractors.

---

[36] http://www.hendriksbouwenontwikkeling.nl/
[37] http://www.root-bv.nl/
[38] https://www.tno.nl/en/
[39] http://raamac.cee.illinois.edu/

14.1.2 Coordination model and Subcontractor models

The models delivered by the subcontractors are coordinated in the file: <u>Schependomlaan Nijmegen.tbp</u>. This file can be opened using Tekla BIMsight[40] and contains all the subcontractor models. Some subcontractors couldn't deliver their products in BIM/.IFC. Those designs are aligned with the BIM using .DWG formats.

14.1.3 Planning and as-planned models

As-planned models are created by use of Synchro[41] software. The design IFC is connected by use of the .XML planning. The native Synchro file is available <u>Synchro project Schependomlaan.sp </u>. From Synchro as-planned models are exported to .IFC. A model containing all the planning information is exported called: <u>IFC Schependomlaan incl planningsdata.ifc. </u> As part of the study also as-planned models on specific dates had to be exported. So the following models have been made:

- <u>Week 26 26 june IFC Schependomlaan incl planningsdata.ifc</u>

- <u>Week 27 3 july IFC Schependomlaan incl planningsdata.ifc</u>

- <u>Week 28 10 july IFC Schependomlaan incl planningsdata.ifc</u>

- <u>Week 29 17 july IFC Schependomlaan incl planningsdata.ifc</u>

- <u>Week 30 24 july IFC Schependomlaan incl planningsdata.ifc</u>

- <u>Week 37 11 sept IFC Schependomlaan incl planningsdata.ifc</u>

Note that at the moment of writing Synchro does not have optimal IFC exporting options, as a result the data structure in IFC is changed during export. The elevation tree is changed and the property sets are changed. As can be seen in the figures below. One has to take this into account when working with this dataset.

