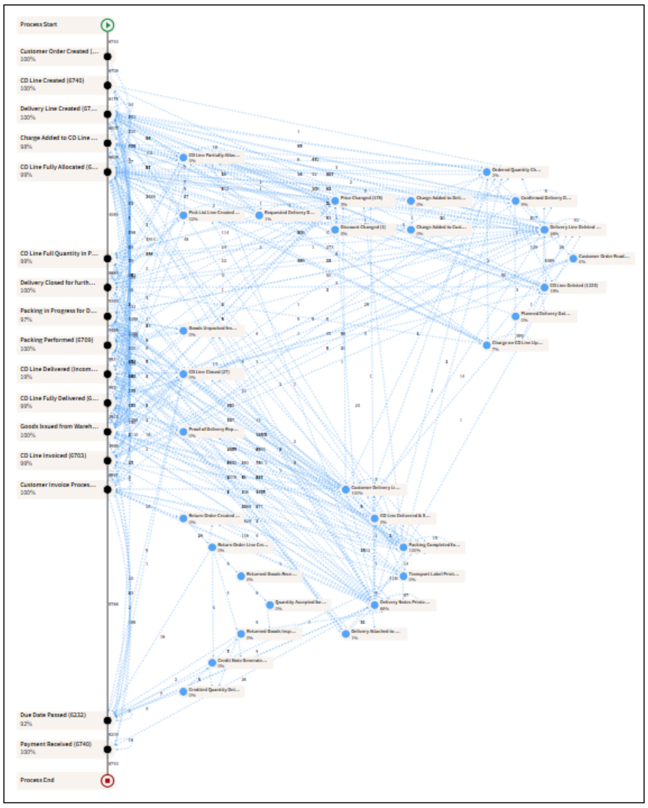
CHALMERS



Analyzing Order-to-Cash Using Process Mining

A Case Study in Collaboration with Paulig

Master's Thesis in the Supply Chain Management Programme

JENS ADOLFSSON
DILAN SALEH

DEPARTMENT OF TECHNOLOGY MANAGEMENT AND ECONOMICS
DIVISION OF SUPPLY AND OPERATIONS MANAGEMENT
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden 2026
www.chalmers.se

Analyzing Order-to-Cash Using Process Mining

A Case Study in Collaboration with Paulig

JENS ADOLFSSON
DILAN SALEH

Department of Technology Management and Economics
Division of Supply and Operations Management
Chalmers University of Technology
Gothenburg, Sweden 2026

Analyzing Order-to-Cash Using Process Mining
A Case Study in Collaboration with Paulig
JENS ADOLFSSON
DILAN SALEH

Academic Supervisor & Examiner: Patrik Jonsson, TME Chalmers
Industrial Supervisor: Johan Bystedt, Meridion

Department of Technology Management and Economics
Chalmers University of Technology
SE-412 96 Gothenburg
Sweden
Telefon +46 31 772 1000

Cover: A so-called spaghetti bowl illustrating the different ways a process is executed.

Gothenburg, Sweden 2026

Analyzing Order-to-Cash Using Process Mining
A Case Study in Collaboration with Paulig

JENS ADOLFSSON
DILAN SALEH

Department of Technology Management and Economics
Chalmers University of Technology

Abstract

Organizations increasingly rely on digital data to understand and improve their business processes. Process Mining is a data-driven approach that uses event logs from information systems to visualize actual process behavior and identify inefficiencies. This thesis investigates how Process Mining can be applied in practice to analyze the Order-to-Cash process, with a particular focus on the use of pre-defined reference process models and backward-looking analytical techniques.

The study is conducted as a case study in collaboration with Paulig, using Infor's Process Mining solution integrated with the ERP system M3. Through a combination of Process Mining analysis, interviews, workshops and shadowing sessions, the thesis evaluates how well a pre-defined industry-specific process model reflects an organization's actual Order-to-Cash process and how inefficiencies and bottlenecks can be identified. The reference process model proved to be a strong baseline for understanding the overall process structure, while the analysis revealed bottlenecks related to master data issues that cause unnecessary manual interventions and longer cycle times.

The results demonstrate that Process Mining can support improvements in both administrative processes and physical logistics flows by revealing systematic issues that are difficult to detect through traditional qualitative methods alone. The study also highlights the importance of combining Process Mining insights with domain knowledge and stakeholder involvement to correctly interpret results.

Keywords: Process mining, Order-to-Cash, Process discovery, Conformance checking, Business process management.

Acknowledgments

This master's thesis was completed during the autumn semester of 2025 at Chalmers University of Technology, within the Department of Technology Management and Economics. The project was carried out in collaboration with Meridion, Paulig and Infor.

We would like to thank our academic supervisor and examiner, Patrik Jonsson, for his valuable guidance, feedback and support throughout the thesis process. We also wish to acknowledge our industrial supervisor Johan Bystedt at Meridion for his commitment and support during the course of the project. A special thanks to Andreas Alftrén at Paulig and Hallgeir Øvrebust at Infor for making this project possible.

Furthermore, we are grateful to all interview and workshop participants for generously sharing their time and expertise, which provided essential input to this study. Lastly, we would like to thank our opponent group, Filip Kitevski and Mohamed Arkavazi, for their constructive feedback that contributed to improving the quality of the thesis.

Jens Adolfsson & Dilan Saleh
Gothenburg, January 2026

List of Acronyms

Below is the list of acronyms that have been used throughout this thesis listed in alphabetical order:

3PL	Third-Party Logistics
ATP	Available-to-Promise
BBD	Best Before Date
BPMN	Business Process Model and Notation
CO	Customer Order
CRM	Customer Relationship Management
CS	Customer Service
EDI	Electronic Data Interchange
ERP	Enterprise Resource Planning
GenAI	Generative Artificial Intelligence
IMP	Issue Management Portal
MTS	Make-to-Stock
O2C	Order-to-Cash
P2P	Procure-to-Pay
RPA	Robotic Process Automation
WMS	Warehouse Management System

List of Figures

List of Tables

Contents

1

Introduction

The following chapter presents the background, purpose and scope of the thesis. The context of the study is introduced through a description of Process Mining and its role in organizational process analysis, together with an overview of the case company. The purpose of the thesis and the research questions that guide the study are defined. The limitations and conditions of the project are also outlined, along with the intended contribution of the project.

1.1 Background

This section provides an introduction to Process Mining and a brief overview of Paulig, the case company.

1.1.1 Managing Process Complexity

The current global environment is characterized by increasing complexity, driven by major technological advancements, volatility and speed [1]. This escalating complexity impacts the markets in which companies operate, necessitating that organizations constantly strive for innovation and digital transformation to maintain success and stay ahead of the competition. Remaining competitive requires companies to continuously improve operational efficiency and excellence, becoming more agile in adapting their strategies, structures and processes to actual conditions [1, 2]. Digitalization itself can be described by two critical words, speed and flexibility [1]. These attributes are essential for survival in fast-moving industries experiencing disruption.

A significant challenge arises from the sheer scale and huge variety of processes within modern organizations, which severely increases operational handling and complexity [1]. Large organizations for example, have to deal with thousands of suppliers and millions of purchase orders, customer orders, deliveries and financial transactions. Even smaller companies face challenges in understanding their actual business processes and organizational complexity in order to continuously improve efficiency and reduce transactional cost. The large variation often occurs not only between individuals performing the same tasks, but also in how a single employee carries out the same task at different times [3]. When dealing with this complexity and variety, traditional methods of process analysis, which rely on extensive observation and manual documentation, only draw a rough picture of reality that is open to biased interpretation [1]. This environment has created a need for full transparency and control over business operations, as transformation cannot be managed without a thorough understanding of current processes and complexity drivers.

1.1.2 Introduction to Process Mining

Process Mining has emerged as an exciting technical innovation and a fundamental technology for the digitalization of organizations, providing an innovative path toward simplifying complexity [1]. It is considered a key lever to address the complexity challenge because it allows an understanding of even the most complex processes. Process Mining converts the data produced during daily work into objective insights, thus creating an end-to-end experience and transparency. This allows extraction and visualization of the actual process flow, providing transparency regarding the sequence of activities as they have actually taken place. Process Mining visualizes the actual "as-is" process flows with all deviations and variations which is typically much more complicated than the ideal, theoretical "to-be" process. For example, analyzing complex processes like Order-to-Cash (O2C) can reveal thousands of different process variants, making analysis challenging [2]. This insight is critical because it allows for fact-based process management, leading to objectivity and shifting discussions away from perception-based interpretations [1].

To give an example for the application of Process Mining, a study looking to identify and eliminate inefficiencies in IT system usage from a lean perspective in the healthcare sector, states that impeded workflows can have a negative impact on safety and effectiveness of care delivered to patients [4]. 241 workflow impediments were found through observations and interviews with a large Dutch university hospital. The inefficiencies arised due to misalignment between the standard workflow of the Electronic Health Record (EHR) and the actual clinical workflows, meaning users perform inefficient steps or create workarounds, increasing error risk and reducing efficiency. It is stated that the digital workflows are less visible and more difficult to track than the physical ones, further incentivizing and increasing the relevance of Process Mining to visualize and create transparency in these digital workflows with limited insight, while also bridging the gap between designed and actual workflows.

Process Mining enables quick identification of improvement and automation areas within core processes such as O2C and Purchase-to-Pay (P2P) [1]. The strength of Process Mining lies in its ability to transform underutilized data into actionable insights. Many organizations today collect vast amounts of process related data but fail to exploit it beyond simple reporting. Process Mining addresses this gap by supporting continuous improvement initiatives in a transparent and repeatable way. Unlike traditional improvement programs that often require long lead times, Process Mining enables continuous follow-up and iterative development of processes through constant access to real-time data. As a result, organizations can gain a clearer understanding of their workflows, improve compliance with standards and enhance administrative efficiency. Its utility is also not confined merely to classical administration processes but also extends to areas like production, development and sales processes, but also to other industries such as the healthcare sector as previously stated. The transparency provided by Process Mining is of high importance as technology progresses at an increasing pace. It offers guidance for automation potentials, forming a crucial building block alongside the increasing synergy between Robotic Process Automation (RPA) and Artificial Intelligence (AI).

Process Mining serves as the foundation for the successful transition to a digital organization [1]. It supports the concept of a Digital Twin of an Organization (DTO), which is a dynamic software model leveraging operational data to understand how the organization executes its business model and responds to changes. By enabling full transparency and insight into activities as they actually happen, Process Mining provides the tools and methods necessary to embed digital transformation into day-to-day operations. It is expected to become compulsory for ensuring Business Process Hygiene (BPH), making it necessary for process owners to justify why they are not using Process Mining for continuous process optimization and screening.

In the last decade, Process Mining has been established in the operational environments of many companies, who have adopted its principles to industrial requirements in order to create economic and ecologic value [1]. A large amount of process variants can be reduced and the remaining ones optimized, thus reducing operational inefficiency. Companies like Siemens started as early as 2011 to adopt this technology, and by 2019 large organizations like BMW and Telekom were using it throughout organizations as a standard tool for digital transformation and identification of process inefficiencies. Over time, an increasing number of use cases have been applied along value chains.

Given the increasing environmental challenges and the need for a sustainability revolution, technological innovations like Process Mining could also play a role in solving these challenges and provide ecologic value [1]. Apart from identifying process inefficiencies in general, it could optimize flows within production and logistics to for example reduce energy consumption and material waste, but also increase transport efficiency and optimize inventories. The goal of this sustainable economy is achieved by reducing transactional costs and the environmental footprint of logistics operations, thus creating both economic and ecological value.

The Process Mining technology operates by analyzing event logs, which are digital traces captured from systems like Enterprise Resource Planning (ERP), Customer Relationship Management (CRM) or Warehouse Management Systems (WMS) [1]. By utilizing these system-generated event logs, Process Mining allows organizations to create a transparent and detailed picture of their workflows. Earlier approaches relied heavily on subjective interviews and manual process mapping [5]. The analysis can be performed through three main perspectives, discovery, conformance and enhancement:

- Discovery: Where a process model is automatically generated solely based on event logs without prior knowledge [5].
- Conformance: Where the actual process execution is compared to a pre-defined standard to identify deviations and understand their causes [6].
- Enhancement: Where the aim is to improve existing processes by introducing best practices or optimizing the current workflow [6].

While the digital trace data offers transparency into what actions take place, it does not necessarily explain why those actions occur [7]. For this reason, Process Mining is best used as a diagnostic tool, a way of conducting a current state analysis of processes that can later serve as input for further optimization methods like RPA or Generative AI (GenAI) solutions [1]. Process Mining techniques allows filtering and segmentation to analyze specific subprocesses and business units to uncover inefficiencies and bottlenecks. The tool can also be applied to identify positive deviants, which are instances where processes perform better than the established standard.

There are several reasons why processes do not follow their intended flow. For instance, disruptions such as a supplier delaying a delivery or a customer changing an order quantity with short notice can create deviations from the standard sequence. These kinds of variations, combined with differing work practices, multiple system usage and a lack of standardized procedures, make it challenging to achieve consistent outcomes. This can lead to both inefficiencies and bottlenecks in an organization's workflow. While bottleneck analysis has traditionally focused on production settings [1], there is increasing interest in applying it to administrative and service processes, where efficiency is equally critical but harder to monitor and manage.

1.1.3 Case

This master's thesis is carried out in collaboration with Meridion, Infor and Paulig. Meridion is an ERP consultancy firm specializing in supply chain management and supporting organizations in system implementation and data-driven improvement. Infor is a software vendor and developer of the ERP system M3 and the Process Mining tool used in this study. Paulig is a major actor in the Food & Beverage industry, with headquarters in Finland and owning multiple brands across Europe.

Paulig's portfolio includes multiple business units, where the Tex-Mex brand Santa Maria is well known to most Swedish consumers. This thesis, however, focuses on the Finnish coffee brand, which is one of the company's most established categories with its strongest markets in Finland and the Baltic region. Paulig has historically operated with several different ERP systems across brands, regions and production sites. To address this fragmentation, the company is currently undergoing a large-scale transformation called *One Paulig*, which aims to harmonize processes and transition the entire organization into a single ERP system, M3. Within this transformation, the coffee division acts as the pilot group. This means that the coffee organization is not only the first to migrate to M3, but also responsible for establishing standard ways of working that the other brands soon will follow. Because of this role, understanding their administrative processes is crucial for ensuring a smooth transition and avoiding costly discrepancies when rolling out the system to the rest of the organization.

To support this effort, Paulig is interested in objectively mapping how its O2C process currently operates, identifying deviations from expected behavior, uncovering bottlenecks and highlighting where procedures vary across customers and locations. Infor's Process Mining tool is intended to facilitate this by using Paulig's real event data from M3 to reconstruct the process, show how cases flow through different steps and point out inefficiencies along the way. A key characteristic of Infor's solution is that it begins from a pre-defined industry-specific reference model, in this case tailored to companies in the Food & Beverage sector.

1.2 Purpose

The purpose of this project is to examine Process Mining as a technique for analyzing and diagnosing administrative processes within a Food & Beverage company. The study focuses on understanding whether a Process Mining tool can reveal bottlenecks and deviations within Paulig's O2C process and how these insights can support the organization during its transition to a unified ERP environment.

Since Infor's Process Mining solution is used in the project, particular attention is given to how well the pre-defined Food & Beverage reference model represents Paulig's actual operations. Part of this project is therefore to analyze how well Infor's generic Food & Beverage model aligns with Paulig's real operations and to understand where adaptations may be needed. To ensure an objective perspective, the study includes a broad literature review covering different Process Mining approaches. This creates a foundation for comparing Infor's method with alternative techniques and helps evaluate the strengths and limitations of the specific tool tested in this case.

The study also aims to assess the broader opportunities and limitations of Process Mining as a technique, independently of any specific software. By reviewing academic literature and results from the project, the study evaluates how Process Mining performs and what general insights and constraints companies should be aware of when adopting the technique.

1.3 Research Questions

Process Mining is still a relatively new and rapidly developing analytical technique, and many questions remain about how well it works in real organizational settings. Based on the case context and the purpose of this thesis, three research questions have been formulated to explore both the practical use of Infor's Process Mining tool and the broader implications of adopting Process Mining in administrative processes.

A central starting point is to understand how accurately the tool represents Paulig's actual O2C process, since reliable process models are essential to gain meaningful insights and guide the *One Paulig* transformation. Therefore, the first research question is:

RQ1: *How accurately do pre-defined process models reflect actual O2C processes?*

Another important motivation is to evaluate whether Process Mining can reveal bottlenecks and inefficiencies that may not be visible through traditional qualitative methods such as interviews, workshops or manual process mapping. This leads to the second research question:

RQ2: *Which inefficiencies and bottlenecks in the O2C process can be identified through Process Mining?*

It is also essential to assess the broader potential of Process Mining beyond the specific tool used in this project. By understanding both opportunities and limitations, organizations can form realistic expectations about what Process Mining can achieve. The third research question is therefore:

RQ3: *What are the opportunities and limitations associated with the application of Process Mining tools in organizational process analysis?*

1.4 Limitations

This project includes a number of limitations that shape what is examined and how the results are interpreted.

Paulig has so far only implemented the M3 system in the Finnish and Baltic markets, which means that all event logs originate from these regions. The warehouse and logistics activities are mainly located in Finland, whilst the office functions are split between Finland and Estonia. A central part of the study aims to explore how Paulig's actual O2C process differs from the Food & Beverage industry standard map provided by Infor. Since this investigation is limited to one business unit and two geographical locations, it is not possible to evaluate how these differences vary between other parts of the company.

Tool availability also constrains the project. Access to Infor's Process Mining environment was limited to five weeks, which reduces the time for deeper exploration and iteration. The project only includes Infor's interpretation of Process Mining, meaning that a broader evaluation of the technique must rely on academic literature when addressing the research questions related to general opportunities and limitations.

The scope of accessible data limits the thesis as well. Only information related to the O2C process is available which means that other business processes such as procurement and production are not included in this analysis. The event log data was generated before the study began meaning that the authors had no influence over how the data was recorded and structured. If certain activities or attributes are missing, it is technically possible to add them at a later stage, but doing so requires effort from multiple stakeholders and is therefore only undertaken when the information is expected to contribute significantly to the results.

Activities that take place outside M3 are not captured in the database and cannot be analyzed with the Process Mining tool. Examples include emails and transport bookings in external portals. To obtain a complete understanding of the O2C process, these parts of the workflow are instead explored through interviews and shadowing sessions.

1.5 Conditions & Contribution

As part of this study it is important to note that Infor's Process Mining tool is still in an early stage of adoption, especially in Sweden. This thesis is only the second project in the country to work with the tool and the first to apply it to the O2C process since the other project focused on the P2P process. As a result, there are limited local resources and best practices available to support the work. An objective of the study is therefore to evaluate how well the tool performs in practice and to identify areas where it can be improved.

There is currently no established standard for how bottlenecks should be identified with Process Mining tools, partly because every company operates with different systems and process structures. A secondary aim of this thesis is therefore to develop a structured way of working with Infor's tool that enables quick identification of process areas that require attention and further analysis.

Being early adopters also means that the project includes collaboration with Infor's development team. The authors have the opportunity to provide feedback on usability and functionality as well as suggestions on how the tool can better support the analysis of administrative processes. This adds a developmental dimension to the thesis that is separate from the main research questions but valuable for the ongoing refinement of the tool.

2

Theoretical Background

This chapter presents the theoretical background that forms the basis for the study. It introduces the O2C process, the role of information systems in business operations and the main concepts and techniques of Process Mining that are relevant for the analysis in this thesis.

2.1 The Order-to-Cash Process

The O2C process can briefly be described as the sequence of activities beginning with the receipt of a customer order and ending with the collection of payment for the corresponding invoice [8]. It is one of the core end-to-end processes in supply chain and financial management and plays a central role in both operational efficiency and liquidity performance. Between order receipt and payment, organizations typically perform a series of activities such as order confirmation, order picking, delivery and invoicing. The shorter and more efficient the O2C cycle is, the quicker the organization converts sales into cash, thereby improving its working capital position and reducing administrative transaction costs. The activities in the O2C process are cross-functional and include multiple departments such as Sales and Customer Service (CS) who are responsible for negotiating orders and maintaining customer interactions, while Finance and Accounting manage invoicing and cash reconciliation.

In Figure 2.1, a typical O2C process is presented, illustrating the distribution of responsibilities across departments. The process typically begins in Customer Service, where the order is received, validated, confirmed and sent to the warehouse [9]. Many organizations also perform several pre-order activities prior to order release. These include the creation and maintenance of accurate customer master data, such as customer information, payment terms and pricing agreements. Complete and accurate documentation is essential to prevent later errors in delivery or invoicing [10]. Credit management is also often performed before the order is accepted, where the customer's creditworthiness is assessed and credit limits are set. Such pre-order checks reduce financial risk and help ensure that the order can move through the process without unnecessary credit blocks or manual interventions. Once the order is accepted, it is transferred to warehouse operations for picking, packing and staging for dispatch. Order fulfillment may also involve transportation planning, carrier selection, shipment documentation and the communication of tracking information to the customer. Proof of Delivery (POD) confirms successful delivery and may be contractually required before invoicing can occur, but can vary depending on the terms of the companies.

