FAKULTÄT FÜR !NFORMATIK
Faculty of Informatics

Inter-organizational Performance Analysis from EDI Messages

PhD THESIS

submitted in partial fulfillment of the requirements of

Doctor of Technical Sciences

within the

Vienna PhD School of Informatics

by

Worarat Krathu
Registration Number 0928102

to the Faculty of Informatics
at the Vienna University of Technology

Advisor: Univ.Prof. Dipl.-Ing. Dr.techn. Hannes Werthner

Wien, 02.07.2014

(Signature of Author) (Signature of Advisor)

Declaration of Authorship

Worarat Krathu
Donaufelder Straße 54/2210, A-1210 Vienna, Austria

I hereby declare that I have written this Doctoral Thesis independently, that I have completely specified the utilized sources and resources and that I have definitely marked all parts of the work - including tables, maps and figures - which belong to other works or to the internet, literally or extracted, by referencing the source as borrowed.

(Place, Date) (Signature of Author)

Acknowledgements

There are many individuals and organizations involved in the contributions of this thesis. First of all, the research presented in this thesis has been conducted in the course of the EDImine project which has been funded by the Vienna Science and Technology Fund (WWTF) through project ICT10-010. The EDImine project is jointly conducted by the Vienna University of Technology and the Eindhoven University of Technology. Furthermore, I, the author, was also supported by the Vienna PhD School of Informatics.

Individually, the first important person that I would like to thank is my PhD advisor, Professor Hannes Werthner, who gave me the opportunity for pursuing my PhD. He is respectful, very supportive and hardworking. Since I was a young woman having little knowledge about research he supported me for almost five years. Today, I am capable for conducting research on my own. Thanks to his patience! Moreover, thanks to my project leader, Marco Zapletal, who enabled me to participate in the project and my colleague, Christian Pichler, who closely supported me. Both of them guided me when I got lost and never left me alone when I had a hard time. Without them, I may not have come this far. There are also other colleagues who contributed to this thesis. This includes Robert Engel, Professor Christian Huemer as well as R.P. Jagadeesh Chaandra (JC) Bose and Professor Wil van der Aalst, colleagues from Eindhoven University of Technology, who gave me constructive comments for better improvements. Thanks also to Julia Neidhardt, Maria del Carmen Calatrava Moreno (Mamen), and Nataliia Pobiedina from the Electronic Commerce group for being my consultants in the application of network analysis methods. Thanks to Guohui Xiao from the Knowledge-Based System Group who is currently working at KRDB Research Centre for Knowledge and Data at the Free University of Bozen-Bolzano. He supported me in the research on semantics and reasoning. Apart from research, they are my good friends who always encouraged me as well. I also would like to acknowledge the support from companies which provided us with sample data for conducting the case studies. In addition, I would like to also thank the BOC group for providing us their ADOscore tool for research purposes. We do appreciate and acknowledge their contributions to our research community.

Last but not least, thanks to my family in Thailand and their encouragement from the beginning until today. Also, thanks to people from the School of Information Technology, King Mongkut's University of Technology Thonburi, where I worked and studied before pursuing my PhD. Furthermore, I want to thank Professor A Min Tjoa for supporting and introducing me to the Vienna PhD School of Informatics. Finally, I would like to thank my friends, especially people from the Vienna PhD School of Informatics and the Electronic Commerce group, for always being cheerful and supportive. With all of them I can always smile.

Abstract

The evaluation of Inter-organizational Relationships (IORs) is important in today's businesses for increasing competitiveness and business potential, especially in the Business-to-Business context. However, the evaluation of IORs is often based on success factors, such as trust and flexibility, which are difficult to be measured quantitatively. This motivates us to seek for an approach for evaluating IORs in a concrete manner. Nowadays, Electronic Data Interchange (EDI) widely enables information exchanges between business partners via electronic business messages. Therefore, it is considered as one potential data source for deriving knowledge about IORs. In this research, we aim at providing an approach for evaluating IORs from EDI messages by means of inter-organizational performance analysis. Considering this ultimate aim as a main research question, we further derived three relevant sub-research questions which are necessary to accomplish the main goal. This includes (i) extracting business information from EDI messages, (ii) identifying inter-organizational Key Performance Indicators (KPIs), and (iii) lifting the performance evaluation to the strategic level. In addressing the problem of business information extraction from EDI messages, we defined an ontological approach to conceptualize EDI data into business information concepts by user-defined mappings. Thereby, EDI data can be represented at any abstraction level which eases further querying tasks. Furthermore, to obtain inter-organizational KPIs from EDI messages, we firstly conducted a literature review on inter-organizational success factors together with their measurements and interdependencies among them. We further investigated EDI messages for identifying inter-organizational KPIs and grouped the KPIs according to success factors found in the review. Finally, we built a framework for inter-organizational performance analysis from EDI messages based on the aforementioned works. The framework integrates (i) a bottom-up approach supporting the identification of KPIs from business information, event logs, as well as process models derived using process mining techniques, and (ii) a top down approach for measuring business performance on the strategic level based on the Balanced Scorecard (BSC) method. For demonstrating the framework, we presented two case studies on inter-organizational performance analysis of a beverage manufacturing company and a consumer goods manufacturing company. The case studies show that the framework enables (i) the derivation of quantifiable KPIs from operational data and (ii) the alignment of KPIs with business objectives allowing an evaluation of IORs on the strategic level. The main contributions of this research are the artifacts developed for addressing the relevant research questions which include (i) an approach for business information extraction from EDI messages, (ii) a set of inter-organizational success factors and a corresponding cause and effect model, (iii) a set of inter-organizational KPIs and a method for deriving them from EDI messages, and (iv) an inter-organizational performance analysis framework.

Contents

CHAPTER 1

Introduction

Inter-organizational collaboration impacts an organization's performance [VM06; Dau+06]. In particular, organizations require inter-organizational collaboration to increase profitability, competitiveness, and growth. This is especially relevant in the Business-to-Business (B2B) context, where partnership is important to increase business potential.

During collaboration, business activities and transactions are executed within the network of organizations forming an inter-organizational business process, which we refer to as *choreography*. These choreographies of inter-organizational business processes specify the exchange of business documents, i.e. actual business information, as part of business processes. In particular, an inter-organizational business process comprises one or more messages exchanged between companies in order to conduct an electronic business transaction. The specifics of inter-organizational business processes require not only focusing on the executed activities, but also on the business information that is exchanged.

Nowadays, the exchange of business information is supported by Electronic Data Interchange (EDI). EDI is considered as one of the driving tools fostering collaboration between business partners, such as in supply chain management [Jan11]. During the evolution of EDI, different standards have been developed. This includes the traditional delimiter-based EDI standards such as UN/EDIFACT and ANSI X12 [KD10; The+01] as well as the XML-based document standards which are used in a Web Service environment. In recent academic research for Web Services and business process modeling, lots of emphasis is placed on modeling choreographies [BDO06]. In practice, however, neither explicit choreography modeling nor Web Services are widely employed in electronic business transactions. Rather, traditional EDI standards, especially, UN/EDIFACT standards, are still widely applied in B2B scenarios [Ber94; VGS07]. According to a survey [VGS07] conducted in 2007, an estimated 85%-90% of the total volume of electronic B2B transactions were carried out using traditional EDI standards at this time. Similarly, other surveys and studies including [Nur08; Son11; Jan11; Cho+11; MSC11] reveal that UN/EDIFACT standards will continue being one of the accepted standards in the B2B market due to their long establishment in the cross-industry context.

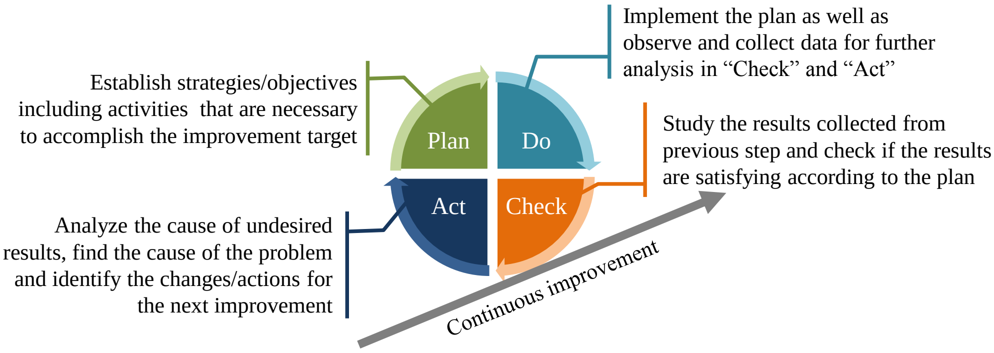


Figure 1.1: Management cycle Plan-Do-Check-Act [SPP10]

We believe that knowledge on Inter-organizational Relationships (IORs) can be derived by investigating inter-organizational business processes as well as the business documents exchanged in the course of the processes. This business information is valuable for analyzing IORs since it allows understanding business relationships and provides benchmarks to assess the “value” of business partners [PS08]. Based on the electronic messages exchanged during business collaboration, this work aims at evaluating IORs by means of inter-organizational business performance analysis. It is important to evaluate and analyze inter-organizational performance since the analysis supports continuous improvement for business management. For example, as shown in Figure 1.1, the Plan-Do-Check-Act management cycle is an iterative management process consisting of four main activities: plan, do, check, and act [SPP10]. In each cycle, performance evaluation plays an important role especially in the “Check” and “Act” activities. Particularly, performance evaluation helps an organization and its business partners to evaluate their achievement against their plan as well as to identify and analyze problems, faults, or weaknesses that affect their performance results. Consequently, the analysis results lead to necessary changes or actions for the next improvement cycle which in turn benefit the entire collaboration network.

Overall, the work described in this thesis is part of the *EDImine* project [Eng+11; Eng+12a], whose main objective is to extract inter-organizational knowledge from electronically exchanged messages. In the following, we describe the business context as well as provide an overview on the EDImine project in order to introduce the reader to the context of the work. Subsequently, the motivation and the research questions addressed in this work are introduced.

1.1 Background

The focus of the work presented in this thesis is on deriving inter-organizational knowledge from EDIFACT messages exchanged between business partners in the course of business transactions. We concentrate on UN/EDIFACT standards since it is, as mentioned earlier, a widely applied format for EDI documents. Figure 1.2 provides an overview on business partners communicating by electronically exchanging business documents during their business activities. Figure 1.2 illustrates an example where a manufacturer communicates with the supermarket partners. The

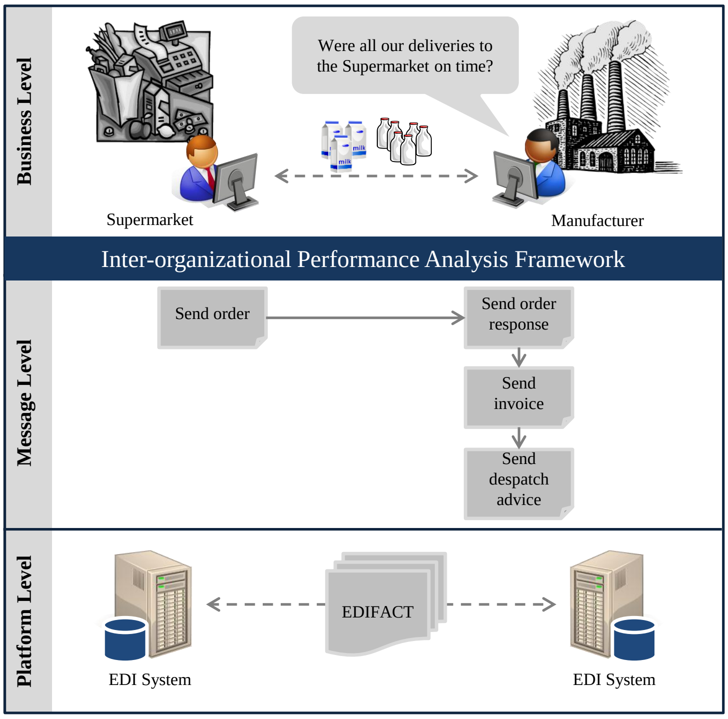


Figure 1.2: Business context

business level illustrates a purchasing processes between the manufacturer and the supermarket. The actual communication about activities is realized by exchanging electronic business documents. As depicted at the message level, such business documents are sent and received among business partners. In the example, the purchasing process starts when the supermarket sends an order to the manufacturer. Consequently, the manufacturer returns an order response to the supermarket. After goods are delivered an invoice is sent by the manufacturer. Finally, the manufacturer sends the supermarket a despatch advice in order to inform about detail of the delivery. This message interchange is technically supported by EDI systems of both parties as shown at the platform level. Particularly, the business documents such as orders and invoices are represented by EDI messages where the information is stored in a certain structure of a specific message type. The structure of EDI messages depends on EDI standards that are in use. As depicted at the platform level, those business messages on the message level are stored as EDIFACT messages and exchanged between EDI systems of both business parties.

Based on different contributions for performance analysis presented in this thesis, the overall contribution of this thesis is an inter-organizational performance analysis framework. As shown in Figure 1.2, the framework aims at supporting the analysis of IORs by means of business performance. It is designed to derive business information from the technical level (i.e., platform level and message level) for answering business-related questions on inter-organizational performance on the business level.

1.1.1 Inter-organizational Performance

In the context of the work presented, we define the term inter-organizational performance as organizational performance having two additional characteristics: (i) it is (partially) visible for the involved business partners and (ii) it is influencing or being influenced by IORs. Hereby, we do not distinguish between intra- and inter-organizational performance. We rather perceive inter-organizational performance as organizational performance which has the aforementioned characteristics.

The performance results derived from EDI messages meet the first aforementioned criteria of inter-organizational performance since the information conveyed in EDI messages is visible for message-exchange participants. The second criteria depends on whether the performance measures and their results affect IORs. For example, delivery performance can be perceived as one kind of inter-organizational performance. This is because (i) it is seen by business partners, at least by the one who participates in the delivery activities, and (ii) it influences the relationships between business partners in the way that if delivery performance is poor (e.g., unreliable deliveries, late deliveries, etc.), it will influence the decision of future collaboration (e.g., ordering, manufacturing, ect.) of the business partners.

However, the question whether a specific performance result influences or is influenced by IORs depends on the type of relationship. Considering the cost of product manufacturing and the relationship between the business partners is buyer and seller, then the cost may not be considered as an inter-organizational performance indicator, even though it can be seen by all participants. In this case the cost of product manufacturing may be considered as a burden solely for the seller who produces products, but not for the buyer. Based on the cost of product manufacturing there might be no effect on an ordering decision by a buyer unless the price of products is raised according to the cost. In contrast, it can be seen as an inter-organizational performance indicator in the case of stakeholder relationships where business partners corporate or invest together in product manufacturing.

1.1.2 Research Context EDImine

The research project EDImine aims at deriving inter-organizational knowledge from EDIFACT messages exchanged electronically during the collaboration of business partners. The EDImine project is jointly conducted by the Vienna University of Technology and the Eindhoven University of Technology. Furthermore, the EDImine project is funded by Vienna Science and Technology Fund (Wiener Wissenschafts-, Forschungs- und Technologiefonds, WWTF[1].

Figure 1.3 illustrates the overall framework of the EDImine project. In the preprocessing stage a collection of EDIFACT messages is preprocessed and stored as ontologies and event logs. In particular, EDIFACT messages are parsed into EDI/EDIFACT ontologies and knowledge bases which semantically describe UN/EDIFACT standards, message types, message instances and their contained values. EDIFACT ontologies provide means for the semantically enriched interpretation of actual values by considering qualifiers and codes. Furthermore, EDIFACT data is also conceptualized as generic business information (BI) concepts for reaching a higher level of abstraction for EDIFACT data.

[1] `http://www.wwtf.at` (visited March 01, 2014)

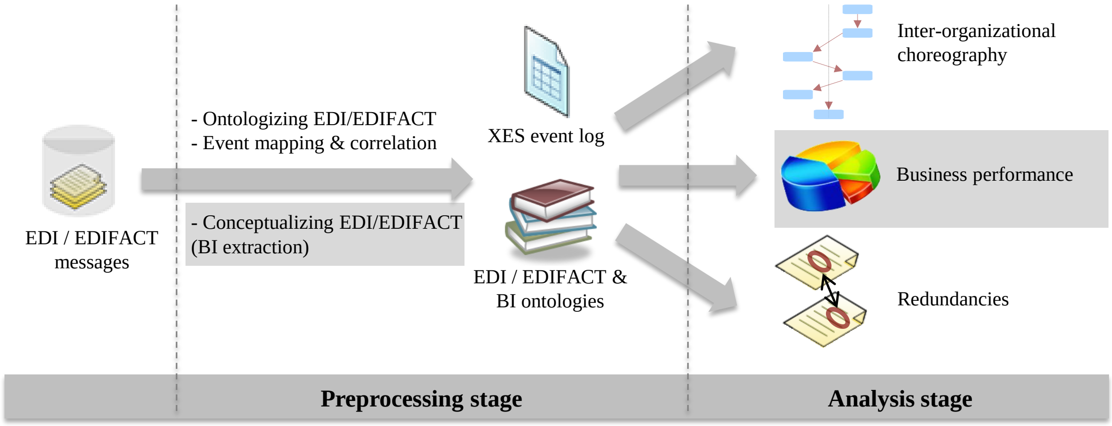


Figure 1.3: The overview of EDImine architecture [Eng+11]

Furthermore, process instances are derived from a collection of EDIFACT messages by mapping the messages and values contained in the actual messages to events and, consequently, correlating those events into process instances. The resulting process instances are stored in the *eXtensible Event Stream* (XES) format [Ver+11]. XES is an XML-based format for storing event logs and the standard input format for ProM[2] (as of Version 6), a well-known process mining tool. ProM is developed at the Eindhoven University of Technology and is the most prevalent tool in the area of process mining. The architecture of ProM has been designed with extensibility in mind by means of plug-ins. In the EDImine approach, we build upon state-of-the-art process mining techniques [Aal11; Aal+07], which we extend for inter-organizational systems realized by means of EDI. Hereby, we leverage the extensibility mechanisms of ProM by providing appropriate plug-ins for addressing our research questions.

The preprocessed XES log and EDIFACT ontologies serve as a basis in the analysis stage focusing on three main research goals: (i) deriving inter-organizational choreographies, (ii) analyzing business performance, and (iii) identifying information redundancies. The prepared XES data serves as a database for mining the inter-organizational choreographies. EDIFACT ontologies are used for the identification of redundancies by matching the business data contained in the individual EDI messages with regard to their conceptual belonging. Finally, the resulting mined process model as well as EDIFACT ontologies serve as information sources for business performance analysis.

As highlighted in Figure 1.3, the core contribution of this thesis is the second research goal of the EDImine project, i.e., the inter-organizational performance analysis. This also includes the preprocessing task of information extraction by generating BI ontologies which conceptualize EDI/EDIFACT messages into BI concepts. In the following, we explain the motivation of this work as well as provide an introduction to the detailed research questions addressed in this thesis.

[2] http://www.processmining.org (visited March 01, 2014)

1.1.3 Motivation and Research Questions

To understand the impact of business collaboration and to provide a benchmark for judging business partners, the evaluation of *IORs* is required [PS08]. However, the evaluation of IORs is typically difficult and ambiguous since it is usually measured by high-level *success factors*. Such success factors, including trust, knowledge sharing, and others, are difficult to be measured quantitatively since they are often intangible. Thus, the interpretation of these success factors into some quantifiable measures or *Key Performance Indicators (KPIs)* depends on the experience of business professionals. This problem motivates us to seek for an approach for enabling the evaluation of IORs in a concrete manner by making success factors more tangible by translating them into quantifiable KPIs derived from EDI messages.

Main research question. By investigating EDIFACT messages exchanged in a network of companies we aim at evaluating IORs by using the inter-organizational KPIs derived from the information contained in the exchanged EDIFACT messages. Furthermore, we also aim at lifting the evaluation, which uses KPIs gained on the IT level, to the strategic level. This leads to the main research question addressed in this thesis: *"How can we evaluate inter-organizational performance based on EDIFACT messages?"*. In addressing this main research question, we further define relevant sub-research questions which need to be solved in order to be able to answer the main research question.

Research question 1. Business information is required for supporting performance analysis tasks. Such analysis tasks are usually performed by management people who are aware of the information needed for evaluating business performance. However, EDIFACT messages are designed for enabling and automating business transactions. In particular, the information in EDIFACT messages is stored in a certain structure and parts of the information are encoded. Interpreting such EDIFACT messages requires in-depth knowledge on EDIFACT message structures. It is therefore difficult for management people to extract their interested business information from such messages. Consequently, support in extracting business information is necessary for obtaining business information in a human-understandable manner for facilitating the analysis task. This results in the first research question: *"How can we extract relevant business information from EDIFACT messages?"*.

Research question 2. The second research question relates to inter-organizational KPIs which is: *"What are inter-organizational KPIs that can be derived from EDIFACT messages?"*. The evaluation of business performance requires the KPIs for indicating the success of business. However, as mentioned earlier the evaluation of IORs is typically described through success factors which broadly describe measurement aspects without specifying concrete quantifiable measures. In order to provide a concrete way for evaluating IORs, the relevant quantifiable KPIs need to be identified. This includes identifying inter-organizational success factors and their measurements. Moreover, EDIFACT messages have to be investigated for deriving relevant KPIs which in turn can be used for measuring inter-organizational success factors.

Research question 3. Although the KPIs derived from EDIFACT messages are quantifiable and tangible for performance evaluation, and concretely reflect business performance, such KPIs are considered as low-level information and still cannot reflect business performance on a strategic level. This leads to the third research questions: *"How can we lift the evaluation of business performance to the strategic level?"*. This research question focuses on providing

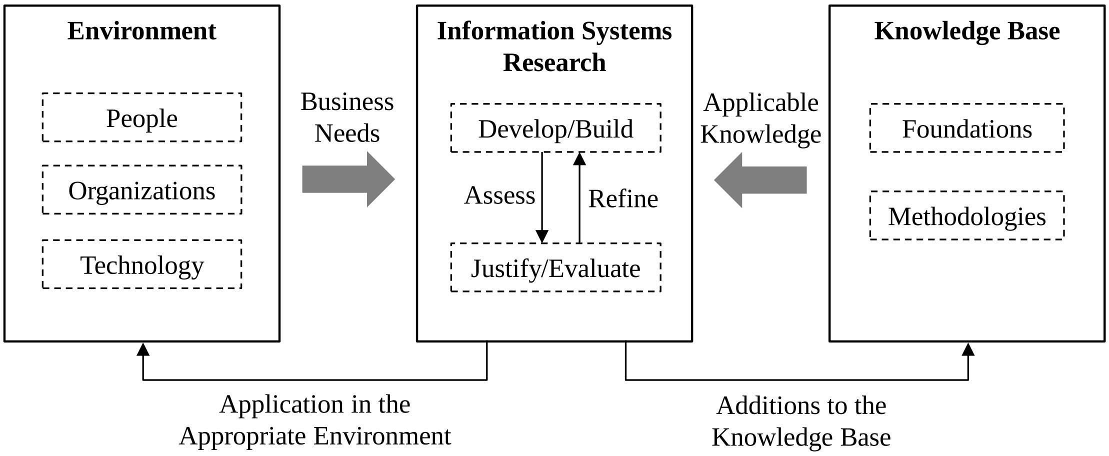


Figure 1.4: Information systems research framework [Hev+04]

an inter-organizational performance analysis framework which supports the identification and calculation of KPIs based on information derived from EDIFACT messages. Moreover, the framework provides means to connect such KPIs with business strategies which enables the evaluation of IORs on the strategic level.

1.2 Research Methodology

The research presented in this thesis follows the Design Science methodology introduced by Hevner et al. [Hev+04]. Figure 1.4 depicts the information system research framework presented as part of the design science methodology. In this research paradigm, the information system research is driven by business needs coming from business environments such as organizations, people and current technology. Such business needs drive researchers to develop artifacts (and/or theories) based on existing foundations and methodologies for supporting or fulfilling those needs. Driven by business needs and by building upon existing knowledge, artifacts are developed iteratively. In particular, the developed artifacts must be evaluated for justifying their utility. Several assessment methods can be applied including case studies, experiments, or simulations, etc. After the assessment of artifacts, refinement is required if artifacts still have weaknesses or if they are not yet developed to the point where business needs are satisfied. The resulting artifacts should be able to be applied as a solution to an appropriate environment having the same kind of business needs or may be able to be applied to more generic problems. In addition, the results of the design-science research can also be added to the existing knowledge base for further research applications. Hevner et al. proposed seven guidelines for an effective design-science research [Hev+04]. Those guidelines should be addressed when conducting research. The degree to which research satisfies those guidelines indicates the completeness of the research. These guidelines are elaborated on in the following.

1. **Design as an Artifact.** In design-science research, the resulting artifact must be viable. The artifact in the sense of the design-science paradigm includes instantiations (e.g., sys-

tems, products), constructs (e.g., concepts), models (e.g., representation, semantics/syntax) and methods (e.g., algorithms, techniques) applied in the development as well as the use of information systems.

2. **Problem Relevance.** The knowledge and understanding acquired from the design-science research must lead to technical-based solutions for important and relevant business problems.

3. **Design Evaluation.** The resulting artifacts must be well-evaluated in the sense of proving or demonstrating the usability, quality, and efficacy. The evaluation methods typically include: the *Observational method* such as case studies or field studies, the *Analytical method* such as static analysis, the *Experimental method* such as simulations or controlled experiments, the *Testing method* such as functional testing or structural testing, and the *Descriptive method* such as scenarios that are constructed to demonstrate the utility of the artifact. The artifact is complete and effective when the evaluation shows that it satisfies the requirements and constraints of its relevant problems.

4. **Research Contributions.** The effective design-science research must provide clear contributions within one or more of these areas:

   a) *The Design Artifact.* The artifact is the resulting contribution of the research conducted. Moreover, the artifact must provide a solution to the problem addressed as indicated through the bottom-left arrow of "Application in the Appropriate Environment" in Figure 1.4. In addition, the artifact may extend or apply existing knowledge in a new or innovative way.

   b) *Foundation.* The development of an artifact may lead to new knowledge or extend existing knowledge. Such knowledge can be added to the existing knowledge base (cf. the bottom-right arrow of "Additions to the Knowledge Base" in Figure 1.4). Therefore, it can be identified as a contribution of the research conducted.

   c) *Methodologies.* The creative development and use of evaluation methods and new evaluation metrics can also be considered as a contribution. Such new evaluation methods as well as new metrics can be added to the existing knowledge base for further evaluation of design-science research (cf. the bottom-right arrow of "Additions to the Knowledge Base" in Figure 1.4).

5. **Research Rigor.** The rigor in design-science research refers to rigorous methods for the construction and evaluation of artifacts. Researchers should effectively apply methods in the existing knowledge base including theoretical foundations and research methodologies for constructing and evaluating the artifacts.

6. **Design as a Search Process.** The iterative development of an artifact is considered as a search process. In particular, the design of an artifact is iteratively implemented to search for the best or optimal solution for the problem addressed.

7. **Communication of Research.** The research must be well presented to both technical-oriented and management-oriented audiences. For the technical-oriented audiences, the

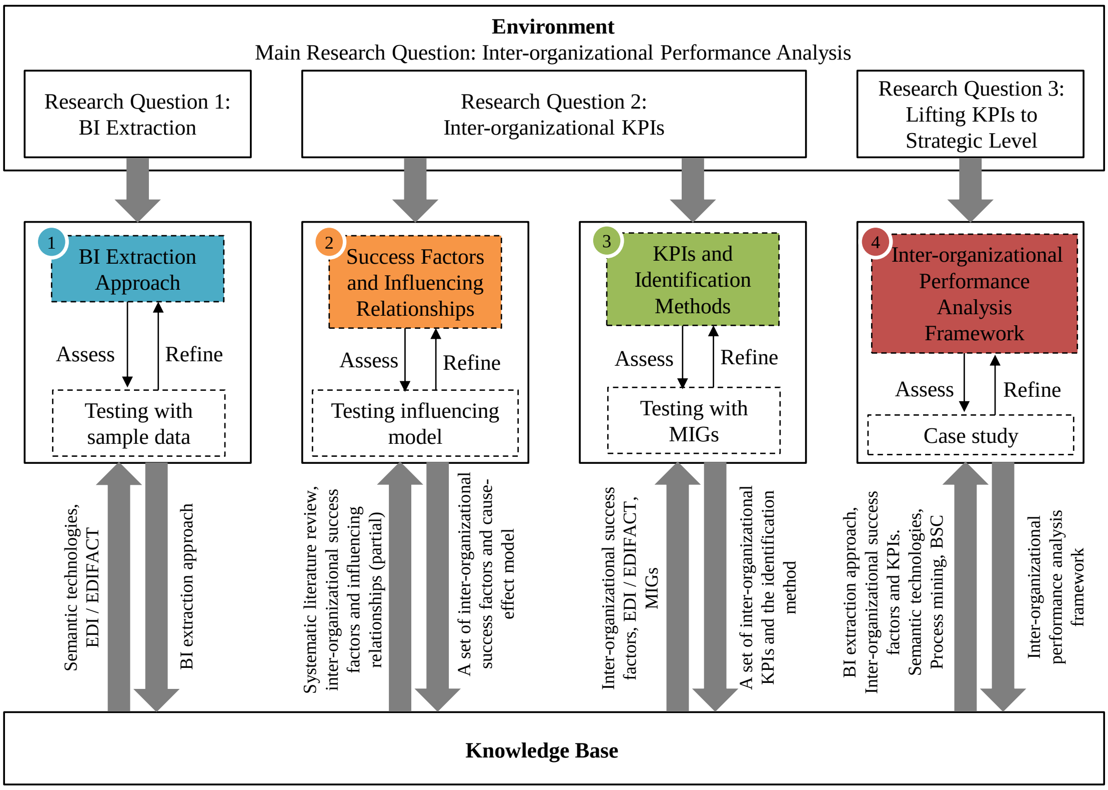


Figure 1.5: Design science process applied in this research

technical detail should be provided in sufficient detail for understanding the implementation of an artifact. This enables further re-usability as well as extension to the knowledge discovered in the research. Furthermore, the presentation must provide sufficient detail for management audiences in order to allow organizations to be able to acquire all requirements as well as knowledge for employing the artifact.

The Research Methodology applied in the Context of this Thesis. The research conducted and presented in this thesis is motivated by the need of evaluating IORs based on business documents exchanged electronically between business partners. Figure 1.5 illustrates design-science process applied to our research. The research is conducted for addressing three research questions presented in 1.1.3 which leads to the solution addressing the main research question. The research is implemented sequentially. In other words, we employ the design-science process for each research question in sequential order.

For each research question, the contributed artifact or knowledge has been iteratively developed and tested against requirements. During the development process, existing and relevant knowledge is applied properly. Once the developed artifact satisfies its requirements, the development process ends and the artifact and/or knowledge found are added to the existing knowledge base. The knowledge as well as the contributions of prior processes are further applied in other following processes.