---

[40] https://www.teklabimsight.com/
[41] https://synchroltd.com/

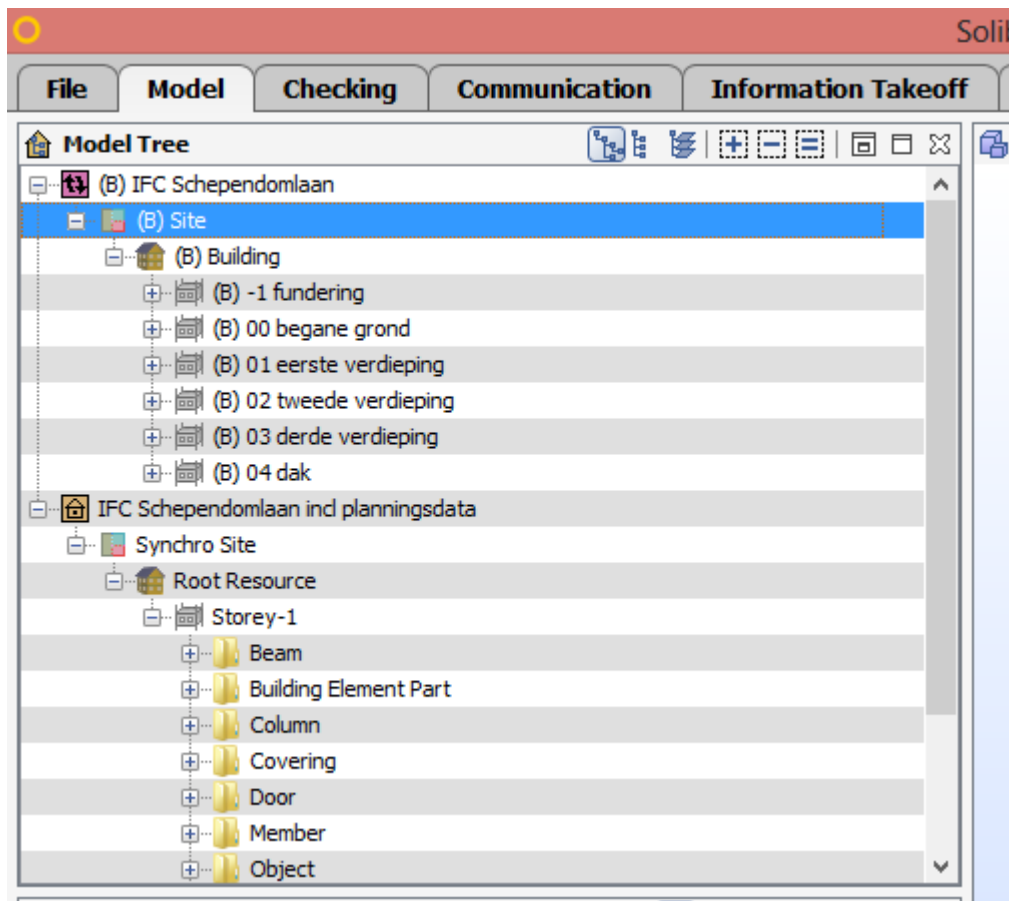Figure 49 Changed IFC structure elevations. Upper model is the original, the lower model is the as-planned model with the changed elevation structure.
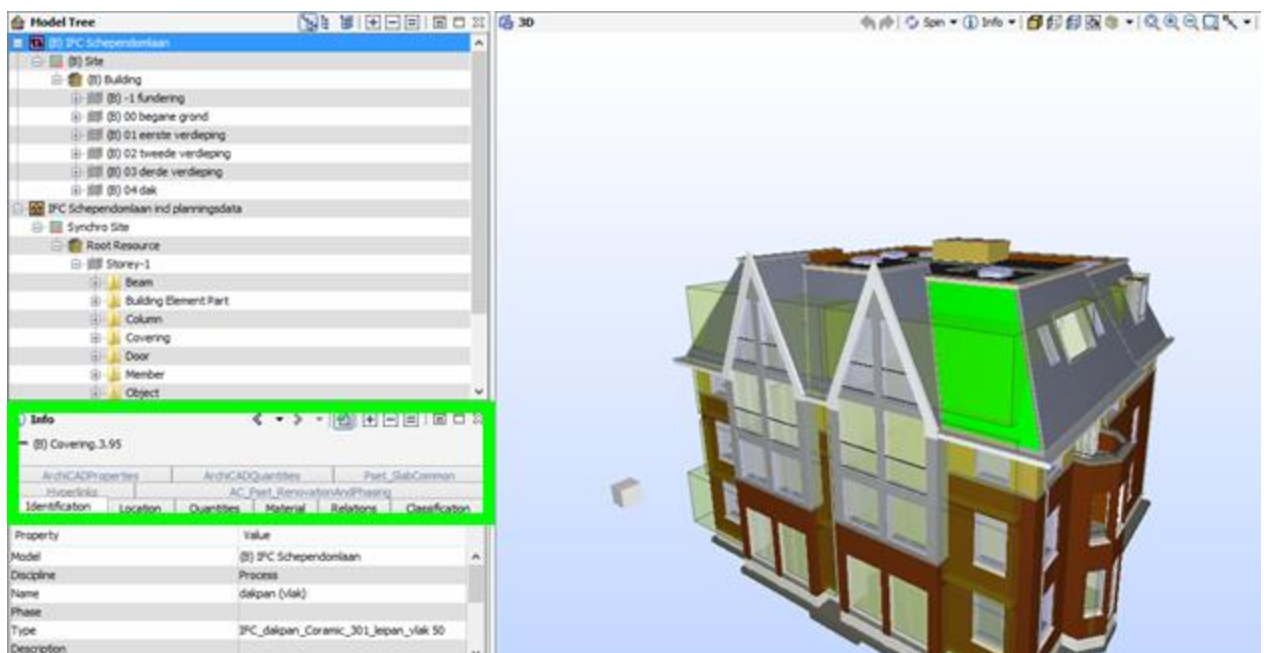