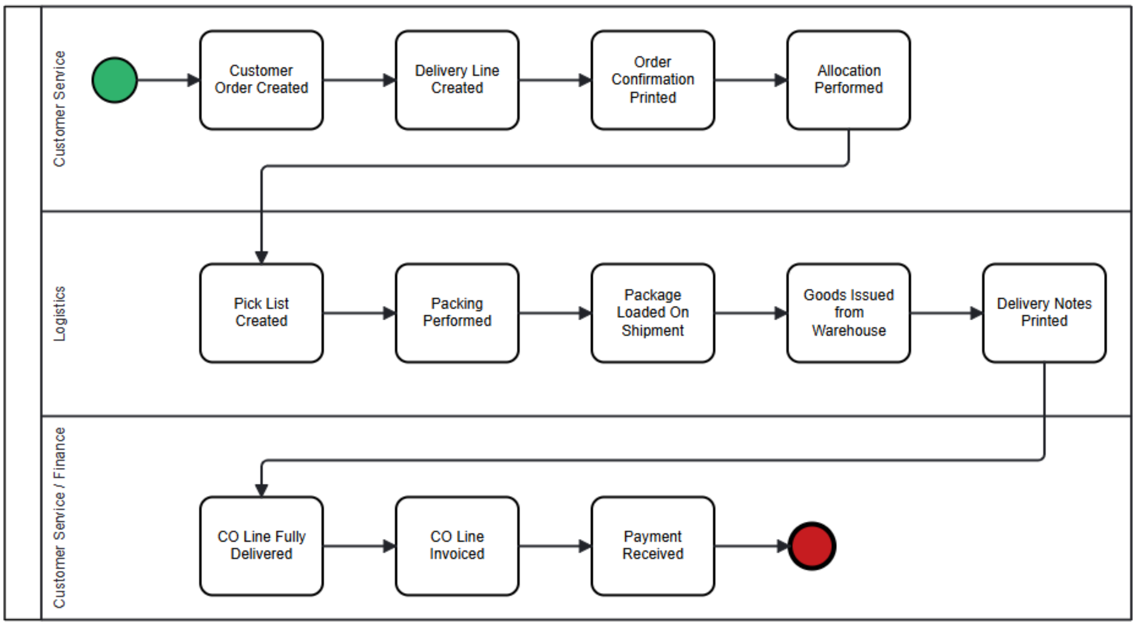


Figure 2.1: A basic O2C process showing the major steps from customer order creation to received payment.

The invoicing process follows the physical delivery of goods or completion of services. Accuracy in invoicing, related to for example quantities, prices, discounts, taxes and delivery details, is important as discrepancies often lead to customer disputes, delayed payments and increased administrative workload [9]. When disputes occur, organizations typically employ a formal dispute management process that investigates discrepancies and issues adjustments where appropriate. Effective dispute management not only accelerates cash flow but also contributes to customer satisfaction and long-term relationship quality. After invoicing, the process continues to accounts receivable and collections, where Finance monitors open invoices and escalates overdue payments when necessary. When payment is received, it must be accurately matched to outstanding invoices, a task known as cash application. Depending on the organization's level of automation and the quality of remittance information, this step can range from highly automated to heavily manual. Effective cash application is important for maintaining accurate accounting records and ensuring that customer credit limits and account statuses are up to date. Once payment is matched and posted, the O2C cycle is formally complete.

In addition to the forward flow of orders, many organizations also manage a reverse logistics flow which involves a return flow [10]. Customers may return products due to various reasons such as incorrect shipments or quality issues. This reverse flow begins with the customer sending a complaint, followed by logistics handling and inspection of the returned items and the issuance of a credit note or refund where applicable.

2.2 Information Systems

Today's companies rely on a range of information systems that capture and maintain digital records of business operations. These systems provide detailed data about everything from transactions, customers and suppliers to production and inventory levels [11]. Since different types of information systems are specialized in different organizational areas, companies usually employ several complementary systems to ensure that all parts of their operations are covered digitally. While some of these systems are designed to exchange data and operate most effectively when integrated with one another, others are more isolated or function-specific, which results in fragmented information flows. Because the collective information stored in the systems represents a digital reflection of how the organization actually operates, it also forms the essential data foundation for Process Mining. Table 2.1 presents three of the most commonly used information systems with a short description of their main characteristics and functions.

Table 2.1: Main types of information systems.

Information System	Description
ERP	A centralized information system that combines a shared database with a range of applications supporting the management and coordination of business activities. It is divided into various modules and covers functions such as customer order handling, production planning and procurement [12]. While an ERP system in principle can cover most organizational needs, it is often complemented by other systems that handle more specialized tasks, for instance CRM and WMS [13].
CRM	A system used to manage and improve a company's interactions with existing and potential customers. It supports activities such as marketing, sales and after-sales. It gathers and connects customer information from various channels and departments to create a complete understanding of each customer. By analyzing and understanding customer behavior, a CRM system helps organizations increase satisfaction and improve long-term profitability [14].
WMS	A system used to manage and control warehouse operations by tracking the flow and storage of goods. Simpler versions are primarily focused on registering stock levels and locations. It also supports storage and picking activities. More advanced systems can plan and coordinate resources and material flows. A WMS provides detailed tracking of products, destinations and movements [15].

Although many types of information systems contribute to the digital operations of a company, the ERP system most often serves as the central platform where data from different areas is gathered [11]. Over time, ERP systems have developed from simple material planning tools into comprehensive solutions that integrate most aspects of business management [12]. Companies also use Electronic Data Interchange (EDI) as a standardized way to share information across organizations. Through EDI, documents like purchase orders and invoices can be exchanged automatically between systems, which reduces manual work and helps ensure consistent and accurate data.

2.3 Process Mining

This section provides an overview of Process Mining. It covers its background and development, how event data is used to understand real processes, the different types of Process Mining, guidance on how to start a project and how Process Mining relates to other tools.

2.3.1 Background & Development

Process Mining is a field that combines ideas from data science and process management with a focus on how digital records can be used to understand and improve business processes [16]. It can be applied to study both administrative processes such as order handling and invoicing as well as the physical flow of products. Figure 2.2 presents some of the main topics in data and process science that intersect within the field of Process Mining, illustrating how it integrates data-driven analysis with process-oriented thinking [16, 17].

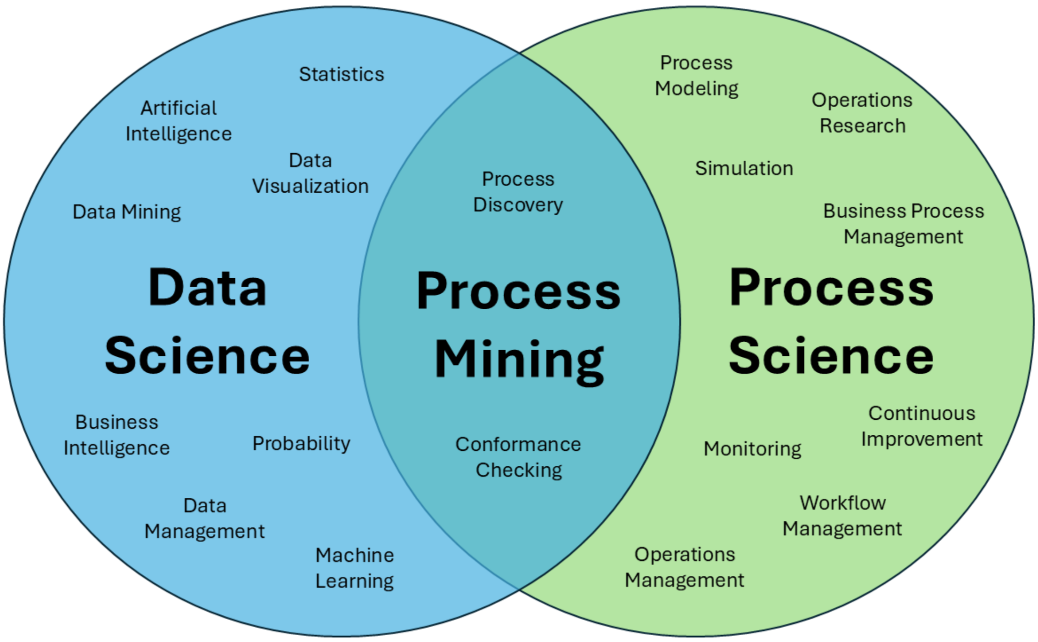


Figure 2.2: Process Mining is the intersection between data and process science. Image adapted from [16].

Research on the topic has been ongoing since the late 1990s [18]. At that time numerous barriers existed that hindered the implementation of Process Mining tools in industrial contexts. The main issues were the limited availability of structured event data and the unrealistic assumptions the early models were based on [1]. During this period most organizations were still in the beginning of transitioning from manual and partially paper-based workflows to digitalized business processes. Information systems such as ERP and CRM were introduced, but were neither widespread nor fully mature [11]. The early forms of ERP systems were far from comprehensive and mainly focused on specific operational areas such as Material Requirements Planning (MRP) and lacked the integration needed to represent complete process flows. Since most administrative business activities were still carried out outside of digital systems it was nearly impossible to analyze entire end-to-end processes [1]. Consequently, the type of standardized and consistent data required for Process Mining was rarely available.

As digitalization advanced and information systems became more established the situation changed gradually. These systems were able to record large amounts of event data that reflected how activities were actually carried out [11]. This created a foundation for data-driven process analysis methods in practice such as Process Mining [19]. In the early 2010s, Process Mining began moving from research context to industrial use as commercial tools became available. In 2018, there were more than 30 software vendors offering Process Mining tools [1]. Combined with greater computing power and improved visualization techniques, it is now possible to extract data from information systems to visualize and analyze processes. But in comparison to the 1990s, today's problem is instead that the large variety of information systems within individual companies makes it challenging to gather all relevant data in a single place and in the same format [20, 21].

2.3.2 Purpose of Process Mining

The main idea behind Process Mining is to bridge the gap between how organizations believe their processes are executed and how they actually work in practice. By extracting digital event data from information systems, organizations can visualize the "as-is" process flow. Based on this, it is possible to perform an analysis to detect inefficiencies and uncover improvement opportunities. At this stage, it also becomes possible to identify deviations from internal guidelines and user manuals [22].

Process Mining relies on event data that is automatically generated when activities are executed within information systems. This data forms an objective representation of how processes unfold over time and enables organizations to base their improvements on empirical evidence rather than assumptions or interviews. Compared to traditional process modeling methods, which are often based on workshops and manual documentation [1], Process Mining captures the actual behavior of how a process is performed rather than the idealized version of it.

Process Mining is mainly suitable for processes that are repetitive and follow a reasonably stable structure across many cases. Research and industry experience show that organizations most often begin with core administrative processes such as P2P and O2C, as these are typically well standardized, generate clear processtraces and offer substantial opportunities for improvement [1]. Procurement and sales related processes have repeatedly demonstrated quick wins, both in terms of cost savings and increased revenue, since they produce high quality event data and follow well defined patterns from initiation to completion [16].

2.3.3 Event Logs & Process Variants

Event logs form the core data foundation of Process Mining as they capture the digital traces when activities are executed in information systems. An event log can be understood as a collection of recorded events, where each event represents a single step in the execution of a business process and is linked to the specific case in which it occurred [1, 23]. Since most information systems store data in relational databases rather than in ready-made process formats, relevant information must often be extracted from different tables and converted into a flat event log structure [23]. Despite this, the basic requirements for an event log are simple and consistent across systems [1]. Based on the event data, Process Mining automatically derives process variants that reflect the different paths cases have taken through the process in practice. At a minimum, each event record must contain three key attributes that describe what happened and when:

- **Case ID:** Uniquely identifies the process instance to which the event belongs. Examples include customer orders, customer order lines or purchasing orders.
- **Activity label:** Specifies the action or step that was executed. Examples are *Customer Order Created*, *Packing Performed* or *Payment Received*.
- **Timestamp:** Records the exact moment the activity took place, including the date and exact time down to milliseconds.

Additional attributes of a case can be included to support a deeper analysis [23]. Digital traces arise whenever a user performs an action in an information system. For example, when an order line is created, when a delivery date is updated, when a quantity is changed or when an invoice is generated. Across systems such as ERP, CRM and WMS, each of these actions updates the underlying database and produces an event that becomes part of the event log [1]. Automatic activities are also recorded and generate event data in the same manner, which means that they can be included in the analysis even though no user actively performs them. As a result, the event log provides an objective and chronologically ordered representation of how processes are actually executed. Figure 2.3 presents an example of a digital event log, showing the sequence of recorded activities for a single case together with the basic attributes.

A strength of Process Mining is its role of both being a visualization technique and an analytical method. The visualization component provides a process model derived directly from event logs, commonly presented as a flow diagram of the most frequent paths. In Figure 2.4, a specific execution path is shown, illustrating how the sequence of activities can be visualized directly from the event log to reveal the actual order in which the process unfolded. This is a fairly simple path that moves from the creation of a customer order to the final receipt of payment. Even in this straightforward case, all three functions Customer Service, Logistics and Finance are involved in carrying out the process.

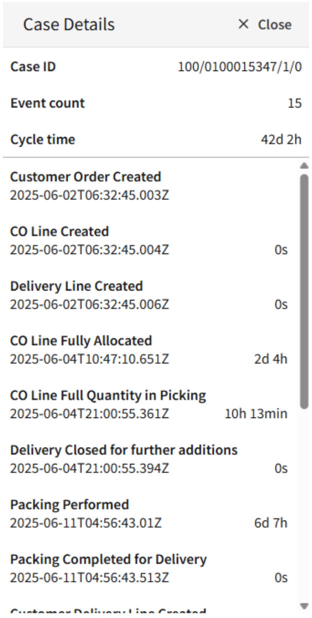


Figure 2.3: Example of an event log.

Figure 2.4: Example of a process variant.

The discovered process model forms the foundation for performance measurements such as throughput times and waiting times. Analytical features enable organizations to evaluate different process variants and identify undesired activity loops in the standard procedure [1]. Event logs and process variants can be used to understand how the process actually operates and to identify actionable improvement opportunities.

2.3.4 Reference Process Models

Process Mining builds on the idea that event data can be compared with a reference process model that represents how the process is expected to operate [16]. The reference process model can be a manually designed process map, for example a Business Process Model and Notation (BPMN) diagram used by the organization or a mathematically precise model such as a Petri net that is discovered directly from the event data. In both cases the model functions as a point of reference when analyzing how the real process behaves and where differences between the expected and observed flow occur [16].

BPMN is a widely adopted standard for describing business processes in a clear and accessible way. It uses a set of graphical elements such as activities, business rules and arrows indicating the sequence flow to form a business process diagram [24]. A central purpose of BPMN is to create a notation that is understandable to business analysts, managers and technical developers, which helps bridge communication gaps during process design and improvement work [24]. BPMN diagrams often include swimlanes that represent different functions or roles in the organization, which makes responsibilities easy to interpret. BPMN models describe how the organization believes the process should operate, rather than the full variation present in real executions.

Petri nets provide a more formal and analytical way of representing processes. They consist of elements that show how activities depend on each other and how steps may occur in parallel. In Process Mining, Petri nets are closely linked to several important techniques. Many discovery algorithms create Petri nets directly from event logs and they are also commonly used when checking how well the observed execution matches the expected behavior [25]. There are many different forms of Petri nets and most of them can express the same workflow patterns that are also found in BPMN models [25]. Because of their formal structure, Petri nets are useful for detailed analysis, even though they can be more difficult to interpret for business users.

Both BPMN and Petri nets support Process Mining, but they contribute in different ways. BPMN is intuitive and easy to understand, which makes it suitable for communicating how a process is intended to work and for involving business stakeholders in improvement efforts. Petri nets provide a more precise and formal representation of the behavior found in event data, which is helpful when analyzing the logic and structure of a process in detail. Although they are less accessible to non-technical users, they offer a deeper level of insight. Most commercial Process Mining tools therefore present BPMN-style diagrams to the user while internally converting these models into a Petri net representation to handle the underlying process logic and ensure that complex behavior can be analyzed correctly [26]. In this way, the two notations complement each other with BPMN offering clarity and Petri nets providing analytical precision.

2.3.5 Different Types of Process Mining

Process Mining techniques are typically divided into two main categories, backward-looking Process Mining and forward-looking Process Mining. Backward-looking techniques focus on understanding how the process has behaved in the past and are used to investigate issues such as bottlenecks, deviations and variants. This category is commonly divided into four sub-types that together describe the "as-is" state of the process and form the diagnostic part of Process Mining [16]. Forward-looking techniques aim to anticipate future behavior by predicting outcomes, an example is to estimate the remaining processing time of a running case or identify where a bottleneck is likely to occur based on the path the case has taken [1]. Apart from these two main categories, there is also an action-oriented type of Process Mining that focuses on turning insights into real actions by triggering workflows or responses when certain situations occur in the process [16].

2.3.5.1 Backward-looking Process Mining

Process discovery is the most fundamental application of Process Mining and refers to the automatic generation of a process model solely based on event data without relying on pre-defined descriptions or user assumptions [16]. It allows organizations to gain an objective representation of how processes are actually performed. Through the discovery part it becomes possible to visualize the most common paths, parallel ongoing activities and the overall structure of the process [6]. Process discovery serves as the foundation for further analysis and evaluation [16].

Conformance checking focuses on comparing a reference process model with the actual execution recorded in event data [6]. Its purpose is to assess how well the documented or assumed process aligns with reality by making deviations visible. Since process models are often handmade and influenced by assumptions, they may represent an outdated or idealized version of the process [27]. Conformance checking helps identify where the model and the event log differ by highlighting:

- **Wrong sequenced activities:** Showing where the execution order deviates from the reference model.
- **Missing activities:** Indicating that expected steps were never carried out.
- **Additional activities:** Revealing tasks that appear in the event log even though they are not part of the reference model.

Locating these discrepancies allows organizations to evaluate the quality of their models and consider necessary redesigns. It also helps to determine where the real process may require adjustments or improved support [27].

Performance analysis focuses on identifying timing and efficiency-related issues in a process such as long waiting times, missed deadlines or recurring quality problems [16]. By linking event data to the process model, it becomes possible to examine how often activities occur, how long they take and where delays or loops appear. A method called token-based replay is frequently used to relate event data to the process model and to highlight discrepancies in timing and execution behavior [28]. Using the timestamps contained in the event log, performance analysis enables the calculation of indicators such as throughput times, waiting times, response times and service times. This information can also be used to evaluate compliance with Service Level Agreements (SLA) and to identify areas where performance expectations are not being met [16].

Comparative Process Mining focuses on analyzing and contrasting multiple event logs that represent the same process executed under different conditions such as across locations, time periods or customer segments [16]. Comparing these logs makes it possible to identify both common patterns and differences in how the process behaves in each context. Root cause analysis is often used to explain these differences by linking variations to factors such as workload, resource availability or case characteristics. A method of particular interest is the detection of execution gaps, which are recurring problems that appear more frequently in one log than another, for example longer waiting times or additional rework [16].

2.3.5.2 Forward-looking Process Mining

Predictive Process Mining focuses on anticipating how a running process case is likely to unfold and on forecasting outcomes that have not yet occurred [29]. This includes predicting the next activity, estimating remaining processing time or identifying whether a case is at risk of deviation. These predictions help organizations respond to changes in their operations, manage resources more effectively and take preventive actions before performance or compliance problems occur. A range of Machine Learning (ML) methods can be applied to support predictions, drawing on insights from the backward-looking analysis to recognize patterns in past behavior and use them to forecast future events [16].

2.3.5.3 Action-oriented Process Mining

Action-oriented Process Mining focuses on turning analytical insights into concrete interventions that actively support or adjust the running process. While backward-looking techniques explain what has happened and forward-looking techniques predict what may happen next, action-oriented Process Mining uses these diagnostics to trigger actions during execution [16]. Examples include:

- Reallocating resources when waiting times increase.
- Notifying managers when important checks are skipped.
- Blocking suppliers whose behavior indicates potential risk.

An important development within this area is the growing connection between Process Mining and RPA. RPA is used to automate repetitive and rule based tasks, while Process Mining contributes by identifying which tasks are suitable for automation and by monitoring how the tools perform once they are in use [1]. Research further shows that Process Mining can uncover repetitive work patterns and user-interface interactions that can be automated, allowing organizations to find additional opportunities for improvement [16]. In addition to task automation, action-oriented Process Mining can also support automatic management interventions, sometimes referred to as Robotic Process Management (RPM). In these situations, diagnostic insights are translated into actions such as alerting supervisors, informing customers about delays or assigning additional staff when it looks like it is needed [1]. Through these capabilities, action-oriented Process Mining supports the translation of analytical insights into targeted improvements in operational processes.

2.3.6 How to Start a Process Mining Project

One of the questions raised for Process Mining is not knowing how to start [1]. A successful Process Mining project requires clear preparation, realistic expectations and reliable data. Many initiatives fail because teams are not fully engaged, goals are vague or data is incomplete. Three foundational factors, known as the 3P:s [1], need to be aligned from the start to create the right conditions:

- **Purpose:** A clearly defined business objective that explains why Process Mining is being used and what improvement it should support.
- **People:** Stakeholders with the right mix of process knowledge, technical understanding and willingness to drive change.
- **Processtraces:** Complete and accessible event data form the factual basis for the analysis.