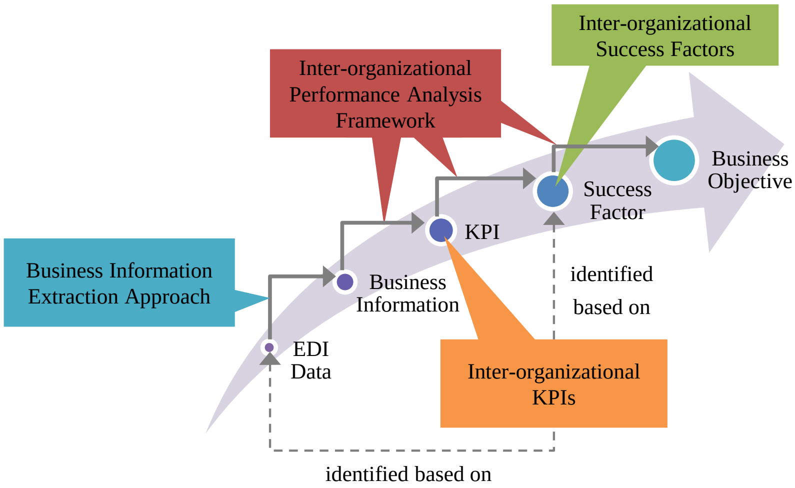


Figure 1.6: Overview of contributions in this thesis

In the last step of our research (i.e., the fourth process), all prior contributions are integrated as a final artifact representing a framework supporting the evaluation of IORs. Consequently, such a combination serves as a solution for the main research question. The feasibility of the final artifact is shown using case studies. In the following, the contributions derived from this research and their underlying challenges are described in details.

1.3 Contributions of this Thesis

There are four main contributions presented in this thesis: (i) a business information extraction approach, (ii) inter-organizational success factors, (iii) inter-organizational KPIs, and (iv) an inter-organizational performance analysis framework. The main focus of the contributions in this thesis is to link EDI data to business objectives for supporting the performance analysis and monitoring on a strategic level, as depicted in Figure 1.6. The first contribution of the business information extraction approach allows extracting and representing EDI data as generic business information concepts (e.g., order quantity, invoiced amount, etc.). The contributions of inter-organizational success factors and KPIs (i.e., the second and the third contribution) focus on identifying KPIs and success factors related to the success of IORs. For connecting business information concepts, KPIs, and success factors to business objectives, the inter-organizational performance analysis framework is introduced as the last contribution. The framework enables the connection between business information and business objectives through KPIs and success factors. In other words, the framework allows defining KPIs based on extracted business information as well as linking those KPIs to success factors and business objectives.

We define the maturity of these contributions by using the knowledge contribution framework of the design-science research introduced in [Hev11]. As illustrated in Figure 1.7, the knowledge contribution framework indicates maturity of the contributions by two dimensions:

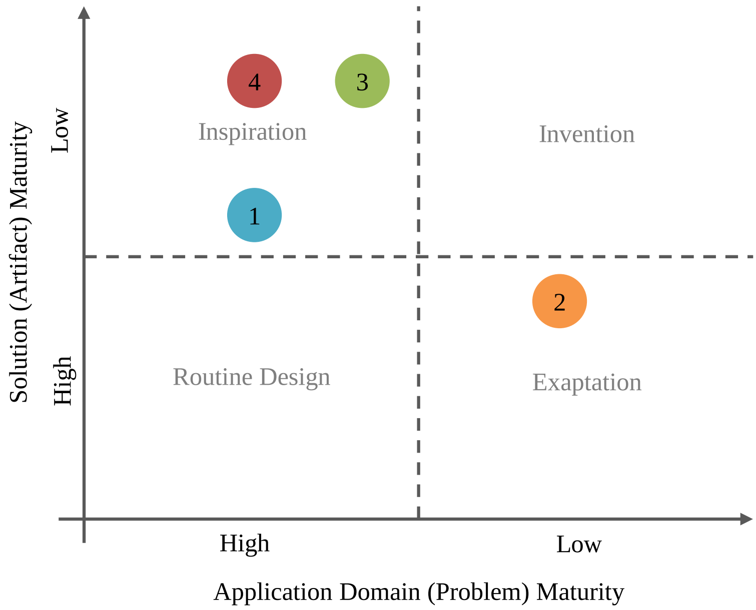


Figure 1.7: Contributions maturity in knowledge contribution framework [Hev11]

(i) the maturity of application domain (i.e., problem) and (ii) the maturity of solutions (i.e., artifact). The framework divides maturity into four main quadrants as follows:

- **Routine Design.** The first quadrant refers to contributions which apply existing solutions to known problems. In particular, the framework considers that the more well-known the problem is, the more mature the problem is. Similarly, if the solution already exists, the maturity of the solution is considered high. The more the solution is applied, the more mature the solution is. The *routine design* quadrant refers to contributions where their focused problem and their proposed solution are very mature.

- **Inspiration.** The quadrant of *inspiration* refers to contributions solving mature problems with new solutions. In this quadrant, the research opportunities are new artifacts that are developed for tackling known problems.

- **Exaptation.** Extending or adapting known solutions to new problems is classified into the *exaptation* quadrant. The core of contributions lying in this area is to raise new challenges and to address those challenges with existing solutions. The adaptation and the extension of existing knowledge for solving unknown problems can be considered as a research opportunity.

- **Invention.** The research opportunity in the *invention* quadrant lies in introducing new problems as well as providing unknown solutions for them. The artifacts located in this quadrant can be called as inventions, since they introduce totally new contributions to society. This may lead to the emergence of a new branch of knowledge, a new paradigm, or a new realm of particular domain rather than extending existing state-of-the-art.

The contributions of this thesis, which are described in the following, are classified into the inspiration quadrant (i.e., the first, third and fourth contribution) and the exaptation quadrant (i.e., the second contribution). In the following, all challenges addressed in the context of this thesis as well as the corresponding contributions are explained in detail.

1. **Business Information Extraction from EDI Messages**

   Challenge. Due to the long establishment and wide utilization of EDI technology, many different standards have been developed. In addition, individual standards generally comprise multiple different versions. Moreover, EDI standards are typically afflicted with a lot of optionality. For instance, names of data elements as well as their positions in containing segments may be changed from version to version. For example, one EDI standard/version might use the term "Document identifier", while another might use the term "Document/message number" for the same data element (i.e., the data element describing identification number of EDI messages). Furthermore, codes with different interpretations may actually refer to the same meaning. For example, a code representing the term "seller" may be considered as semantically similar or identical to a code representing the term "supplier" for certain purposes. In order to allow generic and automated business information extraction from EDI messages, the employed extraction mechanism must be able to deal with different standards/versions of EDI documents as well as optionality. Hence, mappings from data elements of EDI standards to more generic business information concepts are required. It is necessary that such mappings can be specified in a way so that both, multiple syntaxes as well as subtle differences in the semantics of data elements and codes, can be flexibly accounted for.

   While current state-of-the-art EDI systems typically allow for accurate information extraction from specific subsets of EDI standards (generally by using hard-coded interpretation logic), the automated and accurate interpretation of arbitrary EDI messages still poses a challenge. In particular, the challenge is to arrive at an information extraction mechanism that allows for a sufficient degree of automation while at the same time being flexible enough to deal with different EDI standards/versions and the pitfalls of accurately determining the semantics of concrete data element instances. Hence, ontologies that describe and unify the semantics of different EDI standards are needed.

   Contribution. In the course of EDImine project [Eng+11; Eng+12a], an approach interpreting EDI data and storing it in EDI ontologies has been introduced. The approach allows automatically resolving coded data as well as qualified data of arbitrary EDI standards/versions, and storing EDI data in EDI ontologies which describe both syntax (i.e., structure of messages) and semantics (i.e., interpreted information). However, full interpretation of EDI data may also rely on the position of data elements which implicitly specify the exact meaning. Moreover, mapping EDI data to generic business information concepts is still necessary for facilitating further analysis tasks which are typically implemented by management people who are unaware of the technical background on EDI.

   To overcome the described challenge, we developed an ontological business information extraction approach based on existing EDI ontologies, which have been developed in the course of the EDImine project, for supporting user-defined mapping of EDI messages and generic BI concepts. In particular, the approach allows the mapping definitions by using ontological rules. Based on the mapping, the underlining mechanism generates BI ontologies on top of EDI ontologies. The detailed explanation of this approach is further elaborated on in Chapter 3. The work on this business information extraction approach has also been published in [Kra+12a].

The problem of business information extraction from EDI messages is a classic problem which has been identified since decades. Several related solutions have been introduced. Best to our knowledge, the current state-of-the-art of EDI data extraction is typically realized by hard-coded interpretation. However, our approach aims at providing an automation for interpreting arbitrary EDI messages in any different standards as well as allowing the flexibility of generating generic business concepts specialized in any specific domain by defining ontological rules. Since our approach improves existing business extraction methods, we assign this contribution to the inspiration quadrant of the knowledge contribution framework (cf. Fig. 1.7, Mark 1).

2. **Inter-organizational Success Factors and their Influencing Model**

   Challenge. The information conveyed in EDI messages is valuable for analyzing inter-organizational businesses. Analyzing inter-organizational business performance helps to assess the value of business partners and the performance of collaboration, and hence lead to the understanding of the impact of IORs [PS08]. Most studies concerned with the evaluation of IORs (e.g., [Cas08], [SQ03]) tend to build upon the analysis of success factors having an impact on IORs. For example, trust [SBS07; ZH06; Sau+04], information sharing [LL06; Che11] and joint working [KHT07; Joh+04; Duf+12] are mentioned as such factors, which are difficult to measure. The underlying reason is that success factors are generally broadly defined.

   In this work, we proposed using quantifiable KPIs derived from EDI messages to improve quantifiability and explicitness of evaluating IORs. Nevertheless, inter-organizational success factors are still important as they provide a limited scope of one particular aspect of measuring an IOR. The limited scope supports the identification of relevant KPIs within one particular success factor. Therefore, success factors, their measurements, and their inter-dependencies need to be investigated beforehand.

   In order to identify all relevant success factors in the entire context of IORs, all relevant knowledge found in the studies must be extracted and integrated. This is challenging since most of the studies on inter-organizational success factors and their influencing relationships focus on some particular success factors. Moreover, some success factors are named differently but semantically equal in terms of meaning. Therefore, the integration and simplification of such knowledge for further applications pose a major challenge.

   Contribution. In addressing this challenge we conducted a systematic literature review on inter-organizational success factors. The review concentrated on identifying success factors having an impact on the success of IORs and their influencing relationships. The main contributions of the review is (i) a set of inter-organizational success factors organized in a hierarchical structure, and (ii) a cause and effect model of influencing relationships between those success factors. The cause and effect model can be used for inferring further influencing relationships which, though not found in the review, implicitly exist in the knowledge. For making this hidden knowledge apparent we defined and applied ontological rules.

The knowledge of inter-organizational success factors and their relationships is derived for understanding the success of IORs. The objective of this work is to identify success factors related to the success of IORs as well as expressing the influencing relationships through a cause and effect model. The model serves as a basis for supporting decision making within an organization. Existing and relevant studies mostly concentrate on developing an understanding of a dedicated set of success factors and their influence on IORs. In contrast, our work aims at developing knowledge of inter-organizational success factors by integrating those studies and deriving an implication from it. For acquiring such knowledge, we implemented a systematic literature review to identify success factors. Furthermore, we adopted semantic technologies for representing this knowledge as well as for deriving inferences by using ontological rules. This resulted in knowledge on the impact of success factors on the success of IORs regardless of any specific business or application domain. The objective of deriving an entire understanding of success factors in IORs is still considerably new. By applying and adapting an existing knowledge for accomplishing this objective, the contribution is located within the exaptation quadrant of the knowledge contribution framework.

3. **Inter-organizational KPIs and the Method of KPI Identification from EDI Messages**

Challenge. As mentioned earlier, we aim at identifying inter-organizational KPIs from EDIFACT messages. Nevertheless, UN/EDIFACT standards define different message types (e.g., order message, invoice message, etc.) and each of them contains a variety of data elements. Although data elements appearing in each message type are already predefined in the standard, most of them are optional. In other words, the appearance of data elements depends on the customization performed by an organization itself. Consequently, due to the large number of message types and the high number of data elements as well as their optionality, the KPI identification from EDIFACT messages becomes a major obstacle.

Contribution. To this end, we studied different EDIFACT message type specifications in various EDIFACT releases (ranging from D96A to D10A) and identified sets of KPIs. These KPIs reflect the success factors identified in the aforementioned literature review. Thereby, we considered the frequencies of data elements as well as the semantics of both data elements and message types. We applied this method on a data set of real-world industry Message Implementation Guidelines (MIGs) and derived a set of inter-organizational KPIs as well as defined guidelines for their calculation based on concrete EDIFACT data. The KPIs identified are then used for quantitative measurements of inter-organizational success factors. Details on the employed method for KPI identification and the resulting set of identified KPIs are published in [Kra+13].

The identification of inter-organizational KPIs is challenging, especially in the context of performance analysis. Best to our knowledge, however, the solutions provided in the existing state-of-the-art focus on defining success factors rather than tangible and quantifiable KPIs. In addition, determining potential KPIs for indicating business performance depends on the interest of organizations. Nevertheless, in this work we aim at providing a set of KPIs derived from EDIFACT data which may be potentially applied for measuring inter-organizational business performance. Furthermore, we also provide a method for KPI identification which can be applied to other EDI standards. Therefore, this contribution can be classified in the inspiration quadrant. The maturity of this focused problem is considered lower than the first contribution (i.e., business information extraction). Although the problem of KPI identification is already known, the problem is less advanced when comparing to the business information extraction problem. Nevertheless, the solution presented in this thesis is novel since there exists no related solution.

4. **Inter-organizational Performance Analysis Framework**

   Challenge. The work of business performance analysis presented in this thesis covers both, business-related information and process-related information. The KPIs applied in the analysis tasks are derived from the information in both perspectives. Business-related information can be extracted by using our business information extraction approach, whereas process-related information such as choreographies or process models as well as time-related process information can be discovered by applying existing process mining techniques. Although KPIs obtained from both types of information can reveal insights on the operational performance of business transactions, the bottom-up KPI identification approach does not directly reflect inter-organizational performance on the strategic level. In particular, process mining is a bottom-up performance analysis approach where event logs are mined for deriving process models. Subsequently, performance analysis is done based on the mined models. ProM 6 [Ver+11], the most prevalent tool in process mining, provides several plug-ins supporting analyzes based on low-level log data (e.g., ILP Miner [Wer+08], α-Miner [AWM04], performance analysis through process mining [Hor07]) as well as business data (e.g., data-aware process mining [LA13]). Results from process mining can also be applied for in-depth analysis of business processes for answering specific business-related questions. However, a drawback of bottom-up approaches is that they usually fall short of accurately reflecting business success on the strategic level.

   In contrast, top-down approaches for performance analysis exist which enable performance evaluation on a strategic level. Top-down approaches suggest to base the performance evaluation on business objectives and translate these objectives into measurements. Balanced Scorecards (BSC) [KN92] are a widely applied top-down measurement system [Eck06]. There are also several works on applying BSC in inter-organizational contexts such as Supply Chain Management (SCM). For instance, Brewer et al. [BS00] and Bullinger et al. discussed the interrelationship between BSC and the SCM field and introduce approaches for supply chain performance analysis based on the BSC method. Kleijnen et al. [KS03] and Chia et al. [CGH09] studied examples of KPIs commonly used for measuring supply chain performance following the BSC paradigm. However, top-down approaches are difficult to implement since business objectives and/or strategies are often too broadly defined and, hence, too ambiguous to identify appropriate KPIs. Nevertheless, top-down and bottom-up approaches complement each other. The challenge is how to leverage the advantages from both approaches for improving the performance evaluation method and for lifting the evaluation from bottom-up identified KPIs to the strategic level.

Contribution. In addressing this challenge, we developed an inter-organizational performance analysis framework where top-down (i.e., BSC method) and bottom-up (i.e., process mining) performance analysis approaches are integrated. We suggest that the top-down and bottom-up styles complement each other in that a bottom-up definition of KPIs based on real data facilitates the calculation of KPIs in concrete use-cases and that a top-down alignment of business objectives with KPIs enables the evaluation of business performance on the strategic level.

The framework is implemented by using semantic technologies. It consists of (i) the BSC ontology and (ii) a knowledge base describing success factors and KPIs identified from the previous contributions (i.e., inter-organizational success factor and KPI identification). The BSC ontology conceptually describes BSC elements such as business objectives, success factors, and KPIs. Using the BSC ontology, KPIs can be modeled and aligned with relevant business objectives. The predefined set of success factors and KPIs allows for the automated suggestion of potential KPIs with regard to concrete instances of input data. This is supported by using the influencing model of success factors and ontological rules for inferring related KPIs. The framework and its detail have been published in [Kra+].

Since the problem of inter-organizational performance analysis already appears in the scientific community, we position its maturity as high. However, the solution we provide in this contribution is considerably advanced. Although there exist several top-down and bottom-up performance analysis approaches, those approaches have their limitations. In this work, we address these limitations by integrating both approaches in order to utilize the advantages of both. Therefore, this contribution is located in the inspiration area where the problem is quite mature and the related solutions are immature.

1.4 Outline of this Thesis

The presentation of this thesis is structured as illustrated in Figure 1.8. In the upcoming chapter we provide the background on the necessary knowledge which this thesis is builds upon. The detail of each challenge and contribution introduced above is presented within a dedicated chapter. In other words, the chapters, addressing a particular challenge and contribution, provide their related works as well as their assessments and a final summary. The related work given in each chapter emphasizes on the studies related to a specific contribution, whereas the state-of-the-art provided in Chapter 2 refers to the main fundamental knowledge on which this research relies (e.g., EDI, balanced scorecard, process mining, etc.). In addition, the last contribution of this research, i.e., inter-organizational performance analysis framework, and its demonstration are presented separately. The framework builds upon the combination of prior contributions, hence

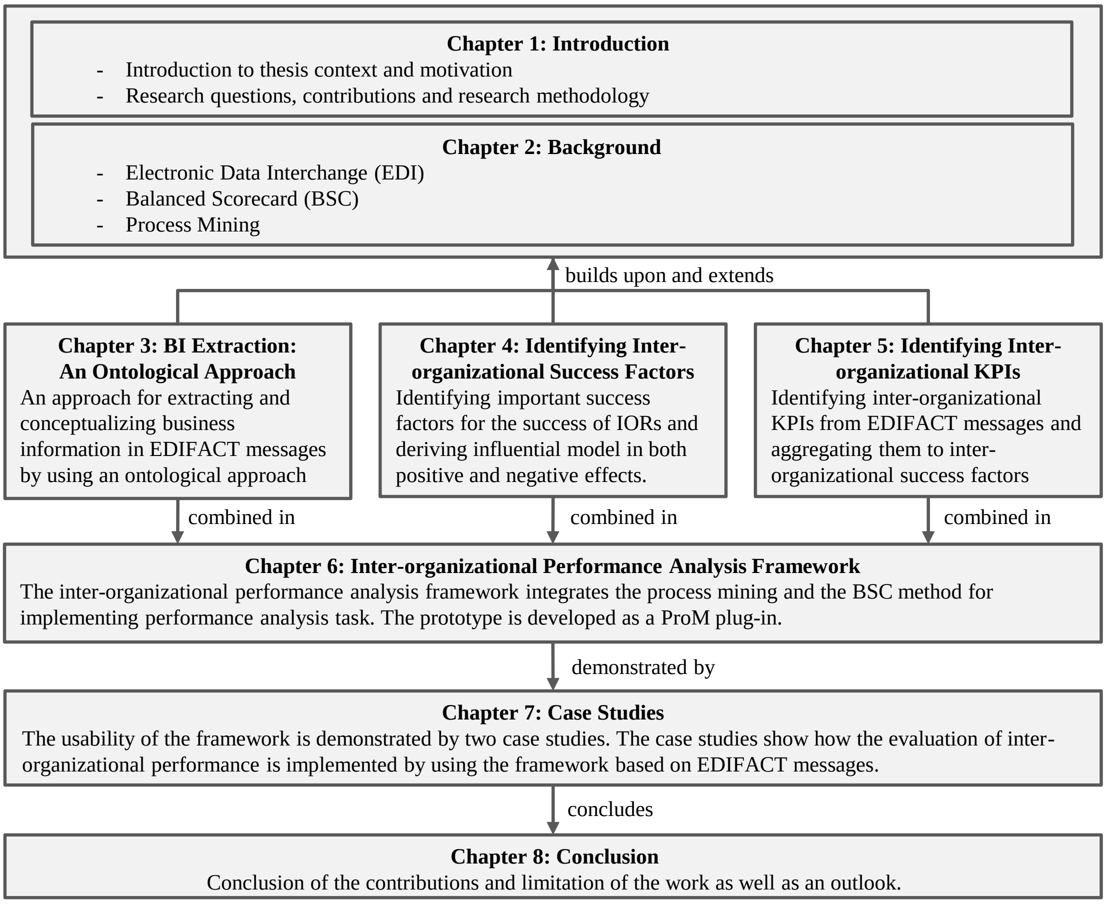


Figure 1.8: The outline of the thesis

it consequently addresses the main research question. Therefore, we provide a demonstration of the framework with two case studies as its final assessment. Furthermore, the contributions have already been published and submitted to scientific conferences and workshops. Therefore, the corresponding publications are pointed to where applicable.

Chapter 2. The upcoming chapter provides the theoretical background necessary for the work presented in this thesis. This includes a solid background of B2B and how it relates to EDI. As mentioned earlier, this work concentrates on UN/EDIFACT standards. Therefore, we provide the background on the standards in both, theoretical and technical aspects. Furthermore, process mining techniques and the BSC method, both used for performance analysis, are discussed. In the course of this thesis, both approaches are integrated for enabling quantifiable performance evaluation on a strategic level.

Chapter 3. In this chapter, we present an ontological approach for extracting business information from EDIFACT messages addressing the first research question. The content of this chapter is based on [Kra+12a]. Section 3.3 describes the approach in detail. This includes (i) the Meta-BI ontology used for mapping EDIFACT data to generic BI concepts by defining mappings with ontological rules, (ii) the BI ontology generation, and (iii) the classification of EDIFACT data into BI concepts by reasoning on the BI ontology.

Chapter 4. As described in Figure 1.8, this chapter focuses on the identification of inter-organizational success factors. In addressing the research question of KPI identification in an inter-organizational context, the relevant success factors which influence the success of IORs are investigated at first. Chapter 4 presents the results of the literature review on inter-organizational success factors. In particular, the inter-organizational success factors as well as their influencing relationships are identified. Moreover, the success factors and their influencing relationships are simplified and presented as a cause and effect model presented in Section 4.4. Finally, implications based on the findings are derived by applying network analysis methods. This work is currently under submission [Kra13].

Chapter 5. In addition to inter-organizational success factors, inter-organizational KPIs are identified. In this chapter, the method for identifying KPIs from EDIFACT messages is introduced and elaborated on. Based on the success factors found, the identified inter-organizational KPIs are grouped according to the corresponding success factors. The content of this chapter has been published in [Kra+13].

Chapter 6. The final artifact of this work is the inter-organizational performance analysis framework. The framework integrates the process mining approach and the BSC method for enabling bottom-up KPI definition as well as top-down alignment of KPIs and business strategies. This enables performance evaluation on a strategic level. Hence, it provides a solution to the third research question. The framework builds upon artifacts resulting from prior contributions as illustrated in Figure 1.8. In particular, the business information extraction approach presented in Chapter 3 is used for data preprocessing. The knowledge on inter-organizational success factors and KPIs presented in Chapter 4 and Chapter 5 is stored as a knowledge base for supporting the performance analysis tasks performed through the framework. Moreover, semantic technologies are employed as the underlying technologies for implementing the BSC method. In particular, the BSC ontology is applied in the framework for modeling and aligning BSC elements together with the employed BSC calculation method, which is discussed in Section 6.3. The content described in this chapter has been published in [Kra+12b].

Chapter 7. In this chapter two case studies are presented for showing the application of the framework introduced in Chapter 6. The aim of the case studies is to demonstrate the technical as well as the practical feasibility of the framework. We perform an inter-organizational performance analysis based on sample EDIFACT messages collected from EDI systems of a consumer good manufacturing company and a beverage manufacturing company. The first case study of the beverage manufacturing company is presented in Section 7.2. The second case study of the consumer good manufacturing company is discussed in Section 7.3. In particular, the case studies demonstrate the framework's utility starting from plain EDIFACT messages and ending by delivering performance analysis results.

Chapter 8. The conclusion of this work is presented in Chapter 8. This includes a brief summary of the work about, the developed artifacts, the contributions, as well as a discussion on their benefits and limitations. Finally, an outlook on future works is provided.

CHAPTER 2

Background

In this chapter, the theoretical foundations which are adopted and/or extended in this research are discussed. Since the research focuses on deriving inter-organizational performance information from EDI data, we provide a background on EDI including the UN/EDIFACT standard in Section 2.1. This also includes a discussion on conceptualizing EDI data. The following sections also provide an overview on performance analysis approaches the research in this thesis builds upon. This includes process mining, a bottom-up approach for performance analysis, and the balanced scorecard method, representing a top-down performance analysis approach. The background on process mining and the balanced scorecard method is provided in Section 2.3 and Section 2.2.

2.1 Electronic Data Interchange

Electronic Data Interchange (EDI) refers to a way of transmitting structured business documents electronically [Emm90, p. 5] such as orders, invoices, and receiving advices. It is widely applied in the B2B context [Tim01] where communication between business partners plays an important role. The initial purpose of EDI is to replace paper-based communication between business partners with electronic documents transferred between the systems of business partners [Emm90; HF89]. The instant advantage of replacing paper-based business documents is the elimination of processing delays and data reentry. Furthermore, unlike other types of electronic documents such as PDF (Portable Document Format) or emails, EDI documents are structured. This makes EDI documents machine-processable, and hence it allows the automation of data entry. Such an automation task helps speed up business processes in the way that business information can be retrieved and stored quickly. Therefore, processing delays caused by manual implementation can be eliminated. Furthermore, employing EDI in an organization supports a lot of long-run improvements such as cost saving, improving operations, and improving customer responsiveness [Emm90].

However, all business participants needs to agree on technical issues such as the specific data and document format as well as on the technical environment in order to exchange their elec-

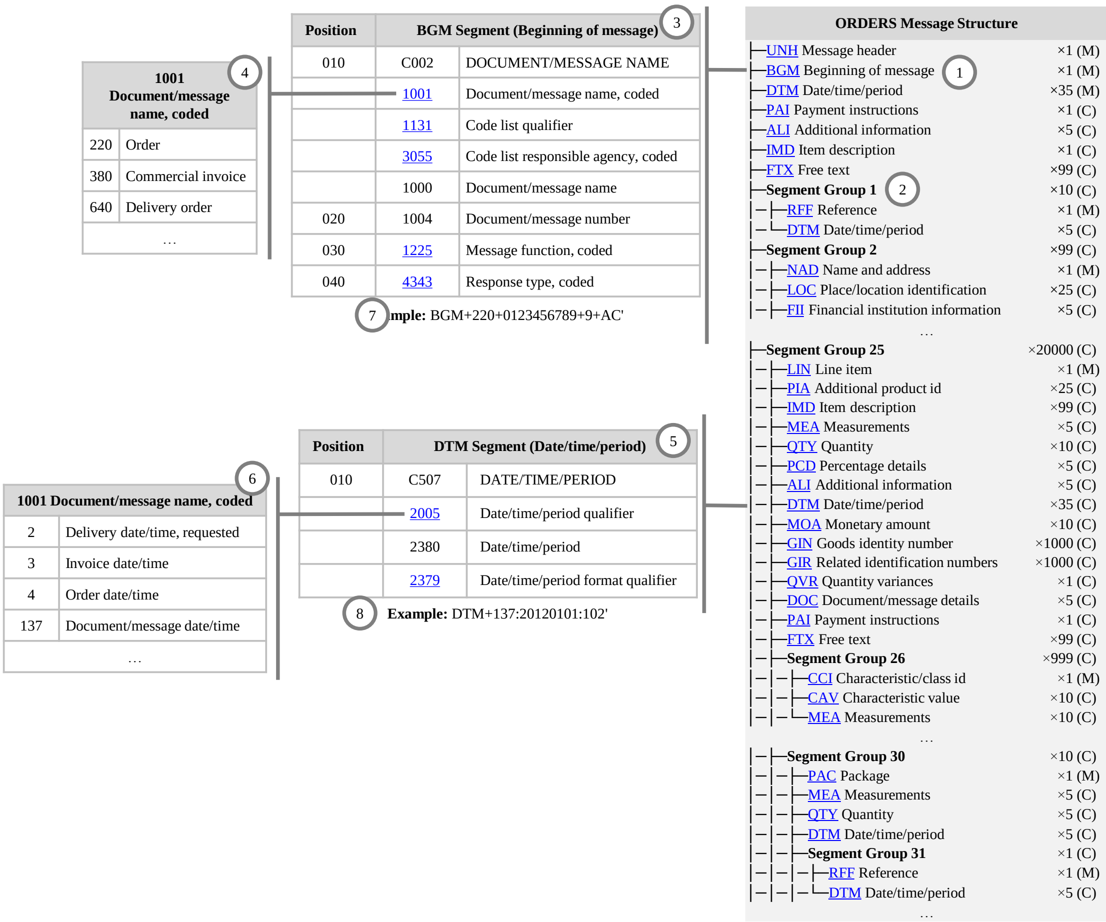


Figure 2.1: An order (ORDERS) message structure (excerpt)

tronic documents [HF89]. The agreement is originally only among a specific business network, but due to the widespread use of EDI in today's business common agreement is required. In particular, EDI is currently widely applied since business collaboration becomes important for companies' competitiveness, and hence EDI is required for supporting information exchange for successful collaboration. In order to foster the widespread use of EDI, several EDI standards (e.g., ANSI X12, UN/EDIFACT, etc.) have been developed. EDI standards define the structure as well as the semantics of electronic documents. In other words, these standards are used as common agreements on the format and the structure of business documents that are exchanged among business partners. In the following, we present the foundation of the UN/EDIFACT standard since the implementation of this research is based on it. The reason is that the UN/EDIFACT standard is still playing a dominant role in inter-organizational business scenarios [Ber94; VGS07] due to its long establishment [Nur08; Son11].

2.1.1 UN/EDIFACT Standard

The United Nations Electronic Data Interchange for Administration, Commerce and Transport (UN/EDIFACT) is an international EDI standard developed and published by the United Nations Economic Commission for Europe[1]. It comprises of a set of syntax rules describing how to structure data. In particular, pieces of data are formed as data elements. The UN/EDIFACT standard defines a specific structure for different message types (e.g., order message, invoice message, etc.).

Figure 2.1 shows an example of an order message structure (ORDERS)[2]. The order message consists of a sequence of segments and/or segment groups (cf. Fig. 2.1, Mark 1 and 2 respectively). For each segment or segment group the maximum number of occurrences and whether the segment group is mandatory or optional, is indicated by the letters *M* and *C* accordingly. A segment consists of data elements and/or composite data elements consisting of multiple data elements. For example, the *BGM* (Beginning of message) segment (cf. Fig. 2.1, Mark 3) is comprised of a composite data element *C002 DOCUMENT/MESSAGE NAME* and three additional data elements including: *1004 Document/message number*, *1225 Message function, coded*, and *4343 Response type, coded*. The composite date element *C002 DOCUMENT/MESSAGE NAME* includes the data elements *1001 Document/message name, coded*, *1131 Code list qualifier*, and *3055 Code list responsible agency, coded*. Some data elements are coded which means that their data is represented through codes. For these coded values the EDIFACT standard provides code lists specifying their semantics. For example, the data element *1001 Document/message name, coded* (cf. Fig. 2.1, Mark 4) can have any value from a specific code list. In order to interpret these coded data elements, the value must be resolved using the corresponding code list. Furthermore, some data elements or segments have qualifiers which are coded as well. A qualifier data element provides specific meaning to its qualified data elements or segments. For example, the segment *DTM* (Date/time/period) consists of a composite data element *C504 DATE/TIME/PERIOD* (cf. Fig. 2.1, Mark 5) which includes the data element *2005 Date/time/period qualifier*, the data element *2380 Date/time/period*, and the data element *2379 Date/time/period format qualifier*. The data element *2380 Date/time/period* is the data element qualified by the *2005 Date/time/period qualifier* and the *2379 Date/time/period format qualifier*. The data element *2005 Date/time/period qualifier* indicates the exact meaning of date/time data depending on the used value listed in its code list, such as requested delivery date/time, and order date/time (cf. Fig. 2.1, Mark 6). In addition, the data element *2379 Date/time/period format qualifier* specifies date/time format (e.g., “CCYYMMDD” (Calendar date: C = Century ; Y = Year ; M = Month ; D = Day.), “DDMMYY”, etc.). In the syntax of the UN/EDIFACT standard, the plus sign (+) is used as a separator between positions of (composite) data elements and the colon (:) is used as a separator between data elements inside composite data element. The apostrophe (’) is used at the end of a segment.