Figure 50 Parameter structure before Synchro export, notice the property sets in the green rectangle.
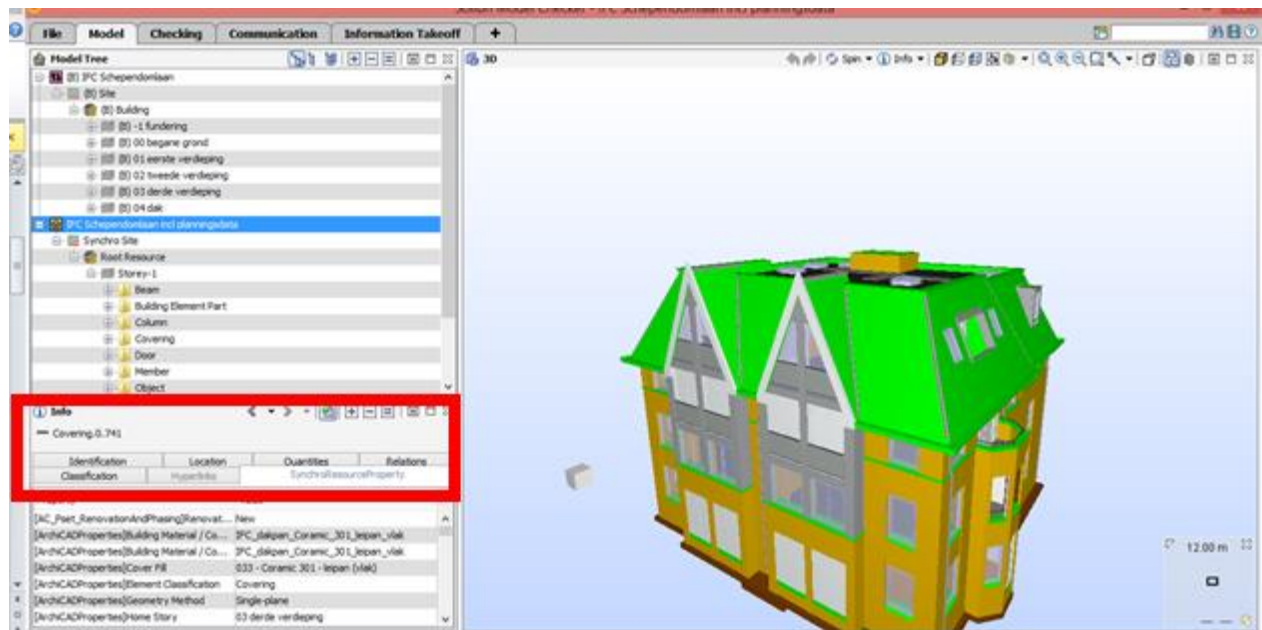
Figure 51 Parameter structure after ifc export, notice the red rectangle: original parameter sets are merged into one SynchroResourcePropertyset

14.1.4 As-built Point clouds

The construction site is filmed with drones over days during the project. The movies can be seen at the Youtube chanel of the drone pilot:

https://www.youtube.com/channel/UCHsKVCxqYaZn39JNZH7v9LQ

Movies of week 26, 27 , 28, 29 and 30 are used to create as-built point clouds using structure from motion techniques. The point clouds are available in .PLY, .ASCII and some in .LAS files.

14.1.5 Comparison as-planned as-built

The as-built point clouds are compared with the as-planned IFC using a matlab algorithm by RAAMAC. As a result a comparison sheet is created. See the file: comparison_all weeks.xlsx

14.1.6 Event logs

Part of the research was to enable construction process mining. In order to do so event logs had to be created. Those files can be analysed in process mining tools. The as-planned event log is in the file eventlog_IFC_schependomlaan.csv . The event log containing as-built information is the file: Schependomlaan eventlog as planned data incl as built data.csv .
As part of an experiment also actors are added to an event log. Those actors are fake and not realistic. But to proof that social network analytics can be done with those data structures this file is made: Schependomlaan eventlog as planned data incl as built data incl Actors FAKE.csv .