These three elements must support each other. A strong purpose and motivated people cannot compensate for missing or low-quality data, and high-quality data will not create impact without clear goals or committed stakeholders [1].

Defining the intended process requires input from several parts of the organization. For example, a typical O2C process involves functions like Customer Service, Logistics and Finance, each with its own responsibilities and viewpoints. Representatives from different levels within these functions must participate, as operational staff, team leads and managers often see different parts of the process and highlight different challenges [1]. Technical expertise is also required to understand where the relevant information is stored and how to extract it in a structured way. Someone must also define what to examine in the process so that the reference model is structured correctly and captures the intended flow. This collaboration is essential in order to create a representable BPMN model that can serve as a reliable baseline for the following Process Mining analysis.

Findings from previous Process Mining projects indicate that it is effective to begin with a small and simple pilot. Focusing on one well-structured process like P2P or O2C helps generate quick results because these processes are standardized, produce high-quality event logs and offer clear improvement potential [1]. Process Mining should be viewed as a continuous activity, since processes change over time and insights need to be refreshed accordingly [1]. In this sense, Process Mining serves as the administrative counterpart to continuous improvement practices in production environments, providing a data-driven approach to identify inefficiencies, monitor changes and support iterative process optimization.

2.3.7 Process Mining in Relation to Other Tools

Process Mining is often discussed together with simulation and digital twin technologies, since all three aim to support process understanding and operational improvement. Each approach contributes with a different perspective on how processes behave, are monitored and how potential future outcomes can be evaluated. Together they form a broader toolbox for evidence-based process analysis [30, 31].

Simulation is a model-driven approach where a pre-defined process model is used to generate simulated behavior. This makes it suitable for exploring alternative de-signs, capacity changes and potential future outcomes. The results strongly depend on how accurately the model reflects the real process [30]. Building such a model often requires expert knowledge and assumptions that may be difficult to validate. Process Mining begins with real event data extracted from information systems and reconstructs the actual process flow based on this information. This provides a factual view of how the process behaves but is primarily backward-looking and does not by itself support the evaluation of hypothetical redesigns. When combined, Process Mining and simulation complement each other. The discovered models can serve as a realistic basis for simulation and the simulated event logs can be analyzed with Process Mining to verify whether a simulation model behaves realistically [30].

Digital twins are virtual representations of physical processes that are continuously updated with real-time data. Process Mining supports the development and operation of digital twins by extracting event data and deriving process information that reflects how the system actually functions [31]. This helps the digital twin stay aligned with real executions and reduces the risk that the virtual representation drifts away from what happens in practice. It also improves the ability of the twin to provide timely predictions and detect early signs of inefficiencies or deviations [31]. Process Mining and digital twins together combine real-time monitoring with data-driven insights into how the process unfolds, creating a strong foundation for operational decision support. This integration strengthens the analytical capabilities of the twin and supports the continuous improvement of complex systems during runtime [31].

2.4 Theoretical Application in the Project

The O2C process provides the foundation for the thesis and defines the scope of the empirical analysis. It establishes a reference for evaluating whether Paulig's O2C process reflects a typical end-to-end flow and clarifies how different activities interact. The focus on cycle times highlights the importance of fast conversion from order creation to payment and guides which process data is included in the study.

Information systems form the digital basis for the capture of business processes. In this thesis, the ERP system M3 is the main data source as it contains most of the O2C activities analyzed with the Process Mining tool. CRM and WMS are discussed to illustrate activities that may fall outside the scope of this project. The historical perspective also clarifies why Process Mining is a relatively new field and why ERP maturity is central to this case with regard to data availability.

Process Mining supports the thesis objectives by enabling a data-driven analysis of how administrative processes are actually executed, rather than how they are assumed to work. Reference process models provide a baseline for comparing expected and observed behavior, where BPMN based models are used on the user level while Petri net logic supports systematic detection of deviations in the background. The study applies backward-looking Process Mining techniques to identify inefficiencies and non-conforming behavior in Paulig's O2C process. Forward-looking and action-oriented approaches are discussed to frame how insights could be used for proactive decision making and operational improvements in the future.

The methodological framework of purpose, people and processtraces ties the theoretical concepts together and explains how the project was structured. The framework motivates the choice of a focused O2C pilot and clarifies how stakeholder involvement and data selection support the overall analysis.

3
Methodology

The following chapter describes the research design and methodology used in this study. The aim of the chapter is to explain how the study was conducted and to motivate the methodological choices made in relation to the research questions. By presenting the approach, data collection and analysis methods, the chapter provides an understanding of how the results of the report were produced.

The study is mainly based on a qualitative case study approach, as the purpose is to explore how Process Mining can be applied in organizational process analysis and to identify opportunities and limitations. This analysis is combined with quantitative elements through the results of the Process Mining tool. Since Process Mining is highly dependent on the organizational context and the availability of structured data, a study of real processes within a company was considered suitable. The empirical material consists of event log data extracted from Paulig's ERP system. This data is analyzed using Infor's Process Mining tool in order to discover and analyze actual process flows. Insights from interviews, workshops and presentations with employees involved in the O2C process were used to support the analysis and interpretation of the results. Aspects such as data quality and limitations of the chosen methodology are addressed at the end of the chapter.

3.1 Method Design

The study follows a case study–based research design, focusing on the application of a Process Mining tool in a real organizational setting. Since Process Mining relies on event data but also requires an understanding of how processes are intended to function in practice, a combination of qualitative and quantitative methods is applied. This approach enables a more complete analysis by combining data-driven insights with contextual knowledge from the organization.

To structure the methodological approach, the study is guided by the 3P:s previously presented, which includes purpose, people and processtraces. The purpose is defined by the research questions and determines what insights the analysis aims to generate. The people dimension is represented by process stakeholders and domain experts at Paulig, who contribute essential knowledge about how the O2C process is executed and where the boundaries between processes lie. The processtraces consist of event log data extracted from Paulig's M3 environment, which forms the technical foundation for the Process Mining analysis.

As illustrated in Figure 3.1, the methodology follows a stepwise process. In the early stages of the project, the focus was primarily on qualitative methods such as semi-structured interviews, workshops and review of company documents. This phase was necessary to build an understanding of both administrative and physical O2C processes, establish process boundaries and reduce the risk of misinterpreting system data. At this stage, Paulig's process experts played a key role in providing contextual knowledge, while the technical use of the Process Mining tool was still developing.

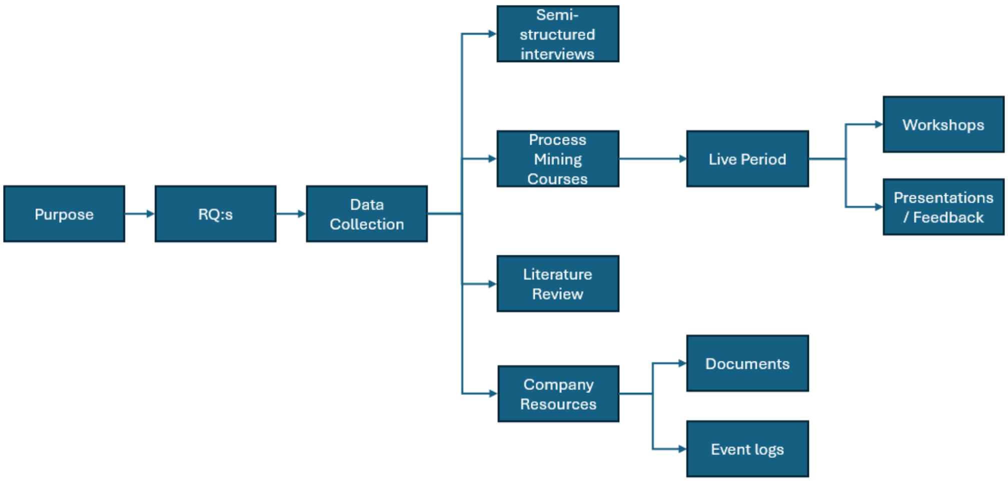


Figure 3.1: Overview of the methodology used in the study.

As the project progressed, improved understanding of Process Mining enabled a more detailed analysis of processtraces during the live period. At this point, the authors became more knowledgeable in Process Mining and how to use Infor's tool, while Paulig contributed with process-specific knowledge to interpret findings at a detailed level. Workshops and presentations were used to validate insights, explain deviations and ensure a shared understanding of the results. The O2C process is extensive and consists of several subprocesses with owners from different departments. Customer Service and Logistics play central roles, while Sales and Finance are also involved. For this reason, the presentation sessions were important to bring together stakeholders from different functions, as shown in table 3.1, to create a shared understanding of the whole O2C process. They also enabled discussion of detailed issues that are often difficult for individual roles to fully oversee.

Table 3.1: Involved actors from each company.

Company	Role
Paulig	• Head of Customer Service and Commercial Quality • O2C Development Manager • Head of IT - Supply Chain • Customer Service Specialist • Customer Service Manager CSC • ERP Architect • IT Solution Specialist • Head of Logistics • Supply Chain Development Manager - Logistics
Infor	• VP Advanced Services • Senior Business Consultant, Advanced Services • Consultant Process Mining

Different types of data were required to conduct the study. Quantitative data in the form of event logs was used to discover and analyze the actual execution of the O2C process, identify bottlenecks and measure performance. However, event log data alone is not sufficient to explain why deviations occur or how exceptions are handled. To bridge this gap, qualitative data from interviews, workshops and company documentation was used to capture process logic and contextual factors that are not visible in the system data.

Company documents, such as process maps and work instructions, represent the intended process and serve as a reference point when comparing it with the process models discovered through Process Mining. By combining these data sources and methods, the study applies triangulation to strengthen the reliability and validity of the findings. Qualitative methods support the interpretation of Process Mining results, while quantitative analysis provides objective data-driven insights into how the O2C process is executed in practice. The following sections describe the method design, data collection and analysis procedures in more detail.

3.2 Data Sources

This section describes the data sources used in the study and how they form the empirical basis for the subsequent analysis. Since Process Mining relies on both event data and contextual process knowledge, the study uses a combination of qualitative and quantitative data sources. The qualitative data provides an understanding of how the O2C process is intended to function and how it is handled, while the quantitative data enables an objective analysis of how the process is actually executed based on historical event logs. These data sources form the foundation for the Process Mining analysis and the interpretation of its results.

3.2.1 Qualitative Data

Qualitative data was collected to gain a deeper understanding of the O2C process at Paulig and to provide context for the analysis performed in the Process Mining tool. The qualitative data was collected through interviews, workshops, company documentation, literature review and Infor's Process Mining courses. The material was used both to support the analysis and to validate findings derived from the quantitative data. Table 3.2 displays where the data is derived from and how it was collected, including which roles participated in the different sessions. Each session had a pre-decided purpose shared with the involved actors prior to the meetings.

Table 3.2: Overview of data collection activities.

Data Collection			
Date	**Participants**	**Purpose**	**Outcome**
Interviews / Shadowing			
2025-10-06	- CS Specialist	Shadow the O2C process in M3.	How an order enters the system. How it is planned for transport. Common problems/deviations in the O2C process.
2025-10-14	- O2C Development Manager	Interview about the O2C process. Aimed to understand the O2C and Exceptions Management process better.	More insights regarding how an order is received in the ERP system, while also learning what the most common types of deviations are and why they happen.
Workshops			
2025-10-28	- Industrial Supervisor - Consultant Process Mining	Short orientation in the tool with a consultant to do some sanity checks and confirm how the tool should be used.	Ensured correct data load and how to proceed with the analysis.
2025-10-31	- Industrial Supervisor - ERP Architect - O2C Development Manager - Head of Customer Service and Commercial Quality	Together with Paulig try to find interesting insights and identify problems in the processes using the Process Mining tool.	Tested the Process Mining broadly. Looked closer into customer order line charges, but could not find any systematic issues.
2025-11-05 2025-11-07 2025-11-14	- O2C Development Manager - CS Specialist - IT Solution Specialist	Look at the current insights together in the Process Mining tool and try to explain and understand the reason for some of the deviations. Present three cases with unique dashboards which were created for each workshop session.	Verified Process Mining findings and their potential reasons. More understanding about the link between the discovered process and reality, and what is defined as non-conformance.

3.2.1.1 Semi-structured Interviews

As a first step to understand the standard administrative O2C process at Paulig, semi-structured interviews were conducted with the Customer Service (CS) team. The CS team is responsible for handling customer orders, creating invoices and ensuring deliveries are fulfilled. They are therefore experienced in the O2C process, and the standard process was mapped with their assistance and knowledge. The semi-structured interviews enabled follow-up questions to clarify and complement the information gathered. This resulted in a better understanding of the standard procedures and reduced the risk of misunderstandings.

Interviews with the CS team were conducted to be able to understand the O2C process. The interviews were scheduled through mail contact with the Head of CS and Commercial Quality and were held on Microsoft Teams. The first interview was a combination of shadowing and interview questions, where a CS Specialist showed the typical customer order management steps from order receipt to invoicing. Since the study examines the various process steps in O2C, understanding the individual steps in the ERP application is important for the later analysis of data in the Process Mining tool. Questions were prepared based on the process maps and working instructions that were received prior to the interview and were asked while the CS Specialist showed the steps. Follow-up questions were asked to understand the reasons for executing some of the steps. The outcome of the interview was a better understanding of how the order flow is executed and which steps are included. Some common problems and deviations during the process were also shared.

The second interview was conducted together with the O2C Development Manager. It was scheduled after being in contact with the Head of CS and Commercial Quality. This interview helped understand the provided O2C process maps and order management further, while also increasing knowledge about how exceptions are handled. The interview was carried out in a similar manner, but was only conducted as a semi-structured interview without any major shadowing. Follow-up emails were sent afterwards for further clarification on the processes.

3.2.1.2 Workshops

During the live period of the project, a series of workshops was conducted to support the iterative analysis using the Process Mining tool. One workshop was held together with an Infor consultant, followed by four workshops with Paulig representatives.

The initial workshop with Infor focused on verifying that the correct event log data had been loaded into the tool and reviewing the default dashboards to ensure that the relevant process activities were included. The second workshop concentrated on defining appropriate filters in the Process Mining environment in order to isolate the coffee business. This included excluding non-coffee products such as coffee machines, spare parts and services, focusing on the Finnish warehouses, ensuring coverage of the correct time period and narrowing the scope to primarily manual customer orders. This workshop was conducted with Paulig, where both an ERP architect and process owners participated.

The last three workshops were used to identify and analyze concrete insights related to inefficiencies and bottlenecks in the O2C process. These sessions followed a sprint-like and iterative approach, where a set of three to five cases was prepared between the workshops and then discussed in detail during each session. Paulig's representatives contributed process-specific knowledge and used M3 to compare data from completed customer order lines with the observed patterns in the Process Mining tool. This made it possible to validate the findings and distinguish between system-recorded behavior and actual process execution. After each workshop the most relevant cases were documented for a later presentation to a broader group of stakeholders at Paulig. For these cases, more detailed dashboards were developed to complement the standard dashboards provided by Infor, allowing a targeted analysis of selected issues identified during the workshops.

3.2.1.3 Presentations

Several presentations were held during the project where representatives from all functions involved in Paulig's O2C process participated. These sessions enabled joint discussions regarding the findings in the Process Mining tool and the overall process. Given the size and complexity of the O2C process, it is not feasible for a single individual to have detailed knowledge of all aspects. Bringing together process owners, super users and managers therefore proved valuable for developing a shared understanding of both the overall process and specific problem areas. In the largest session, more than 20 participants with different areas of expertise took part, allowing both the Process Mining tool and process-specific details to be discussed. This helped establish a clear link between the intended and actual process execution.

Alongside the sessions conducted with Paulig, the project also included several internal presentations at Meridion. These sessions functioned both as an introduction to Infor's Process Mining tool and as a forum for discussion, where colleagues could share feedback. The discussions primarily addressed which parts of the O2C process were most relevant for further analysis and helped strengthen the contextual link between Process Mining insights and M3.

3.2.1.4 Literature Review

To get a better knowledge of the subject of Process Mining, articles from various journals have been read and gathered in the report. This includes information about what Process Mining is, its use applications, opportunities and limitations. Key concepts have been included to establish a fundamental base for the study. The sources for the theoretical framework have been critically evaluated to ensure high academic value, and included articles from reputable journals and databases, with authors who are experts in the field.

3.2.1.5 ERP & Process Mining Courses

To enable background knowledge and facilitate understanding of the ERP system, training sessions were conducted together with the industrial supervisor in M3. The purpose of the training was to increase familiarity with the system and covered how customer orders are created, allocated, released and invoiced within the O2C process. It also included an overview of how orders are handled in the warehouse, including allocation logic and different allocation rules. The training supported a clearer understanding of how the process is executed in the ERP system.

To understand how the Process Mining tool works and is applied, courses were taken within the field. The courses are based on Infor's approach to Process Mining and are part of a course package including tools for RPA and GenAI. However, since the report only addresses Process Mining, the courses have been taken with a larger focus on the Process Mining tool. The content of the courses included both the background to Process Mining and its purpose, but also more practically regarding the features of the Process Mining tool and how it is used. The courses covered all aspects and functions of the tool, including how to define, load and analyze data. The course package set the foundation and understanding for how the tool functions and increased preparation for the live access period.

3.2.1.6 Documents

An important part for the success of the project was to get a good understanding of the O2C process at Paulig. Through meetings and mail contact, relevant documents of the process were provided by representatives from Paulig. This included process maps of the O2C process along with work instructions for the order management process. Representatives from Logistics provided process maps of the logistical part of the flow. These process maps were analyzed and served as a base for the later comparison with the actual process map, discovered from the Process Mining tool. The user manuals and process maps enabled a better understanding of the O2C process and how the tasks are carried out in detail.

3.2.2 Quantitative Data

Quantitative data was used to analyze the actual execution of the O2C process based on M3 event logs. This data enabled objective measurement of process flows, cycle times, variants and deviations. The event log data was extracted from Paulig's ERP system and formed the basis for discovering and analyzing the "as-is" O2C process. Quantitative analysis allowed identifying patterns, bottlenecks and non-conforming behavior, which were later interpreted and validated using qualitative insights from the organization.

3.2.2.1 Event Log Data

Event log data, which is collected by the ERP system Paulig uses, was extracted and used in the Process Mining tool. The data was collected through M3 with the help of consultants from Infor for a duration of roughly 6 months prior to the start of the thesis project. The event log data includes information about various actions that users take in the ERP system such as creating sales orders, creating invoices or changing the order quantity. The data collection is pre-configured by Infor, selecting which activities to monitor based on the O2C process at Paulig. Access to event log data in combination with the Process Mining tool was scheduled between weeks 44 and 48. Once access was granted, the data needed to be filtered and cleaned. An analysis of how the data was connected to the different process steps in the O2C flow was carried out to understand the underlying logic and to ensure data quality and accuracy.

3.3 Methodology for Each RQ

To answer the research questions, relevant methods have been used for each question. In Table 3.3, the methods for answering the three research questions are displayed to give an overview of the approach.

Table 3.3: An overview of the methodology used to answer each research question.

Research Question	Methodology
How accurately do pre-defined process models reflect actual O2C processes?	- Process Mining tool with provided event log data - Process Mining courses - Provided process map documents - Interviews and shadowing
Which inefficiencies and bottlenecks in the O2C process can be identified through Process Mining?	- Process Mining tool with provided event log data - Interviews and workshops
What are the opportunities and limitations associated with the application of Process Mining tools in organizational process analysis?	- Process Mining literature - Working in the tool

3.4 Live Access Period

The following section presents the approach to the Process Mining tool during the live access period. It covers how the tool was configured, used and what delimitations and inputs were needed in order to analyze Paulig's data. Access to the tool was granted for a duration of 5 weeks.

3.4.1 Start-up & Prerequisites

Various formal requirements were needed in order to gain access to the Process Mining tool. The provisioning of the tool required coordination between all involved actors. Infor activated the module for the tool in Paulig's system, hence access to Paulig's M3 environment was also needed. After this was in place, access to enter the Process Mining tool in Paulig's environment was made possible.

3.4.2 Delimitations

In order to conduct a more focused study, certain delimitations had to be made based on input from Paulig. For example, due to the company's recent transition to M3, the full O2C process was not operational until the start of June 2025. At the start of the live access period, which was at the end of October 2025, it was also decided to only include full months of data. Therefore, the end date range of the data was set to the end of October 2025. Through the interviews and workshop sessions that were conducted with the O2C Development Manager, it was also recommended to only look at a selection of warehouses and order types. Their main in-house warehouse is designated as A01, and their main third-party logistics (3PL) warehouses are designated as A04 and A06. These three warehouses were in focus for the analysis. Furthermore, since the main focus was on the manual flow, the order type MAN was used as it involved the highest level of manual tasks. In the discovered process, a return flow is included. In this study however, this flow was not examined in detail as the number of cases was relatively limited. Additionally, to enable a focused and comparable study, the focus was solely on coffee products, not coffee machines or spare parts.

3.4.3 Loading Data

Paulig's data had already been logged by Infor, meaning it was already ready to be loaded into the tool. However, some circumstances had to be taken into account in order to load more accurate and relevant data. The first consideration was the date range of the data. It was known from communication with Paulig that the ERP system is new for the company, and was at the time of this thesis being successively adopted throughout the organization. For this case study, data has been logged since the start of Spring 2025, meaning that data points from the test period was included in the initial load as well. Therefore, even though the start date could be set in 2024 in the tool, it was decided and instructed by an Infor consultant to set the start date on Mar 1st and end date on September 30th in order to cover the most relevant data. After a short period of working in the tool and receiving new information from Paulig's O2C Development Manager, this date range was changed to Jun 1st - Oct 31st to cover more recent and relevant data.