Figure 2.1, Mark 7 shows an example of the *BGM* segment: BGM+220+0123456789+9+AC’. Separated by the plus sign, the first part indicates segment type. In the example, it is *BGM* which indicates that this is the beginning of a message segment. The next position is the position of the

[1] `http://www.unece.org/cefact/edifact/welcome.html` (visited March 01, 2014)

[2] The structure described in Fig. 2.1 is from the UN/EDIFACT standard version D96A which is obtained under: `http://www.stylusstudio.com/edifact/D96A/ORDERS.htm` (visited March 01, 2014).

Order No. 0123456789	
EDIFACT Message Content	**Translated Message Content**
UNH+000000001+ORDERS:D: 96A:UN:EAN008'	This is an ORDERS message type in the version of 96A with message reference number 000000001.
BGM+220+0123456789+9+AC'	This is an order having a number 0123456789.
DTM+137:20140101:102'	The message/document date/time is 20140101 (in a CCYYMMDD format).
DTM+2:20140115:102'	The requested delivery date/time is 20140115 (in a CCYYMMDD format).
NAD+BY+1111111::9+BUYERCOMPANY'	The participant in this message is a buyer named BUYERCOMPANY with the identification of 1111111.
NAD+SU+2222222::9+SUPPLIERCOMPANY'	The participant in this message is a supplier named SUPPLIERCOMPANY with the identification of 2222222.
LIN+1'	Line item 1
PIA+5+0000001234:IS'	Product identification is 0000001234 which is a type of ISSN number.
QTY+21:100'	Order quantity is 100 units.
LIN+2'	Line item 2
PIA+5+0000009876:IS'	Product identification is 0000009876 which is a type of ISSN number.
QTY+21:50'	Order quantity is 50 units.
UNS+S'	The following is a summary section.
CNT+2:2'	Number of total line items in this message is 2 line items.
CNT+1:150'	Number of total quantity values in line items in this message is 150 units.
UNT+16+000000001'	Number of total segments in this message is 16 segments and the message reference number is 000000001.

Figure 2.2: An example of EDIFACT order message

composite data element *C002 DOCUMENT/MESSAGE NAME*. In this position there is only one value which is “220”. In this case, the value “220” belongs to the data element *1001 Document/message name, coded*. By consulting its available code list of this data element, the value “220” refers to *Order* message (cf. Fig. 2.1, Mark 4). In this example, the remaining data elements (i.e., *1131 Code list qualifier*, and *3055 Code list responsible agency, coded*) are not used. Next position is the data element *1004 Document/message number*, therefore the value “0123456789” refers to the identification number of this order. In the position of the *1225 Message function, coded*, the value “9” refers to *Original* which means this order message is the initial message of a given transaction.

Figure 2.1, Mark 8 shows another example of the *DTM* segment: DTM+137:20120101:102’. *DTM* indicates the date/time segment. As mentioned earlier, the *DTM* segment is composed of a composite data element *C504 DATE/TIME/PERIOD*. In this example, “137:20120101:102” is the value that belongs to this composite data element. Separated by the colon, the first value “137” refers to document/message date/time (cf. Fig. 2.1, Mark 6). The second value “20120101” refers to date/time value. The last value “102” refers to date/time format of “CCYYMMDD”. As explained earlier, the data element *2380 Date/time/period* is qualified by the *2005 Date/time/period qualifier* and the *2379 Date/time/period format qualifier*. Therefore, date/time data provided in this example can be interpreted that this document/message is issued on 1st of January, 2012. An extended example of an order message according to the EDIFACT standard as well as an interpretation of the message is provided in Figure 2.2.

2.1.2 Conceptualizing EDI Data

As described earlier, EDI data is structured and typically represented as coded data for enabling automated data processing. However, in order to allow the use of data for analysis tasks it is necessary that business information contained in EDI messages must be identified and extracted. Since EDI data is designed to be machine processable, it is difficult for humans to understand and interpret EDI messages directly. Recently, works on conceptualizing EDI messages have been developed for interpreting and representing EDI data on a conceptual level. The purpose of conceptualizing EDI data is to interpret and present EDI messages (i) in a way that humans can understand, as well as to (ii) maintain the structure of EDI data to automate machine processing. In the following, we provide the background on conceptualizing EDI messages as well as discuss its benefits and limitations.

Traditionally, EDI systems interpret EDI data by using hard-coded interpretation logic or mapping techniques [BHM07; DW09]. Those techniques require the knowledge of EDI structures and the available codes for generating mappings between EDI data and business information accurately. In the current state-of-the-art, there exist approaches for conceptualizing EDI data for representing the data at a conceptual level. In other words, these approaches aim at semantically abstracting EDI data as generic business information which is human-understandable. Approaches to conceptualize EDI standards include the works conducted as part of the Tripcom project[3] [FB05; FB06]. The Tripcom project provides an ontological infrastructure for business processes and business data. The purpose of the provided infrastructure is to enable B2B business process integration necessary due to the heterogeneity in the different EDI formats. An approach for ontologizing EDI applied in the Tripcom project is presented in [FB05; FB06]. The approach is based on manually defined semantic templates serving as a basis for deriving syntax and semantics from EDI standard specifications [Eng+12b]. In particular, the Tripcom project utilizes the approach for ontologizing UN/EDIFACT standards in both, syntax and semantics, as a basis for B2B processes integration in Web Services environments. Regarding syntax, Tripcom conceptualizes data elements and data segments specified in the UN/EDIFACT standard. In terms of semantics Tripcom focuses on conceptualizing message type structure. Moreover, the purpose of the ontological infrastructure of the Tripcom project is to support the communication for web services for integrating B2B processes. Technically, the ontologies are realized by using the Web Service Modeling Language (WSML)[4].

In addition to the Tripcom project, EDI ontologies have been developed in the course of the EDImine project with the aim of ontologizing EDI standards for different purposes, such as identifying redundant business information or business performance analysis. From a technical perspective, unlike the Tripcom project, the approach for ontologizing EDI presented in the EDImine project builds upon the Web Ontology Language (OWL)[5] [Eng+12b]. An overview of an approach for ontologizing EDIFACT as well as EDIFACT ontologies are depicted in Figure 2.3. EDIFACT ontologies consist of (i) the EDIFACT Standards Ontology (cf. Fig. 2.3, Mark 1), (ii) the EDIFACT Message Ontology (cf. Fig. 2.3, Mark 2), (iii) the EDIFACT Message Types Knowledge Base (cf. Fig. 2.3, Mark 3), and (iv) the EDIFACT Messages Knowledge Base (cf. Fig. 2.3, Mark 4). The EDIFACT Standards Ontology and the Message Ontology are manually designed. The EDIFACT Message Types and the Messages Knowledge Bases (KB) are automatically generated by storing EDIFACT messages and their contained data elements according to the aforementioned standard and message ontologies.

The EDIFACT Standards Ontology conceptualizes the meta-structure of UN/EDIFACT standards. It describes core concepts that individual EDIFACT standards are built upon includ-

[3] `http://tripcom.org/ontologies` (visited March 01, 2014)
[4] `http://www.w3.org/Submission/WSML/` (visited March 01, 2014)
[5] `http://www.w3.org/TR/owl2-overview/` (visited March 01, 2014)

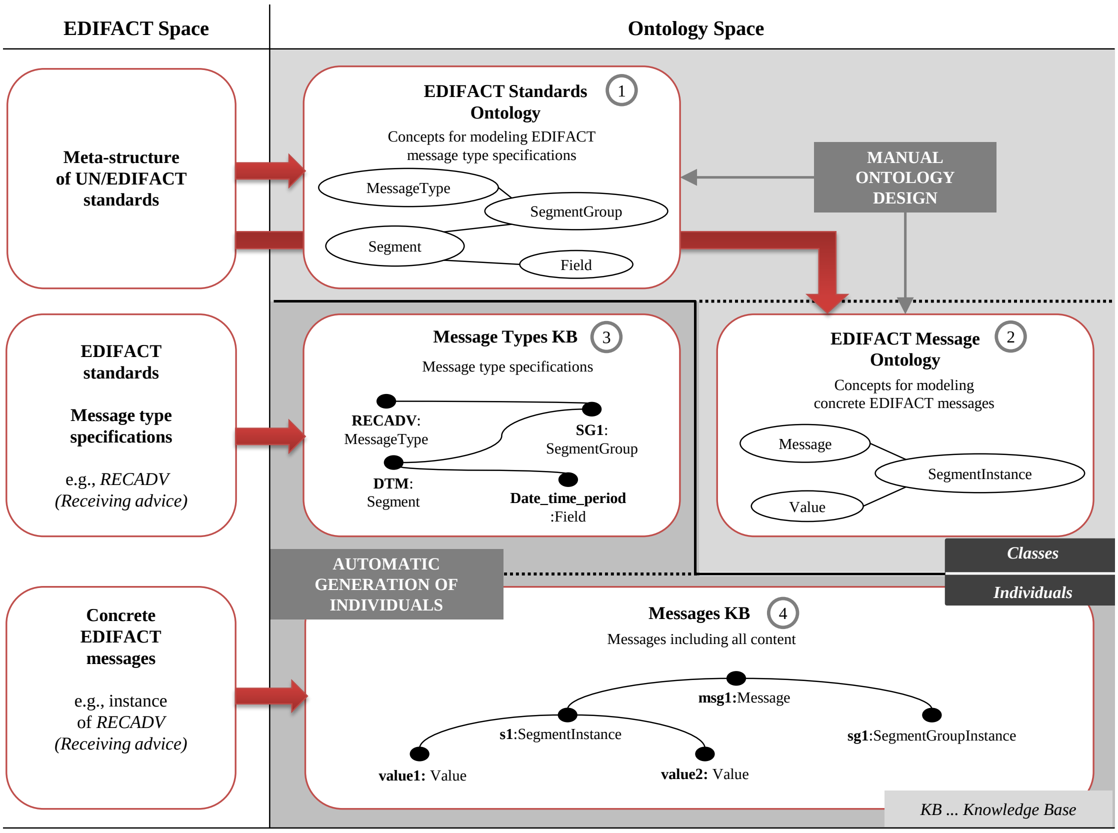


Figure 2.3: Overview of the approach for ontologizing EDIFACT [Eng+12b]

ing message type (`MessageType`), segment groups (`SegmentGroup`), segment (`Segment`), data fields (`Field`) which refers to composite data elements or data elements contained in a segment, components of fields (`Component`) which refers to data elements contained in a composite data element. Moreover, the concepts in EDIFACT standards also include code lists (`CodeList`) and their available codes (`Code`). These concepts are organized as a hierarchical structure defined in the standards. In the ontology they are connected by object properties such as `hasSegmentGroup`, `hasSegment`, `hasField`, etc. as described in Figure 2.4, Mark 1. Furthermore, the ontology is also designed to deal with *qualification relationships*. As described earlier, segments and data elements may be qualified for specifying their exact meaning. The qualification relationships can be classified into: semantic qualification and format qualification as previously described in the example of the *DTM* (Date/time/period) segment (cf. Fig. 2.1, Mark 5). These qualification relationships are represented by object properties `qualifies` which are further classified into: the `qualifies_semantics_of` and the `qualifies_format_of` property.

The EDIFACT Message Ontology describes the concepts used for the composition of an individual EDIFACT message. These concepts include messages (`Message`), segment group instances (`SegmentGroupInstance`), segment instances (`SegmentInstance`) and concrete

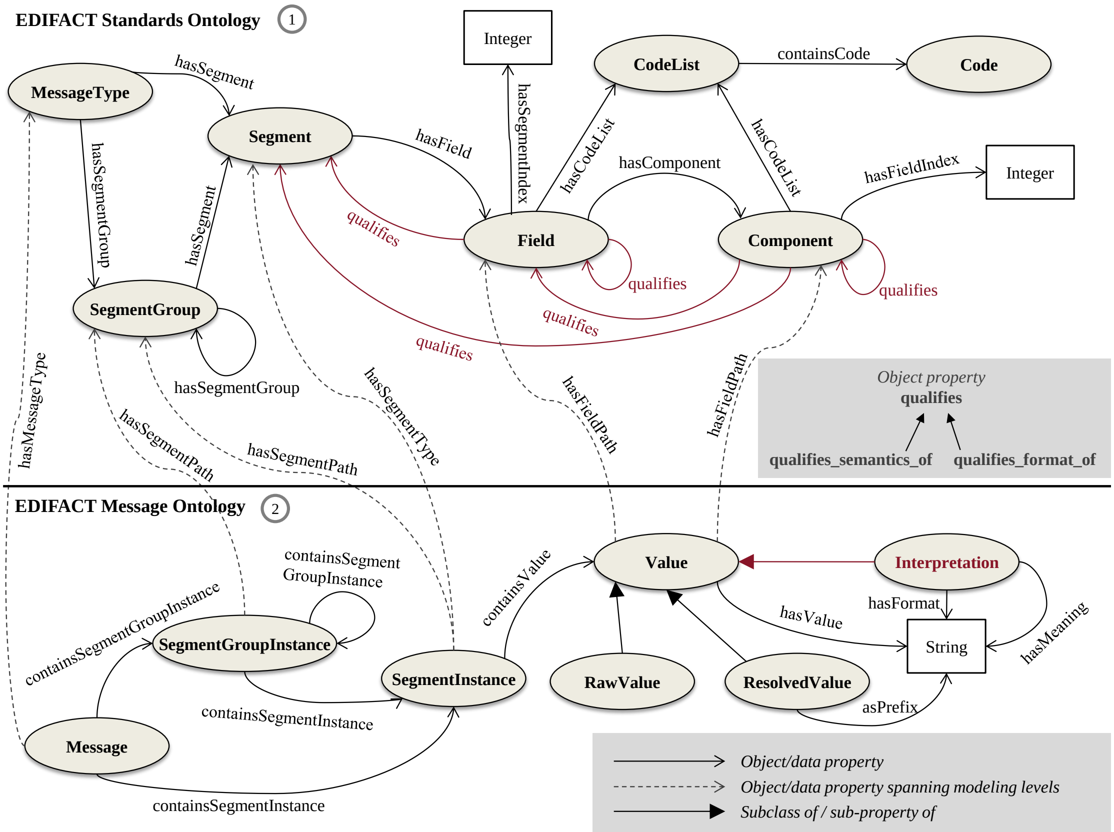


Figure 2.4: EDIFACT Standards Ontology and EDIFACT Message Ontology [Eng+12b]

values (`Value`). Values are divided into: raw values (`RawValue`), resolved values (`Resolved-Value`), and interpretations (`Interpretation`). Raw values represent values contained in EDIFACT messages without considering additional qualifiers and codes. Resolved values represent the values resolved from coded value, whereas interpretations represent the values interpreted from qualified values by considering qualification relationships. Segment group instances, segment instances, and values are connected to the structure specifications defined in the EDIFACT Standards Ontology by the object property `hasSegmentPath` and `hasFieldPath` as illustrated in Figure 2.4, Mark 2.

The Message Type Knowledge Base defines concrete specifications of individual EDIFACT message types as individuals of the concepts in the EDIFACT Standards Ontology (cf. Fig. 2.4, Mark 1). By parsing message type definitions from the official UN/EDIFACT directories, these individuals of class `MessageType` and its contained segment groups, segments, fields, and components are instantiated as individuals of their corresponding concepts automatically. Figure 2.5, Mark 1 shows an example of a message type RECADV (Receiving advice) along with its contained *DTM* segment. In the example, `RECADV` is instantiated as an individual of the concept `MessageType`. The `RECADV` has `Date_time_period` (i.e., the segment *DTM*) segment which consists of the field `DATE_TIME_PERIOD` (i.e., the composite data element *C504*

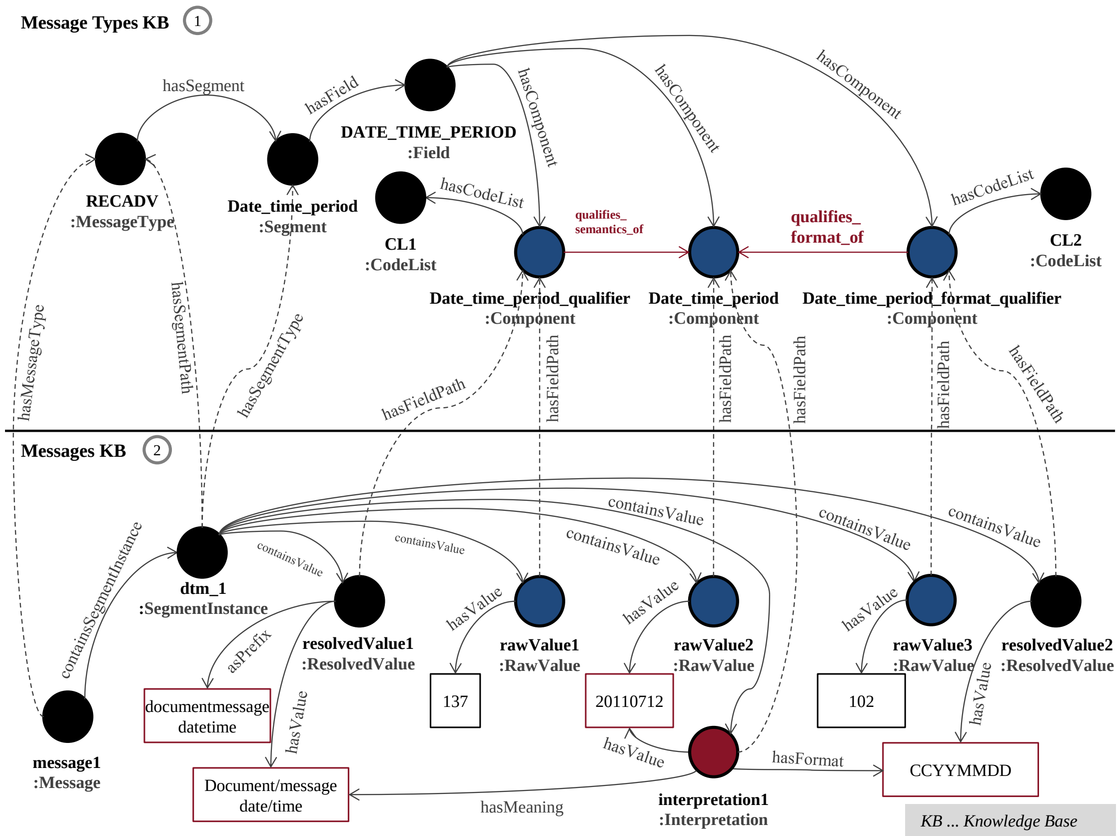


Figure 2.5: Sample knowledge base (excerpt) [Eng+12b]

DATE/TIME/PERIOD) composed of three components: `Date_time_period_qualifier`, `Date_time_period`, and `Date_time_period_format_qualifier`. Furthermore, it also describes the qualification relationships between components, fields, or segments through object properties: `qualifies_semantics_of` and `qualifies_format_of` as defined in the EDIFACT Standards Ontology.

The Message Knowledge Base stores concrete EDIFACT messages and their contained values. Each concrete EDIFACT message is instantiated as an individual under the concept `Message` described in the EDIFACT Message Ontology (cf. Fig. 2.4, Mark 2). All values contained in the message are parsed and stored as individuals of the concept `Value`. In addition, coded values are resolved and stored as type `ResovedValue`. Similarly, qualified values are interpreted and stored as the type `Interpretation` by considering the qualification relationships specified in the Message Type KB (cf. Fig. 2.5, Mark 1). Figure 2.5, Mark 2 illustrates a sample message knowledge base. The example shows a RECADV message represented as an individual `message1` of the `Message` concept. `message1` connects to `RECADV` message type by `hasMesageType` property. The `message1` contains segment instance `dtm_1` which is of type `Date_time_period` segment. The segment instance `dtm_1` contains three raw values which are the values of the following components: `rawValue1` (i.e., `Date_time_period_quali-`

fier having value “137”), rawValue2 (i.e., Date_time_period having value “20110712”), and rawValue3 (i.e., Date_time_period_format_qualifier having value “CCYYMMDD”). Date_time_period_qualifier and Date_time_period_format_qualifier are coded values after resolving their concrete value by consulting the associated code list (cf. Fig. 2.5, Mark 1) their values are stored as resolvedValue1 and resolvedValue2 respectively. Furthermore, according to the Message Type KB (cf. Fig. 2.5, Mark 1) there are qualification relationships between the components in Date_time_period segment. Therefore, the qualified date/time value (i.e., rawValue2) is interpreted and stored as interpretation1 with the attached meaning and format information.

The approach for ontologizing EDI makes the implicit semantic relationships between data elements in EDIFACT formats explicit. By modeling these relationships explicitly, it enables an accurate automatic interpretation of business data contained in EDIFACT messages. The research presented in this thesis builds upon these EDIFACT ontologies. In particular, EDIFACT ontologies are extended with the Business Information (BI) ontology where data elements are classified into generic business information concepts. This serves as a data infrastructure for supporting performance analysis tasks. The detailed discussion of the approach for generating BI ontology is provided in Chapter 3.

2.1.3 Benefits and Limitations of Conceptualizing EDI Data

The purpose of conceptualizing EDI is to represent EDI data on a conceptual level as well as to maintain its structure for further automated processing tasks. In other words, the benefits of conceptualizing EDI data are twofold. Firstly, EDI data is semantically interpreted and presented through human-understandable concepts. In the approach introduced in the Tripcom project [FB05; FB06], EDI data elements, codes and qualifiers are mapped to terms and relations that describe their meaning. Similarly, the approach for ontologizing EDI data developed in the course of the EDImine project also provides means for interpreting and storing EDI data as well as additional information about relationships between data elements [Eng+12b]. Secondly, EDI data can be organized and structured in an ontology which allows automated data processing. For example, EDI ontologies, developed from both the Tripcom project and the EDImine project, also describe data structures based on EDI standards (e.g., structure of message types, segments, composite data elements, etc.). Since EDI data is still structured in ontologies, it remains machine processable and in turn enables further automation.

Nevertheless, semantic templates, such as concepts and relations, provided in the EDI ontologies of Tripcom are manually defined [Eng+12b]. Although the approach in the EDImine project for ontologizing EDI data can automatically generate EDI ontologies, the interpretation of EDI data is still limited for coded and qualified data elements. In particular, the approach can automatically interpret coded and qualified data. However, the interpretation as generic terms or specific domain concepts is not yet considered. To address this shortcoming, we introduce the business information extraction approach (cf. Chapter 3) that extends the EDI ontologies with a BI ontology that contains generic business information concepts and their mapping to concrete EDI data.

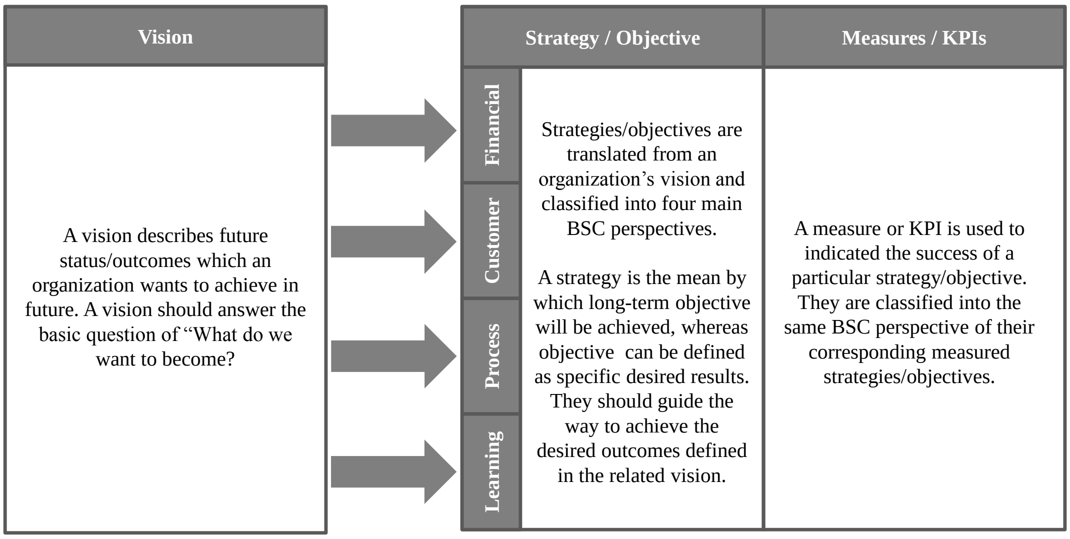


Figure 2.6: Balanced scorecard method (definitions of vision, strategy and business objective are from [Dav99, pp. 43,13])

2.2 Balanced Scorecard

The concept of Balanced Scorecard (BSC) has been introduced by Kaplan and Nortan in [KN92]. Originally, it has been designed as a method for analyzing and monitoring an organizations' performance by aligning business goals with measurements. This includes sets of measures, or so-called *Key Performance Indicators (KPIs)*, reflecting financial performance as well as a complementary set of operational measures that further drive financial measures. These operational measures are grounded on customer satisfaction, internal processes, and organizations' innovation and improvement capabilities. This yields a balanced presentation of both, financial and non-financial measures. Thereby, organizations' managers are provided with the information in many aspects of business performance supporting their business decisions.

2.2.1 Background and Foundations

As depicted in Figure 2.6, the BSC method is a top-down approach for performance monitoring and analysis. The method starts considering from the level of an organization's vision/strategy to the lowest level of operations [Eck06]. The implementation of BSC includes the steps from designing a scorecard until learning and adapting the result for appropriately adjusting business strategies [KN04; Eck06]. In particular, it begins with the step of designing a scorecard which links measurements to strategies. This involves the activities such as translating business strategies into operational objectives and assigning appropriate measures for each objective. In the original BSC method, an organization's strategies and the corresponding assigned measurements are divided into four main perspectives:

1. **Financial Perspective.** The financial perspective focuses on financial performance. It contains measures related to the indicators of the financial status such as profitability, income, revenue growth, etc. The measurement of this perspective intends to indicate whether organizations achieve their bottom-line result. Thus, most of the measures applied in this perspective are lagging indicators which are used to indicate output of past activities.

2. **Customer Perspective.** This perspective includes measures that reflect factors that matter to customers. Its objective is to reflect an organization's performance from the eye of their customers. The examples of relevant measures are percent of sales, ranking of organizations, number of orders, etc. Moreover, it also includes measures about the quality of products/services since they have an influence on customer satisfaction.

3. **Process Perspective.** The measurement of the process perspectives concentrates on internal processes including product development, manufacturing process, delivery and service process. It covers any activities and processes which are required to meet customers' expectation. The measures should be derived from business processes which have an impact to customers' satisfaction such as on-time deliveries, inventory turnover, number of defects, response time, etc.

4. **Learning and Growth Perspective.** The measurement of this perspective focuses on internal resources which are crucial for driving the other three perspectives such as employees and technologies. Since the ability to learn, innovate, and improve are important factors for an organization's competitiveness. Organizations having employees with those skills and supporting technologies are able to improve their processes/products, and launch new innovations. This may result in continuous improvement and growth of organizations. Therefore, it is important to monitor and analyze the performance related to this perspective. The example of relevant measures includes employee turn-over rate, training hours, number of new products, etc.

The scorecard is used for monitoring and analyzing organizations' business performance. The insights obtained from an analysis are further incorporated in future business planning for improving an organization's operations and strategies.

2.2.2 Balanced Scorecard in an Inter-organizational Context

Although the initial application of the BSC method was the performance analysis and strategic management within an organization, some studies about the applications of BSC in inter-organizational context have been developed. Most of those studies focus on employing the BSC method for supply chain performance analysis. For instance, Bullinger et al. [BKV02] proposed a measurement method for supply chain analysis based on the BSC. Bhagwat and Sharmar [BS07] developed the BSC method for supply chain management called balanced Supply Chain Management (SCM) scorecard. They conducted a review on performance metrics in the SCM domain and divided those metrics into four main perspectives of BSC forming a balanced SCM scorecard. Similarly, Kleijnen and Smits [KS03] and Chia et al. [CGH09] studied examples of KPIs commonly used for measuring supply chain performance following the BSC paradigm. In addition, Brewer and Speh [BS00] discussed the interrelationship between BSC and the SCM field and introduced approaches for supply chain performance analysis. Particularly, they presented the modified version of BSC for SCM performance framework. Four main perspectives of SCM performance have been developed which relate to the goals of SCM to customer satisfaction, firm financial performance, and the potential for an organization's learning and growth. Their proposed framework offers several benefits. The most important advantages are that it provides a "balanced" management approach for supply chain partners, and it helps managers focusing on achieving goals beyond performance measures within their organizations.

2.2.3 Benefits and Limitations of Balanced Scorecard Method

The BSC method enables performance monitoring and analyzing against business strategies across different perspectives. Monitoring performance in different perspectives provides a complete and comprehensive view of performance information. In particular, by monitoring financial measures managers can evaluate business success from the outcomes of their interested measures. In addition to financial measures, the BSC also provides managers an information of non-financial measures in different perspectives (i.e., customer, internal process, leaning and growth) that drive the success of financial outcomes. Thereby, managers can have a quick understanding of the relationships between outcomes and operational performance as well as can identify shortcomings for future improvements.

Furthermore, the alignment of measures and business strategies enables organizations to evaluate performance on a strategic level [KN04; KN96]. Organizations are not only able to assess the success of their strategies, but are also able to manage their strategies and make business decisions effectively based on the insights derived from the result of the BSC analysis. The presentation of BSC is like a cockpit where several sources of information are visualized simultaneously for supporting management and decision making. Due to the aforementioned advantages the BSC method is being widely applied. The BSC is not only applied as a performance analysis method, but it is also applied as a strategic management tool for organizations [Eck06; JBT07] in both, intra- and inter-organizational context.

However, the BSC method does not suggest how to identify measures or KPIs for measuring an organization's strategies. This makes the BSC method difficult to implement since business strategies are usually too wide and sometimes too ambiguous to identify appropriate KPIs reflecting those strategies. Regarding this difficulty, the BSC method best practices suggest aligning business strategies with KPIs based on critical success factors [KN04]. In other words, success factors are used for reflecting an organization's strategies. For example, Figure 2.7 shows a simplified example for applying the BSC method utilizing success factors. In the first column of Figure 2.7, the four different perspectives of the BSC method are shown. The subsequent three columns show business objectives, critical success factors, as well as KPIs. The KPIs are used for measuring success factors, which in turn are used for measuring business objectives. Considering KPIs from success factors is easier than considering KPIs from business objectives, since success factors focus on particular aspects which are more precise than business objectives.