15 Appendix Algorithm Eventlog service

Generating an event log from an IFC model can probably be done in several ways. In this appendix the used algorithm for the Eventlog service is explained in general. The source code which uses this algorithm in the Eventlog service can be found at:

https://github.com/opensourceBIM/BIMserver/blob/master/DemoPlugins/src/org/bimserver/demoplugins/service/EventLogService.java

Using the Eventlog service asks the user to give the location of the element code and the material. In this sample algorithm the parameters Element code and Material are filled in as follows:

- Element code (i.e. Nl-Sfb): [ArchicadProperties]Layer
- Material:[ArchiCADProperties]Building material / Composite / Profile

1. Start
2. Load IFC
3. Make .csv with columns:
    a. *GUID*
    b. *ifcClass*
    c. *Nl-sfb*
    d. *Material*
    e. *Task*
    f. *Resource*
    g. *TaskName*
    h. *TaskStart*
    i. *TaskFinish*
4. List all elements in IFC based on GUID
5. Start with GUID1
6. Take GUID1 export to *GUID* column in .csv list
7. Search GUID 1
    a. Find:IFCRELDEFINESBYPROPERTIES
    b. Take 'ifcPropertySetDefinition #number'.
8. Search 'ifcPropertySetDefinition #number'
    a. Find IFCPROPERTYSET
    b. Take 'HasProperties'
9. Search all 'HasProperties' for:
    a. Element classification
        i. Take IFCLABEL
        ii. Export to column *IfcClass*
    b. [ArchiCADProperties]Building material / Composite / Profile
        i. Take IFCLABEL
        ii. Export to column *Material*
    c. [ArchicadProperties]Layer
        i. Take IFCLABEL
        ii. Export to column *Nl-sfb*
    d. Name

          i. Take IFCLABEL

          ii. Export to column *Resource*

10. Search GUID1

    a. Find IFCRELASSIGNTOPROCESS

    b. Take #number RelatingProcessSelect

        i. Find IfcTask

           1. Take name

           2. Export name to column *TaskName*

           3. Take TaskId

           4. Export TaskId to column *TaskName*

        ii. Take #number IfcTask

    c. Search #number IfcTask

        i. Find IFCRELASSIGNTASKS

        ii. Take #number TimeForTasks

    d. Search #number TimeForTasks

        i. Find IFCSCHEDULETIMECONTROL

        ii. Take #number ScheduleStart

           1. Search #number ScheduleStart

              a. Find IFCDATEANDTIME

              b. Take #number DateComponent

           2. Search #number DateComponent

              a. Find IFCCALENDARDATE

              b. Take DayComponent,MnthComponent,YearComponent

              c. Export components to column TaskStart

        iii. Take #number ScheduleFinish

           1. Search #number ScheduleFinish

              a. Find IFCDATEANDTIME

              b. Take #number DateComponent

           2. Search #number DateComponent

              a. Find IFCCALENDARDATE

              b. Take DayComponent,MnthComponent,YearComponent

              c. Export components to column TaskFinish

    e. For each task related to GUID1 create new row and copy columns *GUID, IfcClass, Material, Resource.* Then add task related information tot he specific row.

    f. After GUID 1 take GUID 2, GUID 3, all other GUIDs etc, and repeat the algorithm as described above.

    g. Sort output on TaskStart time.

    h. Output as .csv string.

16 Appendix Algorithm Planning consult service

Generating an planning advice from an event log and an IFC model can probably be done in several ways. The algorithm of the Planning consult will be described in general in this appendix.

The source code which uses this algorithm in the Planning consult service can be found at:
https://github.com/opensourceBIM/BIMserver/tree/master/DemoPlugins/src/org/bimserver/demoplugins/service/planner