3.4.4 Filtering Data

The data can be displayed in several ways in the Process Mining tool. The tool offers the filtering options stated in Table 3.4. These filters can be applied either on a global level, regional (tab) level or local (widget) level. If the filters are applied on a global level, it will be applied to all tabs on the page. If the filters are applied on a specific tab, all widgets and tools under that tab will have those filters applied. Filters applied to individual widgets will only affect those widgets. This means that a page can display different data visualizations with separate filters.

Table 3.4: Available filters in Infor's Process Mining tool.

Available Filters	Description
Attribute	Select cases based on column values (OrderType, Warehouse, CustomerName etc.).
Event	Select cases including or excluding certain activities (Customer Order Created, Picking Performed etc.)
Process Flow	Select cases with specific flow. Ensures cases include an event and a consecutive event, either followed directly, eventually or not at all.
Flow Start & End	Select cases that start and/or end with a specific event.
Predefined Metrics	Select cases with specific predefined metrics (e.g. conforming/non-conforming cases).
Date Range	Select cases with first event timestamp in date range.
Case Cycle Time	Select cases where the duration of the case cycle time is faster/slower than a defined period of time.
Events Cycle Time	Select cases where the duration between two events is faster/slower than a defined period of time.

Filtering the provided data is critical to facilitate a comparable analysis. As previously mentioned, defining the date range is the base for the entire analysis. This ensures the same data is being analyzed and that the Paulig's O2C process after their transition is fully in place by not setting a too early start date. In Table 3.5, the base case filters used in the project are shown, including the date range, order type, warehouses and defined start and end events. The start and end events are set to ensure that only full O2C cycles are included in the analysis.

Table 3.5: Base case filters.

Filter	Set Value
Date Range	June 1st - October 31st
Order Type	MAN (manual)
Warehouses	A01, A04, A06
Event Start & End	Customer Order Created –> Payment Received

3.4.5 Working in the Tool

The analysis in the Process Mining tool followed an iterative and exploratory approach, where the level of detail was gradually increased as understanding of both the data and the process improved. Since a large amount of information is available in the tool, the analysis was initially conducted on a high level to gain an overall understanding of the O2C process execution.

In the first phase, the analysis focused on general process indicators such as overall KPI:s, number of cases, number of variants and average cycle times. The overall process flow was also examined using visualizations such as the process map, spaghetti bowl view and the most common paths. Both conformant and non-conformant flows were reviewed. The purpose of this initial analysis was to understand the "as-is" process and to verify whether the data-driven process model corresponded with the understanding obtained from interviews and standard process maps.

After this initial validation, the analysis became more exploratory. Different tabs and views in the insights section of the tool were examined, and various filters were applied to observe how the process behavior changed under different conditions. This included drilling down into specific cases and event sequences to better understand deviations and patterns observed at a higher level. Particular attention was given to activities related to customer order handling and manual interventions. Events such as manually changed quantities or delivery dates, blocked orders and other manual adjustments were identified and analyzed. These activities were defined as non-conforming and the frequency and context in which they occurred were examined to understand why they happened and how they affected process performance.

An analysis of the cycle times was also performed during the study. Including the *Payment Received* event added several weeks to the overall case duration due to varying payment terms. To allow for a more meaningful comparison of operational performance, the analysis was therefore divided into two parts. This consisted of a full cycle view, covering the events from *Customer Order Created* to *Payment Received* and an operational cycle view, covering the events from *Customer Order Created* to *Customer Order Line Invoice Processed*. This separation made it possible to distinguish between operational inefficiencies and inefficiencies across the whole flow. To support this analysis, the data model was extended by adding attributes to the customer order line level. This included attributes such as *Payment Term* and *Reason Code*, which were not initially available. The *Payment Term* attribute was used to compare Invoice-to-Payment cycle times and to assess how well actual payment behavior conformed to the agreed terms.

As workshops with Paulig were conducted and process understanding improved, more precise and relevant filters could be applied. The analysis was narrowed down to specific order types, with a focus on manual orders, as well as relevant date ranges starting from the point when the ERP system was fully operational. Additionally, the analysis was limited to selected warehouses in Finland that were considered most relevant for the study. With these filters in place, the analysis became more focused, allowing for detailed examination of selected cases and deviations. Particularly interesting cases were documented and for some of these, individual dashboards were created within the tool to visualize the findings more clearly and enable easier comparison between scenarios.

3.5 Reliability & Validity

Reliability concerns the consistency and stability of the data and the extent to which the same findings would be obtained if the study were repeated under similar conditions [32]. For the qualitative part of the study, reliability is influenced by the scope of interviews and observations. Only one CS representative was interviewed in depth, and only one full shadowing session was conducted. Interviewing additional CS representatives could have provided a broader perspective and increased the robustness of the qualitative findings. Warehouse-related activities were primarily analyzed through Process Mining and workshops, rather than through individual interviews with logistics personnel. Although logistics representatives participated in some workshops, the lack of dedicated interviews limits the level of detail regarding warehouse operations. This led to a stronger focus on order handling and administrative activities compared to physical warehouse processes.

To strengthen reliability, similar questions and observations were discussed with multiple participants during workshops. This allowed recurring issues to be compared across different roles and functions, reducing the risk that findings were based on isolated opinions or individual interpretations. However, many examples discussed during interviews and workshops referred to past cases. This introduces a risk of recall bias, since it can be difficult to remember details accurately over time.

For the quantitative data, the event logs were generated automatically by the ERP system prior to the start of the project. The authors did not participate in the data collection process, which limits the ability to verify data accuracy at a technical level. Nevertheless, the overall structure and content of the data were reviewed together with the CS team and was considered reasonable and representative of actual process behavior.

Validity refers to whether the study measures what it intends to measure and whether the findings meaningfully reflect the real processes under investigation [32].

Face validity was supported by close collaboration with Paulig employees. The pre-defined process models, identified deviations and interpreted bottlenecks were reviewed and discussed with process owners and operational staff, who confirmed that the results aligned well with their experience of the O2C process. This aligns with the concept of face validity, where expert judgment is used to assess whether the measurements appear reasonable and relevant [32].

At the same time, validity is affected by data scope limitations. Only events recorded within the ERP system were available for analysis. Activities performed outside the system are not captured in the event logs. These missing elements may explain certain deviations or delays observed in the data. The qualitative methods were therefore essential to contextualize the quantitative findings and reduce the risk of misinterpretation.

3.6 Use of AI tools

In line with current recommendations for transparency in academic research, this thesis made limited use of the AI tool ChatGPT, developed by OpenAI. The tool was used solely as a support during the writing process to refine language, improve sentence clarity and enhance the overall structure and readability of the text. ChatGPT was not used to generate content, empirical material, analyses or conclusions. All conceptual, analytical and substantive work presented in this thesis was produced by the authors. The authors also take full responsibility for the accuracy, coherence and integrity of the final text, including all revisions made with the support of AI.

4

Empirical Data

This chapter presents the empirical data used in the study. It introduces Infor's approach to Process Mining and how the tool is positioned in relation to other analytical solutions on the market. Based on interviews and shadowing sessions conducted with Paulig's process owners, a conceptual understanding of the company's O2C process was developed, with a primary focus on the standard order flow and exception management. The chapter concludes with an initial overview of how Paulig's actual O2C process appears in the Process Mining tool.

4.1 Infor

Infor is a global enterprise software provider offering business applications for more or less all industries. The company has several ERP systems, where M3 is one of the most widely adopted in northern Europe. M3 supports core operational processes such as order management, procurement and production [33].

4.1.1 Infor's Process Mining Approach

Infor has expanded its portfolio in recent years with advanced analytics and automation capabilities. One of these additions is a Process Mining tool, which together with RPA and GenAI forms *Infor Velocity Suite* [34]. Within this package, the Process Mining component provides the diagnostic foundation that enables automation and optimization initiatives.

Since Infor provides both the ERP system and the Process Mining application, the event logs generated in M3 are directly integrated into the analytical environment. This leads to several advantages:

- The event logs follow a consistent structure.
- Data objects and process maps are pre-defined, enabling simple and automatic matching to the event logs.
- Analysis and visualization occur within the same ecosystem.

This speeds up the setup and reduces the complexity of gathering data for customers whose processes mainly run in M3. Event logs from other systems can also be added, but it requires more preparation to make the data fit the M3 structure. Process Mining within *Infor Velocity Suite* corresponds to the backward-looking Process

Mining described in the theoretical background part of the report.

4.1.2 Pre-defined Process Maps

Infor uses an Industry Process Catalog (IPC), which is a pre-configured set of industry-specific process maps that includes the Food & Beverage sector [35]. These maps serve as pre-defined BPMN models for Process Mining and outline the typical workflows expected in each industry. Since the maps are designed to suit many different companies within the same sector, they remain relatively general and may include steps that are not used by every organization. They are therefore intended to be scaled down and adapted to better represent the customer's operations. To support this customization, Infor categorizes process activities into three groups:

- **Core activities:** Standard activities that most companies in the industry use. They can be applied as they are and make the implementation fast and straightforward.
- **Differentiator activities:** Activities that are common but need small adjustments to fit the customer's way of working. They require some extra time and effort to set up.
- **Unique activities:** Activities that are specific to one company and do not follow industry norms. They need their own tailored solution and take even more time to implement.

The advantage of this structure is that most parts of the workflow can be analyzed with minimal setup since the process is relatively standardized, while still allowing company-specific steps to be added if that level of detail is needed.

4.1.3 Standard Dashboards

Infor's Process Mining tool includes a set of standard dashboards that give a general overview of the process and support backward-looking Process Mining. These dashboards make it possible to compare processes between warehouses, item groups and other attributes. The dashboards also include common KPI:s such as cycle times. As a starting point, they help users quickly understand how the process works and indicate where deviations occur. Since these dashboards are designed to be applicable for a wide range of companies within the Food & Beverage industry, they are intentionally generic and aim to cover common process patterns rather than company-specific details. As a result, they offer limited depth when it comes to identifying more specific challenges in the process. Their role is therefore not to provide final answers but to guide attention toward areas that may require further investigation.

Some standard views are useful as a starting point for a deeper analysis. The event list makes it possible to identify the most frequent non-conforming events, while process discovery illustrates how activities are connected and how different execution paths emerge in practice. These views provide an initial understanding of where inefficiencies or unexpected behavior may be present.

The main value of the tool is realized when the insights from the standard dashboards are used as input to create more detailed dashboards. These custom analyses allow a closer focus on specific subprocesses, customer groups and operational conditions that are relevant to the organization. Examples of such custom dashboards and the motivation behind them are presented in Chapter 6 about bottlenecks and inefficiencies, where the analysis moves from high level observations to more detailed and actionable insights.

4.1.4 Comparison to Other Process Mining Approaches

Most Process Mining tools on the market are designed to work independently of any specific ERP system. This means that process models are typically built from scratch and the event data is gathered from multiple information systems [36]. For these tools to function, all data must be mapped, cleaned and transformed into a common structure, which often requires a considerable amount of preparation. The benefit of this approach is flexibility, since organizations can include data from many systems and create highly customized process models that reflect their full end-to-end workflows. The downside is that merging data from various systems often extends the implementation time.

Infor's approach differs from most other Process Mining solutions by focusing on efficiency within the M3 environment. Because the underlying data structure is already defined, insights can be produced quickly with relatively little setup. This advantage depends on how much of the customer's process actually runs inside M3. When organizations use additional systems or have many company-specific activities, extra work is required to prepare external data so it can be analyzed together with the events extracted from M3.

4.2 O2C Process Flow Chart

This section presents a conceptual description of the O2C process flow in Paulig's operations in Finland and the Baltics. Figure 4.1 displays Infor's pre-defined O2C reference model, which is designed to be representative of typical O2C processes within the Food & Beverage sector. This reference model serves as the baseline against which Paulig's actual process execution is later compared in the analysis.

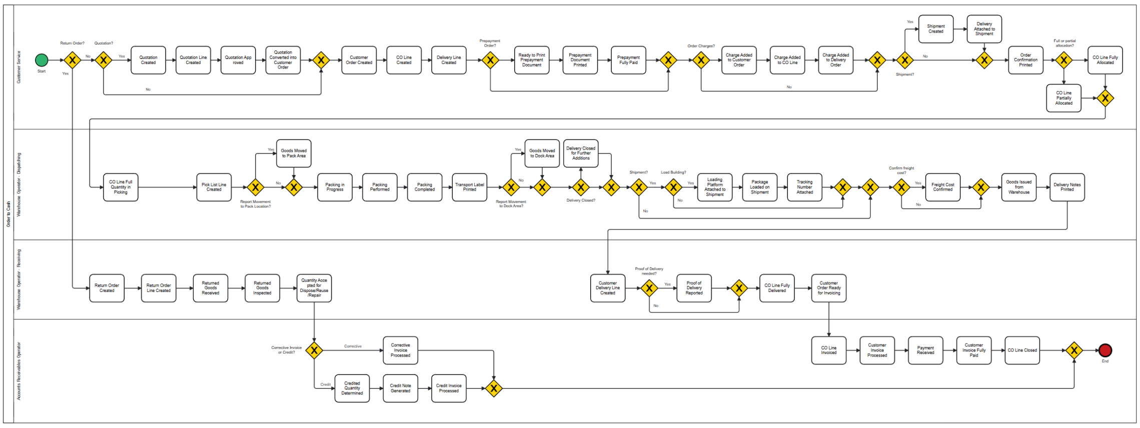


Figure 4.1: Infor's pre-defined O2C process for the Food & Beverage industry.

The reference model is divided into four swimlanes. The upper one represents the initial order handling activities performed by the CS team. The two middle swimlanes illustrate activities typically carried out by the Logistics function, while the bottom field relates to invoicing activities performed either by the CS team or the Finance department.

The pre-defined model consists of a main flow and a return flow. In the context of Paulig's coffee business, customer returns occur very rarely and were therefore excluded from the scope of this study. The main flow represents the standard order process, describing how customer order lines are intended to progress when orders are executed according to defined procedures and no changes are made to delivery dates, quantities, or similar parameters after order creation. This standard flow is largely automated. When changes do occur, they are handled through exception management routines, which are generally not automated and require manual intervention. It is primarily within these exception-handling situations that Process Mining can be used to identify inefficiencies and bottlenecks in the process.

4.2.1 Standard Order Process

This section describes how customer orders are received and handled at Paulig. The information presented is primarily gathered from the interviews, shadowing and workshops with the CS representative and O2C Development Manager.

The interviews showed that almost all orders are received into the ERP application automatically through an integration with another application. In this case, Paulig and their customers use an EDI application where the orders are entered by the customer and then sent to Paulig's ERP system. Some orders are sent from the customers by email, which could include Excel- or PDF-files. These are in turn automatically scanned with another tool and sent to the ERP system, where the corresponding customer order is created.

If all details in the order are correct, such as delivery date, quantity and item number, the order is released for picking by the CS representative to one of the warehouses. This is usually done one day before the orders are scheduled to be dispatched. The goods are then picked, packed and prepared for delivery. The CS representative plans the transportation for each order through different booking portals, based on which carrier is used, and a pickup is booked the day after picking. For many customers, the shipments are also booked or pre-advised in their own portals. Every order is checked with the Sales team, who then verify and confirms the order to CS, allowing them to release it. For example, Sales verifies that the customers do not order too much in relation to what has been forecasted. The Sales department does not have access to M3 and can therefore not perform any order changes. Instead, they need to communicate changes to the CS representative who then updates the orders accordingly if needed. The communication regarding orders between CS and Sales is mostly done by sending the customer order as a PDF by email to the Sales team, who then checks the order and later confirms it to the CS department.

The transports to the Baltics are booked by the CS team in Estonia, with approximately one day delivery time to Estonia and two days delivery time to Latvia and Lithuania. Depending on the destination of the goods, different carriers are used. After shipment, the order is automatically invoiced through another EDI system. The status of the invoices can be seen in M3. Invoices are either sent as electronic invoices or as PDF invoices by email. Usually, the invoices are not checked by the CS representative, since this is an automatic process and should not be necessary. Some customers also request a so-called CMR, which is an established document used for road transport and is printed through the ERP system. What differentiates the CS team in Estonia from other CS teams is that they handle transportation and claims as well, while the other CS teams only handle customer orders.

4.2.2 Exception Management

In this section, the deviations and exceptions during the standard order process will be presented. One of the first errors that occurs in the order flow is during the EDI transfer to the ERP system. Some orders receive an error status, which could be caused by for example a customer putting an incorrect order date. Orders may also go to error because of missing customer master data, missing item codes or not sufficient stock of an item. During the interviews, some issues related to specific customers were also highlighted. For example, when an attempt to change an already created order was made, it resulted in an error in the system.

The CS representative explained that Available-to-Promise (ATP) checks are not routinely performed. ATP refers to the process of verifying whether sufficient inventory or production capacity is available to fulfill a customer order within the requested timeframe. In cases involving exceptionally large order volumes, the Sales team notifies the Forecasting team to assess feasibility. If an order deviates significantly from typical order quantities, either by being unusually large or unusually small, the Sales team may request the customer to adjust the ordered quantity accordingly. Such situations occur infrequently. When an order cannot be fulfilled in the requested quantity, it is either partially reduced or fully cancelled, and backorders are not created.

A specific exception which the CS representative encounters when handling the customer orders is that certain items have the wrong packing or item code when they enter the ERP system for some customers. For some reason, the customer cannot place the correct item code when sending the order. To handle this, the CS representative needs to manually change the item code every time this specific item is ordered. The background to the problem is that the item enters the system in a single-piece format, even though it should be in a half-pallet format. The CS representative handles the issue by copying the order line, changing the item code and then deleting the original order line.

With regard to pricing during campaigns, prices are generally communicated accurately by the Sales team. However, there have been instances in which the CS representative have needed to adjust prices on certain orders prior to the start of a campaign, for example due to calculation errors. To mitigate such issues, the CS representative reviews and verifies all prices in advance whenever a campaign is scheduled, ensuring that pricing information is accurate before implementation.

As previously described, the majority of the invoicing process is automated. A current issue in the EDI system however prevents invoices from being correctly transmitted to customers in Latvia and Lithuania. To address this issue, the CS representative manually prints the invoices from M3 and sends them via email. Some customers may request PDF copies of invoices, for instance if the original invoice has been misplaced or due to other administrative issues. These requests are also handled manually via email. The CS representative reports that certain customers have specific requirements regarding the information included on the invoice. For example, some customers request a preliminary text containing details such as the number of pallets included in the order. Although this information is typically already provided on the delivery note that accompanies the physical shipment, some customers also request it to be included on the invoice. The underlying rationale for this request remains unclear.

According to the O2C Development Manager, most inquiries and complaints from customers are received by email, while some are received in the ERP system. If there is no stock for an order, it will appear as an alert in M3, notifying the user that the order cannot be fulfilled. The same logic is applied if there is no suitable stock available, such as customer requirements on Best Before Date (BBD) not being met. In these cases, the customer is contacted to see if they decide to continue with the order or cancel it.

Customer claims are currently managed through the Issue Management Portal (IMP), a module within M3 primarily intended for handling quality-related issues. Some entities also use the IMP to register transport-related claims. As a result, the system is currently used for multiple purposes. In the future, the intended approach is to limit the use of the IMP exclusively to quality-related claims. Transport-related issues, including failed deliveries and other non-conformances not associated with product quality, are planned to be managed through alternative processes outside the IMP.

The most common types of exceptions and reasons for claims at Paulig are the following:

- Wrong item
- Picking error
- Wrong BBD
- Customer rejections
- Customer returns
- Pricing error
- Missing discount on invoice
- EDI failure

There have been instances in which customers initially accept an order but later return it. In such cases, the underlying reason for the return is not immediately apparent and requires further investigation in collaboration with relevant departments, such as warehousing, transportation or quality assurance. Some exceptions are related to pricing discrepancies, including incorrect invoice pricing or the absence of agreed discounts.

Paulig currently uses two systems to receive customer orders, which work in different ways. The first system receives orders through emails with PDF:s, scanning them and sending the data to M3. The second system receives the orders through EDI, which are then sent to M3. The usage of each system varies between geographical regions. The UK and Belgium use the first system more frequently, since the customer base consists of smaller wholesalers. For the Baltics, the EDI system is used to a larger extent. During the interview with the CS representative, it appeared that some orders go into error in the EDI order system. The O2C Development Manager was not aware of this situation and the cause for it, as the interviewee lacked experience in that particular system. Nevertheless, it was mentioned that there may be several causes for orders going into error:

- No price setup
- Wrong delivery date
- Exceeded credit limit
- Master data issues

As previously mentioned, it appeared that the O2C Development Manager was not aware of CS daily interaction with Sales to confirm orders. The O2C Development Manager stated that they were not aware of the extent of constant contact with the Sales team on every order, describing this level of interaction as slightly alarming. They were aware that Sales was contacted during promotions to verify stock availability and manage call-offs for specific quantities, but not for confirming regular orders.