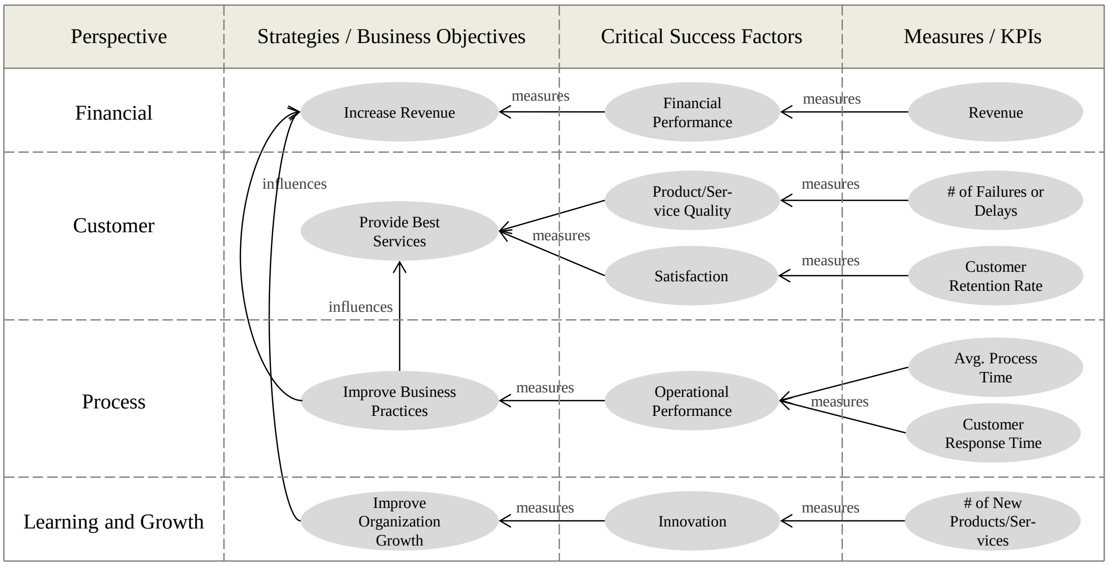
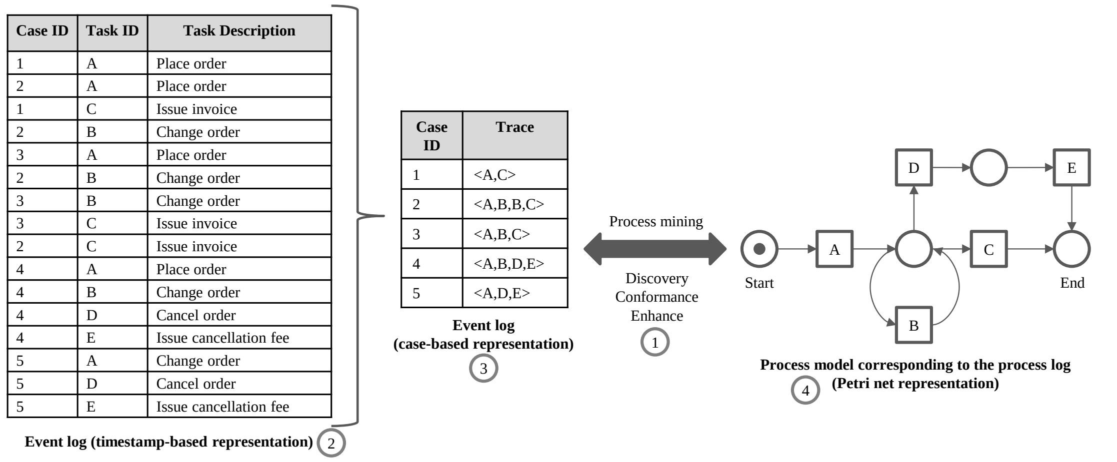


Figure 2.7: Example of balanced scorecard having the alignment between business objectives and KPIs through success factors

For example, in Figure 2.7, the customer perspective is comprised of the business objective "Provide Best Services". This business objective is linked to the success factors "Product/Service Quality" and "Satisfaction". The success factor "Product/Service Quality" focuses on the quality of products and services. Therefore, the KPIs concentrate on measuring product/service quality such as the number of failures or delays. The KPIs reflecting the success factor "Satisfaction" are derived in a similar manner. The other perspectives are measured similarly by defining proper business objectives, critical success factors, as well as KPIs.

2.3 Process Mining

Process mining is considered as a relatively young research area which aims at discovering, monitoring and improving real processes (derived from reality, not from ideal process models that are predefined) by extracting knowledge from event logs [Aal11]. Process mining lies between the two research areas of (i) data mining and machine learning, as well as (ii) process modeling and analysis. In particular, it extends the scope of data mining and machine learning for discovering knowledge from event logs in process perspectives (e.g., discovering process model from event log) which in turn fosters or/and enhances process modeling and analysis. In the following, we provide the foundations of process mining as well as the necessary background that this thesis builds upon.



Case ID	Task ID	Task Description
1	A	Place order
2	A	Place order
1	C	Issue invoice
2	B	Change order
3	A	Place order
2	B	Change order
3	B	Change order
3	C	Issue invoice
2	C	Issue invoice
4	A	Place order
4	B	Change order
4	D	Cancel order
4	E	Issue cancellation fee
5	A	Change order
5	D	Cancel order
5	E	Issue cancellation fee

Case ID	Trace
1	<A,C>
2	<A,B,B,C>
3	<A,B,C>
4	<A,B,D,E>
5	<A,D,E>

Figure 2.8: Process mining concept [AW04; Aal11]

2.3.1 Background and Foundations

The main idea of process mining is to investigate and derive process-related knowledge from low-level event logs of (business) process activities. Figure 2.8 illustrates underlying concepts of process mining. In particular, process mining is divided into three types: discovery, conformance, and enhancement (cf. Fig. 2.8, Mark 1). The objective of discovery is to derive process models from event logs without a-priori information [Aal11]. Event logs are an important ingredient for process mining. The process discovery takes event logs as input and derives the corresponding process model by investigating and analyzing event sequences within the log. Figure 2.8 illustrates an example of a log containing five cases (cf. Fig. 2.8, Mark 2) ordered by timestamps. In Figure 2.8, Mark 3 the log is represented in a case-based perspective. In the case-based representation, events are grouped according to a case identifier which results in *traces*, or *process instances*. The log shows that each process has been executed in a different sequence resulting in different cases. In the first case, task A and C are executed, while the next two cases (2 and 3) task B has been executed in between. In the last two cases (4 and 5), task D and E are executed instead of C. Based on the example, applying process mining techniques can discover a process model such as the model shown in Figure 2.8, Mark 4. Several algorithms have been developed for process model discovery including: α-algorithm [AWM04], genetic algorithm [AAW05], and heuristic algorithm [WAA06].

In addition to discovery, process mining for conformance checking focuses on comparing an existing process model with event logs of the same process to check if data in the log conforms to the model and vice versa. In other words, it measures the alignment between event logs and process models [RA08; Aal+08]. Conformance analysis helps to detect, locate, identify, and explain the deviation of unexpected process executions against a desired model. For instance, one of the applications of conformance analysis is to detect security violations as presented in [AA05]. Furthermore, *Log replay* is one of the conformance checking techniques that has be introduced recently [AAD12]. The idea is to replay event logs on top of process models. By replaying a log, discrepancies between the log and the model can be detected and quantified [Aal11].

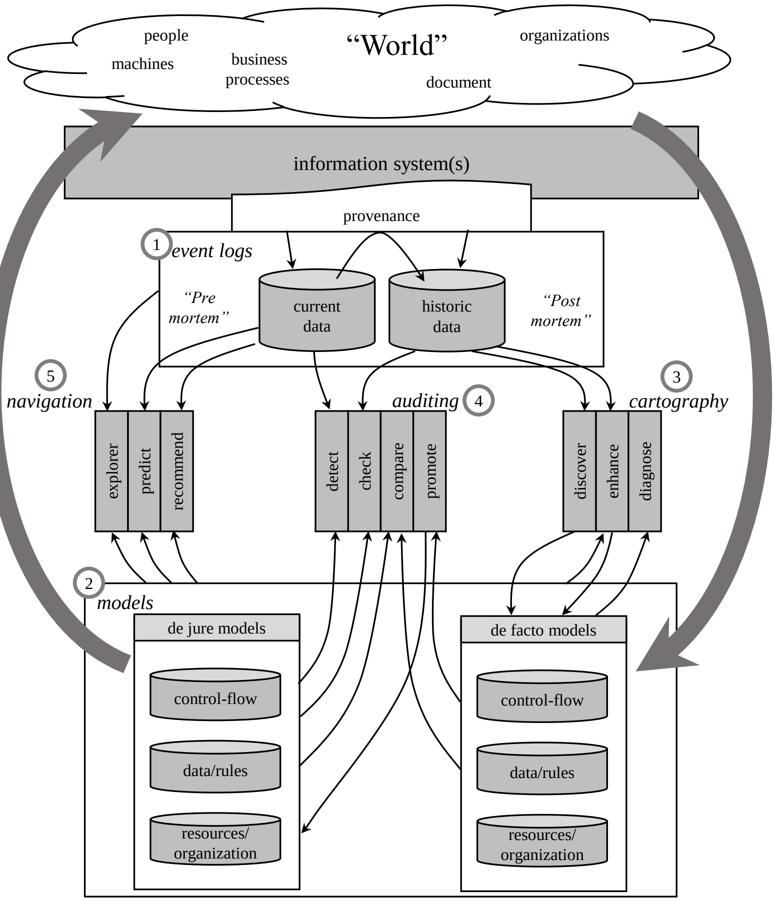


Figure 2.9: Process mining framework [Aal11, p. 242]

While conformance checking aims at verifying the alignment between a log and a model, enhancement aims at extending or improving an existing process model. Similarly, the enhancement is based on event logs and process models [Aal11]. Burattin et al. introduced an approach for discovering organizational roles [BSV13]. Their approach enhances a process model by means of extending the model with the information of organizational roles associated with activities. To *extend* (e.g., by adding new perspective to a model), *repair*, or *modify* a model, it can also be considered as one type of enhancement.

Figure 2.9 shows the process mining framework. In detail, the main input for process mining are event logs, which can be obtained from information systems. Event logs can be divided into "pre mortem" and "post mortem" logs (cf. Fig. 2.9, Mark 1). "Pre mortem" logs contain current cases that are still in progress and not yet finished, whereas "post mortem" logs contain past cases or historic data of cases that are already completed. Furthermore, models are also classified into two different types according to their purpose. These are "de jure models" and "de facto models" (cf. Fig. 2.9, Mark 2). The purpose of the former type is to specify how things are done. In contrast, the latter focuses on capturing reality.

The aforementioned types of process mining techniques can be applied on event logs for several activities such as discovering process models, detecting problems, deriving predictions, etc. These activities can be classified into three groups: cartography, auditing, and navigation. Activities in the group of cartography (cf. Fig. 2.9, Mark 3) concentrate on process models that are abstracted from reality (i.e., "de facto models"). Such activities include discovery, enhancement, and diagnosis. In addition, the auditing group (cf. Fig. 2.9, Mark 4) refers to activities that aim at checking whether the process execution in reality follows requirements, rules, restrictions, and other things defined by the "world" (e.g., managers, organizations, etc.). This type of activity typically uses both, "de jure models" and "de facto models" for comparing things that are expected to be done against reality. In contrast to the other two, the purpose of navigation (cf. Fig. 2.9, Mark 5) is to look into the future. Therefore, activities in this group tend to use the information from current data in order to predict the future of running cases or to recommend further actions.

2.3.2 Benefits and Limitations of Performance Analysis with Process Mining

In the area of performance analysis, process mining also plays an important role for model-based performance analysis. In particular, it introduces techniques that allow deriving models from reality following a bottom-up approach. These models may then be compared, monitored, and analyzed against the planned models. Consequently, the better models reflect reality, the more sense these kinds of analyzes make sense [Aal11]. Therefore, process mining aims at eliminating *a lack of alignment between hand-made models and reality* [Aal11, p. 57] by providing connections between models and event data directly coming from actual process execution.

The performance analysis through process mining centers around the process perspective. Currently, the KPIs focused in process mining are mostly concerned with time aspects such as throughput times of processes, waiting times, service times, synchronization times, etc. [AAD12; Hor07]. Moreover, other additional information (e.g., organizational roles or costs) can be taken into account as well. For example, models can be extended by organizational roles discovered from event logs through process mining techniques [Aal+12]. In addition, recent studies of process mining provide data-aware process mining techniques for enhancing models by discovering decision points based on business information [LA13]. Those enhanced models can be used for supporting performance analysis tasks. Nevertheless, best to our knowledge, currently the focus of KPIs in process mining are still centered around the time aspect. Furthermore, the connection between KPIs and business strategies for allowing performance analysis at strategic level is not yet considered.

CHAPTER 3

Business Information Extraction: An Ontological Approach

As introduced in Chapter 1, the ultimate research goal of this thesis is the evaluation of IORs based on EDI messages by means of inter-organizational performance analysis. Analysis tasks are usually performed by management people who aim at optimizing an organization's performance and its profit. Management people are typically aware of the information needed for analyzing IORs. For extracting business information from EDI messages, in-depth knowledge on EDI message structures is required. However, it is not always the case that management people are aware of the technical knowledge required for understanding EDI messages and extracting information from the same. Therefore, for management people, extracting business information from raw EDI messages becomes a complex, ambiguous, and error-prone task.

To close this gap, this chapter presents an ontology-based approach for extracting business information from EDI messages [Kra+12b]. The main idea of the approach is to link EDI data to business information concepts (cf. Fig. 1.6). In this thesis, we distinguish information into: business information and process information. Business information focuses on the content of business documents exchanged among business partners. In particular, we refer to business information as a certain kind of information that is required to be communicated for executing business transactions such as ordered quantities, order identification, delivery date, etc. In addition, process information focuses on processes of transactions which can reflect process efficiency. Process information includes process duration, activity timestamps, process model, etc. As mentioned earlier, our presented approach aims at extracting only business information from the content contained in EDI messages. The implementation of the approach focuses on UN/EDIFACT standards since, as mentioned earlier, they are currently still widely applied in B2B business scenarios [Ber94; VGS07]. In particular, the business information extraction approach presented in this chapter is one of our contributed artifacts addressing the first research question of extracting business information from EDIFACT messages (cf. Chapter 1.1.3). Having the extracted business information at hand supports management people in the evaluation of

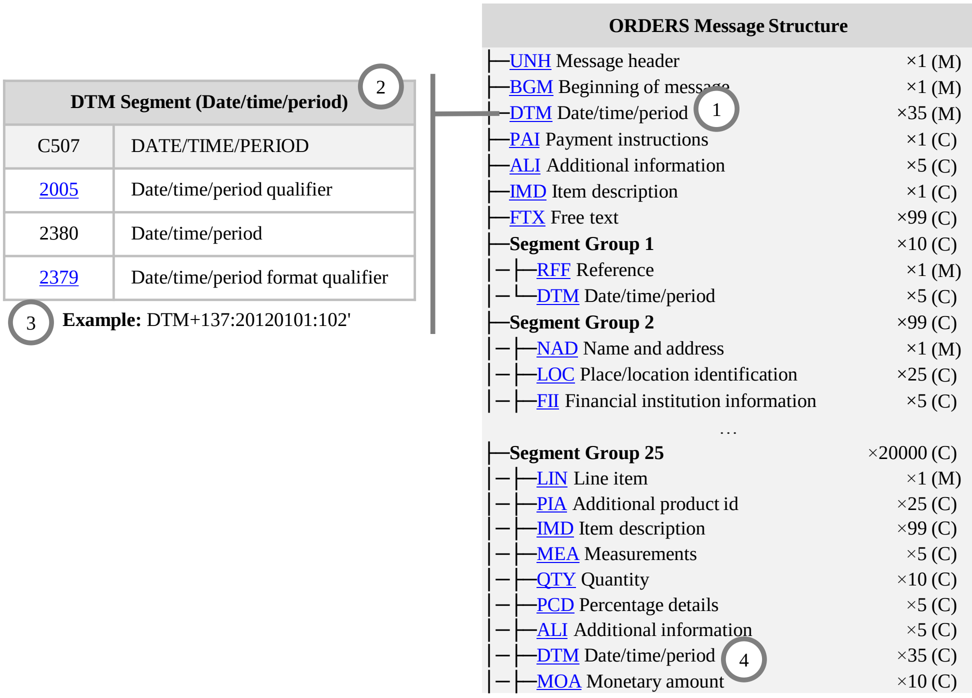


Figure 3.1: Excerpt of the purchase order EDIFACT message type (ORDERS) of D96A version

IORs. In the larger context of the framework presented in this thesis (cf. Chapter 6), the business information extraction approach is applied for semantic preprocessing of the raw EDI data.

3.1 Motivation

In the following, the ambiguous task of extracting information from EDIFACT messages, due to the complex structure of EDIFACT message types, is introduced and accompanied through an excerpt from the EDIFACT message type *Purchase order (ORDERS)* (cf. Fig. 3.1). Let's consider retrieving date/time information from the ORDERS message type. In particular, the date/-time information is located in the *DTM* segment at the beginning of the message (cf. Fig. 3.1, Mark 1). According to the EDIFACT standard, the *DTM* segment contains a date/time qualifier, a date/time value, and a date/time format (cf. Fig. 3.1, Mark 2). In other words, the date/time value is located at the second position of the *DTM* segment. Corresponding to the example illustrated in Figure 3.1, Mark 3, the actual date/time information of the message is “20120101”. Furthermore, as defined in the standard, the *DTM* segment also contains a qualifier having additional semantic meaning which is typically encoded through code lists. For interpreting such qualifiers, the corresponding code lists defined in the standard need to be interpreted. In the example shown in Figure 3.1, Mark 3, the qualifier “137” defines that the date/time segment represents the date/time information of the actual message. As illustrated in Figure 3.1, Mark 4, another *DTM* segment is located within segment group 25. However, due to the location of the segment within the message, the date/time information has a different semantic meaning. Consulting the EDIFACT standard reveals that the *DTM* segment within segment group 25 represents date/time information related to a specific product. This shows that the position of a segment is also relevant for the semantically correct interpretation of EDIFACT data. To summarize, the example demonstrates that for the correct interpretation of business information in EDIFACT messages requires in-depth knowledge of the EDIFACT standard.

For these reasons, we provide a business information extraction approach for EDIFACT messages which (i) is based on ontologies, (ii) provides business information (BI) concepts which in turn (iii) facilitates retrieving business information from EDIFACT messages. The idea is to generate BI concepts on top of EDIFACT data. In particular, we extend EDIFACT ontologies introduced by Engel et al. [Eng+12b] (cf. Chapter 2, Section 2.1.2) with BI concepts which can be used for classifying EDIFACT data. In other words, in our approach concrete EDIFACT data is derived in a bottom-up approach by utilizing ontologies. The extracted EDIFACT data is linked to BI concepts defined by domain experts in a top-down manner.

The remainder of the chapter is organized as follows: Section 3.2 provides an overview on related work. The business information extraction approach is described in Section 3.3. Section 3.4 provides a discussion on the advantages of the approach. Finally, the summary of this chapter is given in Section 3.5.

3.2 Related Work

EDIFACT message types describe the message structure and its semantics which includes the position of data elements within the overall message structure. For extracting business information from such messages there have been attempts based on mapping techniques [BHM07]. However, as shown earlier, the correct interpretation of EDIFACT messages also requires considering qualifiers and encoded values. Therefore, for the interpretation of EDIFACT messages the following three aspects must be considered: (i) the position of data elements, (ii) optional qualifiers of a certain data element, and (iii) coded values of data elements. This may be achieved by creating specific mappings and/or by creating dedicated querying statements (see, for instance, [DW09]). However, opposed to existing works, the approach presented in this chapter is based on EDIFACT ontologies for abstracting from dedicated mappings or specific querying statements.

The introduced approach builds upon EDIFACT ontologies [Eng+12b] where EDIFACT standards, message types, message instances and their contained values are described in ontologies and corresponding knowledge bases. The existing ontologies provide means for the semantically enriched interpretation of values by considering qualifiers and codes. In particular, coded values are resolved and stored as resolved values. Similarly, qualified values are interpreted by resolving their qualifiers and storing them as so-called *interpretations*. While EDIFACT ontologies already provide means for accurate semantic interpretation of isolated coded and qualified values, semantic variances of data elements determined by their position in the overall message structure are not considered. In the work presented in this chapter, we extend EDIFACT ontologies by linking the positions of individual data elements within the overall message structure to corresponding BI concepts. Furthermore, having the ontological approach at hand allows as-

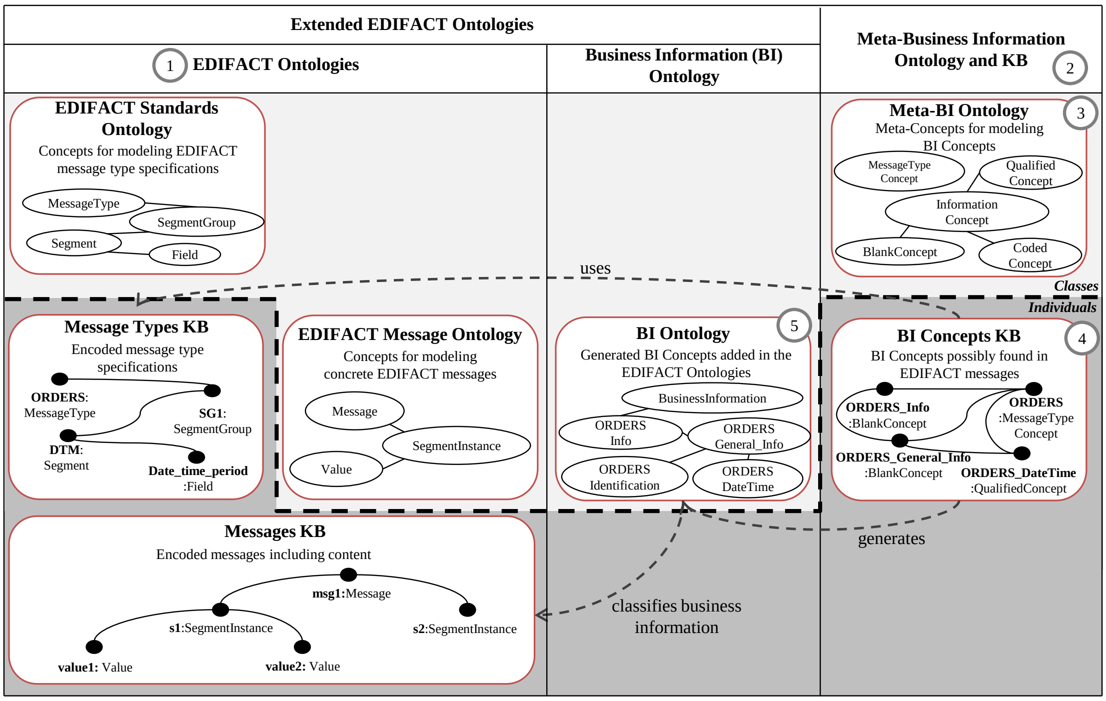


Figure 3.2: Overview on ontology architecture of the business information extraction approach

signing relationships between different concepts, such as hierarchical orders among BI concepts, which further supports the inferences required for querying tasks.

3.3 Business Information Extraction Approach

In the following, the ontology architecture of our business information extraction approach (cf. Fig. 3.2) is presented which builds upon our existing works on EDIFACT ontologies (cf. Fig. 3.2, Mark 1) [Eng+12b]. The basic elements of EDIFACT standards are represented as an ontology, namely the EDIFACT Standards Ontology. For example, the ontology specifies that each `SegmentGroup` contains `Segments`, each `Segment` has `Fields` and composite `Fields` may contain `Components`. Concrete specifications of individual message types are defined based on the original EDIFACT standards, referred to as the Message Types Knowledge Base (KB). Furthermore, the EDIFACT Message Ontology contains necessary concepts for representing individual EDIFACT messages and their content. The concepts of the ontology include messages (`Message`), segment group instances (`SegmentGroupInstance`), segment instances (`SegmentInstance`) and concrete values (`Value`). `Values` are further classified into `RawValues`, `ResolvedValues` and `Interpretations`. `RawValues` represent the values in EDIFACT messages which are interpreted without considering additional qualifiers and codes. Furthermore, some values are coded. In EDIFACT ontologies, these values are resolved (i.e., the meaning of the code is looked up in the corresponding code list) and stored as

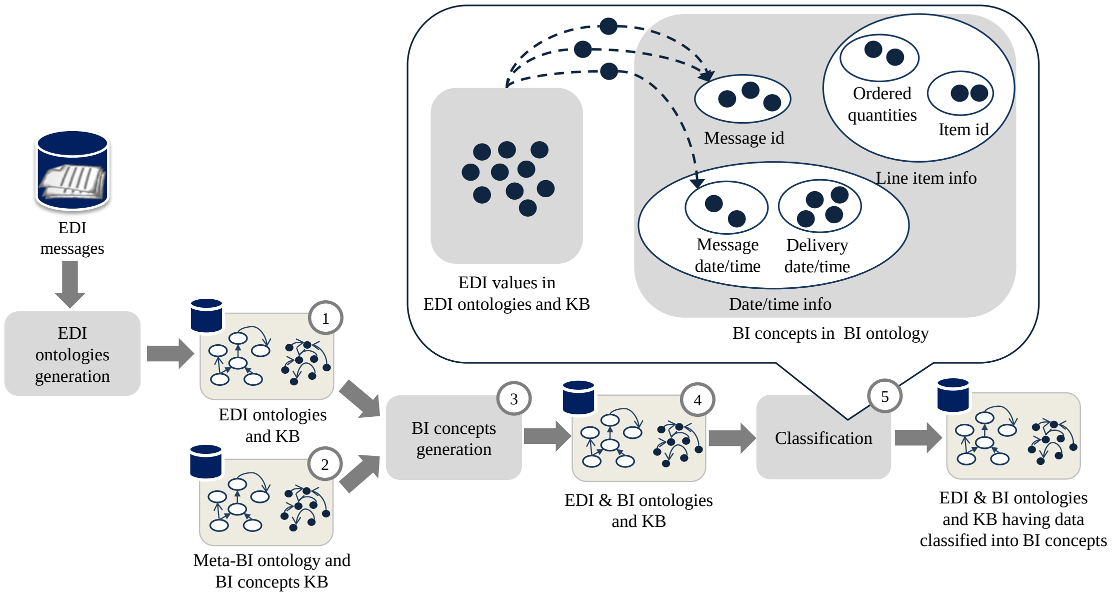


Figure 3.3: The business information extraction approach

`ResolvedValues`. Moreover, some values are qualified. The qualifiers add additional meaning to values. In EDIFACT ontologies, this additional meaning of qualified values is captured in `Interpretations`. Concrete EDIFACT messages and values are parsed and stored in knowledge bases (Messages KB) as individuals of the aforementioned concepts.

As mentioned earlier, the semantically correct interpretation of data requires considering the position of data elements as well. However, currently, the interpretations in EDIFACT ontologies only consider information on qualifiers and codes. To close this gap, we define the Meta-BI ontology (cf. Fig. 3.2, Mark 3) which defines an ontological schema for BI concepts. Concrete BI concepts are defined as instances of the Meta-BI ontology in the BI Concepts KB (cf. Fig. 3.2, Mark 4). The business information extraction approach generates mappings between BI concepts and the corresponding paths to data elements in individual message types. Such mappings are represented in a separate, automatically generated ontological schema, the BI ontology (cf. Fig. 3.2, Mark 5). The concepts in this ontology are defined by means of *class equivalent expressions* that reflect the mapping rules between data elements in individual message types and BI concepts.

Figure 3.3 depicts the business information extraction approach presented in this thesis. Please note that the symbol of ellipses and filled circles, appearing in figures presented in this thesis, represent concepts in an ontology and individuals (i.e., instances of concepts) in a knowledge base respectively. For instance, EDI/EDIFACT ontologies and KB in Figure 3.3 Mark 1 is represented by both, ellipses and filled circles. This is because ellipses represent concepts in EDI/EDIFACT ontologies, whereas filled circles represent instances in the EDI knowledge base.

The approach relies on (i) the existing EDI/EDIFACT ontologies containing concepts of EDIFACT message types and the parsed EDIFACT data (cf. Fig. 3.3, Mark 1) and (ii) the Meta-BI ontology together with the BI Concept KB (cf. Fig. 3.3, Mark 2). The EDI/EDIFACT ontologies are obtained by applying the ontologizing approach introduced in [Eng+12b] and as described in Chapter 2, Section 2.1.2. The Meta-BI ontology and the BI Concepts KB are created manually. In particular, the Meta-BI ontology is designed for describing types of BI concepts (e.g., qualified concepts, coded concepts, valued concepts, etc.) as well as mapping concepts at a meta level. The BI Concepts KB contains BI concepts which are manually defined as concrete instances of the Meta-BI ontology. Those BI concepts are defined based on business information found in EDIFACT message types. Furthermore, they contain the mapping information (i.e., corresponding positions of concrete EDIFACT data elements) for the BI ontology generation. Section 3.3.1 provides a detailed explanation on the Meta-BI ontology and the BI Concepts KB.

According to the definitions indicated in the BI Concepts KB, the BI concepts generation mechanism automatically generates BI concepts and their class equivalent expressions (i.e., mapping rules for data classification) in the BI ontology on top of the EDIFACT ontologies. This results in the extended EDIFACT ontologies (cf. Fig. 3.3, Mark 4). The details on the BI concepts generation are provided in Section 3.3.2. Finally, actual values from EDIFACT messages are classified into their corresponding BI concepts by reasoning over the extended ontologies, as illustrated in Figure 3.3, Mark 5. The classification step is elaborated on in Section 3.3.3.

3.3.1 Meta-BI Ontology and BI Concepts Knowledge Base

Figure 3.4, Mark 1, illustrates the Meta-BI ontology in detail. At its core, the Meta-BI ontology contains the concepts `InformationConcept` and `MessageConcept`. The `MessageConcept` refers to EDIFACT message types which contain `InformationConcepts`. `InformationConcepts` are either (i) `DirectConcepts`, (ii) `QualifiedConcepts`, or (iii) `CodedConcepts`. `DirectConcepts` are further classified into `BlankConcept` and `ValuedConcept`. For interpreting data in EDIFACT messages, the following aspects need to be considered. First, all BI concepts whose data elements can be interpreted directly through their positions are classified as `ValuedConcepts`. Second, BI concepts describing qualified data elements are represented as `QualifiedConcepts`. Third, BI concepts representing coded data elements are classified as `CodedConcepts`. Fourth, `BlankConcepts` are used for grouping all other concepts. The `InformationConcept` may have several mappings to specific data located in EDIFACT messages in different versions. This is achieved by associating the `Mapping` concept to the `InformationConcept`. Thereby, several sub-concept of `Mapping` can be assigned in order to group concrete mappings into a particular UN/EDIFACT version.

Based on the Meta-BI ontology, concrete BI Concepts are modeled manually for individual message types. Depending on the type of BI concepts (i.e. `BlankConcept`, `ValuedConcept`, `QualifiedConcept` and `CodedConcept`), the concepts refer to either `RawValues`, `Interpretaions` or `ResolvedValues` of EDIFACT ontologies (cf. Fig. 3.2, Mark 1). In order to allow the mapping of BI concepts to data elements, BI concepts are linked to individual mappings by using object property named `mappingOf`. Such mappings contain corresponding position information through data properties named `fieldPath` and `segmentPath`.