1. Input IFC:
   a. Read IFC:
      i. Parse IFC model.
   b. Make object model (with EMF).
2. Input .csv event log:
   a. Read .csv:
      i. Parse .csv file.
   b. Make an object model from events in .csv.
   c. Arrange events on start date.
   d. Analyse of the code, go through all events:
      i. Make a planning for each event.
      ii. Track mapping all materials and GUIDs.
      iii. Look at GUID from object belonging to event.
      iv. Create task with TaskName if it is not done before.
      v. Look if task is according to schedule or not (failure data).
      vi. Update counters:
         1. Total
         2. On Time
         3. Too late
      vii. Add task to planning if it is not done before.
3. Give suggested planning per material and make mapping material to planning advice:
   a. Look at all IfcProduct objects.
   b. Get material from all products:
      i. Take materialParameter which is given by the user.
      ii. Simplify material if it is given by the user.
   c. Check is there is a suggested planning for the material:
      i. If yes: Count update.
      ii. If no: create new suggested planning.
   d. Search materials in mapping event log:
      i. First original material name.
      ii. Second simplified material.
         1. No found list: Database count 0.
         2. Yes found list: Database count amount of GUIDS.
   e. Go through list of GUIDS.
   f. Search for planning from mapping planning by GUID.
      i. If found: Add planning variant.

ii. Within planning advice store all planning.
iii. Keep all unique planning variants.
4. Print all info into a structure document file which is readable and understandable (for example to .HTML).

17 Appendix Planning consult Schependomlaan

This appendix contains the result of the Planning consult software tool as used in the case study Schependomlaan.

The following elements are not recognized in the database:

Product type	Amount of elements in model
M_Single-Flush: 0915 x 2134mm	1
Level: 8mm Head	2
Basic Roof: Generic - 400mm	1

Planning Consult Results

Breedplaat

In the model are found 2 elements of this type of product.

Planning advice

Amount of elements in the database: 58

In the planningsdatabase the following process variants are found

Variant 1: 43 % of the elements from this product type are planned according to these steps:

Task number	Task name	Percent on time	Percent too late	Percent unknown
1	ST00560 Leggen breedplaat incl. Veiligheid	100%	0%	0%
2	ST00600 Randkist	100%	0%	0%
3	ST00660 Installateurs	100%	0%	0%
4	ST00630 Stalen balken	100%	0%	0%
5	ST00210 Wapening	100%	0%	0%
6	ST00230 Stort	100%	0%	0%
7	ST00540 Opstellen	100%	0%	0%

Variant 2: 57 % of the elements from this product type are planned according to these steps:

Task number	Task name	Percent on time	Percent too late	Percent unknown
1	ST00540 Opstellen	100%	0%	0%
2	ST00560 Leggen breedplaat incl. Veiligheid	100%	0%	0%
3	ST00600 Randkist	100%	0%	0%
4	ST00660 Installateurs	100%	0%	0%
5	ST00210 Wapening	100%	0%	0%
6	ST00230 Stort	100%	0%	0%

Kanaalplaat

In the model are found 2 elements of this type of product.

Planning advice

Amount of elements in the database: 46

In the planningsdatabase the following process variants are found

Variant 1: 100 % of the elements from this product type are planned according to these steps:

Task number	Task name	Percent on time	Percent too late	Percent unknown
1	ST00350 Kanaalplaatvloer leggen (lev. week 16)	100%	0%	0%

Stelkozijn

In the model are found 9 elements of this type of product.

Planning advice

Amount of elements in the database: 135

In the planningsdatabase the following process variants are found

Variant 1: 28 % of the elements from this product type are planned according to these steps:

Task number	Task name	Percent on time	Percent too late	Percent unknown
1	ST00970 Plaatsen kozijnen	84%	15%	1%

Variant 2: 72 % of the elements from this product type are planned according to these steps:

Task number	Task name	Percent on time	Percent too late	Percent unknown
1	ST00940 Stelwerk buitengevel	85%	14%	1%
2	ST00970 Plaatsen kozijnen	84%	15%	1%

Baksteen

In the model are found 4 elements of this type of product.

Planning advice

Amount of elements in the database: 357

In the planningsdatabase the following process variants are found

Variant 1: 55 % of the elements from this product type are planned according to these steps:

Task number	Task name	Percent on time	Percent too late	Percent unknown
1	ST00940 Stelwerk buitengevel	85%	14%	1%
2	ST01050 Metselwerk	92%	7%	1%

Variant 2: 45 % of the elements from this product type are planned according to these steps:

Task number	Task name	Percent on time	Percent too late	Percent unknown
1	ST01050 Metselwerk	92%	7%	1%