Regarding transportation charges, the payment responsibility often depends on the agreement and Incoterms. This includes terms such as DAP (Delivery at Place) or EXW (Ex Works), with approximately 90% of Paulig's business applying DAP, meaning the selling company pays for the transportation. These charges are generally included in the product price and are not frequently added to the orders. While service charges might be added manually as an exception, in this case about 10% of the time, charges are often automated and can be pre-set in the customer setup in the ERP application.

The use of full trucks is prevalent in Finland and Estonia. In contrast, the UK sees much fewer full trucks, as some customers only order 1, 2 or 3 pallets. These smaller UK orders are often grouped together into bulk shipments going to similar geographic areas, functioning like a "milk run". For exports, grouping deliveries is standard practice. Orders destined for Latvia and Lithuania are typically consolidated and leave on Friday for a delivery on Monday, and UK exports to Ireland are consolidated to leave on Thursday for a delivery on Monday. The mechanism for ATP and BBD checks is currently undergoing enhancement within the ERP system. For items produced in the UK, the ATP check is currently managed by a 3PL using various external tools. The goal within M3 is to use an ATP tool that flags orders that cannot be fulfilled, although current functionality lacks adequate notifications to the workspace or email. The BBD requirement is managed by the ATP tool, which flags issues if the available stock does not meet the customer's BBD requirement, such as requiring stock with 50% or 182 days remaining of expected BBD. CS is instructed to check this tool multiple times a day. If flagged, CS must take action, such as canceling part of the order. This could occur if for example 200 items are available instead of 300, otherwise the customer needs to be contacted to secure acceptance of stock with a slightly shorter BBD. Regular ATP issues are identified via order statuses. For example, if an order remains in status 22 (reserved) in M3 and has not been allocated 5 days before the delivery date, it indicates a stock issue that requires investigation. CS must then check the material plan or contact production or supply planners to determine if stock will be available before loading.

The interval for receipt of customer orders is generally between two to five days before shipping. The standard procedure requires orders to be released to the 3PL or warehouse two working days before delivery. While the standard is preparation two days before delivery, export customers may require more time, occasionally up to a week before shipping. Regarding forecasts, the system relies on weekly forecasts used by production. All coffee products are generally handled as Make-to-Stock (MTS). Stock levels are often high for products with a long shelf-life, though managing stock for short shelf-life is considered more complicated. Should stock shortages occur, there are established rules dictating how products are allocated to customers. For campaigns, there is a requirement for customers to commit to a certain quantity through sales. Campaigns, such as those involving special Mother's Day packaging, occurring four times per year, are typically limited to a specific time period and stock quantity until they are sold out.

4.3 Process Mining Findings

In this chapter the discoveries from using Infor's Process Mining tool will be presented. Based on the methodology and previous setup and decided parameters, the cases are shown in various visualizations and formats. The displayed data is the result of using the base case filters, which includes the date range, order type, warehouses and defined start and end events. The first visualizations shown are suggestions or standard layouts by Infor, acting as an example of how the data can be displayed. These may not be relevant or sufficient for all companies, depending on the scope of interest. However, the tool supports creating own insights and visualizations, further adapting it to the user's needs and enables a more detailed analysis. These insights have been created and will be showcased later in the results.

4.3.1 First Overview

The first visualization that appears when entering the insights, shown in figure 4.2, is the Overview tab. This tab includes information regarding how many cases are present, cycle times, number of events and variants, order amount and number of customers. The order amount and cycle time are censored for integrity purposes. The result of the base case filters shows that 6,740 cases are available to analyze, with a mean end-to-end cycle time of roughly a month. The number of events logged is a total of 192,090, meaning that each case consists of almost 30 activities on average.

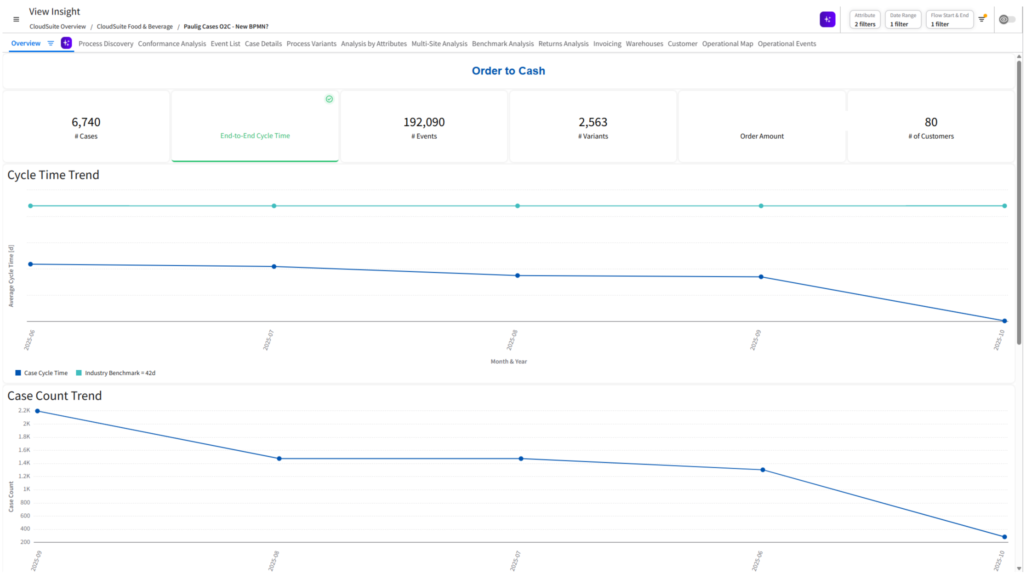


Figure 4.2: The first overview in the Process Mining tool.

4.3.2 Discovered Process Flow

The Process Discovery tab displays the most common conformant path of events in a similar manner to the later described Process Variants tab. The main difference is that, instead of selecting specific variants to display, the visualization can be controlled by specifying the proportion of total events to include and the number of connections to show between events. In Figure 4.3, these parameters are set to their maximum values, resulting in a visualization that includes the largest number of events and connections, often referred to as a "spaghetti bowl". This representation shows the complexity of the processes and provides information on which activities occur in practice and to what extent.

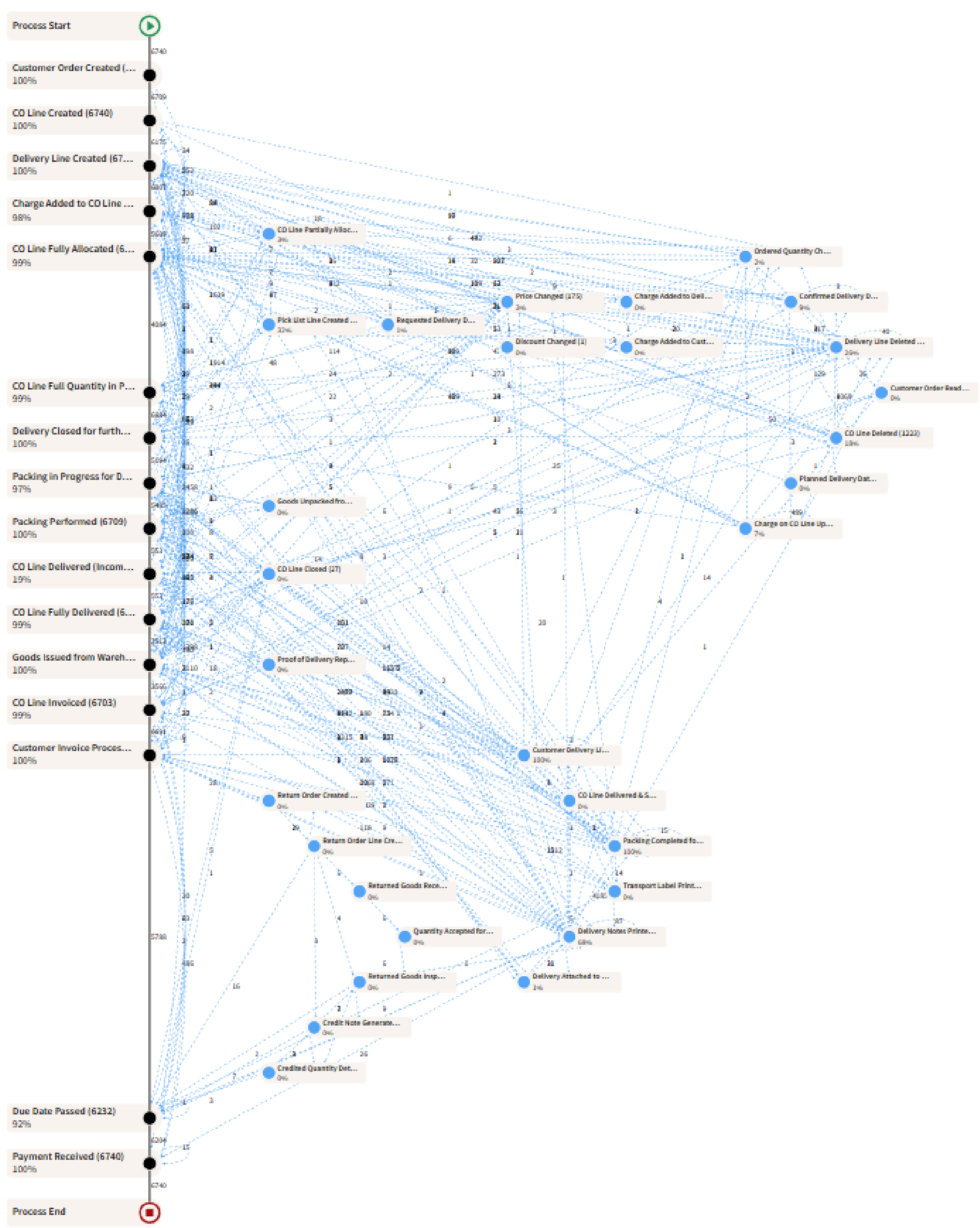


Figure 4.3: A spaghetti bowl of Paulig's processes, showing all events and connections.

4.3.3 Process Variants

In Figure 4.4, the 15 most common conformant paths of events for the O2C process at Paulig are shown. In this example, there are 2,081 cases out of the 6,740 that follow one of the paths to the left. Almost all of these variants are not unique processes in the sense that they follow different events, but instead include all the events of the most common path but with a variation of the ordering. There is one variant that has an additional event that includes *Pick List Line Created*, which is not present in the other variants.

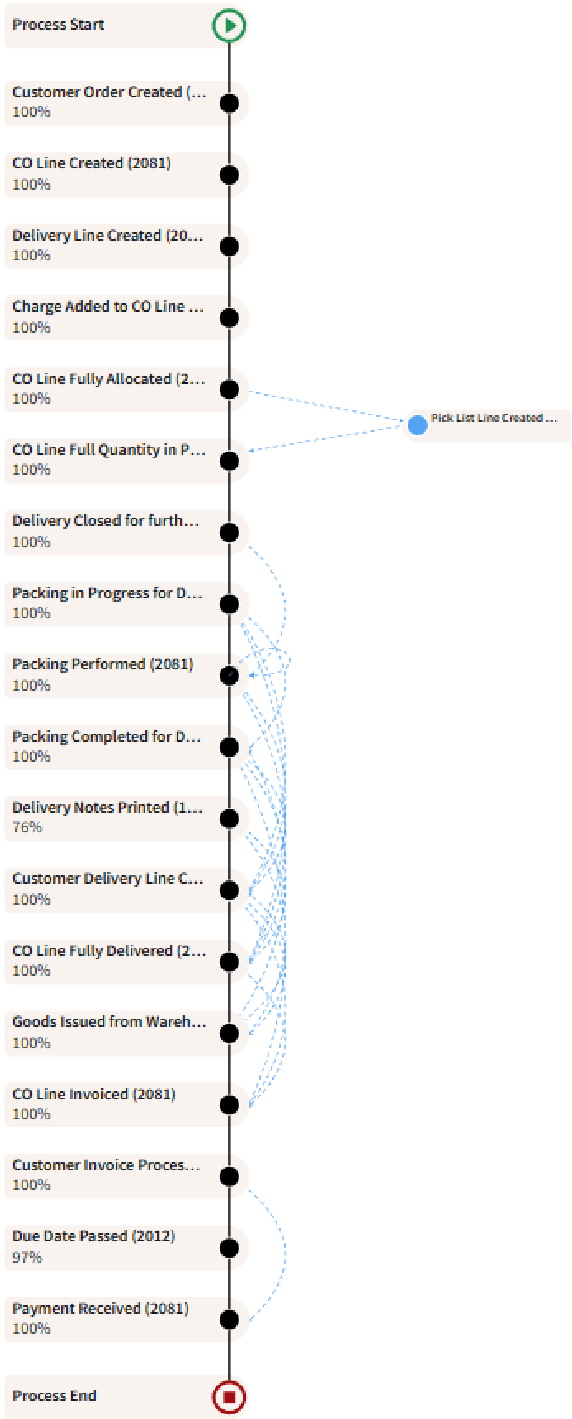


Figure 4.4: The most common conformant case including the top 15 process variants.

4.3.4 Conformance & Non-conformance

In the Conformance Analysis tab, the different paths that each variant takes throughout the O2C process are displayed. In Figure 4.5, the most common conformant case is shown. There are 69 cases out of the 365 conformant cases that follow this exact path between *Customer Order Created* and *Payment Received*. The visualization not only shows in which order the activities occur, but also the average time between activities, in practice indicating how much time an activity takes to be performed.

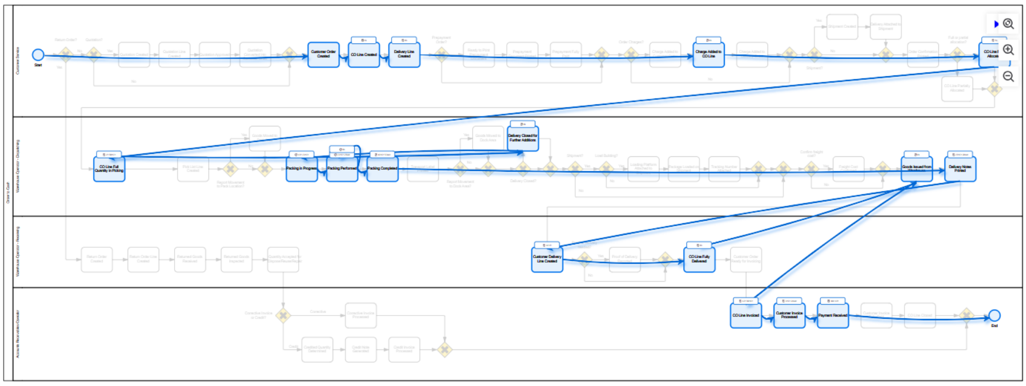


Figure 4.5: The most common conformant case.

Out of the 6,740 total cases, 6,375 cases are marked as non-conformant. Cases are defined as non-conformant when they include events that are not part of the standard pre-defined process map. Usually these are unwanted events that should indicate some kind of error or deviation. In Figure 4.6, the most common non-conformant case is displayed. 481 cases follow this exact path and deviate because of the *Due Date Passed* event, which shows that the receivables invoice has not been paid on time.

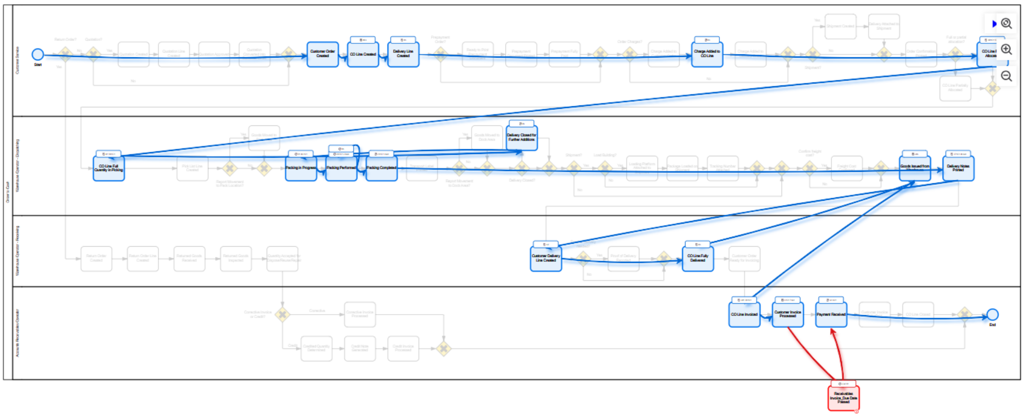


Figure 4.6: The most common non-conformant case.

The *Due Date Passed* non-conformant event is analyzed further in chapter 6, where it emerges that this discrepancy is not cause for any major concern.

5

Pre-defined Process Model vs Actual

This chapter presents the results and discussion related to the first research question. To remind the reader, the question investigated is the following:

RQ1: *How accurately do pre-defined process models reflect actual O2C processes?*

The analysis focuses on how well Infor's pre-defined reference process model represents the actual execution of the O2C process at Paulig. The findings are structured around the backward-looking Process Mining techniques, with primary emphasis on process discovery and conformance checking. Process discovery is used to examine whether the overall structure and sequence of activities captured by the tool align with the observed O2C flow, while conformance checking is applied to assess deviations between the reference model and actual process behavior. These techniques provide insight into the strengths and limitations of using a standardized reference model as the basis for analyzing real organizational processes.

5.1 Process Discovery

The process discovery analysis compares the pre-defined Food & Beverage O2C reference model with the actual process execution in Paulig's ERP system. The focus is on identifying similarities and differences between the standard flow and real behavior, providing an initial assessment of how well the reference model represents the actual O2C process, highlighting similarities and differences.

5.1.1 Similarities

Overall, the process discovery functionality reflects the actual O2C process at Paulig with a relatively high degree of accuracy. The tool provides a clear overview of the activities recorded in the database and can be seen as a direct representation of what is executed in the ERP system. The discovered end-to-end flow aligns well with Paulig's defined O2C process, starting from customer order creation and continuing through logistics and shipping to invoicing. Most of the relevant activities in the O2C process are included, together with important attributes such as order numbers, order types and customer names. Attributes that are not included by default can be added when needed. In this study, the attribute *Reason Code* was added to better understand why certain changes were made during the process. It was however not entirely successful at this stage, but the function remains relevant.

In several concrete examples, the discovered process clearly reflected how users actually work in the ERP system. One such example was the repeated change of delivery dates by a CS representative for a specific customer. This behavior could be observed directly in the tool, indicating that the discovered model accurately captures recurring manual interventions and does not only reflect the intended or designed process.

The analysis of the most common process variants further supports the view that the O2C process at Paulig is relatively standardized. The fifteen most frequent variants were very similar to each other, with differences mainly related to the sequence of activities rather than their content. These differences can largely be explained by warehouse operations being logged in different orders across warehouses. Given that a large part of the O2C flow is automated, the similarity between variants is not unexpected. However, it was also observed that even a simple change in the order of two events results in a new variant, despite the overall process remaining essentially the same. This can make the interpretation of variants somewhat confusing.

One notable observation is that the activity *Pick List Line Created* does not appear among the most common variants until variant fifteen. From a logical perspective, this activity should occur for every order, which raises questions about why it is missing from the earlier variants. This issue was not examined in detail in the study and therefore no definitive explanation can be provided. One possible explanation could be that pick lists are often released at the order header level rather than at the line level, which may affect how the event is recorded and shown in the discovered variants. Another reason could be an issue in the tracking of events, or possibly a failed auto-release of orders leading to a pick list not being created.

Some limitations and inconsistencies were also identified in the discovered process. For example, certain logistics activities appeared as loops, such as *Packing Performed* occurring multiple times within the same case. Due to limited knowledge of the detailed warehouse operations, it was not possible to verify whether this behavior reflects actual practice. One potential explanation could be that picking for a single order line is carried out in multiple warehouses with split quantities. Another observed issue was that some events appeared in a logically incorrect order, such as charges being added to a customer order line before the line itself was created. This seems to be caused by events occurring exactly at the same time and then being sorted alphabetically by the tool. While such issues may reduce precision at the event level, they do not significantly affect the overall understanding of the process or the comparison between the pre-defined process model and the actual execution of the O2C process.

5.1.2 Discovered Differences

Despite the strong alignment between the discovered process and the standard O2C flow, several important differences were identified when looking more closely at how the process is actually executed. One of the significant differences emerged from the interviews rather than from the Process Mining results. The CS representative described that they consistently need to interact with the Sales department for every customer order in order to have it confirmed and validated. In practice, this often involves sending the order confirmation to Sales for review before continuing the process. This interaction could not be identified in the discovered process model, as it is not recorded as an event in M3. This finding highlights a clear limitation of Process Mining, namely that it can only capture activities that leave a digital trace in the system, which is further described in Chapter 7. Certain parts of the process, especially informal coordination and communication between departments, remain invisible and can only be uncovered through interviews or observations. What makes this finding particularly interesting is that neither the O2C Development Manager nor the Head of CS and Commercial Quality were aware that this validation step was taking place. According to them, this way of working should not be necessary and represents an unnecessary manual step in the process, which explains their surprise when it was described by the CS representatives.

One potential way to make this order validation with Sales visible in the Process Mining tool would be to log the event *Print CO Confirmation*, as this action often precedes the interaction with Sales. However, even then it would still be difficult to determine why the confirmation was printed and what happened to it afterwards, since it could be printed for other reasons as well. More importantly, this approach requires prior knowledge of the deviation in order to configure the tool accordingly. This partly undermines the purpose of using Process Mining as an exploratory technique, since the intention is to discover deviations from the standard process rather than to pre-define what should be monitored.