An example of a BI Concepts KB is illustrated in Figure 3.4, Mark 2. A `BlankConcept` named `ORDERS_Info` (cf. Fig. 3.4, Mark 3). It is an empty individual without any data properties which is used for grouping related BI concepts contained in purchase order (ORDERS) messages as a hierarchical structure. In addition, another `BlankConcept` named `ORDERS_Docu-`

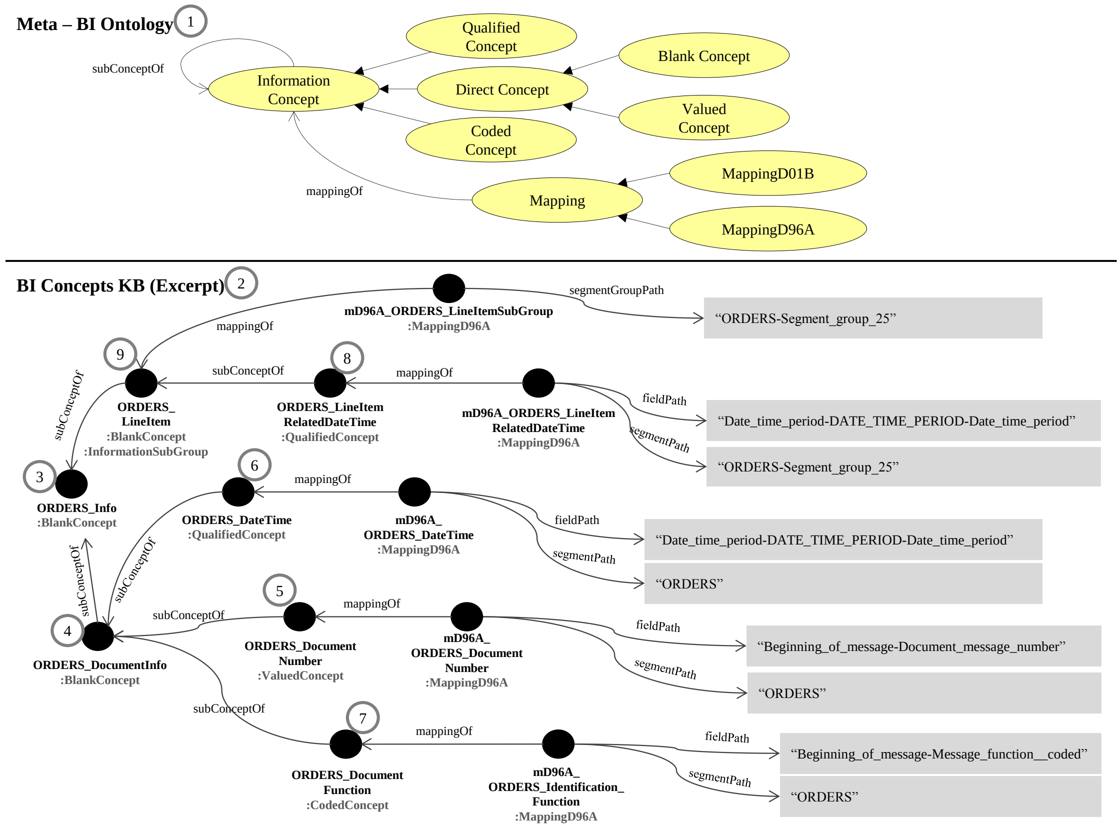


Figure 3.4: Meta-BI ontology and BI concepts knowledge base

`mentInfo` is defined (cf. Fig. 3.4, Mark 4). Other BI concepts may be linked to this `BlankConcept` via the object property `subConceptOf` resulting into one single concept aggregating all message identification information. Figure 3.4, Mark 5 shows the example of a `ValuedConcept` named `ORDERS_DocumentNumber` which corresponds to the data element located in ORDERS messages in the *BGM* (Beginning of Message) segment (cf. ORDERS message structure in Fig. 3.1). This is expressed by connecting the `ORDERS_DocumentNumber` to the mapping `mD96A_ORDERS_DocumentNumber` through the object property `mappingOf`. More precisely, the mapping `mD96A_ORDERS_DocumentNumber` contains the location of document number appeared in an ORDERS message type of D96A version. The information of location in this mapping is stored using the `fieldPath` and `segmentPath` data properties. Similarly, this information is stored in the same way for the mappings of qualified BI concepts and coded BI concepts. The example for modeling qualified and coded BI concepts is illustrated in Figure 3.4, Mark 6, namely the `ORDERS_DateTime` concept, and in Figure 3.4, Mark 7, the `ORDERS_DocumentFunction` concept.

Considering the example on date/time information introduced earlier in Figure 3.1, Mark 1 and 4, the *DTM* segments are represented in the BI Concepts KB in the following manner. For the `ORDERS_DateTime` (cf. Fig. 3.4, Mark 6), related to the whole message, the

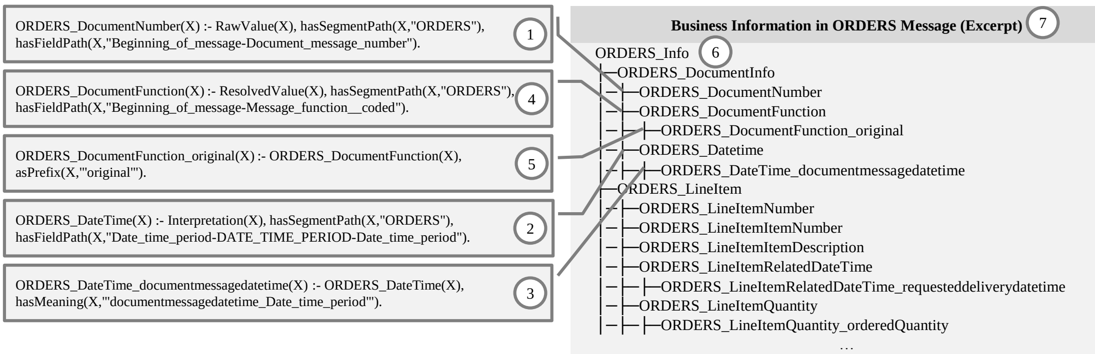


Figure 3.5: Example of extracted business information concepts

corresponding mapping `mD96A_ORDERS_DateTime` leads to the path of "ORDERS". For the `ORDERS_LineItemRelatedDateTime` (cf. Fig. 3.4, Mark 8), related to a specific line item, the path is indicated to "ORDERS-Segment_group_25" through the mapping `mD96A_ORDERS_LineItemRelatedDateTime`.

However, according to the UN/EDIFACT standard segment groups may contain several segments as well as nested segment groups. For example, the ORDERS message type may contain several line items. Each line item has its related information such as the ordered product, the ordered quantities, the monetary amount, etc. In case of querying information within the same line item (e.g., querying for ordered quantities of product A), all line items containing related information must be identified. Considering the query of ordered quantities of product A, the query mechanism must identify line items referring to product A. Searching within the same line item, the query mechanism must be able to retrieve the correct ordered quantities of product A. For the purpose of such queries, BI concepts can optionally be of type `InformationSubGroup`. For example, the `ORDERS_LineItem` (cf. Fig. 3.4, Mark 9) is not only a `BlankConcept` but also an `InformationSubGroup`. Its mapping `mD96A_ORDERS_LineItemSubGroup` points to the segment group 25 which is referred to as the group of line items (according to the UN/EDIFACT standard of ORDERS message type of the version D96A). This indicates that all data elements related to a particular line item are contained within the `ORDERS_LineItem` instance. In other words, all data elements under an instance of segment group 25 are related to each other under the same line item.

3.3.2 Generating BI Concepts

Based on the BI concepts KB, the BI concept generation mechanism automatically generates the BI ontology (cf. Fig. 3.2, Mark 5) and extends EDIFACT ontologies with concepts contained in the BI ontology (cf. Fig. 3.3, Mark 3 and 4). In terms of technical implementation, EDIFACT ontologies as well as Meta-BI ontology are represented in Web Ontology Language (OWL). Hence, the mechanism applies SPARQL queries for querying the BI Concepts KB. However, for optimizing reasoning performance, in our implementation, the resulting EDIFACT and BI ontologies are transformed to a set of Datalog [MSS05] rules. The generation mechanism is based on the following two steps which are detailed in the following: (i) generating concepts and (ii) generating hierarchies.

1. Generating Concepts. For each BI concept defined in the BI Concepts KB, corresponding concepts are generated in the BI ontology. The names of the generated concepts in the BI ontology mirror the names of the BI Concept individuals in the BI Concepts KB. Depending on the type of the BI concept, the mechanism for generating concepts and their class equivalent expressions in the BI ontology differs as follows:

ValuedConcepts. This type of BI concepts represents data elements in EDIFACT message types that contain uncoded and unqualified values and are mapped to corresponding data elements by means of Datalog [MSS05] rules representing class equivalent expressions. These class equivalent expressions reference `RawValues` and contain further constraints for the field path and segment path of the data elements to be referenced. The mechanism generates these class equivalent expressions by using the following template of Datalog rule:

```
Template #1
BIConceptName(X) :- RawValue(X),
                    hasSegmentPath(X,"segment_path"),
                    hasFieldPath(X,"field_path").
```

In Datalog syntax, a rule consists of two parts: a head and a body. The punctuation `:-` separates the head from its body. In the template, the head is `BIConceptName(X)` and the body is `RawValue(X), hasSegmentPath(X,"segment_path"), hasFieldPath(X,"field_path")`. A body can consists of several atoms which are separated by a comma (`,`). In the template, the body consists of three atoms: (i) `RawValue(X)`, (ii) `hasSegmentPath(X,"segment_path")`, and (iii) `hasFieldPath(X,"field_path")`. An atom represents either class of, or relation between its related variables or constants defined in a parenthesis. For example, the atom `RawValue(X)` specifies a class of a variable `X`. In other words, it specifies that the variable `X` is of type `RawValue`. In addition, the atom `hasSegmentPath(X,"segment_path")` represents the relation named `hasSegmentPath` between the variable `X` and the constant value of segment path `"segment_path"`. Particularly, it means the variable `X` has segment path pointed to `"segment_path"`. Please note that, in our presented templates, we use `X` as a variable representing a constant. Technically, at the time of classification, the variable will be substituted for a constant (i.e., individuals or literals contained in the EDIFACT ontologies) to check whether the constant satisfy the rule (i.e., class equivalent expression). The constant can be a specific individual contained in an ontology or can be a literal. In the template, the constant refers to an individual is presented between double quotes. The constant refers to literal is presented between triple quotes.

According to the syntax, the template for BI concepts of type `ValuedConcept` specifies that if `X` is an individual of type `RawValue`, `X` has a segment path to `"segment_path"`, and `X` has a field path to `"field_path"`, then `X` is also classified into the concept of `BIConceptName`. At the time of concept generation, the italic and underlined text, presented in the template, will be replaced with the name of BI concepts and their related segment path and field path accordingly. In particular, the *`BIConceptName`* will be replaced with the name of the generated BI concept. The *`segment_path`* and the *`field_path`* will be replaced with the related segment path and field path of the generated BI concept.

Concrete example of the ValuedConcept is provided in Figure 3.4, Mark 5 (i.e., ORDER_DocumentNumber). The ValuedConcept named ORDERS_DocumentNumber has the mapping pointed to segment path of "ORDERS" and field path of "Beginning_of_message-Document_message_number". According to this information, the mechanism generates the concept which yields an equally named concept in the BI ontology having a class equivalent expression generated based on the template #1 as follows:

```
ORDERS_DocumentNumber(X) :- RawValue(X),
hasSegmentPath(X,"ORDERS"),
hasFieldPath(X,"Beginning_of_message-Document_message_number").
```

The example shows class equivalent expression of the BI concept named ORDERS_Document Number. It is generated in the BI ontology (cf. Fig. 3.5, Mark 1) based on the aforementioned template (i.e., template #1). In particular, the *BIConceptName* is replaced by the name of BI concept which is ORDERS_DocumentNumber. The *segment_path* and the *field_path* are replaced by the segment path and field path of the BI concept ORDERS_DocumentNumber which are ORDERS and Beginning_of_message-Document_message_number respectively. In total, this class equivalent expression includes all individuals of type RawValue belonging to the Beginning_of_message-Document_message_number field which are positioned in the ORDERS segment path, indicating that the element is located at the root of the ORDERS messages (i.e., not in a specific segment group).

QualifiedConcepts. Class equivalent expressions for QualifiedConcepts are constructed analogously to ValuedConcepts, except that Interpretations rather than RawValues are selected. The rule template for generating class equivalent expression of the QualifiedConcept is as follows:

```
Template #2
BIConceptName(X) :- Interpretation(X),
                    hasSegmentPath(X,"segment_path"),
                    hasFieldPath(X,"field_path").
```

Furthermore, unlike ValuedConcepts, for QualifiedConcepts additional sub-concepts are generated. While the aforementioned class equivalent expressions select data elements qualified by *any* qualifying values, these additional sub-concepts are used to represent data elements qualified with *specific* qualifying values. Hence, the class equivalent expressions of the sub-concepts differ from the class equivalent expressions of their super-concepts by adding additional constraints on the meaning of a qualifier. Therefore, the rule template for generating class equivalent expressions of sub-concepts has to specify the meaning interpreted from a qualifier. This constraint is defined through the data property named hasMeaning provided in the EDIFACT ontologies (cf. Fig. 2.5, Mark 2 introduced in Chapter 2). The template is defined as follows:

```
Template #3
BIConceptName_meaning(X) :- BIConceptName(X),
                            hasMeaning(X,"'meaning_of_qualifier'").
```

As shown in the template, the names of sub-concepts are generated by combining the names of super-concepts and the specific meaning given by the corresponding qualifiers. The example of `QualifiedConcept` named `ORDERS_DateTime` from the BI concepts KB is depicted in Figure 3.4, Mark 6. Based on this BI concept, the super-concept named `ORDERS_DateTime` is generated in the BI ontology (cf. Fig. 3.5, Mark 2) with a class equivalent expression based on the template #2 as follows:

```
ORDERS_DateTime(X) :- Interpretation(X),
hasSegmentPath(X,"ORDERS"),
hasFieldPath(X,"Date_time_period-DATE_TIME_PERIOD-Date_time_period").
```

The generation of sub-concepts depends on the available qualifiers found in the EDIFACT ontologies and KB. In other words, the mechanism will generate sub-concepts only for the qualifiers that are in use. In the example, we assume that the qualifier giving a specific meaning of "documentMessageDateTime" is found. Hence, the mechanism generates a sub-concept called `ORDERS_DateTime_documentmessagedatetime` (cf. Fig. 3.5, Mark 3) with a class equivalent expression referencing the `ORDERS_DateTime` concept and having an additional meaning of "documentmessagedatetime_Date_time_period" as shown below:

```
ORDERS_DateTime_documentmessagedatetime(X) :- ORDERS_DateTime(X),
hasMeaning(X,"'documentmessagedatetime_Date_time_period'").
```

In particular, the mechanism generates sub-concepts by using the template #3 for generating sub-concept of `QualifiedConcept`. The name of sub-concept is the combination of the name of super-concept and the specific meaning of the corresponding qualifier. In this case, the name of sub-concept is `ORDERS_DateTime_documentmessagedatetime` which consists of the name of its super-concept (i.e., `ORDERS_DateTime`) and the meaning interpreted from a specific qualifier (i.e., `documentmessagedatetime`). The remaining body of the rule indicates that this sub-concept is an `ORDERS_DateTime` having a specific meaning of "documentmessagedatetime_Date_time_period".

CodedConcepts. Class equivalent expressions for `CodedConcepts` are constructed analogously to `ValuedConcepts`, except that `ResolvedValues` rather than `RawValues` are selected. Hence, these class equivalent expressions capture coded data elements with no constraints on the specific code value. The rule template for generating class equivalent expression of the `QualifiedConcept` is as follows:

```
Template #4
BIConceptName(X) :- ResolvedValue(X),
                    hasSegmentPath(X,"segment_path"),
                    hasFieldPath(X,"field_path").
```

Similar to `QualifiedConcepts`, additional sub-concepts with corresponding class equivalent expressions are generated to capture coded data elements having *specific* code values. The template defines the constraint of specific code value by using the data property named `asPrefix`. However, in the EDIFACT ontologies, the `ResolvedValue` has two data properties: `hasValue` and `asPrefix` (cf. Fig. 2.4, Mark 2 introduced in Chapter 2). We use the property `asPrefix` because of technical reason. In particular, both, `hasValue` and `asPrefix` property, contain similar values which are the values of specific code, but are stored in different formats (cf. Fig. 2.5, Mark 2 introduced in Chapter 2). The property `hasValue` stores the value of code by keeping the original text format. For example, it allows having upper cases and some special characters (e.g., slash (/), dot (.), comma (,), etc.) in the stored values. In contrast, the property `asPrefix` stores the value of code with lower cases and without any special characters. In other words, the format of data stored in the `asPrefix` property is clean. Hence, it can ensure that the template constructed by using the value from the property `asPrefix` will not be corrupted according to the Datalog syntax. Therefore, the template is defined as follows:

```
Template #5
BIConceptName_meaning(X) :- BIConceptName(X),
                            asPrefix(X,"'meaning_of_code'").
```

The example of `CodedConcept` is provided in Figure 3.4, Mark 7 (i.e., the `ORDERS_DocumentFunction`). Based on the BI concept `ORDERS_DocumentFunction` provided in the BI Concepts KB, the mechanism generates super-concept named `ORDERS_DocumentFunction` by using the template #4 (cf. Fig. 3.5, Mark 4). For generating the corresponding sub-concepts, the mechanism identifies codes which are used in the available data. Let's assume that the code referring to the meaning of "original" is found. Therefore, the mechanism generates sub-concept named `ORDERS_DocumentFunction_original` which, again, combines the name of super-concept and the meaning of the code. Furthermore, the body of the class equivalent expression is constructed according to the template #5 (cf. Fig. 3.5, Mark 5). The class equivalent expressions of the super-concept `ORDERS_DocumentFunction` and the corresponding sub-concept `ORDERS_DocumentFunction_original` are shown as follows:

```
ORDERS_DocumentFunction(X) :- ResolvedValue(X),
hasSegmentPath(X,"ORDERS"),
hasFieldPath(X,"Beginning_of_message-Message_function__coded").

ORDERS_DocumentFunction_original(X) :- ORDERS_DocumentFunction(X),
asPrefix(X,"'original'").
```

BlankConcepts. The `BlankConcepts` do not represent actual values but are used for grouping purposes. Thus, neither class equivalent expressions nor any sub-concepts are added to the generated concepts (cf. Fig. 3.5, Mark 6).

2. Generating Hierarchies. This step is identifying the hierarchical order between the generated BI concepts (cf. Fig. 3.5, Mark 7). Therefore, the mechanism first creates a `BusinessInformation` concept acting as a super concept for all previously generated concepts. Second, for each generated BI concept, the mechanism queries the BI concepts KB for finding the related `subConceptOf` property and generates a corresponding hierarchy in the BI ontology. For instance, the `ORDERS_DocumentFunction` is a `subConceptOf` `ORDERS_DocumentInfo` (cf. Fig. 3.4, Mark 7). Therefore, the mechanism defines `ORDERS_DocumentFunction` as a sub-concept of `ORDERS_DocumentInfo`. For all other BI concepts which do not have a `subConceptOf` property the mechanism defines them as sub-concepts of the concept named `BusinessInformation`.

Table 3.1: Examples of queries before and after reasoning in Datalog syntax

Query #1: All date/time information related to ORDERS messages	
Before classification	answer(X) :- Interpretation(X), hasSegmentPath(X,"ORDERS"), hasFieldPath(X,"Date_time_period-DATE_TIME_PERIOD-Date_time_period").
After classification	answer(X) :- ORDERS_DateTime(X).
Query #2: All document message date/time related to ORDERS messages	
Before classification	answer(X) :- ORDERS_DateTime(X), hasMeaning(X,"'documentmessagedatetime_Date_time_period"').
After classification	answer(X) :- ORDERS_DateTime_documentmessagedatetime(X).
Query #3: All ORDERS messages that have the function name "original" and have a document message date/time equal to "20110407"	
Before classification	answer(X) :- Message(X), containsValue(X,Y), containsValue(X,Z), ResolvedValue(Y), hasSegmentPath(Y,"ORDERS"), hasFieldPath(Y,"Beginning_of_message-Message_function__coded"), asPrefix(Y,"'original"'), Interpretation(Z), hasSegmentPath(Z,"ORDERS"), hasFieldPath(Z,"Date_time_period-DATE_TIME_PERIOD-Date_time_period"), hasMeaning(Z,"'documentmessagedatetime_Date_time_period"'), hasValue(Z,"'20110407"').
After classification	answer(X) :- Message(X), containsValue(X,Y), containsValue(X,Z), ORDERS_DocumentFunction_original(Y), ORDERS_DateTime_documentmessagedatetime(Z), hasValue(Z,"'20110407"').
Query #4: All Supplier IDs in ORDERS messages that have a product difference equal to "0"	
Before classification	answer(X) :- Interpretation(X), hasSegmentPath(X,"ORDERS-Segment_group_2"), hasFieldPath(X,"Name_and_address-PARTY_IDENTIFICATION_DETAILS-Party_id._identification"), Message(M), containedIn(X,M), containsValue(M,Z), Interpretation(Z), hasSegmentPath(Z,"ORDERS-Segment_group_25"), hasFieldPath(Z,"Quantity_variances-QUANTITY_DIFFERENCE_INFORMATION-Quantity_difference"), hasValue(Z,"'0"').
After classification	answer(X) :- ORDERS_DocumentParticipant(X), Message(M), containedIn(X,M), containsValue(M,Z), ORDERS_LineItemQuantityVariance(Z), hasValue(Z,"'0"').

3.3.3 Classifying EDI Data into BI Concepts

The classification of values, contained in concrete EDIFACT messages stored in a messages KB, is performed by applying reasoning techniques on the extended ontologies (cf. Fig. 3.3, Mark 5). In particular, the idea of the classification is to classify EDI values into the empty generic BI concepts generated from the previous step as illustrated in Figure 3.3, Mark 5. In our implementation for reasoning over ontologies, we apply the DReW system [1] [XEH12] that allows the evaluation of the combination of ontologies and logical rules. During reasoning, the class equivalent expressions are used for classifying all values into corresponding BI concepts. This classification eases querying of EDIFACT messages as described in the following.

3.4 Demonstration

Applying the reasoner to the extended EDIFACT ontologies has three advantages for querying business information from EDIFACT messages:

Advantage #1. The approach supports improved **accessibility of information** and **reduced complexity of query statements** meaning that queries may be created without having to spec-

[1] `http://www.kr.tuwien.ac.at/research/systems/drew` (visited March 01, 2014)

ify the position of particular data or code elements. This is because the extraction approach extracts and interprets data by interpreting qualifiers, resolving coded values, and considering their position. Corresponding examples are illustrated in Table 3.1 which show query statements before and after reasoning. The examples are expressed using rules in Datalog syntax [MSS05] to description logic based ontologies [Baa+03]. Considering the query statements before reasoning, the position information, codes and qualifiers have to be specified, which is not required after reasoning. Considering the example of querying document/message date/time, the query can be formulated to directly point to the `ORDERS_DateTime_documentmessagedatetime` concept. Without reasoning, the position information, qualifiers and codes have to be specified (cf. Table 3.1, Query #2). Furthermore, considering the complexity of the different queries illustrated in Table 3.1, it is shown that the query statements after reasoning have much less complexity than before reasoning.

Advantage #2. The second advantage relies on having **super- and sub-concept relationships** at hand which supports querying at a specific level as well as on a generic level. For instance, querying for date/time information contained within a particular message, one may either query for any date/time information appearing without considering the specific semantics of the date/time information (e.g. document message date/time, invoice date/time, delivery date/time, etc). For example, the `ORDERS_DateTime_documentmessagedatetime` concept is a sub-concept of `ORDERS_DateTime`. Therefore, querying for `ORDERS_DateTime` (cf. Table 3.1, Query #1, After Reasoning) also includes every value belonging to `ORDERS_DateTime_documentmessagedatetime`. This results from the inferences created during reasoning time. In contrast, one can also query for the specific concepts such as `ORDERS_DateTime_documentmessagedatetime` (cf. Table 3.1, Query #2, After Reasoning).

Advantage #3. The classification **reduces the search scope** when querying business information. For instance, querying date/time information in the ORDERS message, such as document/message date/time having the value "20120418", the sub-concept relationships allow searching for `ORDERS_DateTime_documentmessagedatetime` directly. Without having these sub-concept relationships at hand it would be necessary to search through all values until the corresponding value is found. The related example is shown in Table 3.1, Query #2.

3.5 Summary

In this chapter, we presented our business information extraction approach which (i) provides BI concepts and (ii) utilizes these BI concepts for extracting business information from EDIFACT messages. In particular, the existing EDIFACT ontologies [Eng+12b] are extended with BI concepts which are then used for conceptualizing EDI data at any abstraction level. By reasoning over EDIFACT and BI ontologies, data contained in EDIFACT messages is classified into generic BI concepts. Having the classification at hand, we can (i) improve the accessibility of business information and, at the same time, reduce the complexity of query statements, (ii) make super- and sub-concept relationships explicit, and (iii) reduce the complexity of the search scope for evaluating query statements.

Consequently, the approach tackles the first research question (cf. Chapter 1.1.3) for extracting business information from EDIFACT messages and representing the same on a conceptual level. The content of this chapter has also been published in [Kra+12a]. In the larger context of the work presented, the work presented in this chapter serves as a basis for data preprocessing for enabling inter-organizational performance analysis (cf. Chapter 6).

CHAPTER 4

Identifying Inter-organizational Success Factors

One of the important elements for evaluating inter-organizational performance are so called *Key Performance Indicators (KPIs)* which quantitatively indicate performance results. Therefore, inter-organizational KPIs are necessary to be identified for supporting the evaluation. This leads to the second research question (cf. Chapter 1.1.3) which is identifying inter-organizational KPIs. In addressing this research question, we (i) first identify inter-organizational success factors and then (ii) investigate EDIFACT messages for deriving KPIs as well as assign them to success factors. These success factors and KPIs are further used as one of the important elements for connecting EDI data to business objectives, as depicted earlier in Figure 1.6. The first task of identifying inter-organizational success factors is presented in the following chapter. The second task of deriving KPIs is provided in the subsequent chapter.

4.1 Motivation

The evaluation of IORs is usually implemented by means of measurements which are based on certain aspects of an IOR. These aspects are typically considered as "*success factors*", such as communication, trust, and information sharing. Success factors are required for deriving measurements and for monitoring IORs in the context of performance analysis and strategic management. One of the most widely applied strategic management frameworks, namely the *Balanced Scorecard (BSC)* method [Eck06] uses success factors as key elements for implementing the scorecard method. As described in Chapter 2, Section 2.2, the BSC best practices suggest aligning business strategies with KPIs based on critical success factors [KN04] since it is difficult to identify KPIs solely from organizations' strategies. Thus, a thorough understanding of success factors, the influencing relationships between them, as well as their influences on each other is required for effectively identifying appropriate KPIs. A lack of understanding may lead to difficulties in evaluating an organization's strategies or business objectives.

Therefore, before the identification of inter-organizational KPIs, we study success factors related to the success of IORs and based on these success factors we further derive KPIs from EDIFACT messages. For identifying inter-organizational success factors and deriving an understanding of their impact on IORs, we (i) conduct a systematic literature review [Bre+07] as well as (ii) analyze the results. The aim of the systematic literature review is not only to identify success factors related to IORs, but also to provide an understanding of the influencing relationships between success factors. Based on the assumption that scientific literature reflects reality, we conduct the review on publications during the last decade. Several studies identify success factors related to IORs and their effects on each other. However, most of them focus on some particular inter-organizational success factors. In order to understand the whole context of IORs we integrate these success factors and their influencing relationships in a cause and effect model. The integration is achieved by grouping the success factors identified and simplifying their influencing relationships as a cause and effect model. In summary, the cause and effect model describes the influencing relationships of success factors having an impact on IORs. Consequently, the cause and effect model is used as a basis for performing an in- and out-degree analysis for describing the most influenced and the most influencing success factors.

The remainder of the chapter is organized as follows: Section 4.2 provides an overview of the research process applied. A detailed description of the implementation of the review is provided in Section 4.3. The grouping and simplification of the results, resulting in the cause and effect model, is described in Section 4.4. Section 4.5 describes the application of an in-degree and an out-degree analysis of the cause and effect model as well as provides and interpretation of the analysis. Finally, we conclude the chapter in Section 4.6.

4.2 Research Method

In the following the research process followed in this work is described. The research process is depicted in Figure 4.1 and consists of three main activities: (i) conducting the systematic literature review, (ii) simplifying the success factors found, which are commonly called *constructs* in the existing literature, as well as their influencing relationships, and (iii) analyzing the constructs as well as their influencing relationships based on an in- and out-degree analysis.

The first step followed is obtaining constructs from relevant literature which relate to the success of IORs. In addressing this step we conduct a systematic literature review for identifying these constructs as well as their influencing relationships (cf. Fig. 4.1, Mark 1). The literature taken into account has to be published during the last decade as well as has to fulfill certain inclusion criteria. Based on the selected literature we extract (i) constructs or success factors together with their measurements, as well as (ii) influencing relationships among them. The process and the results of the systematic literature review are discussed in Section 4.3.

In a consecutive step the constructs and influencing relationships are further simplified (cf. Fig. 4.1, Mark 2). The simplification is divided into two tasks: i) grouping constructs (cf. Fig. 4.1, Mark 2.1), and ii) deriving the minimum set of influencing relationships between constructs (cf. Fig. 4.1, Mark 2.2). Grouping constructs is achieved by organizing the constructs in a hierarchical structure according to their definition and their measurements. This yields to hierarchical relationships between constructs. The influencing relationships in the hierarchical

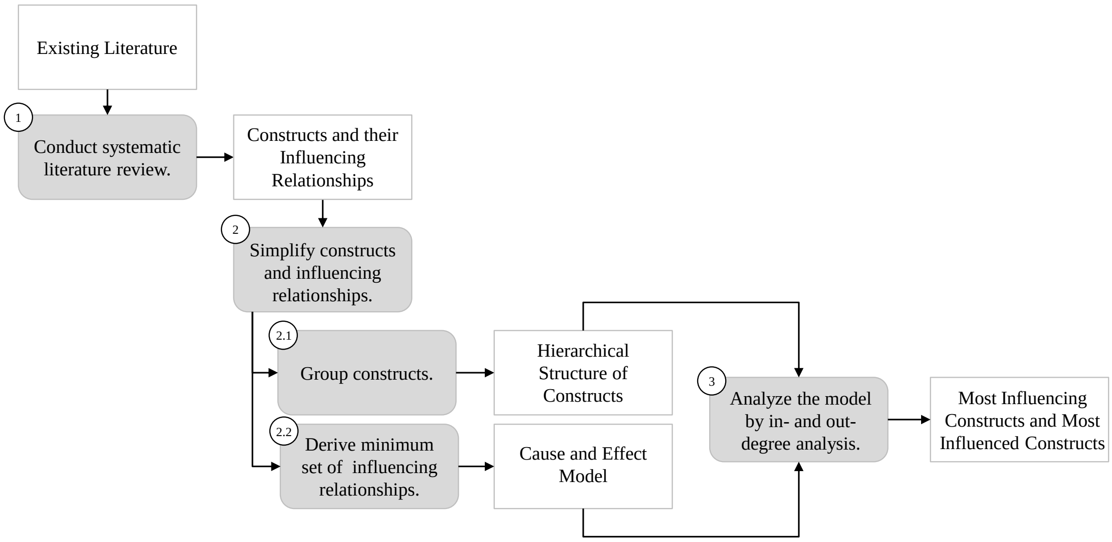


Figure 4.1: The overview of research method

structure are further simplified in order to derive a minimum set of influencing relationships. The minimum set of influencing relationships is described as a cause and effect model. The simplification is described in detail in Section 4.4.

The resulting cause and effect model is then analyzed for deriving new insights (cf. Fig. 4.1, Mark 3). For the analysis we apply an in-degree and out-degree analysis for deriving an implication about the constructs playing an important role in the success of IORs. In particular, we apply an i) out-degree analysis to to identify most influencing constructs and an ii) in-degree analysis to identify most influenced constructs. The analysis as well as the corresponding results are discussed in Section 4.5.

4.3 Systematic Literature Review on Success Factors

The implementation of the review follows the systematic literature review methodology [Bre+07]. The review follows three main phases: (i) defining the review questions and the review protocol, ii) conducting the review and iii) documenting the knowledge. In the following, each of the these phases is discussed in detail.

4.3.1 Review Questions and Review Protocol

The objective of the review is to identify success factors playing important roles in the success of IORs. Consequently, the relevant *review question* of this review is: *"What are the success factors affecting the success of IORs and how do success factors influence each other?"*.

The studies that are of interest in this review directly and indirectly deal with success factors that relate to the inter-organizational context. First, several studies directly investigate inter-organizational success factors and the relationships among success factors. Second, other studies address inter-organizational performance evaluation which typically includes measurement aspects that reflect the success of an organization. This typically implies that those aspects can be perceived as success factors. Third, studies on business partner selection usually provide selection criteria. These criteria imply that having partners matching these criteria lead to the success of IORs. Thus, the selection criteria can be considered as success factors as well.

For answering the review question, we developed a review protocol and search criteria covering the aforementioned topics. The protocol including database, keywords as well as the the search criteria are discussed in the following.

Database

The database used for obtaining the literature reviewed was retrieved using Google Scholar. The literature search was limited to literature published between the years 2000 and 2012.

Keywords

The keywords cover three topics including inter-organizational relationships, success factors, and the evaluation of IORs. These topics correspond to the aforementioned areas of interest. For acquiring literature in these topics different search terms have been used which are listed in the following. Furthermore, we included the supply chain context and considered it as a type of IOR since it is widely studied in terms of measurement and evaluation. Summarizing, the following key words have been used: `inter-organizational relationship(s)`, `inter-organizational relationship success factor`, `measure inter-organizational relationship performance`, `supply chain measure`, `performance measure supply chain`, `performance measurement supply chain`, `inter-organizational measures`, and `B2B performance measures`.

Search Criteria

Inclusion Criteria. The inclusion criteria specified in the following have to be fulfilled by the literature found in order to be included in the review.

1. The literature must be in English and electronically accessible.
2. The literature must be published as a conference paper, journal paper, PhD thesis, masters thesis, or technical report.
3. The main study should
   a) focus on success factors (i.e. constructs) influencing the success of IORs or aspects commonly being used for evaluating IORs or business partners.
   b) focus on the evaluation and measurements for IORs.
   c) mention or highlight some example of measures related to IORs.

Exclusion Criteria. The criteria listed in the following represent exclusion criteria for the literature found. If any of the criteria are met the literature is excluded from the literature review.

1. The literature is not in English and not electronically accessible.
2. The main study focuses on factors influencing the success of inter-organizational systems and applications, as well as the adoption of information technology.

Type of Literature	Area of Study			Total (by type)
	Success Factors in IORs	Performance Analysis in IORs	Business Partner Selection	
Journal	73	80	8	161
Conference	8	4	2	14
Workshop	0	1	0	1
Thesis	1	0	0	1
Total (by area)	82	85	10	177

Table 4.1: Summary of selected literature by area of study

4.3.2 Conducting the Review

The relevant studies are identified according to the review protocol. We selected relevant literature based on their abstract and conclusion that met the criteria. For each keyword combination we implemented the search up to 20 pages since the candidate literature is mostly found up to page 15-17. If there are still some literature on the 20th page which meet the criteria we continue searching on the next page. We keep searching on the next page until we reach the page that has no literature matching the search criteria. Following this process in the review, 177 primary studies are found. In a consecutive step we manually extracted two key information concepts from the selected studies: (i) constructs related to IORs with their measurement items used for evaluating these constructs, and (ii) the influencing relationships between these constructs. The limitation of the review is that it is not controlled. In other words, the review, including the selection of studies and the information extraction, is solely conducted by the author.

4.3.3 Documenting the Knowledge

Four different types of literature are present in the set of the selected studies. These include (i) journal publications, (ii) conference publications, (iii) workshop publications, and (iv) PhD theses. Table 4.1 summarizes the area of study and number of literature focusing on each area.

As illustrated in Table 4.1 the majority of selected literature is of type journal (161 out of 177). The remaining studies are of type conferences (14 out of 177), workshops (1 out of 177), and PhD theses (1 out of 177). The majority of literature in this set focuses on the studies of success factors and performance analysis in IORs. In particular, 82 out of 177 papers study success factors in IORs, and 85 out of 177 papers study performance analysis in IORs. The remaining literature focuses on the study of business partner selection, i.e. 10 out of 177 papers. Figure 4.2 shows number of papers by year. As mentioned, the selected literature had to be published during the years 2000-2012. Around 87% of the studies are published during the last six years (i.e., 2007-2012). Therefore, the selected studies can be considered up to date and still valid.

Based on the set of studies investigated, 88 constructs and 263 influencing relationships between these constructs have been found. The result is further simplified as a cause an effect model describing influencing relationships between constructs. The model is later used for conducting an in- and out-degree analysis for deriving implications about the most influencing as well as the most influenced constructs. The discussion on the simplification and the in- and out-degree analysis is provided in the following sections.

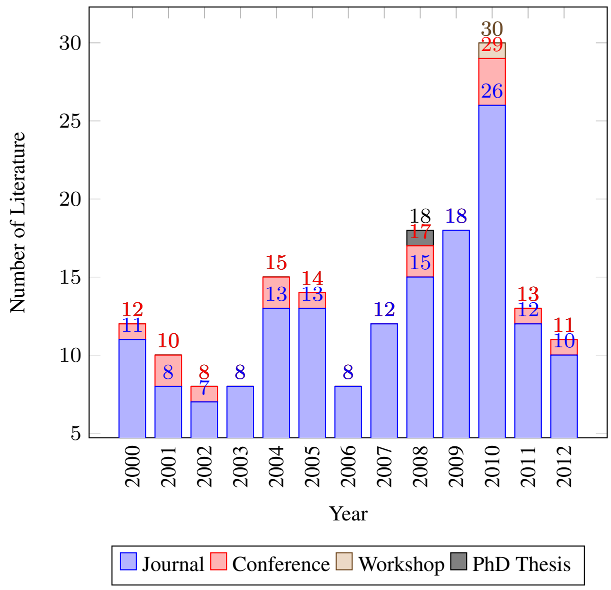


Figure 4.2: The histrogram showing number of papers by year

4.4 Simplification of the Review Results

To simplify the result we implement two main steps: (i) grouping and organizing constructs into a hierarchical structure as well as (ii) applying inference and redundancy checking rules with the ultimate goal of simplifying the influencing relationships. The grouping of constructs is achieved by assigning part-of relationships based on the construct's definitions and measurements. The grouping is also solely implemented by and under the consideration of the author. The outcome of grouping the constructs are 56 constructs organized into a hierarchical structure as depicted in Figure 4.7.

The second step of simplification is to derive a minimum set of influencing relationships. In addressing this task, redundant influencing relationships are removed by applying inference rules and redundancy checking rules. These rules are, again, developed solely by the author. The result is a minimum set of influencing relationships presented as cause and effect model in Figure 4.9. In other words, the result after implementing the second step is the remaining influencing relationships which can be used for further inference by applying inference rules. A detailed explanation on each of these two steps is provided in the following sections.

4.4.1 Grouping Inter-organizational Success Factors

As mentioned before, more than 80 constructs have been found in the studies. For simplifying this information the constructs are grouped according to two different methods. The methods for grouping constructs include (i) grouping based on the definition of constructs as well as (ii) grouping based on the measurement of constructs. The grouping based on these methods leads to either similar constructs or multi-dimensional constructs. Each of these scenarios is discussed in more detail in the following.

Similar Constructs. We found that certain constructs can be considered the same according to their definition and the way the constructs are measured. Thus, constructs sharing a similar definition as well as a similar way of measurement are considered as the same construct.

For example, the construct *Collaboration*, *Cooperation*, and *Integration* have similar definitions which relate to co-working among business partners. In this case, we summarize these constructs with a construct called *Cooperation and Integration*.

Multi-dimensional Constructs. Multi-dimensional constructs rely on assigning "*part-of*" relationships between the construct having other constructs as its dimensions. Some constructs explicitly indicate that they consist of several dimensions while some are found to have overlapping definitions or measures which can be implicitly considered as their dimensions.

For instance, the definition of the construct *Connectedness* is described in [Che11] as "*Connectedness indicates the dependence on each other for assistance, information, commitments or in respect of other behaviors that encourage coordination among individuals, departments or organizations*". Thus, we conclude that Connectedness is part of *Dependency* since according to its definition it is a kind of dependency in terms of behaviors or relationships (cf. Fig. 4.7, Connectedness construct).

Another example is given from the context of *Communication*, *Information Sharing*, and *Information Quality*. The different measurements for measuring these constructs are depicted in Table 4.2. As shown in Table 4.2, the measurement of *Communication* covers the measurement of *Information Sharing* and *Information Quality*. In particular, *Information Sharing* tends to measure if organizations and their business partners keep informing each other about changes or any information that affect their business, whereas *Information Quality* tends to measure if the exchanged information is timely, accurate, complete, adequate, and reliable. The measurements of, both, *Information Sharing* and *Information Quality*, are also found in the measurement of *Communication* (cf. Table 4.2). Therefore, we categorize *Information Sharing* and *Information Quality* as dimensions of *Communication* by assigning part-of relationships between them (cf. Fig. 4.7, Communication construct).

Moreover, the part-of relationship is a transitive relationship. This means that whenever a construct A is a part of B and B is a part of C, then A is also a part of C. For example, *Information Sharing* is a part of *Communication* and *Communication* is a part of *Relational Norm* then we perceive that *Information Sharing* is also a part of *Relational Norm*.

Interpretation of the Hierarchical Structure. After the simplification of constructs, 56 constructs are defined and organized as a hierarchical structure depicted in Figure 4.7. The arrow describes part-of relationships denoted as "partOf". The construct, that is on the arrowhead side, is a construct which consists of the construct on the other end of the arrow. For example, the arrow from *Information Sharing* to *Communication* represents that the construct *Information Sharing* is a part of the construct *Communication*. In other words, *Communication* consists of *Information Sharing*. As shown in Figure 4.7, there are four main groups of constructs: *Relationship Orientation*, *Relational Capital*, *Relational Norm*, and *Atmosphere*. Constructs which do not fit any of the four groups mentioned are grouped into the fifth group, named *Others*. Each of these groups is discussed in the following.

Removal of Redundant Influencing Relationships. Grouping constructs also implies removing redundant influencing relationships. For example, *Flexibility* is grouped with *Adaptabil-*

Communication [CP04]	Information Sharing [Li+06]	Information Quality [LL06]
(1.) We share sensitive information (financial, production, design, research, and/or competition). (2.) Suppliers are provided with any information that might help them. (3.) Exchange of Information takes place frequently, informally and/or in a timely manner. (4.) We keep each other informed about events or changes that may affect the other party. (5.) We have frequent face-to-face planning/communication. (6.) We exchange performance feedback.	(1.) We inform trading partners in advance of changing needs. (2.) Our trading partners share proprietary information with us. (3.) Our trading partners keep us fully informed about issues that affect our business. (4.) Our trading partners share business knowledge of core business processes with us. (5.) We and our trading partners exchange information that helps establishment of business planning. (6.) We and our trading partners keep each other informed about events or changes that may affect the other partners.	The information exchange between our partners and us is: (1.) timely (2.) accurate (3.) complete (4.) adequate (5.) reliable

Table 4.2: Example measurement items of *Communication*, *Information Sharing*, and *Information Quality*. The measurement items of *Communication* cover the measurement items of *Information Sharing* in terms of the frequency of communication and the communication about changes or events. The measurement items of *Communication* cover the measurement items of *Information Quality* in terms of information timeliness, accuracy, completeness, adequateness, and reliability.

ity. Since there is an influencing relationship between *Adaptability* to an additional construct, as well as between *Flexibility* to the same construct in relation with *Adaptability*, we consider one of them as a redundant relationship since *Adaptability* and *Flexibility* are considered as the same construct. Consequently, we remove one redundant influencing relationship. Through performing this grouping and simplification, the hierarchy is simplified to 56 constructs (cf. Fig. 4.7) and 212 influencing relationships.

Relationship Orientation

The relationship orientation represents the tendency of an organization to maintain relationships with partners [CS12]. Relationship orientation is found to be a factor fostering IORs in various dimensions, such as performance and relationship quality.

In long-term relationships business partners tend to have more willingness to share risks and benefits [ZH09; CE93; CPL04]. A long-term relationship orientation may lead to greater commitment and trust [CPL04] as well as may promote collaborative communication. This in turn supports greater cooperation, strengthens relational bond [PLC08; TN99; KMD02; MFN96; PKS96], reduces functional conflict [MH94], improves a firm's performance [CP02], and has a lasting effect on competitiveness of the entire supply chain network [KMD02]. The measurement of relationship orientation, such as in [PC10] and [FKK10], typically measures the attitude of business players towards their relationships in the long run.

Furthermore, relationship orientation could be determined based on relationship benefits, relational proclivity, and connectedness among partners [CS12]. Therefore, the constructs re-

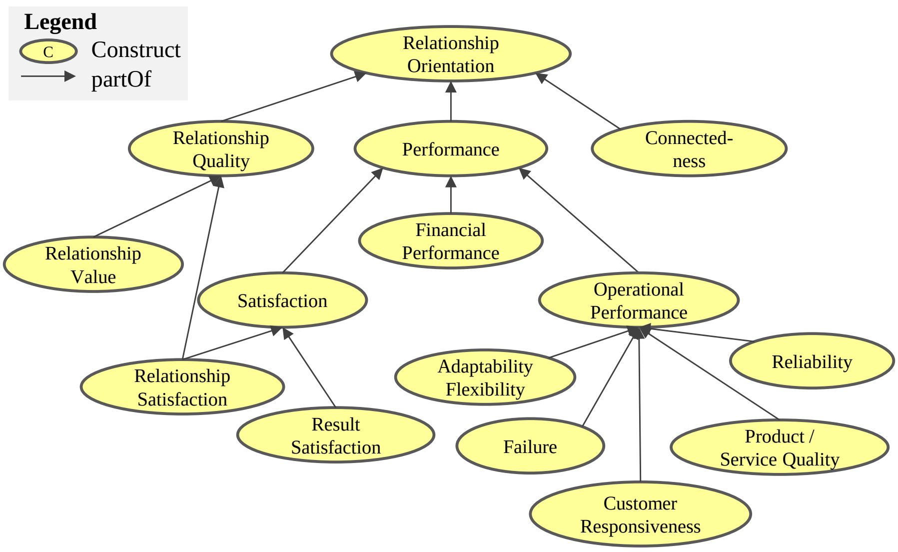


Figure 4.3: The relationship orientation group

lationship benefits, relational proclivity and connectedness are classified as dimensions of relational orientation, as shown in Figure 4.3.

Performance and Relationship benefits are considered as the same construct since the measures of relationship benefits are similar to the measures of performance. The measurements for relationship benefits include financial measures such as profitability and cost, non-financial performance measurements such as delivery time, lead time, product quality and satisfaction [JZ03; CP08; Che11; CS12]. These measurements are similar to performance measurement in general.

In this work, we divide performance into three different dimensions: (i) *Financial Performance*, (ii) *Operational Performance* and (iii) *Satisfaction* since we found that in literature these three dimensions are commonly used to measure performance. Although, it is not always the case that all of them are considered together, performance measurement is still centered around these three perspectives. For example, some of them include satisfaction in performance measurement [Moo11; LLL09; Sep08], some of them refer only to financial and operational performance [WLN10; Wie+10; Joh+04; Koh+12; Liu+12; PLC08; HRS10; Ara+07; PC10; CPL04; Yeu08; SCW00; CGP11; GWI08; FHZ10; Li+06; KHT07; RSK09; CP04; OK03; ZH09; SK10; Cha03], while some of them focus only on either the financial [CGP11; Hsu+08; Möl10; FKK10; Hu+10] or on the operational perspective [AML10; ZB07]. Financial performance typically focuses on outcomes related to economics such as profitability, cost and return on assets, whereas operational performance reflects how well organizations perform in their operations.

There are several dimensions used for evaluating the operational performance, such as, quality [PC10], efficiency [Ara+07], effectiveness [Ash+09], adaptiveness/flexibility [PC10; Ara+07; Ash+09], responsiveness [Ara+07; HSW04; CPL04], productivity [Zha02], delivery [PC10], reliability [HSW04], and failure [CGP11]. For this reason the constructs, found in the review and which are similar to the dimensions mentioned, are considered as a part of operational performance. Therefore, *Adaptability/Flexibility*, *Product/Service Quality*, *Customer Responsiveness*, *Reliability*, and *Failure* are considered as a part of *Operational Performance*.

Adaptability and Flexibility are similar in meaning. [Sin+11] states that flexibility is the ability to deal with uncertainty, modification, or any variety of customer needs which is similar to the definition of adaptability in [JZ03; Ash+09] and [WE04].

However, when considering measurement scales mentioned in literature, they are slightly different. In [JZ03; WE04; AML10] and [CP08], the measurement of adaptation emphasizes more on the ability to adapt the operation or behavior to meet customer needs, while the measurement of flexibility tends to focus more on dealing with unexpected changes and unexpected situations [Joh+04]. Nevertheless, they are considered as the same construct in general since both of them focus on the ability of an organization to adapt itself to changes. In addition, *Product/Service Quality* aims to evaluate the products and services provided by partners. *Customer Responsiveness* focuses on how fast the organization responds to its partners' requests. *Reliability* mostly focuses on delivery performance such as whether the delivery is on time, or how often the delivery is delayed, etc. *Failure* covers any failure in operational processes.

The last dimension of performance is, as mentioned earlier, *Satisfaction*. In general, satisfaction is intangible. It is an emotional response to the difference between what customers expect and what they actually receive [JZ03]. [WLN10] suggests that there are two dimensions of satisfaction: *Result Satisfaction* and *Relationship Satisfaction*. Satisfaction with result focuses on performance issues, whereas satisfaction with relationship focuses on relationship activities, such as decision-making participation, information sharing, and coordination [WLN10].

Relationship Quality is the strength of a relationship as well as the extent to which a relationship meets the need or expectation of partners based on the history of successful or unsuccessful events [Ash+09]. By considering its measurement scales we found that relationship quality consists of *Relationship Satisfaction* and *Relationship Value*. *Relationship Satisfaction* is also considered as a part of *Relationship Quality* because it appears as one of the measurement scales of relationship quality in [CP08]. Moreover, some of the measurement scales of Relationship Quality are matched with the measures of relationship value [CMM10; GFC09] and relational proclivity [Che11; CS12] which try to measure if the relationship supports or motivates the positive outcomes. Therefore, we group the terms relationship value and relationship proclivity as a construct called **Relationship Value** and define it as a part of **Relationship Quality**.

Connectedness is also mentioned to be one dimension of relationship orientation. It indicates the dependence of partners on each other for assistance, information exchange, and commitments or in respect of other behaviors which encourage coordination of departments or organizations [Che11]. To the best of our knowledge, connectedness is a dependence in terms of behavior which is similar to closeness and bond. They refer to the degree of how much the partners are close to or depend on each other in terms of decision making.

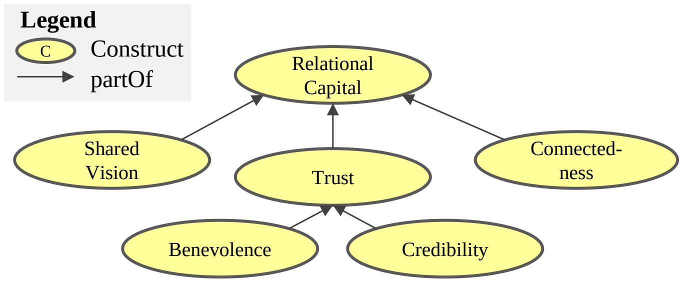


Figure 4.4: The relational capital group

Relational Capital

Relational capital or social capital are described in literature as relationships having dimensions such as trust, shared goals, open interaction, feelings of shared destiny, and togetherness [Koh+12]. Thus, we perceive that relational capital consists of *Trust*, *Shared Vision*, and *Connectedness*. Figure 4.4 shows the structure of the *Relational Capital* group. The definition of *Connectedness* is the same as introduced earlier.

Trust is one of the most widely studied constructs which is shown to have a lot of effects on other constructs. Trust on the inter-organizational level typically refers to many different terms such as credibility, openness, benevolence, integrity, predictability, competence, reliability, etc. [SBS07; ZH06; Sau+04].

In this work, we consider trust in two main dimensions, namely *Credibility* [WLN10; AML10; Moo11] and *Benevolence* [WLN10; AML10; Joh+04; RRG00]. The reason is, that these two dimensions are commonly mentioned in the literature.

Credibility reflects the belief of an organization towards its partners that they perform tasks as expected. From the measurements in [AML10] and [Moo11], credibility focuses on honesty and integrity, reliability, commitment and fulfillment, keeping promises, and the fulfillment of duties. We conclude that credibility tends to focus on the operational perspective in the sense of how well partners perform tasks. We also group reputation with credibility since gaining reputation means gaining credit.

Benevolence reflects the motivation, willingness, and care that business partners tend to have for an organization. It can be seen as a construct against opportunism [Sau+04]. Benevolence tries to measure the willingness to work or help business partners. For example, the measurement scales are: “Does the supplier/partner care for us?”, “Do we see our partner as a friend and vice versa?”, “Can we count on our partner in major issues or problems?” [AML10; Joh+04; RRG00]. This also covers the concept of support [HCV09] which refers to the perception that partners will help with their own free will without having any apparent or immediate benefits to themselves.

Shared Vision is defined in [LL06] as the degree of similarity of shared values and beliefs between partners. It is a kind of compatibility in terms of business objectives and strategies at the strategic level. Generally, this covers the terms shared destination, shared values [HCV09], shared goal, strategy fit [RSK09], goal compatibility, and goal consistency [Kim+10]. The mea-

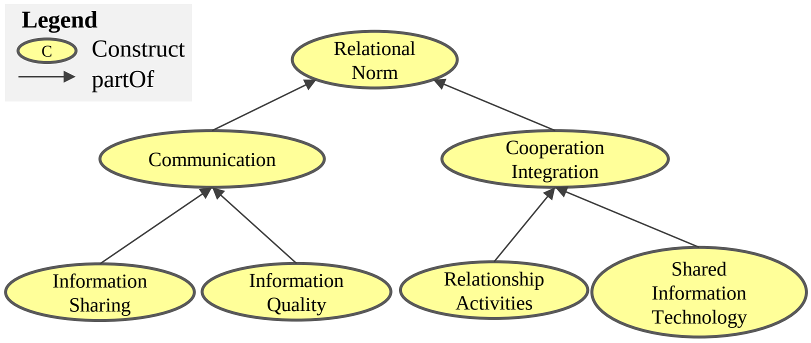


Figure 4.5: The relational norm group

sures or interview questions of shared vision are designed to evaluate if business partners have a similar understanding about the business objectives of each other.

Relational Norm

In literature the concept of Relational Norm is defined ambiguously. Although [LLL09] defines it as the expected norms shared by a group of decision-makers and directed towards collective goals, it is still not clear to what exactly the norm refers to. According to the studies in relational norms, [Dou96] defines relational norm as solidarity, information exchange and role integrity. Similarly, relational norm is defined as solidarity, information exchange and flexibility in [HJ92].

From both definitions, solidarity and information exchange are in common. Solidarity refers to the expectation of partners that they appreciate the relationship and will work to preserve it [Dou96] by establishing activities that help maintaining the relationship such as joint problem solving, meeting obligations, etc. In [HJ92], the solidarity measurement scales consist of items expressing the expected efforts of all partners towards preserving the relationship such as joint responsibility. Thus, solidarity can be considered as behavior of business partners that work together or collaborate to achieve their shared goals and to maintain their relationship. Therefore, we consider solidarity the same as collaboration and cooperation.

Information exchange refers to bilateral expectations that partners provide useful information to each other [Dou96]. The measurement scales of information exchange focus on the frequency as well as the proprietary or importance of the information itself [HJ92]. This in turn could be seen as a whole concept of communication which is similar to the relational norm perceived in [LLL09]. The reason is that the measurement scales of relational norm in [LLL09] focus on communication, information sharing and joint working (e.g. joint problem solving, joint consultations, and discussions). Therefore, we divide Relation Norm into *Communication* and *Cooperation and Integration*, as illustrated in Figure 4.5.

Communication is divided into *Information Sharing* and *Information Quality* since we found that their measurement scales mostly cover the measurement scales of information sharing and information quality. For example, the measure scales for communication described in [JZ03; CP08; Zha02; RSK09; PLC08; CPL04; FKK10] are similar to the measurement scales for information sharing described in [LL06; Che11; CGP11; Li+06]. The measurement scales for information sharing focus on frequency, informality, usefulness of information exchange, and proper communication. Furthermore, the measurement of communication in [PLC08; CPL04; FKK10] and [CP04] also covers information quality which focuses timeliness and accuracy. Some studies even use the same measurement scales interchangeably between communication and information quality, such as [AML10] for communication and [LL06; Li+06; Wie+10] for information quality.

Cooperation and Integration is another dimension of relational norm. It is one of the most widely studied constructs that is found to relate to the success of IORs. This concept includes the terms cooperation [WE04; CP08], collaboration [WE04; RSK09], coordination [Ash+09], integration [PC10; FHZ10] and supply chain linkage [LKS07]. Cooperation, collaboration and coordination clearly refer to the degree of joint working between partners in general whereas integration and supply chain linkage emphasize on the collaboration on the operational level. The collaboration in the operational level includes, for instance, the integration of business processes and activities. In this work, we define the collaboration covering both, the operational and the strategic level.

Relationship Activities and *Shared Information Technology (Shared IT)* are classified as a part of collaboration. Shared IT reflects collaboration in terms of platforms, systems, or infrastructure integration. Relationship activities refer to any activities that are jointly conducted or implemented together with partners. These include the concepts of direct involvement [KHT07], joint responsibility [Joh+04], shared planning [Joh+04], partner engagement in CRM [Duf+12], and social mechanisms [CLS08].

Atmosphere

Atmosphere is studied in [WE04] and described in terms of a relationship that is dependent on power, conflict, cooperation, closeness or distance of the relationship, and mutual expectations between business partners. From the definition, we define *Cooperation and Integration* explained in the last sub-section, *Power* and *Conflict* as dimensions of *Atmosphere*. Figure 4.6 depicts a hierarchical structure of constructs in the *Atmosphere* group.

Power is the ability to (i) influence behavior and decisions and to (ii) cause others to do something they would not have done otherwise [JZ03; Ash+09; Che11]. Power can be perceived in positive and negative senses. Positive power includes mutual or symmetry power which refers to the balance of power, and non-coercive power which refers to the power that comes without forcing but instead it comes from reward, expertise, and legitimacy [Rat00]. The measurement of power symmetry tends to measure if organizations have the ability to influence each other on any decisions or activities [Che11]. Non-coercive power measurements tend to measure if the organization admires their partners or if they are proud to be affiliated with them which in turn makes the organization willing to adapt to partners' requirements [JZ03; AML10]. In contrast, negative power includes coercive power having its roots from the forcing by punishment mechanism. The measurement scales for coercive power described in [JZ03] and [MB00] are likely to help observing if an organization is going to be punished or treated worse when it does not accept the proposals or agreements made. From this reason we divide *Power* into *Non-Coercive Power*, including the concept of mutual power, and *Coercive Power*.

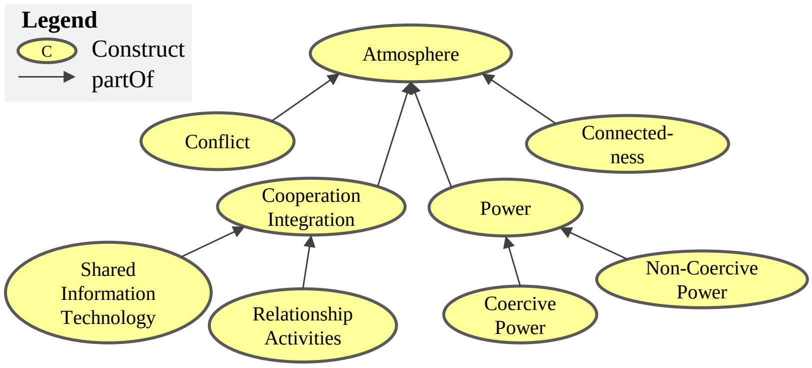


Figure 4.6: The atmosphere group

Conflict or dysfunctional conflict in inter-organizational scenarios or supply chain context typically refers to disagreements that occur in the cooperation relationships or the incompatibility of activities, shared resources, and goals between partners [Che11]. It includes unhealthy behaviors, such as distorting information to harm other decision makers, interacting with each other with hostility and distrust, or forming barriers during the process of decision-making. The measurement of conflict includes interference in decision making, overstating or distorting information to influence partners, and the frequency of disagreements.

Others

Furthermore, there are other constructs that do not have any major similarities and which can thus not be grouped. Hence, the remaining constructs are assigned to the group *Others*. Each of these constructs is described in the following.

Dependency is divided into *Connectedness* and *Operational Dependency*. Connectedness is explained earlier as one kind of dependency in terms of behavior. In contrast, operational dependency refers to the dependency at the operational level. It indicates the extent to which the organization depends upon its partners in terms of resources, tools, manufacturers, products, etc. [JZ03].

Compatibility is defined as the congruence in organizational cultures and capabilities, business objectives, and visions between partners [CMM10]. We divide compatibility into *Shared Vision* as mentioned before, and *Operational Compatibility*. Shared vision implies the compatibility between organizations at the strategic level, while operational compatibility refers to the similarity between organizations' operations, processes, and the way how they conduct business. The measurement of compatibility therefore reflects both, the similarity of an organizations' operations [RSK09; CMM10] and the similarity of business goals [LL06; CMM10; Kim+10].

Uncertainty, or environmental uncertainty, reflects organization's volatility and unpredictability [CMM10]. In [LL06], uncertainty is classified into *Customer Uncertainty*, *Supplier Uncertainty*, and *Technology Uncertainty*. Generally, uncertainty is measured around those three dimensions [RSK09; LL06; CMM10; Moo11]. Customer uncertainty can be measured from unpredictable behavior of customers such as the change of orders and product requirements. Supplier Uncertainty reflects unpredictable behavior of suppliers to their customers which includes unpredictability of product quality, delivery time, operations, or the properties of materials within the same batch [LL06; CP04]. Moreover, technology uncertainty tends to focus on technology changes [LL06; CP04; Kim+10].

Commitment is defined as the willingness of partners to exert effort on behalf of the relationship [LL06; WLN10] and the believe that an ongoing relationship is important to be maintained [MSP11]. There are various items used for measuring commitment. Those measurement items focus on the willingness or the intention to maintain relationships and to devote time and resources such as described in [JZ03; AML10; Zha02; RSK09; LL06; WLN10; MSP11; Liu+12; Möl10; MB00; GFC09; RRG00] and [Vij10]. Moreover, we define *Loyalty* as a behavioral intention as part of *Commitment* since its measures are part of Commitment in terms of the continuation of business or the purchase of products from an organization [GFC09].

Innovation helps organizations survive in the long run since the competition through product and technology innovation is stronger than competition among organizations offering similar products [CF05]. The effect of innovation is not only relevant within individual organizations but also covers the inter-organizational context such as supply chains as discussed in [CF05; Cha03] and [CGP11]. If an organization succeeds in innovation, there is a possibility that it can maintain its position in the market among its competitors as well as that it gains additional market share. Innovation is difficult to be measured since it is not clear to which extent the innovation is successful. However, some studies use the number of new products, the percentage of sales of a new product, the percentage decrease in time or resources when new technology or new inventions are applied.