Another important difference concerns the scope of the data that was available for the analysis. The Process Mining study was limited to data from Paulig's M3, meaning that only activities executed within this system could be included. As a result, the full O2C process is not captured. Several activities are performed outside the ERP system such as transport bookings, which are handled in separate external portals. Based on the shadowing session with the CS representative, these transport related activities account for a substantial share of the total processing time. The pre-defined process model does not capture this and information about this part of this process is therefore lost. Different transport booking portals are used depending on country and customer, which means that this part of the process is not standardized. These variations do not appear in the discovered process model, and without integrating data from the external portals, Process Mining tools cannot be used to analyze these activities. Understanding this part of the process instead requires more traditional methods in the form of interviews, shadowing or additional system integrations. This further illustrates how differences between the pre-defined O2C process and the actual execution are not always visible in Process Mining results, even when the core system based flow appears to match the standard model closely.

Further differences between the pre-defined O2C process and the discovered process can be identified when comparing the process maps and working instructions provided by Paulig with the process model generated by the tool. Several activities described in these documents do not appear in the discovered process. The main reason for this is that the corresponding events are not logged in the system. As discussed earlier, this applies to all activities that do not trigger a change in the underlying database and therefore do not leave a digital footprint that can be captured by Process Mining. Even if such activities were technically possible to include, they might still not appear frequently in the discovered process. Credit checks are one example of this. According to the CS representative, credit checks are generally not performed in their daily work. Due to the limitations of the available data, this part of the process could therefore not be analyzed. A similar situation applies to ATP checks, which were also reported to be used only occasionally, mainly for larger or less common orders. Within the coffee business, this approach is generally considered sufficient, as demand is relatively stable and products are MTS. As a result, strict use of credit limits and ATP checks is not seen as critical in this context. This does not necessarily apply to all parts of Paulig's portfolio. During the interview with the O2C Development Manager, it was noted that ATP checks are more important for certain products with shorter BBD:s. This illustrates that the relevance of specific activities in the O2C process can vary between brands and product categories. It is also important to note that the process manuals and working instructions are written at a relatively generic level and are intended to cover all Paulig brands, which may explain why some described activities are not commonly observed in the discovered process for the coffee business.

Another source of discrepancy between the discovered process and the logical understanding of the O2C flow concerns event timing. Several activities in the process maps are logged as having a duration of zero seconds. One example is the *Customer Order Line Created* event, which often appears to occur at the same time as the *Charge Added on Customer Order Line* event. From a logical perspective, the order line should be created before a charge can be added, and the simultaneous timestamps therefore cause some confusion when interpreting the process model. Similar timing issues were observed in the logistics part of the process, where multiple events also appear with zero second delays. One possible explanation is that these events are automatically triggered in rapid succession once an initial event occurs, causing them to be recorded with identical timestamps. As the logistics flow was not examined in detail in this study, no definitive explanation can be provided. The presence of 3PL providers further increases the complexity of this part of the process. These providers operate their own WMS, meaning that large parts of the logistics flow take place outside Paulig's ERP system and are therefore not visible in the Process Mining analysis. As a result, the discovered O2C process is further narrowed when it comes to logistics activities, reducing transparency in this part of the flow. Despite this limitation, it appears that certain key events related to the 3PL:s can still be captured, such as *Pick List Line Created* and *Goods Issued From Warehouse*. While this provides some visibility into the start and end of the logistics flow, the detailed execution in between remains outside the scope of the discovered process model.

5.2 Conformance Checking

Conformance checking and interviews were used to assess how well the pre-defined reference process model reflects the actual execution of the O2C process. As described in both the theoretical and empirical parts, a pre-defined reference model is not always the starting point in Process Mining studies. In this project, the pre-defined BPMN based reference model was used without modifications when comparing the observed process flow with the expected one. Activities classified as non-conforming were also pre-defined in the tool and not adjusted during the project, even though such customization is technically possible.

The pre-defined reference model and its classification of conforming and non-conforming activities were reviewed together with representatives from Paulig's CS function. This review showed strong alignment between the reference model and how the process is understood and managed in practice. Activities such as delivery date changes, item changes and quantity changes were confirmed to represent deviations from the intended standard flow and were therefore considered appropriate candidates for conformance analysis.

After the live period and access to the Process Mining tool had ended, it became clear that a useful extension of the analysis would have been to distinguish whether the identified deviations were driven by customer behavior or by internal conditions at Paulig. The *Reason Code* attribute that was available in the event data could potentially have been used to support such a separation. By grouping deviations based on whether a change was triggered by a customer modification or by internal constraints such as insufficient inventory, a better understanding of the underlying causes could have been achieved. If deviations were found to originate mainly from customers, this would open up possibilities for targeted communication or guidance toward those customers. If the causes were internal, the findings could instead support improvements related to forecasting accuracy, inventory availability and planning of the physical flow. This type of segmentation was not explored within the scope of this project, but it represents a relevant direction for future studies, especially since the necessary data is logged in the ERP system.

The analysis also revealed that some non-conforming events should be interpreted with caution, as they may represent false negatives rather than true process deviations. Events such as *CO Line Delivered (Incomplete)* and *Delivery Line Deleted* are automatically classified as non-conforming when an order line is split across multiple warehouses. This behavior is a consequence of how the ERP system reallocates quantities when a split occurs, which leads to changes in order line quantities and the creation or deletion of delivery lines. In these cases, the non-conforming classification reflects technical system behavior rather than an operational problem, and the context of the event is important for correct interpretation.

5.3 Main Findings

The pre-defined Food & Beverage reference process model provides a clear representation of the overall structure of Paulig's O2C process, capturing the main activities and their expected sequence from order creation to payment. This supports the view of O2C as a standardized end-to-end process and confirms its suitability for analysis using Process Mining based on event data from information systems. From Infor's perspective, the activities captured in the reference model primarily represent core process elements described in Chapter 4.1.2. These activities are common across organizations and largely supported by standard M3 functionality. While this is sufficient for an initial analysis and high-level process understanding, a real-world improvement project would likely require the inclusion of differentiator and potentially unique activities, such as transport bookings and triggers for the ATP and credit checks to cover more important parts of the O2C process.

A considerable number of variations can be observed in how the process is executed in practice. Additional or missing activities, rework loops and alternative execution paths appear that are not included in the reference model. These deviations follow recurring patterns rather than representing isolated exceptions and are largely linked to operational realities such as exception handling and manual interventions, particularly in the early administrative stages of the process. Some non-conforming cases should therefore be interpreted as expressions of process complexity rather than incorrect execution.

While the reference process model serves well as an initial baseline for understanding the O2C process, it does not fully capture the actual process behavior. A more complete representation requires that the reference model is complemented with contextual knowledge from process owners to cover all necessary activities. This highlights the role of backward-looking Process Mining as a diagnostic tool for bridging the gap between documented processes and real execution. From the 3P perspective mentioned in Chapter 2.3.6, this finding illustrates that even when the process structure and supporting technology are in place, the people dimension remains essential for interpreting deviations and understanding why certain execution patterns occur. Increased transparency through Process Mining makes such deviations visible, but meaningful insights depend on combining this transparency with domain expertise to distinguish between acceptable process variations and true non-conformance.

6

Bottlenecks & Inefficiencies

This chapter presents the result and discussion related to the second research question. To remind the reader, the question investigated is the following:

RQ2: *Which inefficiencies and bottlenecks in the O2C process can be identified through Process Mining?*

The findings are divided into two parts. The first part covers the challenges that Paulig's CS department described during the interviews and shadowing sessions, which were later examined in the Process Mining tool. The second part focuses on some key cycle times in the O2C process and how they are affected by bottlenecks and inefficiencies. These analyses build on the backward-looking Process Mining techniques described in the theoretical framework, with main focus on conformance checking, performance analysis and comparative Process Mining.

6.1 Validating Operational Issues

This section presents the findings related to the operational issues described by Paulig during interviews and shadowing sessions. These issues were examined using Infor's Process Mining solution to determine whether they appear systematically in the event data and how they affect the O2C process.

6.1.1 Initial Order Handling

As described earlier, Paulig has implemented an EDI based order intake to automate the creation and validation of incoming orders. This integration aims to reduce repetitive manual work and optimize the early stages of the O2C process. Since such a solution requires both investment and configuration, manual interventions in these steps create a double negative impact in the form of non-value adding activities and wasting the potential of the automated solution.

The shadowing sessions revealed that incomplete or inaccurate master data for certain customers caused recurring exceptions in the initial order handling. Incorrect articles were occasionally loaded into the ERP system and delivery dates were consistently scheduled one day too early. Although the manual corrections themselves are relatively quick to perform, they introduce several risks. CS representatives may overlook hidden errors because the system does not generate any warnings, meaning that staff rely on experience to know which customers or products are typically affected by these errors. If a correction is missed, downstream issues may occur. This may lead to shipments containing the wrong product or to poor KPI performance related to delivery accuracy, even when Paulig follows the agreed commercial terms. Customer orders may also remain blocked until the missing information is resolved, delaying the progression of the process and increasing cycle time.

To investigate these issues, a set of customers with recurring problems were examined in the Process Mining tool. Figure 6.1 illustrates one of these customers.

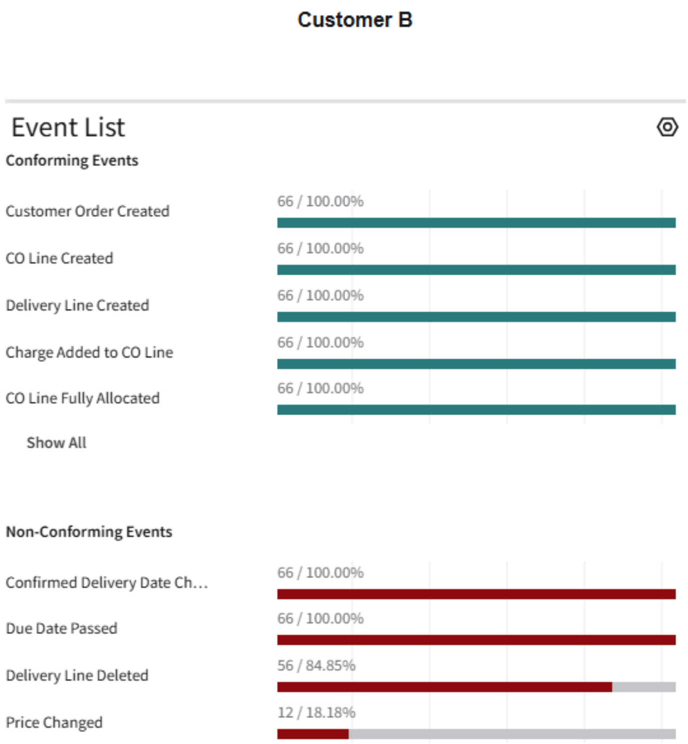


Figure 6.1: Most common events for Customer B.

Looking at the most common non-conforming events for Customer B, the analysis shows that the event *Confirmed Delivery Date Changed* appears in all 66 cases. This indicates a systematic error that repeatedly triggers unnecessary work due to incorrect master data. This customer represents only one example and similar issues likely exist in the master data of other customers. If recurring errors of this kind are widespread, they can result in a considerable amount of unnecessary manual work throughout the O2C process. The Process Mining tool makes these patterns visible by applying backward-looking techniques, clearly showing where master data quality can be improved. This transparency highlights opportunities to strengthen the underlying data in M3 to reduce the non value-adding activities for the CS team and enable the EDI integrations to be used more effectively.

Customer B also demonstrates a second recurring pattern connected to the event

Price Changed. In roughly 20% of all cases, the price on order lines was updated after the customer had placed the order. This suggests that incorrect or outdated prices were stored in the system during order intake. These discrepancies require manual corrections, both by updating the price lists in the ERP system and by informing the affected customer about the adjusted price. As with the earlier issues, this results in avoidable administrative work and reinforces the need for improved master data so that orders can be processed correctly from the beginning.

These findings were identified through a combination of conformance checking and comparative Process Mining. Conformance checking was used to filter out events that deviated from Paulig's intended standard flow, which made it possible to detect both the recurring date changes and the price adjustments as undesired activities in the case of Customer B. After this, comparative Process Mining was applied to assess the scope of the issue. First, the behavior of Customer B was compared with Customer A and Customer C, as illustrated in Figure 6.2 and Figure 6.3 below, to see whether similar deviations occurred for other customers as well. A second comparison was then carried out over time for Customer B alone, confirming that these problems had been present consistently rather than being isolated incidents.

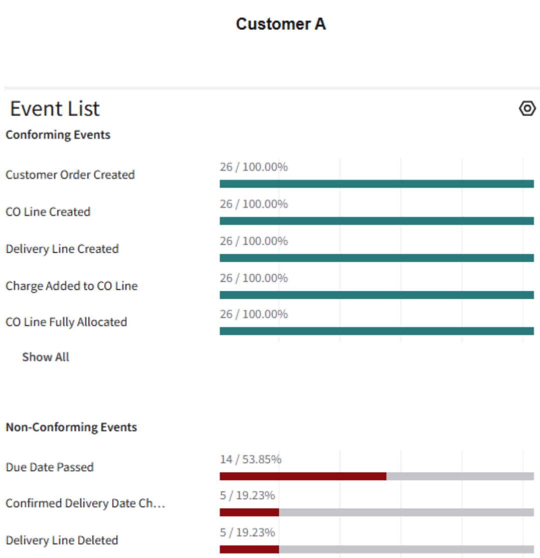


Figure 6.2: Most common events for Customer A.

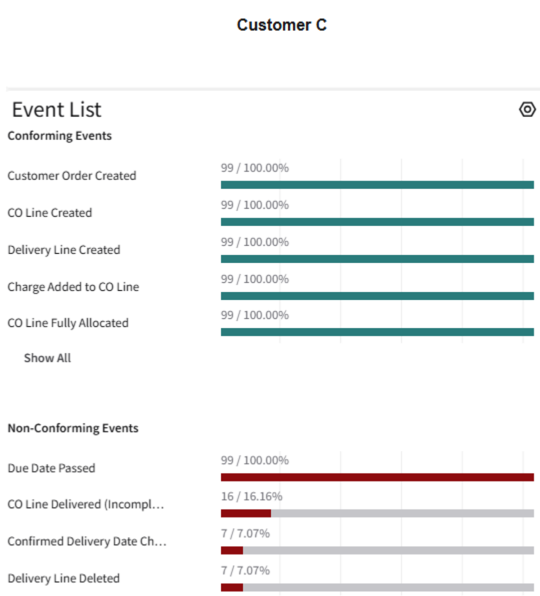


Figure 6.3: Most common events for Customer C.

As seen in the event lists above, neither date changes nor price changes appear as deviations for Customer A or Customer C. This indicates that the issue is specific to Customer B and is related to a local master data problem rather than a broader process-wide challenge.

6.1.2 Logistics

In this project, Warehouse A01, A04 and A06 were included in the scope. Warehouse A01 is operated by Paulig, while Warehouse A04 and A06 are run by 3PL partners. Approximately 70% of all order lines are picked at Warehouse A01, and the remaining volume is distributed fairly evenly between the two 3PL warehouses.

During the workshops in the Process Mining tool, the events *CO Line Delivered (Incomplete)* and *Delivery Line Deleted* stood out as surprisingly frequent across all three warehouses. Through conformance checking, these activities were flagged as non-conforming because they deviate from the intended O2C flow. Discussions with Paulig's O2C Development Manager clarified that such deviations arise when an order line must be split between two or more warehouses or completely reassigned to another location for picking. Figure 6.4, Figure 6.5 and Figure 6.6 present an overview of the most common events for the order lines linked to each warehouse.

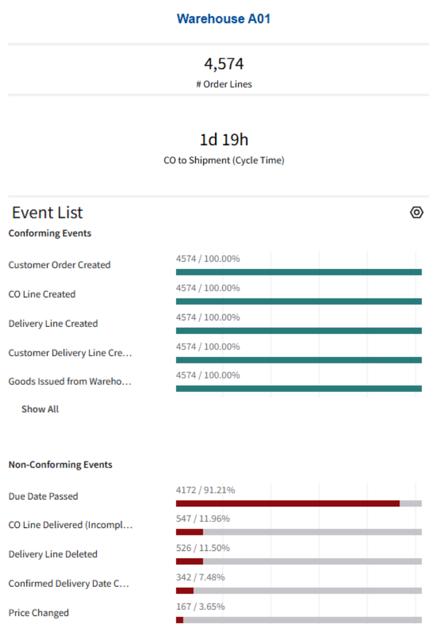


Figure 6.4: Most common events for Warehouse A01.

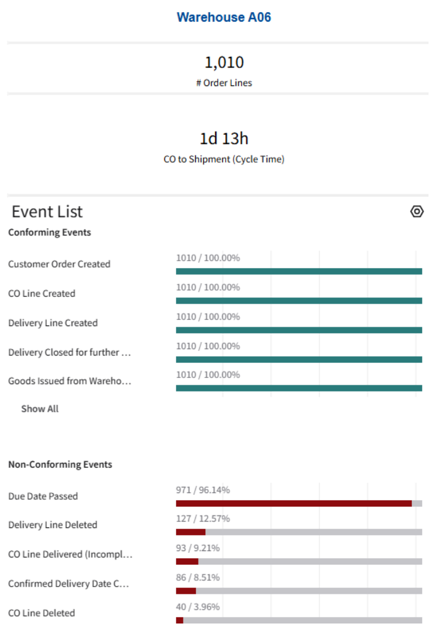


Figure 6.5: Most common events for Warehouse A06.

The event logs show that a large share of orders are affected by these deviations. Across the three warehouses, close to one fifth of all order lines are either split or moved between warehouses. Performance analysis also indicates that these reassignments add extra handling steps that can extend cycle times and reduce transport efficiency. When an order is divided between warehouses, the shipments cannot be consolidated into a single truck since the warehouses are located in different places. This reduces fill rates and can lead to several transports instead of one. The reasons behind these reassignments are typically insufficient stock at the originally planned warehouse or the need to prioritize products with short best before dates or special packaging. These re-allocations do not appear at the moment of allocation in M3 but become visible later in the flow when the delivery lines are created.

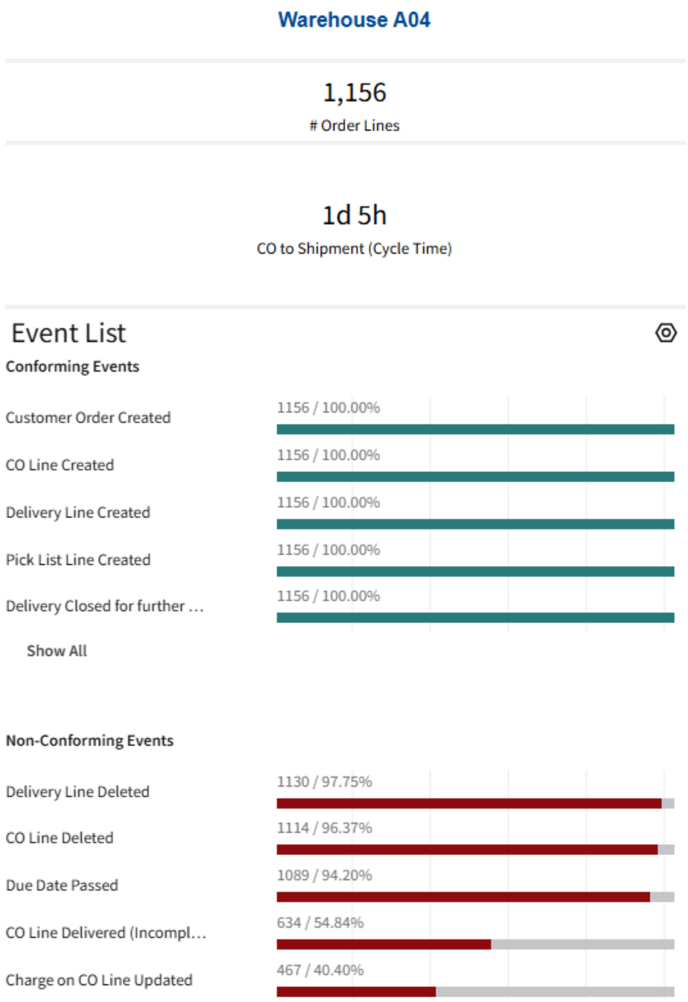


Figure 6.6: Most common events for Warehouse A04.

One result that stood out is the unusually high share of non-conforming events at Warehouse A04. Despite discussions with Paulig representatives and reviews of historical orders in M3, the exact cause could not be fully understood. A likely explanation is that the event logs reflect a pattern where many order lines are first allocated to Warehouse A01 before they are reassigned to Warehouse A04. This would mean that only a small number of lines are directly allocated to Warehouse A04, which causes reallocation events to appear very frequently. A comparison with Warehouse A06 supports this interpretation, since it does not show the same pattern and seems to receive direct allocations in a more consistent way. Comparative analysis therefore strengthens the conclusion that the high number of deviations at Warehouse A04 is linked to how allocations are managed rather than to operational behavior at the warehouse itself.

This analysis shows how Process Mining helps make the real warehouse flow easier to understand. Even when only ERP events are analyzed, the method provides a transparent view of how order handling actually works in practice. In this case, the results show that order splitting is far more common than expected, which affects both cost efficiency and environmental impact because additional transports are needed. The findings give concrete insights into how warehouse allocation and stock distribution influence logistics performance. They also illustrate that Process Mining can reveal operational challenges in the physical supply chain, not only inefficiencies in administrative work.

6.2 Cycle Times

Analyzing cycle times provides a clearer understanding of how non-conforming events affect different parts of the O2C process. The purpose of examining cycle times is to quantify to what extent deviating events influence the total lead time of the process. By breaking the end-to-end process into smaller subprocesses it becomes easier to identify where bottlenecks occur and to understand where targeted improvements are likely to have the greatest effect. This also helps organizations allocate resources more effectively in order to reduce delays and inefficiencies.

To enable a fair comparison, the whole O2C process was divided into two subprocesses. The first subprocess is referred to as the operational cycle and covers the period from *Customer Order Created* to *Customer Invoice Processed*. The second subprocess is referred to as the Invoice-to-Payment cycle and covers the period from *Customer Invoice Processed* to *Payment Received*. This separation is necessary since payment terms differ between customers and would otherwise distort comparisons of total cycle time.

6.2.1 Operational Cycle Time

The operational part of the O2C process was relevant for closer analysis because Paulig aims to deliver products to customers within two to three days after order creation. This target is supported by a make-to-stock strategy, where products are held in inventory and production is based on forecasts rather than individual orders. Customer order quantities are later aligned with these forecasts to regulate availability.

The analysis focused on how non-conforming events influence the duration of the operational flow. While the most frequent deviations were discussed in the previous section, their impact on cycle time had not previously been quantified. By comparing cases that follow the intended process flow with cases that include deviations, a clear difference emerges. As shown in Figure 6.7 and Figure 6.8, cases without deviations complete the operational subprocess in approximately two and a half days on average. Cases affected by one or more deviations take roughly three additional days.

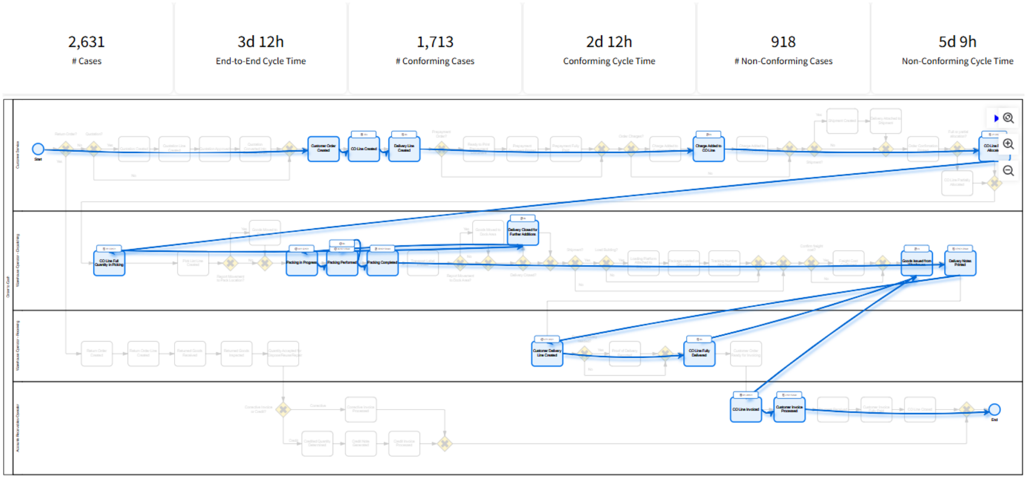


Figure 6.7: Example of conformant case.

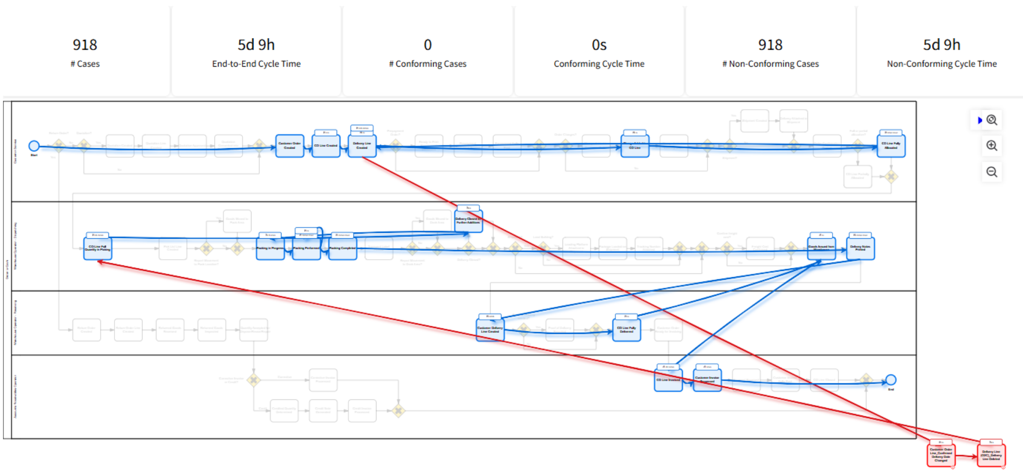


Figure 6.8: Example of non-conformant case.

Discussions with Paulig indicate that delays of this magnitude rarely cause direct negative effects for customers under normal conditions. The impact becomes more significant during promotional periods or when orders involve special packaging related to specific events or seasons. Deviations also affect earlier stages of the flow. Transport plans may need to be adjusted on short notice and additional warehouse activities can be introduced when changes occur close to the planned dispatch date. These findings highlight future potential for forward-looking and action-oriented Process Mining. Historical insights from backward-looking analysis can form the basis for early warnings when certain events occur. This would allow responsible functions to react quickly and adjust plans before delays escalate further.

If additional subprocesses had been analyzed, more insights could likely have been identified. One example is the period from order creation to allocation, which benefits from being as long as possible in order to support production planning, transport coordination and warehouse allocation. Once production is completed, the remaining steps should instead be executed as quickly as possible to convert finished goods into cash. Despite these limitations, the analysis confirms that Process Mining provides strong support for identifying bottlenecks and inefficiencies and for understanding how non-conforming events affect the operational cycle time.

6.2.2 Invoice-to-Payment Cycle Time

As shown in earlier figures, the non-conforming event *Due Date Passed* appears frequently across both customers and warehouses. To better understand this pattern, a dedicated dashboard was created in the Process Mining tool with focus on the Invoice-to-Payment subprocess. The results show that payments are received approximately two days later than their respective due date on average. This pattern was consistent regardless of which payment terms the customer had. Further analysis suggests that these delays are largely explained by practical factors such as additional bank days. The delay seems to be relatively stable across payment terms, being approximately around two days. The deviation is therefore not considered significant from a financial perspective and can be seen as a false negative.

This example illustrates a structured way of working with Process Mining. The initial indication of a potential issue emerged during the process discovery phase. A broader comparison across customers and warehouses showed that the deviation occurred consistently, which motivated a more focused analysis. By isolating the Invoice-to-Payment subprocess and using payment terms as a segmentation attribute, the issue could be assessed independently of the operational flow. So even though no major issues were identified in the Invoice-to-Payment process in this case, the analysis illustrates how a potential deviation identified at a high level can be broken down and examined in more detail. This approach makes it possible to validate whether an observed pattern represents a real problem or simply reflects structural characteristics of the process, and demonstrates how Process Mining can be used to move from overview to detailed investigation when needed.

6.3 Main Findings

Recurring master data issues emerge as a central source of bottlenecks in the O2C process, particularly in the early administrative stages. As described in Chapter 2.1, accurate master data is essential for standardized and efficient order handling, as errors typically lead to rework and increased administrative effort. The empirical findings support this view, as manual adjustments to order lines such as changes to delivery dates and prices introduce non-value-adding work for CS and increase process variability. From a Process Mining perspective, these activities appear as deviations from the reference process and are identified through conformance checking.

Differences in execution performance are further explained in Chapter 2.3.5.1 on performance analysis, which uses event log timestamps to measure cycle times. Cases containing non-conforming events were found to have longer average cycle times than conforming flows, confirming theoretical expectations that deviations negatively affect process performance. Comparative Process Mining also proved useful, as described in the literature, by enabling comparisons across customers and warehouses. This made it possible to identify that several deviations were systematically linked to specific customers and warehouses rather than the general process design.

The findings reflect limitations highlighted in Process Mining literature, namely that event logs show what has happened but not why it has happened. Some non-conforming events were therefore identified as false negatives, where deviations were caused by how the ERP system technically records order splitting rather than by actual inefficiencies. The combined use of conformance checking, performance analysis and comparative Process Mining thus provided a diagnostic understanding of bottlenecks in both administrative activities and the physical flow of goods, while underscoring the need for contextual interpretation.

7

Opportunities & Limitations

In the following chapter, the third research question will be answered and discussed. As a reminder, the question is the following:

> ***RQ3:*** *What are the opportunities and limitations associated with the application of Process Mining tools in organizational process analysis?*

When moving from the accuracy of discovered processes to the broader application of Process Mining, several opportunities and limitations can be identified. The subchapters aim to answer what opportunities and limitations there are with Process Mining tools in organizational process analysis. Many of the insights are from working in Infor's tool provided in the project and may be specific to this case, but many findings can be brought to a more general level as well.

7.1 Opportunities

One clear opportunity is the overall usability of the Process Mining tool. The tool was relatively easy to use and did not require extensive prior knowledge to get started. The basic functionality could be learned quickly, making it accessible even for users without a technical background in process analysis. This simplicity was particularly evident during presentations and workshops where the process models were shared with different stakeholders. The visualizations were intuitive and the audience was generally able to follow and understand what was being shown without requiring detailed explanations. This makes Process Mining a useful communication tool as it allows complex process behavior to be presented in a way that is understandable for both operational and managerial roles.

7.1.1 RPA & GenAI

Since Process Mining is used to diagnose processes and find inefficiencies, the next logical step is to find solutions for these. RPA has the potential to improve and solve inefficiences in processes. In this study, many tasks were found which could be automated. For example as earlier described, a CS representative had to change the delivery date on 66 out of 66 order lines for a certain customer due to master data issues. RPA could potentially be implemented to identify cases like these and suggest possible solutions. This directly relates to action-oriented Process Mining, where analytical insights are turned into actions to support processes. In the case of delivery date changes, RPA could detect and recommend a change in the order or in the master data. Another issue that was encountered by CS is that orders have incorrect information when they enter the system through EDI, and need manual handling every time. RPM has potential to alert supervisors and users when it detects such deviations and suggest a fix.

The value of removing a manual change can be discussed. If it is only necessary in few cases for a single employee, the value is very limited, but if it is a systematic issue across several markets or divisions the value is more significant. The improvement not only saves time, but also prevents orders from being blocked and not being able to continue in the process flow, or even worse, continuing through the process with inaccurate information.

Paulig has already implemented other automated solutions in their O2C process, such as orders being received through EDI and automated invoicing. Therefore, RPA may be more suitable for companies that do not already have similar automated solutions implemented. It is worth mentioning that these automated solutions rely on external applications that require integrations, while the RPA module, provided in the same package as Process Mining for this study, is already integrated into the Infor's environment.

As GenAI progresses, forward-looking Process Mining becomes more relevant. It can be used to help analyze results, but a large utility case would be to, for example, predict the next activity in a process flow. In theory, by seeing that a certain non-conformant event happens for example, it could be possible to predict what happens after. Perhaps it leads to another non-conformant event or manual intervention, and the additional time this would take could be predicted. Even though AI and Digital Twins could play an important role in future Process Mining, an important take from this study is that the tool has value in itself. It gives the opportunity to map processes, create a cohesive and united view in the company and improve workflows by finding and reducing inefficiencies.

7.1.2 Findings & Fast Time-to-Value

The Process Mining tool showed that it could provide value quickly, as it was possible to get an overview of the actual process flow within a few days without much initial setup. The discovered flows give a general picture of how the process behaves in reality and can help identify some early bottlenecks and deviations. These findings are useful as a starting point for improvement discussions and for linking the process to existing KPI:s and dashboards. Since historic data was available, it also became possible to compare how the process has changed over time and to monitor whether improvement efforts have an effect. For organizations that already have a manually documented process model, Process Mining can be used to check whether the pre-defined or intended model reflects what actually happens in the system. In this case, differences and unexpected paths can be detected, which may reveal parts of the process that were previously unnoticed. At the same time, the approach is beneficial for companies that do not have a defined process map, as the tool can quickly generate a visual representation of the actual workflow with relatively little effort required from the users. It is also possible to look more closely at special cases such as return flows, which helps highlight behaviours that do not necessarily appear in a high-level model. This shows that Process Mining can deliver meaningful insights early on and serve as a foundation for both exploratory analysis and more structured improvement work.

The Process Mining tool can be used to potentially detect systematic flaws in the processes. This could for example include customers repeatedly not paying on time, activities where the process tends to stop, customers changing quantities or dates often or using many different warehouses, all of which can suggest an unoptimized flow and underlying issues. Apart from detecting flaws that were not previously apparent, Process Mining can be used for verifying issues you already know about, but want to analyze further in a data-driven way. To bring up the case with CS again when they had to change the delivery date every time for a certain customer, this information from the interview could be verified in the Process Mining tool by filtering on that specific customer and thereby seeing the non-conformant event associated with this. This was just one specific case and could be used to verify and, most importantly, quantify other issues as well to see the extent of the problem.

Another opportunity with Process Mining is the ability to follow up continuous improvement in an administrative environment. If a period of data is analyzed and gives certain insights, initiatives can be made and then followed up later to see what has changed and improved, especially when it comes to following up on, for example KPI:s. Perhaps it can be discovered that someone is doing a task really well, which could be valuable to improve workflows and apply this way of working throughout the function. In contrast to this, Process Mining creates the opportunity and a foundation to implement better routines and user manuals to processes with many deviations and inefficiencies. By filtering on certain customers for example, it is possible to see which customers are the most problematic and if certain deviations are recurring for them specifically. This study focuses on the O2C cycle, but it could also be extended to look at the procurement side and oversee and follow up KPI:s with suppliers as well in a similar way.

7.1.3 Additional Functionalities

Infor's tool includes a pre-defined O2C process map, which fit the actual process quite well. But the tool also enables creating own process maps, or adjusting the pre-defined ones, which is quite easy to do. This is a benefit if an organization has more unique processes that are not mapped in the standard flow. Most other Process Mining tools require starting from scratch and integrating these process maps with the data, but the tool that was used in the study made it very simple to create them and connect each activity in the process to the event logs through the so called "auto-map" function. Furthermore, if an organization, like Paulig, wants to move towards more harmonization and alignment in standard processes, this could be a good tool to create and set these standards and follow up over time. In Paulig's case it is clear that the same O2C process is different between brands. ATP, allocation and the return flow look different in the UK compared to Finland, which is largely due to the differences in the products they deliver. Apart from creating and configuring own process maps, Infor's Process Mining tool also allows creating own dashboards and visualizations to show targeted insights. This is great if a company has its own KPI:s that they wish to track or make certain comparisons between for example warehouses, product groups or order types.

All events and attributes were already pre-configured for the Process Mining, but with that said it is not particularly difficult to add own attributes and events to analyze. In this study, *Reason Code* was added as an attribute as an attempt to get a better understanding of why a customer order line is deleted for example. This attribute was commonly used in the day-to-day processes in M3, hence why it was interesting to examine in the Process Mining tool. It was not fully successful in this case due to reason codes missing for many order lines, and all types of reason codes were not included either. Nonetheless, this proves that the ability to add additional relevant attributes is possible, which is beneficial when further details are needed for the analysis. To add to this, data can also be retrieved from other systems into the ERP system, but requires more work, for example integrating the ERP system with the transport booking portals.

7.2 Limitations

While the tool is easy to use at the surface level, performing a meaningful analysis requires an understanding of both the underlying data and the business process itself. Without knowing where to start, which filters to apply or which questions to ask, there is a risk of either overlooking important insights or drawing misleading conclusions. This suggests that although Process Mining tools lower the entry barrier for process analysis, they still require domain knowledge and analytical skills in order to be used effectively in an organizational context.

7.2.1 General Limitations

A core limitation of Process Mining is what data is available to analyze. If the data for the process that is examined is not available in the data lake, an analysis is not possible. While most activities in Paulig's O2C process occur in M3, some of the process falls outside. For example, after an order is allocated, the transport is booked through an external application. This means that this particular step cannot be analyzed in the Process Mining tool and transparency is reduced, since these data points and events are not logged. This also applies to for example, emails and other types of communication internally in the company. Covering every part of the process, especially one that involves many different systems, is complex and requires more effort to integrate data and give a holistic view. Furthermore, only performed changes or actions trigger a change in the database, meaning a user has to for example change an order quantity for this to be logged. Only opening and viewing an order is not logged, which also reduces transparency in what is happening in the actual workflow. To present a more relevant case, just viewing the credit limit for a customer will not create an event log, unless some sort of button is added where the user has to confirm the credit check has been performed. With this said it is worth noting that too much transparency and details might be overwhelming, and finding a balance and knowing what is important to include is crucial.

Another limitation with event logs is that in order for them to be accurate, they need to occur at the same time as the activity happens in reality. This becomes apparent in logistical steps. For example, if an order is packed and placed ready for dispatch on a location in the warehouse but not marked in the system that it is ready until the truck arrives, this will signal a discrepancy and show that the goods were ready just before dispatch, even though they may have been ready for several hours. It is therefore important that real-life events match the events logged in the system and that timestamps reflect reality.

In the Process Mining tool used in this study, a limitation on the total number of cases was present at 350k, meaning a maximum number of 350k order lines could be loaded. For the data used in this study, it was more than enough but for some companies with larger volumes this may be a limitation and not suitable.

7.2.2 Finding Reasons for Non-conformances

A major limitation with Process Mining analysis is that it is difficult to find the reason for non-conformances and inefficiencies found. In one of the first few cases that were examined, which was related to customer order line charges being added on roughly 90% of all order lines, it initially looked like an inefficiency. Since the *Charge Added to CO Line* event had an average time on it, it was assumed that this might be done manually, which raised the question of inefficiency since it did not seem ideal to need to do this on every order line. Furthermore, a few thousand order lines had missing charges on them, raising questions if this was intentional or something systematically wrong. It was later found out during one of the workshops with Paulig and through sample tests, meaning the logs of some order lines were checked in the M3, that the missing charges are due to internal transactions in the company, or due to certain products such as spare parts. In these cases, charges are not added. When it comes to all the other order lines that have charges, these are done automatically by the system, which was more reasonable but at the same time confusing since there was a delay between the order line being created and the charge being added to it. Therefore, it can be difficult to draw premature conclusions if you do not have background knowledge about the processes.

Another event in which the reasons were difficult to pinpoint was with respect to the quantity and date changes in the customer order lines. In these cases, information is lacking regarding why the changes were made, and if they were made because of internal reasons or because the customer requested it. Deeper analysis is therefore required to find root causes for these cases, necessitating checking individual order logs or emails for example. In many cases, it was also often needed to complement with information from users and process owners, to understand the displayed data in the tool and enable better interpretation. In practice, however, users will likely already have the process knowledge. If you do not know how the software works, such as the ERP system or Process Mining tool, and lack background process knowledge, this will be a limitation and information needs to be complemented.

Apart from not knowing the reasons for the activities happening, another limitation is that some cases that seem to be inefficient or non-conformant are in reality functioning as they should. Usually most orders are received two days before shipment, but for some special items, the orders are placed weeks before and not picked and packed until right before shipment. This may present a false negative and appear as something is wrong when compared to the more standard products, although in reality it needs to be considered that certain articles take longer time and that it is an upside that the orders are received earlier, in order to make room for planning. Another discovered example was when customer order lines are deleted. This appears as a non-conformance, but when Paulig was asked about this, it emerged that a majority of these order lines are placed for a certain warehouse originally but moved to another one later, hence deleting the line from the previous warehouse.

Overall, it can be difficult to immediately try to look for bottlenecks when opening the first overview in the Process Mining tool. It is a good idea to have a plan on how to proceed and what to look at, and potentially have a hypothesis. Even with a plan, it may be difficult to interpret and act on results from Process Mining. The problem needs to be verified and understood, and can be acted on only after this is done.

7.3 Main Findings

The application of Process Mining shows several benefits related to process transparency and data-driven analysis. By using existing ERP data, it is possible to quickly create an objective overview of complex end-to-end processes, which supports discussions based on facts rather than perceptions as explained in Chapter 2.3.2. Visualizing how processes actually operate helps stakeholders from different parts of the organization to analyze the same process and develop a shared understanding of how work is carried out in practice.

Recurring inefficiencies, deviations and rework loops can be identified with relatively short time-to-value, which highlights the usefulness of backward-looking Process Mining as a foundation for continuous improvement initiatives. The findings also indicate that Process Mining can facilitate process standardization and harmonization efforts, particularly in the context of ERP transformations. In addition, insights from administrative process analysis can point to tasks suitable for automation, including rule-based activities and RPA, and may also provide indications of inefficiencies in physical logistics flows when supported by relevant event data.