Top Management Support describes the support from top-level managers having an understanding and seeing the importance of a partner's relationship, especially in the supply chain context [LL06; CP04]. The role of top managers is important for driving the implementation towards successful strategies since they have a better understanding of the needs of an organization's strategies [CP04].

Relationship Learning is considered as a process to improve behavior or joint activities in a relationship by collaboratively creating more value than by doing it individually [CMM10]. The measurement of learning tends to evaluate if there is knowledge created during the collaboration and if knowledge is communicated and applied to create value [CMM10; GFC09].

Equity and Fairness is a perception of an organization that its partners act fairly [HCV09]. The concepts of equity and fairness are similar to the concept of reciprocity of a relationship, as introduced in [Kim+10]. Reciprocity refers to the degree of fairness that the partners perceive about sharing risks, burdens, and benefits [Kim+10]. Inter-organizational justice [Duf+12; Liu+12] is also categorized as fairness since it refers to the perceptions of fairness of business partners. [Duf+12] and [Liu+12] defined four different types of justice which include distributive justice, procedural justice, interpersonal justice, and information justice.

Distributive justice reflects the perceptions of the weaker partners about the fairness of the division of benefits compared to more powerful partners. Procedural justice refers to the perceptions about the fairness of the formal procedures governing a decision process. Interpersonal justice reflects the fairness at the individual level such as if an individual is treated with politeness, dignity, and respect by other individuals. Lastly, information justice focuses on the adequacy of information provided by partners. Besides, reward and cost introduced in [WLN10] are also included in fairness since they focus on sharing benefits between organizations. The measurement of reward and cost tends to measure if benefits are shared equally and if the partners are willing to share unexpected costs. This is similar to the measurement scales of fairness used in [Kim+10].

Internal Information Technology, such as information management systems and decision support systems, has an important role in supporting business collaboration [RRG00; KW05]. In this work, we distinguish information technology between internal information technology and shared information technology. Internal information technology covers any information systems or any technologies supporting internal business processes and activities. In other words, internal informational technology is used only by one organization, while *Shared IT* refers to any information systems or technologies that span beyond organizational boundaries [RRG00]. Such shared information technology systems include communication platforms and information systems that provide a control or monitoring platform to other partners.

Strategy Quality is intangible and ambiguous. However, its measurement scales used in [CS12] tend to focus on the formation, implementation, and comprehensiveness of strategies. A high degree of strategic quality within supply chains could enhance the competitive advantage of the entire supply chain [CS12].

Contract defines the rights and obligations of partners through formal rules, terms, and procedures by explicitly stating how various future situations will be handled [LLL09]. Contract is considered as one mechanism for creating structural systems which all partners must comply with. The measurement of contract is quite straight forward by examining if all agreements and obligations are formally defined.

Supply Chain Practice [Li+06; ZB07] refers to any activities that promote effective management of supply chains, such as strategic purchasing [CPL04; FKK10], supply management orientation [SCW00], and supply chain management strategies [GWI08]. The *Relationship Activities* are also considered as a part of the *Supply Chain Practice* since all collaboration activities between the business partners could result in an improvement of the supply chain management.

Investment, dedicated investment [WLN10], specific investment [LLL09; CGP11; MSP11; Koh+12], or relationship investment [Liu+12] refers to the investment in various ways such as resources and activities made by organizations and dedicated to the relationships with business partners. Investments can be tangible (e.g. manufacturing tools and resources) and intangible (e.g. knowledge, ideas, technology or capability) [CMM10]. Investment has been recognized to have a positive effect on an organization's performance. However, the more specialized investment is, the lower its value is in general use [CGP11].

The studies of investment [LLL09; CGP11; CMM10; WLN10; MSP11; Liu+12; Koh+12] observed in the review describe measures for investment. These describe simple measurement scales that ask if organizations and its partners have made some significant investments in resources, knowledge, and technology.

Complementarity refers to the lack of similarity or the overlap between core businesses or capabilities, and the extent to which the unique strengths and resources of partners are exchanged. Several studies show that it positively correlates with Relationship Learning [CMM10].

More complementarity means that among different business partners there are also different knowledge bases. Therefore, it is more likely that when there is a diversity and non-redundancy in knowledge then organizations and its partners will have learning opportunities.

Environment Dissimilarity covers the diversity of the market environment, sourcing and distribution choices [CMM10]. Thus, it offers flexibility to organizations as well as affects Relationship Learning. Due to the environmental dissimilarity, organizations and its partners are confronted with greater risks and challenges which in turn lead to learning. Environmental dissimilarity can be identified with several items such as government intervention, volatility of regulations, currency exchange rate, overall economic situation, etc. [CMM10].

Competitive Advantage is the extent to which an organization is able to create a defensible position over its competitors [Li+06]. It comprises capabilities that allow an organization to differentiate itself from its competitors. Important competitive capabilities include price/cost, quality, delivery, and flexibility, as described in [Li+06].

Opportunism is the self-seeking behavior with guile (e.g. deceitfulness, a lack of candor or honesty) including hidden information or hidden action [MSP11]. Opportunism includes a wide variety of potentially different behaviors. In an inter-organizational context the behavior is considered to be opportunistic if the behaviors are inconsistent with some prior contract or agreement. In literature on business partner selection, opportunism is also mentioned as a criteria of assessing partners [Möl10]. The measurement scales used to justify opportunism tends to find out if partners perform or act in a opportunistic way to achieve their goals (e.g. alter/detour information, lie, breach agreements).

Success can be interpreted in several ways including the achievement of business goals and the improvement of dedicated aspects. Therefore, we group the concepts of success and improvement which also includes the improvement of Supply Chain Management (SCM) [LL06].

4.4.2 Deriving the Influencing Relationship Model

The hierarchical structure consists of constructs as well as relationships between these constructs. The relationships are either part-of relationships or influencing relationships. Based on this hierarchical structure the second step for simplifying the constructs and their relationships can be applied. This includes applying inference as well as redundancy checking rules. The application of inference rules helps inferring additional influencing relationships which are not explicitly found in literature studies. By applying these inference rules we can obtain a complete model of all constructs including the influencing relationships among the constructs.

Inference Rules

The inference rules consider both types of relationships between constructs for deriving additional knowledge. These include influencing relationships as well as part-of relationships. We developed four inference rules, as illustrated in Figure 4.8. Rule #1 and Rule #2 infer influencing relationships from constructs, i.e. main constructs, which are comprised of either one or more other constructs, i.e. sub-constructs. For instance, the construct *Trust* consists of the constructs *Credibility* and *Benevolence*. In other words, *Credibility* and *Benevolence* are sub-constructs of *Trust*.

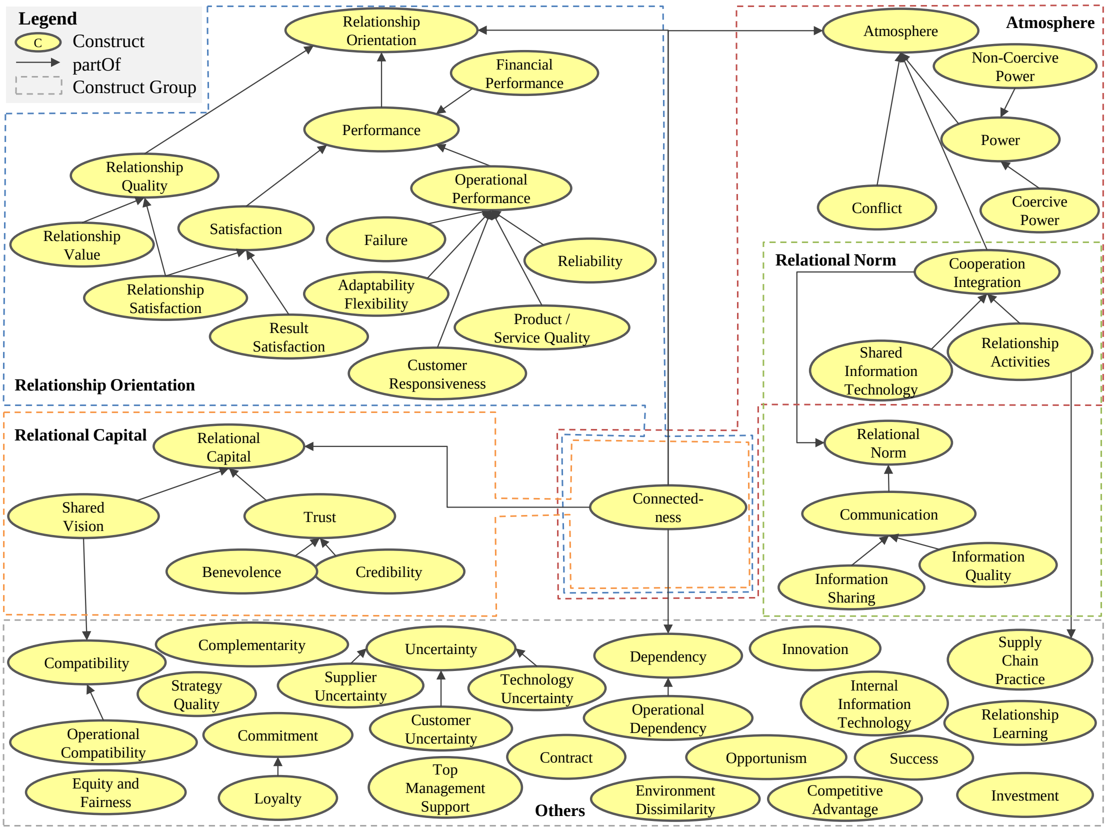


Figure 4.7: A hierarchical structure of inter-organizational success factors (note: this figure illustrates only part-of relationships, but does not show influencing relationships)

In particular, Rule #1 specifies that if an additional construct influences the main construct then this additional construct also influences the sub-construct. For instance, applied to the running example, if the additional construct named *Adaptability* influences *Trust*, then we also perceive that *Adaptability* influences *Credibility* and *Benevolence*. Similarly, Rule #2 specifies that if the main construct influences an additional construct, then the sub-construct also influences the additional construct. For instance, applied to the running example, if *Trust* influences the additional construct named *Adaptability*, then we also perceive that *Credibility* and *Benevolence* influence *Adaptability*. These properties are reflected by rule#1 and #2. Formally, the first two rules are expressed as follows:

Rule #1: If y is a part of x, and z influences x, then it implies that z influences y.

Rule #2: If y is a part of x, and x influences z, thist implies that y influences z.

Rule #3 and #4 are inference rules for deriving additional knowledge. In contrast to rules #1 and #2, the rules #3 and #4 infer additional knowledge from the relationships of the sub-constructs. In particular, for rule #3, it is important to note that only if an additional construct

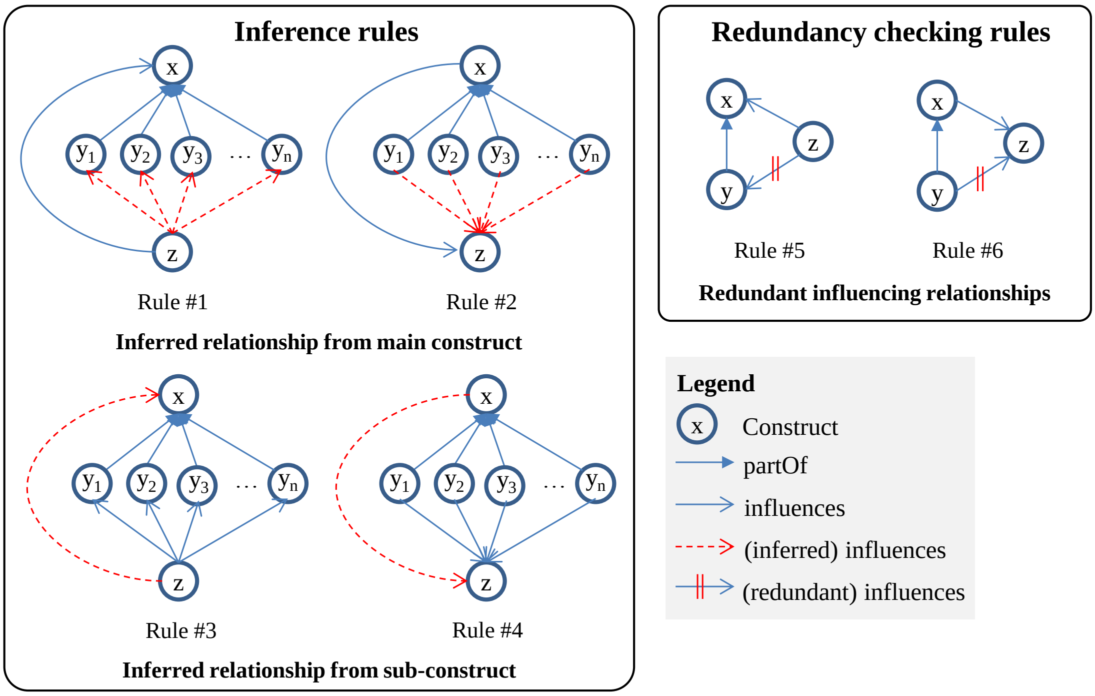


Figure 4.8: Inference rules and redundancy checking rules

influences all sub-constructs then the additional construct also influences the main construct. For example, if *Performance* influences *Credibility* and *Benevolence*, then it infers that *Performance* also influences *Trust*. Similarly, if all sub-constructs influence an additional construct, then the main construct also influences the additional construct. For instance, if *Credibility* and *Benevolence* influence *Performance*, then it also infers that *Trust* influences *Performance*. These properties are represented by Rule #3 and Rule #4.

Rule #3: If $y_1, y_2, y_3 ... y_n$ are a part of x, and for every part of x, there is an influencing relationship from z to each part, then it implies that z influences x.

Rule #4: If $y_1, y_2, y_3 ... y_n$ are a part of x, and for every part of x, there is an influencing relationship from each part to z, then it implies that x influences z.

After applying all inference rules, a complete set of influencing relationships are obtained. This servers as a basis for applying redundancy checking rules, as discussed in the next sections.

Redundancy Checking Rules

The overall aim of this step is deriving a minimum set of influencing relationships which are necessary for inferring the complete set of influencing relationships by applying the aforementioned inference rules. For this reason, we perform a redundancy checking step to remove redundant influencing relationships that can be inferred by the inference rules.

Considering the inference rules, the influencing relationships of a main construct also cover the influencing relationships of its sub-constructs. This means that all influencing relationships of the sub-constructs can be safely eliminated since all of this information can be inferred again based in the influencing relationships of the main construct. For instance, the influencing relationship from *Performance* to *Trust* covers the influencing relationships from *Performance* to *Credibility* and from *Performance* to *Benevolence*. Hence, in this case, the influencing relationships between *Performance* and the sub-constructs of *Trust* (i.e., *Credibility* and *Benevolence*) are considered as redundant.

In order to remove redundant relationships we define two redundancy checking rules as shown in Figure 4.8, Redundancy checking rules. These rules, i.e. Rule #5 and Rule #6, are described as follows:

Rule #5: If y is a part of x, z influences x and z influences y, then z influences y is a redundant relationship.

Rule #6: If y is a part of x, x influences z and y influences z, then y influences z is a redundant relationship.

Implementation and Results

The implementation of these steps is achieved by applying a reasoning mechanism based on the inference and redundancy checking rules introduced earlier. The reasoning mechanism is implemented by the DReW system [1] [XEH12] which is a reasoning engine for the evaluation of the combination of ontologies and logical rules.

In this work, the inputs are (i) constructs (cf. Fig. 4.7), with part-of relationships, which are naturally modeled as OWL ontologies, and (ii) the inference rules and redundancy checking rules in the form of logical rules. The DReW reasoner internally translates all the inputs into logical rules and calls the DLV system [2] [Leo+06] to perform the reasoning tasks. The implementation yields 121 remaining relationships which are important for deriving relationships by using the inference rules. These constructs and the remaining relationships are illustrated in Figure 4.9, representing the cause and effect model of inter-organizational success factors.

In the model an arrow describes a directed influencing relationship. For example, the arrow from *Trust* to *Contract* means that *Trust* influences *Contract*. Moreover, the influencing relationships of sub-constructs can be derived using the inference rules and based on the part-of relationships of the main construct. For instance, applied to the running example, *Trust* influences *Contract*, can imply that *Credibility* and *Benevolence* also influence *Contract* since they are sub-constructs of *Trust* (according to Rule #2).

[1] `http://www.kr.tuwien.ac.at/research/systems/drew` (visited March 01, 2014)
[2] `http://www.dlvsystem.com` (visited March 01, 2014)

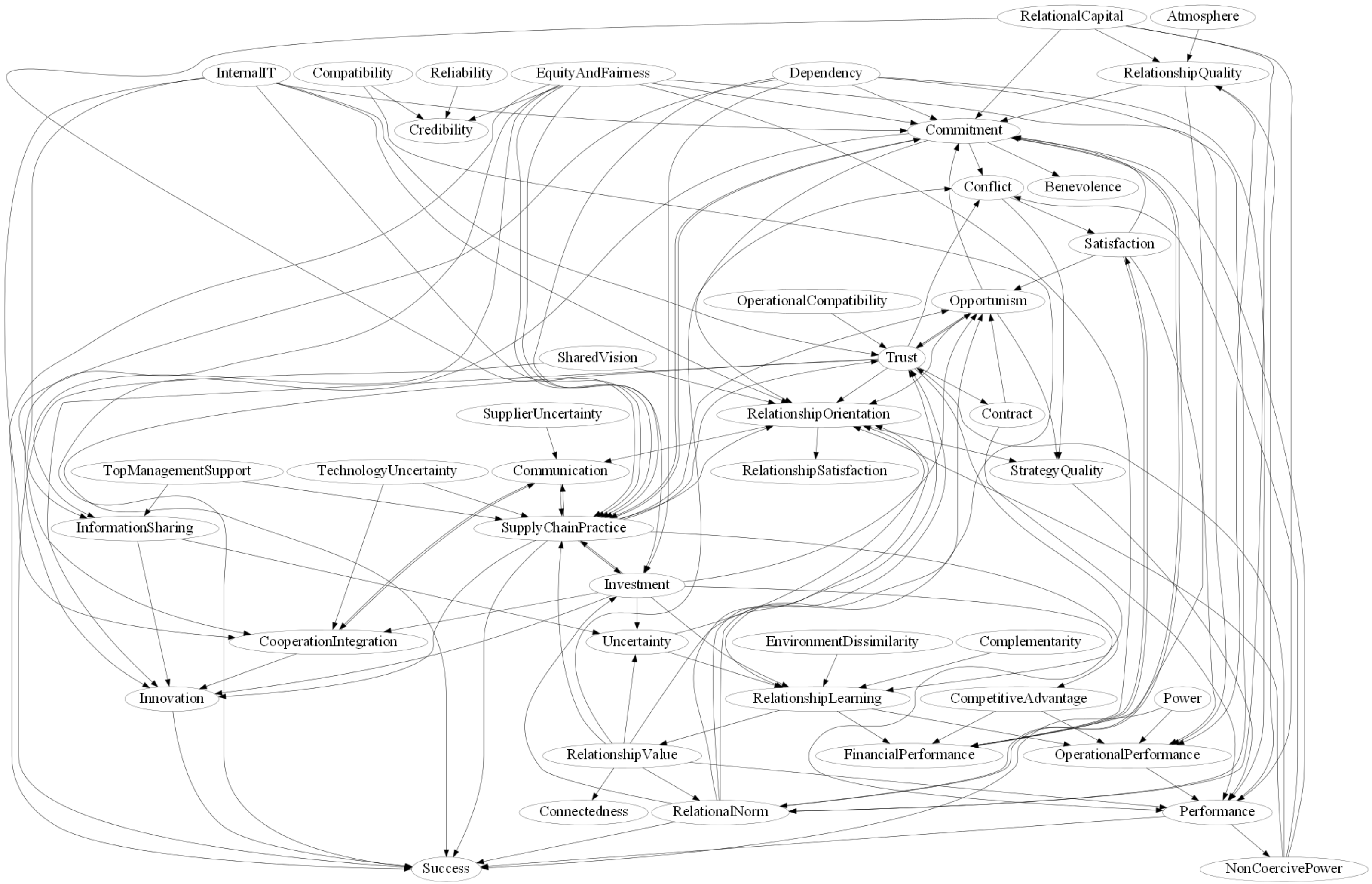


Figure 4.9: Cause and effect model: Influencing relationships between constructs

4.5 Network Analysis

From the essential influencing relationships shown in Figure 4.9, we apply inference rules for deriving the complete set of influencing relationships. Based on the completed model we analyze the most influencing constructs as well as the mostly influenced constructs. This is achieved by performing a network analysis of degree centrality. In this work, we distinguish between out-degree and in-degree centrality since the model is directed. Out-degree and in-degree analysis can provide a straightforward answer of the most influencing and the mostly influenced constructs by looking at the numbers of outgoing and incoming relationships. In other words, the out-degree of a construct is the number of times that the construct appears as a source of the influencing relationship. In-degree analysis investigates the number of incoming influencing relationships or the number of times that the construct appears as a target of an influencing relationship.

However, constructs are defined as a hierarchical structure. In particular, the interpretation of the network analysis applied in this work has to deal with hierarchy. This makes an interpretation of the result from the analysis of other centrality measures complicated and ambiguous. Furthermore, as mentioned earlier, our analysis is based on the complete influencing model, where all influencing relationships are derived by applying inference rules (cf. Fig. 4.8). Hence, the analysis of other centrality measures, such as betweenness (focusing on the number of times a node bridges two other nodes in their shortest path) and closeness (focusing on the total distance to all other nodes), becomes obsolete, since all hidden relationships are already inferred in the complete model. Therefore, we only focus on the out-degree and in-degree analysis to provide insights of the most influencing and mostly influenced constructs.

4.5.1 Out-degree Analysis

The ranking of constructs by out-degree shows the most important constructs which influence other constructs the most. The results of the out-degree analysis are illustrated in Figure 4.10. Each circle represents a particular construct where the label represents the name of the construct. The size of circle visualizes the number of outgoing influencing relationships. That is, the higher the number of outgoing relationships the bigger the circle is. From the result shown in Figure 4.10, the top ranked constructs in terms of out-degree include three main groups which are *Relational Capital*, *Relational Norm*, and *Commitment*. This implies that those three groups of constructs influence many inter-organizational success factors.

Every part of *Relational Capital* is found to have a considerably high out-degree as well. These include *Trust* with its dimensions *Credibility* and *Benevolence*, *Shared Vision*, and *Connectedness*. Similarly, *Commitment* as well as its part *Loyalty* are also in a top rank in regards to out-degree. This shows that trustworthiness and faithfulness in business partners are important for developing IORs since they are the basis of several other success factors.

Furthermore, *Relational Norm* including *Communication* (and its dimensions: *Information Sharing* and *Information Quality*) and *Cooperation and Integration* (including *Shared Information Technology* and *Relationship Activities*) also play an important role as a source to drive other success factors since they appear to have a high out-degree as well. Moreover, the out-degree of

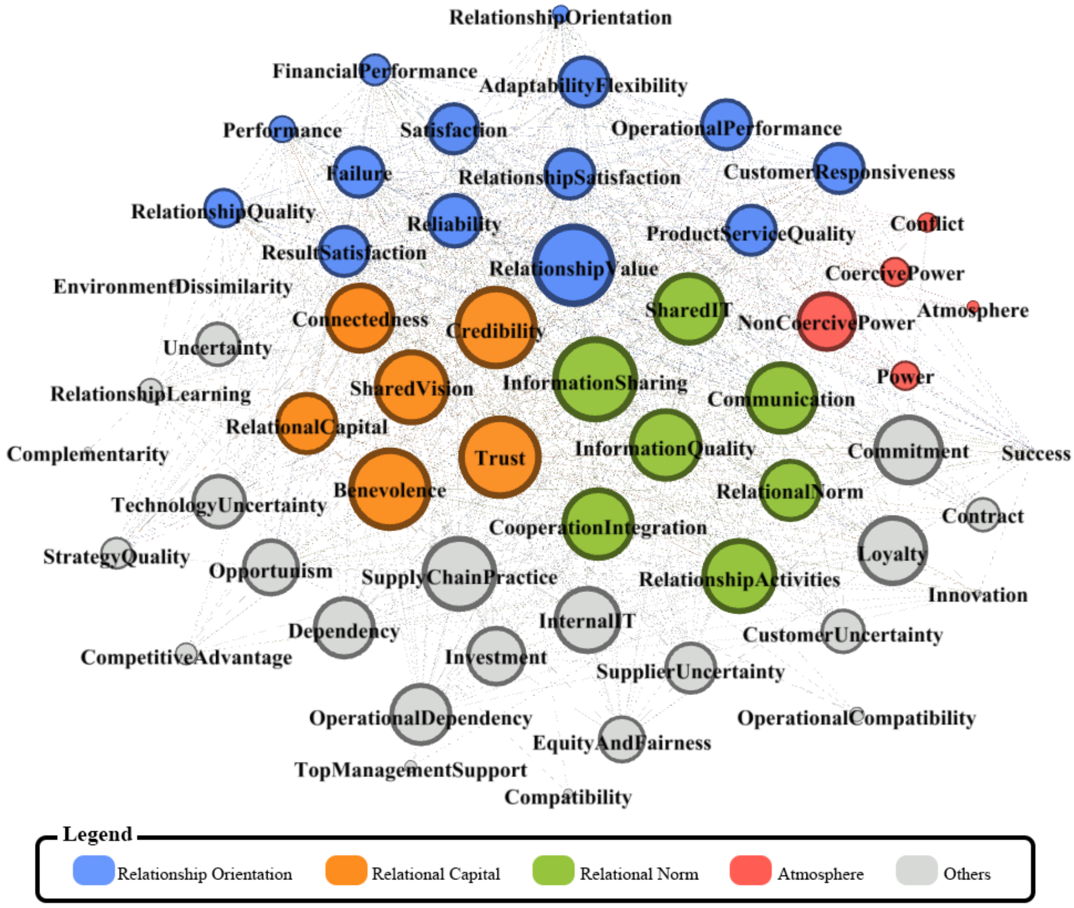


Figure 4.10: The constructs ranked by out-degree

Supply Chain Practice, *Relationship Value*, and *Internal Information Technology* are also high enough to be considered as antecedences of other success factors.

In summary, in order to improve IORs, organizations should consider maintaining or improving success factors in the groups of *Relational Capital*, *Relational Norm*, *Commitment* and the success factor *Information Technology* as well as implementing the *Supply Chain Practice*. Moreover, they should periodically encourage and promote *Relationship Value* among business partners since it can motivate business collaboration.

The result mainly indicates that trustworthiness, loyalty, communication, cooperation and relationship values are necessary for maintaining and improving IORs. If organizations deceive their business partners, this may result in losing trustworthiness and loyalty from their business partners and in turn negatively affect their relationships. In the worst case, when cheating is started, betrayal may continue happening among business partners which eventually leads to a great loss in the entire business chain. Similarly, lack of good collaboration and communication may result in poor business operations. Especially, when a problem arises in a business chain, organizations require effective cooperation and communication in order to quickly iden-

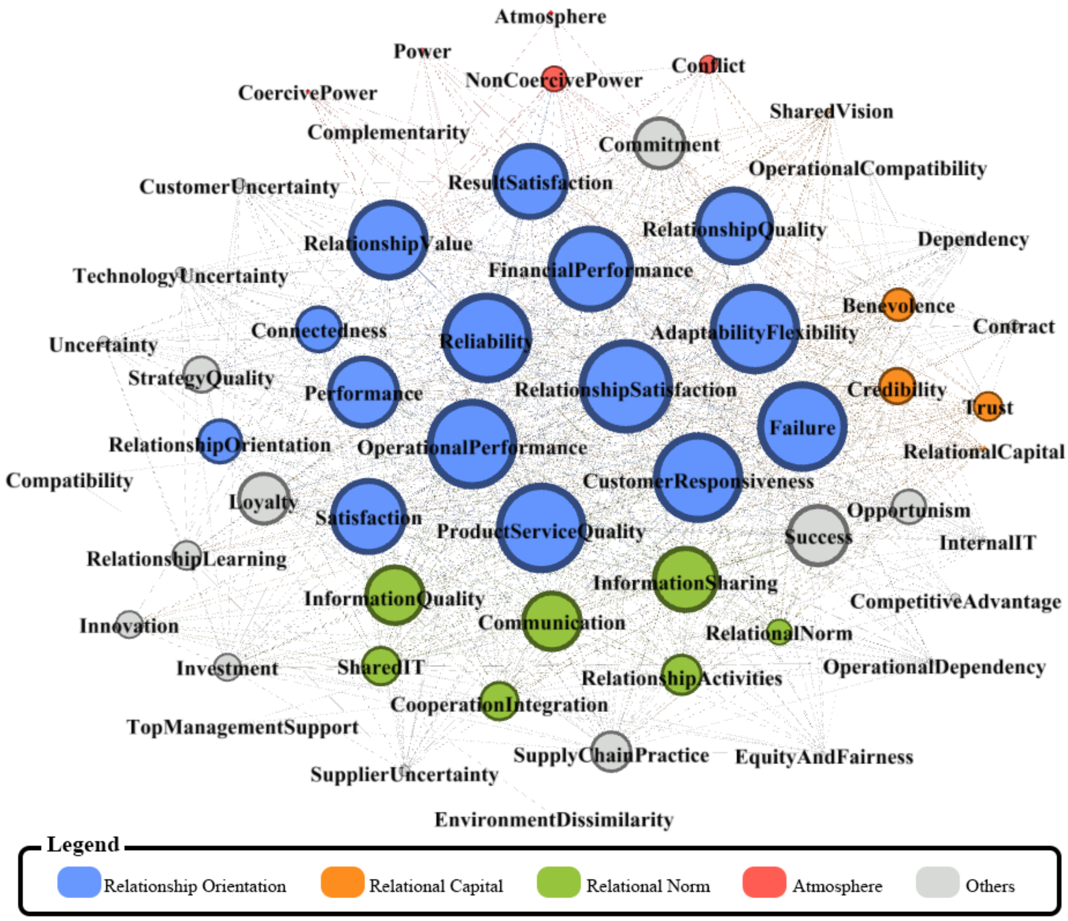


Figure 4.11: The constructs ranked by in-degree

tify the problem and urgently implement the solution. Poor cooperation and communication may cause failure in business operations which consequently leads to business loss and finally harms IORs. Therefore, organizations should do business with their integrity and sincerity as well as continuously maintain and improve their communication and cooperation with business partners. Furthermore, these results are similar to the game theory [NM07] which shows that trustworthiness leads to the cooperation, and cooperation leads to better outcome for all participants in the game [Mea13]. However, without enough trust to cooperate one may start to act only for the benefit of themselves. One way to enable the cooperation is to encourage the value of relationships [Mea13]. Once when participants understand benefits and outcomes from their relationships, business collaboration will be driven.

4.5.2 In-degree Analysis

Figure 4.11 illustrates the ranking of constructs by in-degree. Similar to the out-degree analysis the circles represent constructs. However, in the case of the in-degree analysis, the size of the circles represents the number of incoming influencing relationships. In other words, the bigger the circle, the higher the construct is influenced by other constructs.

Most of the constructs under the group of *Relationship Orientation* are shown to have a high number of in-degree. Those include *Relationship Quality*, and *Performance*. The construct *Performance* contains the sub-constructs *Financial Performance*, *Operational Performance*, and *Satisfaction*. Noteworthy constructs are also the sub-constructs of *Operational Performance*, which include *Adaptability/Flexibility*, *Failure*, *Customer Responsiveness*, *Reliability*, and *Product/Service Quality*. Furthermore, a high in-degree is also shown in the sub-constructs of satisfaction which contains *Result Satisfaction* as well as *Relationship Satisfaction*.