Several limitations affect how Process Mining results should be interpreted and applied. The analysis is highly dependent on the availability, completeness and quality of event log data, meaning that activities performed outside the ERP system remain invisible. While Process Mining provides detailed insight into what happens in a process, it offers limited explanations of why deviations occur without complementary qualitative input. Interpretation therefore requires domain knowledge to avoid oversimplification or incorrect conclusions.

As the analysis is primarily backward-looking, additional techniques are required to support predictive decision-making. Simulation and digital twin approaches can potentially be seen as a complement to Process Mining, enabling organizations to test future scenarios in a controlled environment. By using process models grounded in real execution data, a digital twin of a process could support "what-if" analyses when onboarding new brands, adjusting process configurations or evaluating the impact of system changes before deployment. Such test environments, closely reflecting actual operations, may reduce implementation risks and support more informed decision-making during ERP migrations and process redesign efforts. This perspective remains speculative and has not been validated through practical implementation, but it highlights a possible future direction for Process Mining beyond its current backward-looking use.

8
Conclusion

This chapter concludes the thesis by summarizing the main findings from the application of Process Mining in the O2C process. It presents practical recommendations and general guidelines derived from the project experience as well as reflections on ethical and societal aspects related to the use of Process Mining in an industrial context. The chapter also outlines potential future research areas based on the limitations and insights identified in the study.

8.1 Summary of Findings

This study explored how Process Mining can be applied to understand and improve the O2C process in practice using Paulig as a case company and Infor's Process Mining solution as the analytical tool. The results show that the pre-defined process model provides a solid and accurate representation of the actual O2C flow on a general level. The pre-defined reference model enabled a rapid initial understanding of Paulig's administrative processes and allowed insights to be generated with limited setup effort. This confirms that a pre-defined approach supports fast onboarding and early value creation, particularly when core processes are largely executed within the ERP system. The pre-defined process maps and generic dashboards worked well as an entry point for analysis, providing a clear overview of the O2C process and common deviations. However, the main analytical value emerged when these standard views were used as a basis for creating more focused dashboards tailored to specific questions. This confirms that Process Mining tools are not standalone solutions but rather analytical environments that require active exploration, iteration and contextual understanding to deliver meaningful insights.

The backward-looking Process Mining techniques proved effective in identifying both inefficiencies in administrative work and weaknesses in the physical flow of goods. Event logs provided an objective view of what happened in the process, making it possible to verify issues described during interviews and shadowing sessions. At the same time, the analysis confirmed that event data alone cannot explain why activities occur. This underlines the importance of combining Process Mining with domain knowledge in order to correctly interpret findings and avoid misleading conclusions. In this case, the combination of quantitative event data and qualitative insights was essential for understanding root causes and assessing whether deviations represented real problems or normal process behavior.

The study also shows that Process Mining forms a strong foundation for future automation initiatives. The backward-looking analysis clearly highlights repetitive manual activities and systematic errors, such as recurring master data issues, which represent low-hanging fruit for RPA or rule-based automation. Although Paulig is already far along in its automation journey, the results suggest that further value may lie not in automating more tasks, but in optimizing existing automation and improving data quality. In organizations earlier in their digital maturity, the same insights could have an even greater impact. In addition, the study confirms that Process Mining can play an important role in harmonization efforts. For organizations like Paulig, which operate across multiple countries and brands, Process Mining provides a data-driven way to understand how processes actually differ and where standardization efforts should be focused. By making processes visible and comparable, the tool supports informed decision-making and continuous improvement.

8.2 Recommendations & Guidelines

This section presents practical recommendations derived from the project experience and discusses how Process Mining can be applied in a structured and effective way. It also reflects on ethical and social considerations that should be addressed when using Process Mining as a diagnostic and improvement tool.

8.2.1 Learnings

This project provided the opportunity to act as early users of Infor's Process Mining solution. However, the recommendations presented in this section are intentionally kept general and are applicable regardless of the specific Process Mining vendor or tool used.

Process Mining is still relatively new in industrial practice, and there are currently no widely accepted standards or best practices for how these tools should be applied in practice. One reason for this is that organizations differ significantly in terms of data availability and process complexity. As a result, Process Mining initiatives must be adapted to the specific context in which they are applied rather than following a single pre-defined methodology. A key general recommendation is to begin the analysis at a broad level. During the initial phase of this project, there was a tendency to focus too early on individual cases or specific deviations. A more effective approach is to first analyze the process at an aggregated level to identify recurring patterns and systematic challenges. Even at this high level, Process Mining can reveal potential issues that quickly provide valuable input for further discussions and deeper analysis.

It is also important to understand that Process Mining tools should not be used in isolation. Quantitative insights derived from event data must be complemented with qualitative methods such as interviews and workshops. These activities are essential both for configuring the analysis environment correctly and for interpreting the results in a meaningful way. Without process knowledge from domain experts, there is a risk of misinterpreting deviations or drawing incorrect conclusions from the data. Broad involvement across the organization is also recommended. Engaging many stakeholders reduces uncertainty and concern about monitoring or control while simultaneously enriching the analysis with multiple perspectives. Transparency around the purpose of the initiative and open communication about findings help build trust and acceptance.

Early in the project, it is critical to define a clear purpose and scope and to apply appropriate filters based on these objectives. This helps isolate the relevant data for analysis and prevents unnecessary complexity. For large and cross-functional processes such as O2C, involvement from multiple subprocess owners is required. At least one representative from each key function should participate in setting filters, validating process maps, and discussing analytical results to ensure that findings are grounded in operational reality.

Another important lesson is the need for caution when interpreting early results. Initial findings can easily appear more significant than they actually are. This project showed the importance of remaining humble in the analysis and continuously validating insights through dialogue with process experts. Process Mining tools are generally user-friendly and offer strong support for creating custom dashboards. Users are encouraged to develop their own dashboards tailored to specific KPI:s or subprocesses of interest. This makes it easier to monitor performance over time and supports structured continuous improvement initiatives rather than one time analyses. It is also advisable to start with a smaller pilot focused on a specific process or subprocess. In retrospect, the scope of the project with Paulig was relatively broad given the limited initial familiarity with both the organization's processes and the Process Mining tool itself. A narrower pilot can reduce complexity and create a stronger foundation for scaling Process Mining initiatives over time.

8.2.2 Ethical & Societal Aspects

This thesis has examined Process Mining as a diagnostic tool for identifying improvement opportunities such as deviations and master data issues in administrative processes. The analysis has also shown how Process Mining tools can be used in a similar way to business intelligence solutions, for example to monitor KPI:s and support continuous improvement initiatives. At the same time, the application of Process Mining raises a number of ethical and social questions that need to be addressed before organizations decide to invest in and scale such tools.

One important concern relates to employee integrity and perceptions of monitoring. Since Process Mining relies on event logs generated from daily work in information systems, employees may experience that their actions are being closely observed or evaluated. This can create a feeling of surveillance, particularly if insights are presented at a very detailed level or used to assess individual performance. There is also a risk that Process Mining is perceived as questioning professional judgment, as employees may feel that their expertise and experience is being replaced by automated analysis.

Such perceptions can influence behavior in unintended ways. Employees may become more risk-averse and hesitant to experiment with new ways of solving tasks if deviations from standard processes are automatically highlighted as non-conforming. This can limit learning, innovation and personal development, especially in roles where flexibility and problem-solving are important. To avoid these effects, it is crucial that Process Mining is positioned as a support tool for improvement rather than a control mechanism for individual behavior. Transparency about the purpose of the analysis and how insights will be used is therefore essential.

Another ethical dimension concerns automation and its impact on work roles. Process Mining is often used to identify subprocesses or activities that are sufficiently standardized to be automated or further optimized. As a consequence, some manual tasks may be reduced or eliminated. This can raise concerns about job security if the freed-up time is not redirected toward other value-creating activities. At the same time, there is a societal responsibility to ensure that efficiency gains translate into upskilling, role development and more meaningful work rather than workforce reduction alone. With efficiency as a guiding principle, Process Mining can enable the automation of routine and repetitive tasks. This reduces the share of monotonous work and allows employees to focus on more stimulating, complex and value-creating activities. Over time, this shift has the potential to increase job satisfaction, strengthen expertise and support innovation, both within organizations and in society at large. When implemented with clear ethical guidelines, employee involvement and a focus on learning rather than control, Process Mining can therefore act as an enabler of sustainable and socially responsible process improvement.

In the case of Paulig, the organization is already far ahead in its automation journey, with large parts of the initial order handling and invoicing processes automated through EDI and system integrations. The findings of this study indicate that there is still significant value in maintaining human involvement in administrative processes. Human checkpoints support communication with customers, help resolve exceptions and contribute contextual understanding that cannot easily be captured in system data alone. This balance between automation and human oversight is important to preserve both process quality and employee engagement.

8.3 Future Research

The work presented in this thesis shows that backward-looking Process Mining can be applied across a wide range of processes and organizational settings. Similar analyses could be carried out in other business processes beyond O2C, as well as in environments with different levels of complexity and stability. Studying processes with higher variability than MTS, such as make-to-order or less standardized workflows, could provide a deeper understanding of how deviations should be interpreted when processes are less predictable. There is also potential to apply the same approach across different departments, functions or locations within the same organization in order to better understand how local practices and system configurations influence actual process execution.

The study was influenced by limitations in available data and by the chosen analytical focus. The analysis primarily covered administrative activities and selected logistics-related events, while detailed warehouse operations and other supporting processes were only partly visible in the event logs. Including more detailed operational data would allow a more complete understanding of how administrative decisions and physical execution interact.

Although the analysis in this thesis is backward-looking, the results also point toward possible next steps beyond diagnosis. Repetitive manual work and recurring data-related issues identified in the analysis suggest areas where automation through RPA could be further explored. Rather than viewing Process Mining only as an analytical tool, it may be valuable to examine how insights from backward-looking analysis can be connected to more forward-looking or action-oriented approaches. Exploring this link could help clarify how diagnostic findings can be translated into practical changes in daily operations. Studying the economic effects of such changes would also contribute to a clearer understanding of the value created by Process Mining initiatives.

List of References

[1] L. Reinkemeyer, *Process Mining in Action: Principles, Use Cases and Outlook*. Springer, 2020. https://doi.org/10.1007/978-3-030-40172-6.

[2] G. Park and W. M. P. van der Aalst, "Realizing a digital twin of an organization using action-oriented process mining," in *International Conference on Process Mining*, pp. 104–111, 2021. https://doi.org/10.1109/ICPM53251.2021.9576846.

[3] E. Rojas, J. Munoz-Gama, M. Sepúlveda, and D. Capurro, "Process mining in healthcare: A literature review," *Journal of Biomedical Informatics*, vol. 61, pp. 224–236, 2016. https://doi.org/10.1016/j.jbi.2016.04.007.

[4] V. Blijleven, K. Koelemeijer, and M. Jaspers, "Identifying and eliminating inefficiencies in information system usage: A lean perspective," *International Journal of Medical Informatics*, vol. 107, pp. 40–47, 2017. https://doi.org/10.1016/j.ijmedinf.2017.08.005.

[5] A. H. Roest, "A practitioner's guide for process mining on ERP systems: The case of SAP order to cash," Master's thesis, Eindhoven University of Technology, 2012.

[6] W. M. P. van der Aalst, "Process mining: Overview and opportunities," *ACM Transactions on Management Information Systems*, vol. 3, no. 2, pp. 1–17, 2012. https://doi.org/10.1145/2229156.2229157.

[7] W. M. P. van der Aalst and A. J. M. M. Weijters, "Process mining: A research agenda," *Computers in Industry*, vol. 53, no. 3, pp. 231–244, 2004. https://doi.org/10.1016/j.compind.2003.10.001.

[8] C. de Goeij, R. van Weelden, M. Breijer, J. Jansen, and R. Holman, "Inefficiencies in order to cash processes in the logistics industry: Empirical evidence from the Netherlands," *Vervoerslogsitiekewerkdagen*, pp. 217–230, 2022.

[9] B. Kocaoğlu and A. Z. Acar, "Process development in customer order information systems to gain competitive advantage: A SME case study," *International Journal of Logistics Systems and Management*, vol. 23, no. 2, pp. 209–230, 2016.

[10] D. J. Bowersox, D. J. Closs, M. B. Cooper, and J. C. Bowersox, *Supply Chain Logistics Management*. McGraw-Hill Education, 5 ed., 2020.

[11] F. R. Jacobs and F. C. T. Weston Jr., "Enterprise resource planning (ERP): A brief history," *Journal of Operations Management*, vol. 25, no. 2, pp. 357–363, 2007. https://doi.org/10.1016/j.jom.2006.11.005.

[12] P. Jonsson and S. A. Mattsson, *Manufacturing Planning and Control*. McGraw-Hill Education, 2009.

[13] V. Botta-Genoulaz, P. A. Millet, and B. Grabot, "A survey on the recent research literature on ERP systems," *Computers in Industry*, vol. 56, no. 6, pp. 510–522, 2005. https://doi.org/10.1016/j.compind.2005.02.004.

[14] F. Buttle and S. Maklan, *Customer Relationship Management: Concepts and Technologies*. Routledge, 4 ed., 2019. https://doi.org/10.4324/9781351016551.

[15] K. N. Subramanya, T. M. Rangaswamy, and A. Ramaa, "Impact of warehouse management system in a supply chain," *International Journal of Computer Applications*, vol. 54, no. 1, pp. 14–20, 2012.

[16] W. M. P. van der Aalst and J. Carmona, eds., *Process Mining Handbook*. Springer, 2022. https://doi.org/10.1007/978-3-031-08848-3.

[17] Celonis, "Process mining: The missing link between data science and process science." https://www.celonis.com/blog/process-mining-the-missing-link-between-data-science-and-process-science. Accessed: 2025-12-05.

[18] W. M. P. van der Aalst, *Process Mining: Data Science in Action*. Springer, 2 ed., 2016. https://doi.org/10.1007/978-3-662-49851-4.

[19] A. Tiwari, C. J. Turner, and B. Majeed, "A review of business process mining: State-of-the-art and future trends," *Business Process Management Journal*, vol. 14, no. 1, pp. 5–22, 2008. https://doi.org/10.1108/14637150810849373.

[20] N. Martin, D. A. Fischer, G. D. Kerpedzhiev, K. Goel, S. J. J. Leemans, M. Röglinger, W. M. P. van der Aalst, M. Dumas, M. La Rosa, and M. T. Wynn, "Opportunities and challenges for process mining in organizations: Results of a Delphi study," *Business & Information Systems Engineering*, vol. 63, no. 5, pp. 511–527, 2021. https://doi.org/10.1007/s12599-021-00720-0.

[21] R. Syed, S. J. J. Leemans, R. Eden, and J. A. C. M. Buijs, "Process mining adoption: A technology continuity versus discontinuity perspective," in *International Conference on Business Process Management*, pp. 229–245, 2020. https://doi.org/10.1007/978-3-030-58638-6_14.

[22] C. J. Turner, A. Tiwari, R. Olaiya, and Y. Xu, "Process mining: From theory to practice," *Business Process Management Journal*, vol. 18, no. 3, pp. 493–512, 2012. https://doi.org/10.1108/14637151211232669.

[23] R. Andrews, C. G. J. van Dun, M. T. Wynn, W. Kratsch, M. K. E. Röglinger, and A. H. M. ter Hofstede, "Quality-informed semi-automated event log generation for process mining," *Decision Support Systems*, vol. 132, pp. 1–13, 2020. https://doi.org/10.1016/j.dss.2020.113265.

[24] M. von Rosing, S. White, F. Cummins, and H. de Man, "Business process model and notation — BPMN," in *The Complete Business Process Handbook: Body of Knowledge from Process Modeling to BPM*, vol. 1, pp. 433–457, Elsevier, 2015. https://doi.org/10.1016/B978-0-12-799959-3.00021-5.

[25] W. M. P. van der Aalst, "Decomposing Petri nets for process mining: A generic approach," *Distributed and Parallel Databases*, vol. 31, pp. 471–507, 2013. https://doi.org/10.1007/s10619-013-7127-5.

[26] W. M. P. van der Aalst, "Business process management as the "Killer App" for Petri nets," *Software and Systems Modeling*, vol. 14, no. 2, pp. 685–691, 2015. `https://doi.org/10.1007/s10270-014-0424-2`.

[27] J. Munoz-Gama, *Conformance checking and diagnosis in process mining.* PhD dissertation, Universitat Politècnica de Catalunya, 2014.

[28] A. Rozinat and W. M. P. van der Aalst, "Conformance checking of processes based on monitoring real behavior," *Information Systems*, vol. 33, no. 1, pp. 64–95, 2008. `https://doi.org/10.1016/j.is.2007.07.001`.

[29] V. Pasquadibisceglie, A. Appice, G. Castellano, and D. Malerba, "Predictive process mining meets computer vision," in *International Conference on Business Process Management*, pp. 176–192, 2020. `https://doi.org/10.1007/978-3-030-58638-6_11`.

[30] W. M. P. van der Aalst, "Process mining and simulation: A match made in heaven!," in *SummerSim*, pp. 1–12, 2018. `https://doi.org/10.22360/summersim.2018.scsc.005`.

[31] T. Brockhoff, M. Heithoff, I. Koren, J. Michael, J. Pfeiffer, B. Rumpe, M. S. Uysal, W. M. P. van der Aalst, and A. Wortmann, "Process prediction with digital twins," in *International Conference on Model Driven Engineering Languages and Systems Companion*, pp. 182–187, 2021. `https://doi.org/10.1109/MODELS-C53483.2021.00032`.

[32] K. J. Sileyew, "Research design and methodology," in *Cyberspace*, 2019. `https://doi.org/10.5772/intechopen.85731`.

[33] Infor, "Infor M3: Your industry, our advanced technology." `https://www.infor.com/solutions/erp/m3`. Accessed: 2025-12-04.

[34] Infor, "Eliminate barriers to process innovation with Infor Velocity Suite." `https://www.infor.com/platform/velocity-suite`. Accessed: 2025-12-04.

[35] Infor, "What is process mining?." `https://www.infor.com/platform/data-insights/process-mining/what-is-process-mining`. Accessed: 2025-12-04.

[36] Gartner, "Magic quadrant for process mining platforms." Gartner Research Report, 2025. By T. Srivastava, M. Kerremans, and D. Sugden.

A

Interview Guide

1. Introduction

Brief presentation of the interviewers and the project

- Master's thesis students from Chalmers University of Technology
- Collaboration with Meridion, Infor, and Paulig
- Focus on Process Mining applied to the Order-to-Cash (O2C) process

Purpose of the interview

- To understand how the O2C process is executed in practice
- To identify bottlenecks, exceptions, and system-related issues

Consent

- Permission to record the interview

2. O2C Process

- Can you briefly describe your role and how you are involved in the O2C process?
- What are the main steps in the O2C process at Paulig?

3. Order Entry

- Where do customer orders originate from (email, EDI, quotation, etc.)?
- How many days before delivery does an order enter the system?
- Approximately how many customers send orders via email?
- Is Customer Service responsible for manually entering these orders into M3?

4. Order Validation

- How does the sales order confirmation process look like?
- What information or requirements are needed to validate a sales order?
- How is the order validated?
  - Manually checking several places?
  - Automatically?
- Which checks are performed during order validation?
- Are credit checks performed? What happens if a credit check fails?
- Are ATP checks performed? If not, why is this not an issue?
- What is the role of Sales in the order process?
  - What checks are performed by Sales?

- What are the most common reasons for customer orders going into error?

5. Exception Management

- How is exception management handled in the O2C process?
- Approximately how many claims are handled?
- Are most claims managed within Infor/M3?
- What are the most common types of exceptions?
- Is CRM used in the claims or exception handling process?
- Is it often a problem with forecasting and last-minute changes during campaigns?
- What are the most common reasons for manually changing a customer order?
  - Quantity changes?
  - Delivery date changes?
  - Production delays?
  - Stock availability?
- Are price changes made often?
  - In which situations?
  - Why do these changes occur?
- Is the standard order process the same for all customers, or do some customers have unique flows?
- Are there any other deviations in the order creation, confirmation, or post-confirmation stages?

6. Logistics, Availability, and Campaigns

- Who pays for transportation to the customer and how is this handled in M3?
- Are transportation costs added automatically or manually?
- Are full truckloads commonly used?
- How does ATP work today?
- Are there differences between product categories?
- Are Best Before Date (BBD) requirements important for certain products?
- Are products generally kept in stock, or is production often required?
- Approximately how many orders are handled per day?
- What types of campaigns are run?
- How do campaigns affect allocation, stock availability, and order handling?
- Does the customer have to commit earlier to quantities?

7. Reflections

- Are there any steps in the O2C process that you feel are unnecessary or particularly time-consuming?
- Are there recurring issues that you believe could be reduced or eliminated?
- Is there anything else you think is important for us to understand regarding the O2C process or system usage?
- Are dashboards or workspaces used to monitor orders or performance?

- How well do you feel the current systems support your daily work?

Note: The interviews were conducted as semi-structured interviews. While the same overarching interview guide was used, questions were adapted and follow-up questions were asked depending on the interviewee's role and expertise.

DEPARTMENT OF TECHNOLOGY MANAGEMENT AND ECONOMICS
DIVISION OF SUPPLY AND OPERATIONS MANAGEMENT
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden
www.chalmers.se

CHALMERS