These constructs can be considered as consequences of other inter-organizational success factors. The result implies that the constructs in the group of Relationship Orientation tend to rely on other constructs. In other words, they require support from other constructs in order to be achieved and fulfilled.

Surprisingly, *Success* does not appear as the highest in-degree construct, but instead the constructs under the group of *Relationship Orientation* does. The reason is that best to our understanding most of the studies reviewed often use the constructs of the group *Relationship Orientation* for assessing the success of IORs. Generally speaking, this implies that the success of IORs can be reflected by the constructs under the group of *Relationship Orientation* (i.e. *Performance* and *Relationship Quality* as well as their sub-constructs).

4.6 Summary

This chapter presents a review which aims at identifying inter-organizational success factors and their influencing relationships affecting IORs. We conducted a systematic literature review covering 177 publications for identifying success factors related to IORs. The success factors have been integrated and consolidated in a cause and effect model. Consequently, the resulting cause and effect model serves as a basis for analyzing the most influential and most influenced success factors. The content in this chapter has been submitted as a journal publication and is still in the process of review [Kra13].

The success factors found in literature have been grouped according to their definitions and measurements resulting in a hierarchical structure. This structure serves as a basis for applying inference rules for deriving additional knowledge which is not explicitly addressed in literature. The resulting structure containing additional inferred knowledge has been simplified by applying redundancy checking rules. Thereby, duplicate information is eliminated which resulted in the minimum set of influencing relationships. The resulting constructs as well as their minimum set of influencing relationships are presented in the cause and effect model.

The cause and effect model has further been analyzed in terms of out-degree and in-degree analysis. This allows identifying the most influencing as well as the most influenced constructs. The result of the out-degree analysis shows that the success factors in the groups of *Relational Capital*, *Relational Norm*, *Commitment* as well as the success factors *Supply Chain Practices*, *Internal Information Technology* and *Relationship Value* are the most influencing success factors in IORs. The result of the in-degree analysis shows that the success factors in the *Relationship Orientation* group tend to be the most influenced success factors.

The cause and effect model presented in this chapter explains the existence of the influencing relationships between success factors. However, it does not indicate the tendency of relationships (i.e. positive relationships or negative relationships). Nevertheless, this review provides an understanding of success factors in IORs as well as the directed influencing relationships between success factors.

The identified success factors are used for considering appropriate KPIs derived from EDIFACT messages. The details of the KPI identification based on the success factors found in this work is further provided in the next chapter (i.e., Chapter 5). Furthermore, the knowledge gained in this review may be used by organizations as a strategic guideline for identifying appropriate success factors playing key roles in inter-organizational success as well as for evaluating and monitoring an organizations' achievements.

CHAPTER 5

Identifying Inter-organizational Key Performance Indicators

As mentioned in the previous chapter, the second task for addressing the research question on the identification of inter-organizational KPIs (cf. Chapter 1.1.3) is to investigate EDIFACT messages for deriving KPIs and assigning them to success factors. This chapter elaborates on the method applied for identifying inter-organizational KPIs as well as presents a set of KPIs derived from EDIFACT messages. The objective of this work is to lift EDI data to quantifiable KPIs which are in turn connected to business objectives for supporting the evaluation of IORs on a strategic level (as introduced in Fig. 1.6).

5.1 Motivation

Our main research goal is to enable the quantitative evaluation of IORs. In doing so, quantifiable KPIs have to be defined and mapped to inter-organizational success factors. In this way, we can evaluate the IORs quantitatively through success factors which are measured by quantifiable KPIs. For obtaining such KPIs, we consider EDIFACT messages since they contain relevant data. This is because from a technical perspective inter-organizational collaborations are often supported by means of EDI. Hence, the EDIFACT messages are considered as a potential data source for identifying KPIs. In this work, we identify KPIs based on concrete EDIFACT messages rather than based on the extracted BI concepts which represent EDIFACT values on a conceptual level. This is because BI concepts are user-defined and, hence, more flexible for arbitrary definitions based on users' interest. In other words, the naming of BI concepts and their mapping to EDIFACT values may vary depending on the domain of interest. This makes KPI identification difficult and inaccurate.

In practice, EDIFACT standards are further concretized by bi- or multilateral agreements between collaborating business partners, so-called *Message Implementation Guidelines (MIGs)*. Therefore, in this work we take MIGs of EDIFACT messages under consideration. Based on

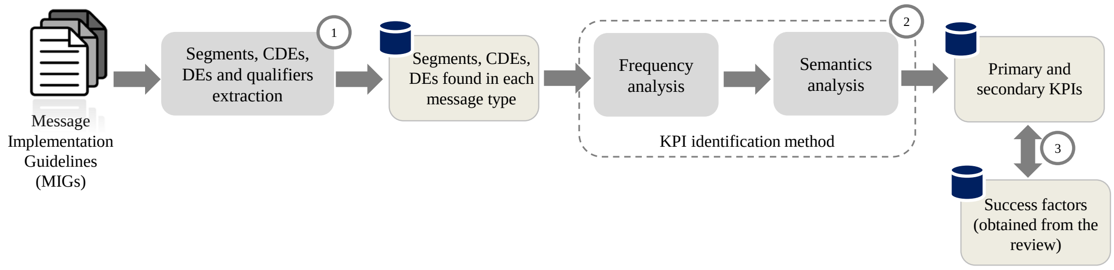


Figure 5.1: Overview of KPI identification method

a data set of real-world MIGs, we derive inter-organizational KPIs and proposed guidelines for their calculation from EDIFACT data elements. Furthermore, we aggregate these KPIs to define quantitative measurements reflecting inter-organizational success factors.

The the chapter is organized as follows: Section 5.2 provides an overview on related work. The method for identifying inter-organizational KPIs from MIGs is described in Section 5.3 including the resulting KPIs. Finally, we conclude with a discussion of the results in Section 5.4.

5.2 Related Work

Due to the ambiguity and intangibility of related measurements, the evaluation of IORs is challenging. Most of the studies regarding the evaluation of IORs (e.g., [Cas08], [SQ03]) tend to build upon the analysis of success factors having an impact on IORs. For example, trust [SBS07; ZH06; Sau+04], information sharing [LL06; Che11] and joint working [KHT07; Joh+04; Duf+12] are mentioned as such factors, which are, however, difficult to measure. In order to be able to define KPIs for IORs, related success factors and ways of measuring them need to be investigated. Hence, we conducted a literature review on factors influencing IORs' success as well as their measurement as presented in Chapter 4. According to the review, 56 success factors are identified. We found that among these success factors ∼89% are measured by measurement items rated by an evaluator and, hence, include a considerably subjective dimension. Only ∼11% of them are measured by quantifiable KPIs, such as financial performance [CGP11; Hsu+08] and operational performance [AML10; ZB07]. This motivates us to investigate whether it is possible to measure success factors of IORs quantitatively.

5.3 Identifying Inter-organizational KPIs

In this section, we present (i) a method for deriving KPIs from EDIFACT messages and (ii) the results of applying this method on a real-world data set of MIGs. Figure 5.1 shows an overview of the presented KPI identification method. We obtain segments, composite data elements, data elements, and qualifiers from MIGs (cf. Fig. 5.1, Mark 1). Based on this extracted data, we identify KPIs (cf. Fig. 5.1, Mark 2) and group them to corresponding success factors (cf. Fig. 5.1,

Table 5.1: MIGs used for identification of inter-organizational KPIs

MIG Set / Industry	Includes / Excludes	Included Version(s)	Based on EDIFACT Release(s)	# MIGs / MSG. types
Odette EDIFACT MIGs (Automotive)	Only official EDIFACT versions of Odette messages available in GEFEG.FX 6.1 were included.	Only latest versions of MIGs have been used if several versions of same MIG were available.	Various releases of EDIFACT (Syntax unknown): D.96A, D.98B, D.03A, D.05B, D.07A	19 / 12
VDA MIGs (Automotive)	Only EDIFACT-based messages available in GEFEG.FX 6.1 were included (VDA 4933-4980, VDA 9001). VDA-messages not based on EDIFACT have been excluded (VDA 4905-4927).	All version 1.0 except VDA 4938 GLOBAL INVOIC (version JAI 3.1, 2010)	D.96A, D.97B, D.04A, D.05B, D.06A, D.07A (all Syntax 3)	32 / 8
JAI Global Messages MIGs (Automotive)	All JAI Global Messages available in GEFEG.FX 6.1 messages were included.	Only Version 2008 (includes JAI 1.1, JAI 2.0, JAI 3.0)	D.03A (Syntax unknown), D.04A (S.3), D.04B (S.3), D.07A (S.3)	6 / 6
EANCOM MIGs (Consumer goods)	All EANCOM messages available in GEFEG.FX 6.1 were included.	Only EANCOM 2002 S4, Edition 2010 (Int.)	D.01B (Syntax 4)	49 / 49
EDIFICE MIGs (Electronics / High-tech)	All EDIFICE messages available in GEFEG.FX 6.1 were included, including messages of non-EDIFACT message types, but based on EDIFACT segments and elements.	Only Version 2011-1.	D.10A (Syntax 4)	24/ 20
Siemens SES MIGs (Technology)	Only EDIFACT-based messages available in GEFEG.FX 6.1 were included. Siemens Y-Messages have been excluded.	Only SES 07/2012.	D.10A (Syntax 4)	23 / 20
EDIFOR MIGs (Transport & Logistics)	All EDIFOR messages available in GEFEG.FX 6.1 were included.	Only latest versions of MIGs have been used if several versions of same MIG were available.	Various releases of EDIFACT (Syntax unknown): D.96A, D.01B	7 / 4
				Total number of MIGs: 160

Mark 3). In the following, we discuss the data set used in this work and provide the explanation on two steps of the KPI identification (i.e., frequency analysis and semantics analysis).

5.3.1 Data Set

While our objective is to identify inter-organizational KPIs that can be extracted from EDIFACT message interchanges in real-world scenarios, it is difficult to obtain real-world message sets from a representative number of organizations that are sufficiently large for research purposes. Therefore, we analyze industry MIGs instead, since we consider them to represent an abstraction and generalization of arbitrary EDIFACT messages exchanged in the real world. MIGs used in industry are usually only available as semi-formal specifications, often in PDF format, which are difficult to be analyzed automatically. However, the commercial software suite GEFEG.FX[1], which is currently the market leader in the field of design tools for EDI standards, provides a formal model of MIGs. In this model, MIGs are defined as subsets of full EDI standards, i.e.,

[1] `http://www.gefeg.com` (visited March 01, 2014)

elements and codes that are marked as optional in the standards may be selectively removed in MIGs. This formalization allows for automated processing of contained segments, data elements and code lists, such as for purposes of set analyzes as conducted in this work. Moreover, the GEFEG.FX software suite also contains a number of concrete instantiations of real-world MIGs that are encoded according to this formal model. For the data considered in this work, we relied on a set of 160 different industrial MIGs that we extracted from MIG sets of various industries contained in GEFEG.FX, Version 6.1 SP1. A detailed description of the data set is shown in Table 5.1.

5.3.2 Method for KPI Identification

Our proposed method for KPI identification consists of two major steps: (i) frequency analysis and (ii) consideration of KPIs based on the semantics of data elements and message types. In the following, these two steps are described in detail.

Analysis of Frequencies. Our proposed method for KPI identification relies on *data elements (DEs)* conveyed in EDIFACT messages. In EDIFACT, such DEs may also be arranged in logical groups called *composite data elements (CDEs)* and *segments*. We assume that DEs that are more frequently used in MIGs are more likely to be available in real-world contexts and, hence, make our results more practical. For this reason, we automatically extract individual DEs, CDEs and segments appearing in the MIGs of the above described data set and count the number of MIGs that contain each individual DE to calculate their number of occurrences. In this chapter, we provide an excerpt example of the complete frequency analysis of all MIGs and use these examples to demonstrate the method for conducting the frequency analysis.

For example, DE *1082 Line item identifier* appears in eight MIGs, therefore the number of occurrences of this DE is eight in total (cf. Table 5.2, row No.3). Similarly, DE *5402 Currency exchange rate* has three occurrences in total (cf. Table 5.2, row No.4). The summary of occurrences for DE *Line item identifier* and DE *Currency exchange rate* are shown in Table 5.3, row No.3 and No.4 respectively. The numbers of occurrences of all DEs are further averaged. These averages are used as a threshold for selecting DEs as candidates for deriving inter-organizational KPIs from them. For example, DE *Line item identifier* is selected as a potential candidate, while DE *Currency exchange rate* is not selected (cf. Table 5.3, row No.3 and No.4, column "Select").

Furthermore, we also extract CDEs and segments. CDEs contain several DEs, such as CDE *C186 Quantity details* (cf. Fig. 5.2a) consisting of DE *6063 Quantity type code qualifier* and DE *6060 Quantity*. Segments may consist of several DEs as well as CDEs. For example, the *NAD* (Name and address) segment (cf. Fig. 5.2b) consists of CDE *C082 Party identification details*, DE *3035 Party function code qualifier*, DE *3039 Party identifier*, DE *3207 Country name code*, etc.

Both CDEs and segments may have so-called *qualifiers* to specify their exact meaning. Qualifiers are DEs which have qualifying functions for specifying the meaning of its qualified CDEs and segments. Consider the following instance of CDE *Quantity details* having as its content two DEs with values "21" and "50" (cf. Fig. 5.2a). This CDE instance contains DE *Quantity type code qualifier* (here: "21") and DE *Quantity* (here: "50"). The DE *Quantity type code qual-*

Table 5.2: Examples of extracted segments, Composit Data Elements (CDEs), Data Elements (DEs) from EDIFACT order message type (ORDERS)

Qualifier Value	Qualified DE	MIG
No.1 CDE: Quantity details (C186), Qualifier: Quantity type code qualifier (6063)		
1 = Discrete quantity	DE 6060: Quantity	EANCOM-2002-S4-Edition2010-International-30-ORDERS
		ODETTE-EDIFACT-latest-versions-14-ORDERS
	DE 6411: Measurement unit code	EANCOM-2002-S4-Edition2010-International-30-ORDERS
		ODETTE-EDIFACT-latest-versions-14-ORDERS
Total occurrences of Quantity type code qualifier “1” = 4		
21 = Ordered quantity	DE 6060: Quantity	EANCOM-2002-S4-Edition2010-International-30-ORDERS
		ODETTE-EDIFACT-latest-versions-14-ORDERS
	DE 6411: Measurement unit code	EANCOM-2002-S4-Edition2010-International-30-ORDERS
		ODETTE-EDIFACT-latest-versions-14-ORDERS
Total occurrences of Quantity type code qualifier “21” = 4		
61 = Return quantity	DE 6060: Quantity	EANCOM-2002-S4-Edition2010-International-30-ORDERS
	DE 6411: Measurement unit code	EANCOM-2002-S4-Edition2010-International-30-ORDERS
Total occurrences of Quantity type code qualifier “61” = 2		
No.2 Segment: Name and address (NAD), Qualifier: Party function code qualifier (3035)		
BY = Buyer	DE 3039: Party identifier	EDIFICE-2011-1-5-ORDERS
		EANCOM-2002-S4-Edition2010-International-30-ORDERS
		EDIFICE-2011-1-6-ORDERS
	DE 1131: Code list identification code	EDIFICE-2011-1-5-ORDERS
		EANCOM-2002-S4-Edition2010-International-30-ORDERS
		EDIFICE-2011-1-6-ORDERS
	DE 3036: Party name	EDIFICE-2011-1-5-ORDERS
		EANCOM-2002-S4-Edition2010-International-30-ORDERS
		EDIFICE-2011-1-6-ORDERS
Total occurrences of Party function code qualifier “BY” = 9		
SE = Seller	DE 3039: Party identifier	EDIFICE-2011-1-5-ORDERS
		EANCOM-2002-S4-Edition2010-International-30-ORDERS
		EDIFICE-2011-1-6-ORDERS
		SES-07-2012-19-ORDERS
	DE 1131: Code list identification code	EDIFICE-2011-1-5-ORDERS
		EANCOM-2002-S4-Edition2010-International-30-ORDERS
		EDIFICE-2011-1-6-ORDERS
	DE 3036: Party name	EDIFICE-2011-1-5-ORDERS
		EANCOM-2002-S4-Edition2010-International-30-ORDERS
		EDIFICE-2011-1-6-ORDERS
		SES-07-2012-19-ORDERS
Total occurrences of Party function code qualifier “SE” = 11		
SF = Ship from	DE 3039: Party identifier	EANCOM-2002-S4-Edition2010-International-30-ORDERS
	DE 1131: Code list identification code	EANCOM-2002-S4-Edition2010-International-30-ORDERS
	DE 3036: Party name	EANCOM-2002-S4-Edition2010-International-30-ORDERS
Total occurrences of Party function code qualifier “SF” = 3		
No.3 DE: Line item identifier (1082)		
-	-	ODETTE-EDIFACT-latest-versions-13-ORDERS
		EDIFICE-2011-1-5-ORDERS
		SES-07-2012-19-ORDERS
		VDA-latest-versions-4-ORDERS
		ODETTE-EDIFACT-latest-versions-14-ORDERS
		EANCOM-2002-S4-Edition2010-International-30-ORDERS
		VDA-latest-versions-19-ORDERS
		EDIFICE-2011-1-6-ORDERS
Total occurrences of Line item identifier = 8		
No.4 DE: Currency exchange rate (5402)		
-	-	EANCOM-2002-S4-Edition2010-International-30-ORDERS
		ODETTE-EDIFACT-latest-versions-13-ORDERS
		ODETTE-EDIFACT-latest-versions-14-ORDERS
Total occurrences of Currency exchange rate = 3		

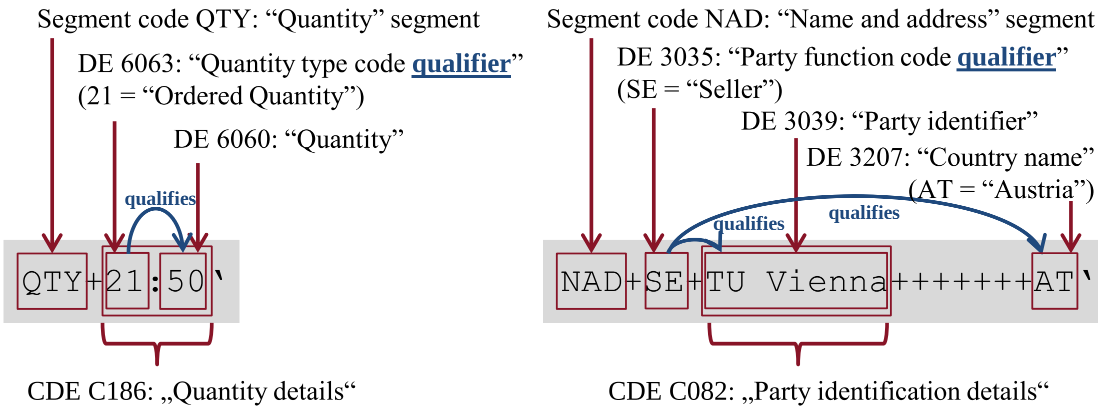
Table 5.3: Selection of data elements based on frequencies of occurrences (using the examples shown in Table 5.2)

Row No.	Segment / Composite Data Element (CDE) / Data Element (DE)	Qualifier	Qualifier Value	Total Occurrences	Select
1	CDE C186: Quantity details	DE 6063: Quantity type code qualifier	1 = Discrete quantity	4	Yes
			21 = Ordered quantity	4	Yes
			61 = Return quantity	2	No
		Average occurrences in ORDERS = 3.33			
2	NAD: Name and address	DE 3035: Party function code qualifier	BY = Buyer	9	Yes
			SE = Seller	11	Yes
			SF = Ship from	3	No
		Average occurrences in ORDERS = 7.67			
3	DE 1082: Line item identifier	-	-	8	Yes
4	DE 5402: Currency exchange rate	-	-	3	No
		Average occurrences of DE without qualifiers in ORDERS = 5.5			

ifier is coded and can be resolved to "Ordered quantity"[2]. In this case, DE *Quantity type code qualifier* qualifies DE *Quantity*. Hence this gives a specific meaning of this DE *Quantity* as an ordered quantity. Similar to qualified CDEs, an example of a qualified *NAD* segment is provided in Figure 5.2b. This segment instance consists of two DEs and one CDE: (i) DE *Party function code qualifier* (here: "SE"), (ii) CDE *Party identification details* consisting of DE *Party identifier* (here: "TU Vienna"), and (iii) DE *Country name code* (here: "AT"). The DE *Party function code qualifier* is coded and resolves to "Seller"[3]. In this case, DE *Party function code qualifier* qualifies the *NAD* segment. This means that every DE in the *NAD* segment is qualified. Hence, in this example the CDEs and DEs contained in this segment instance refer to name and address information of a *seller*. This shows that qualifiers are necessary for considering the semantics of qualified CDEs and qualified segments. Hence, for CDEs and segments found in MIGs we consider qualifiers in our method for identifying KPIs.

Therefore, for each CDE and segment we investigate whether there is a corresponding qualifier based on a heuristic rule: *If a data element stands at the beginning of a CDE or at the beginning of a segment and its name ends with the string "qualifier" (ignoring character case), then it is considered to qualify the remaining data elements of that CDE or of that segment (including data elements embedded in nested CDEs), respectively* [Eng+13a]. According to this rule, we extract CDEs and segments with their related qualifiers appearing in MIGs and select the qualifier based on the frequency of occurrences. Table 5.2 shows examples of CDEs and segments found in MIGs of the purchase order message type (ORDERS). For example, the qualified CDE *Quantity details* is found together with its qualifier, i.e., DE *Quantity type code qualifier* (cf. Table 5.2, row No.1). In the data set of the ORDERS MIGs, the DE *Quantity type code qualifier* is found with the following values: "1" (Discrete quantity), "21" (Ordered quantity), and "61" (Return quantity)". "Discrete quantity" is found to qualify DE *Quantity* and DE *Measurement unit code* in the ORDERS message type in EANCOM MIG set and Odette EDIFACT MIG set. In total, we found two MIGs containing "Discrete quantity" which qualifies DE *Quantity* and two MIGs containing "Discrete quantity" which qualifies DE *Measurement unit code*.

[2] `http://www.unece.org/trade/untdid/d12b/tred/tred6063.htm` (visited March 01, 2014)
[3] `http://www.unece.org/trade/untdid/d12b/tred/tred3035.htm` (visited March 01, 2014)



(a) Qualified CDE

(b) Qualified segment

Figure 5.2: Example of qualification

Therefore, in total the frequency of occurrences of “‘Discrete quantity” equals four times. The other qualifiers (i.e., “21”, and “61”) are extracted similarly, as shown in Table 5.2, row No.1 (i.e. “Ordered quantity” appears four times and “Return quantity” appears two times). Moreover, this extraction method is also applied for segments as depicted in Table 5.2, row No.2 for *NAD* segment. In the example, three values of DE *Party function code qualifier* are found: “BY” (Buyer), “SE” (Seller), and “SF” (Ship from). They qualify DE *Party identifier*, DE *Code list identification code*, and DE *Party name* which belong to *NAD* segments. In summary, there are nine MIGs containing “BY”, eleven MIGs containing “SE”, and three MIGs containing “SF”.

Similar to individual DEs, after extracting qualifiers of CDEs and segments we further select the candidate qualifiers based on their frequencies of occurrence. The frequencies of qualifiers are averaged and qualifiers which have frequencies above or equal to the average value are selected for further consideration. For example, Table 5.3, shows examples for qualifiers belonging to CDE *Quantity details* (cf. Table 5.3, row No.1) and *NAD* segment (cf. Table 5.3, row No.2). Among the qualifiers of CDE “Quantity details”, “Discrete quantity” and “Ordered quantity” are selected for further consideration because they both have frequency of occurrences above the average. Similarly, among the qualifiers of *NAD*, “Buyer” and “Seller” are selected for further consideration.

Analysis of Semantics. From the data processing and selection process, we found 52 message types. Each message type contains different DEs, and qualifiers for qualified CDEs and segments. The consideration of KPIs is based on the semantics of message types and the semantics of DEs, qualified CDEs and qualified segments. The semantics of message types correspond to their purposes. We investigate the purpose of those 52 message types in the EDIFACT standard which can be found in the directories of United Nations Directories for EDI for Administration, Commerce and Transport (UN/EDIFACT)[4]. From the investigation, we derive the message exchange directions as well as business parties playing key roles in the information exchange. The summary of message exchange is shown in Figure 5.3. There are seven main parties identified

[4] `http://www.unece.org/trade/untdid/d12b/trmd/trmdi2.htm` (visited March 01, 2014)

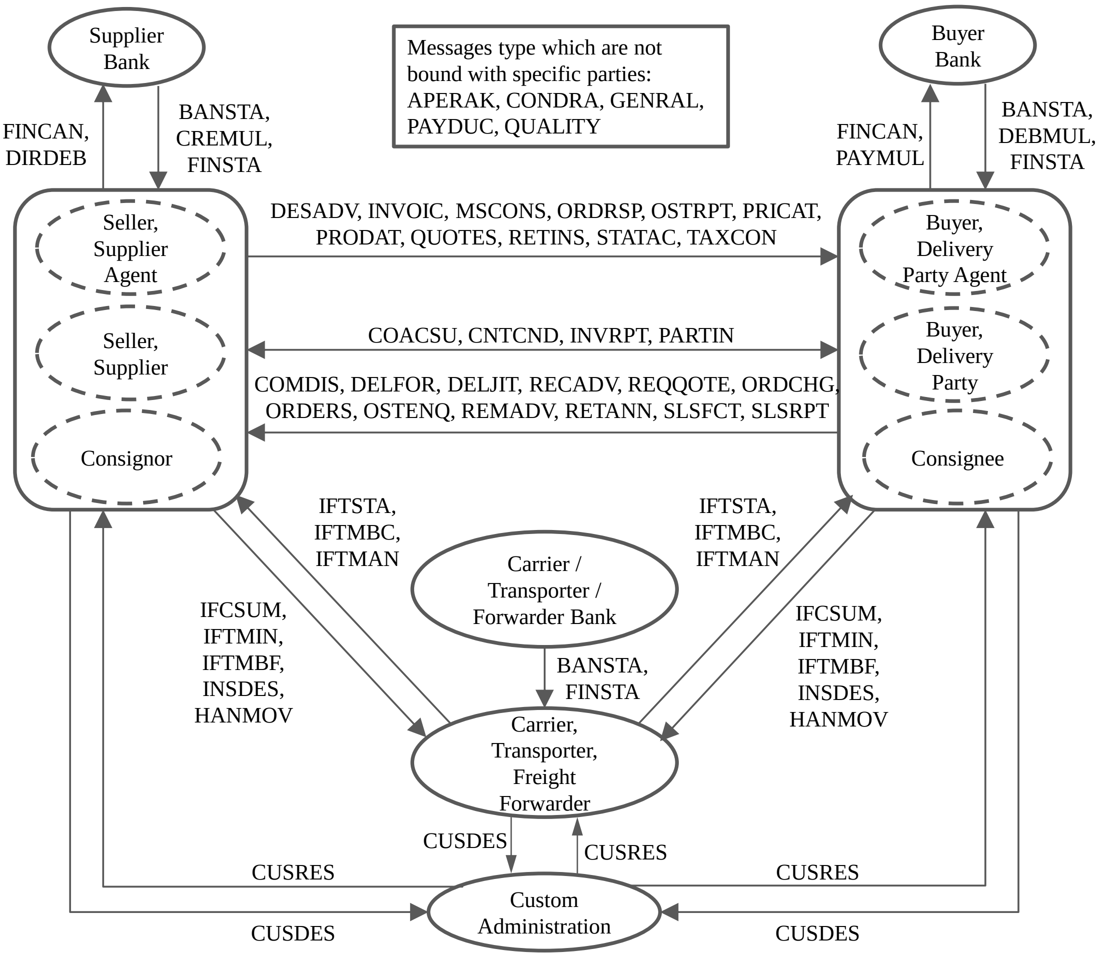


Figure 5.3: EDIFACT message types sent between business parties

from these 52 message types. They are supplier or seller, buyer or delivery party, transporter, bank of supplier, bank of buyer, bank of transporter, and custom administration. The arrows represent the exchange direction. For example, the arrow from supplier to supplier bank indicates the direction of the exchanged *Financial cancellation (FINCAN)* message and *Direct debit (DIRDEB)* message. This means that FINCAN and DIRDEB are usually sent from supplier to supplier bank. In addition, the semantics of DEs rely on themselves, e.g., DE *Line item identifier* simply refers to the number or identity of line item. However, the semantics of CDEs and segments are based on their related qualifier as explained earlier.

Based on the message exchange scenario depicted in Figure 5.3 and the frequency analysis described earlier, KPIs can be defined. We distinguish two types of KPIs: *primary KPIs* directly derived from data elements and *secondary KPIs* aggregated from other KPIs. Primary KPIs are identified by considering the matching of the semantics of the message type and the semantics of DEs, CDEs, and segments. For example, in the ORDERS message, “Ordered quantity” and

"Discrete quantity" are considered (cf. Table 5.3, row No.1). Since the purpose of the ORDERS message is to provide the order information sent from buyer to seller (cf. Fig. 5.3), the "Order quantity" is expected to be present in the messages. In this way, we conclude that the semantics of "Ordered quantity" matches with the semantics of the ORDERS message. Hence, we define "Ordered quantity" as a primary KPI which can be calculated confidentially from "Ordered quantity" in the ORDERS message. In contrast, the "Discrete quantity" is not considered as potential data for calculating the KPIs since the semantics of "Discrete quantity" is ambiguous. In particular, it could possibly refer to any kind of quantity. In this case, we exclude "Discrete quantity" from our consideration. Furthermore, secondary KPIs are derived by considering primary KPIs together with DEs, CDEs, and segments. This depends on personal and professional experience. Finally, we aggregate the resulting KPIs from this work into inter-organizational success factors found in our previous literature review. The result is explained in detail in the following.

5.3.3 Results and Evaluation

By following the method explained above, primary KPIs and secondary KPIs are identified. We provide an excerpt of primary and secondary KPIs, as shown in Table 5.4. The complete set of resulting KPIs can be found in the Appendix A. As mentioned before, we aggregate the identified KPIs into inter-organizational success factors. In order to define inter-organizational success factors we previously conducted a systematic literature review as explained earlier in Chapter 4. According to our consideration from the measurements of identified success factors found in literature review, there are twelve success factors related to IORs which can be measured by the KPIs derived from EDIFACT messages. In the following, these success factors as well as corresponding KPIs for measuring those success factors are discussed. Moreover, the discussion on the evaluation of the identified KPIs are also provided.

Satisfaction or customer satisfaction [CP08; Che11] is an emotional response to the difference between what customers expect and what they actually receive. Satisfaction seems intangible and difficult to measure directly since it is mainly about emotion. One of the most quantifiable KPIs widely applied for measuring this success factor is customer retention rate. It refers to the ratio of the number of retained customers to the current number of total customers. However, ordered quantity as well as return quantity of products can also imply customer satisfaction. Since satisfaction can influence order decisions of customers, the order quantity can indirectly reflect customer satisfaction. Also, the quantity of returned products can be considered as a signal of losing satisfaction.

Financial performance analysis is one of the most widely applied methods for analyzing business performance. It is also applied in the inter-organizational context especially in supply chains [CGP11; Hsu+08]. Financial performance typically focuses on outcomes related to economics, such as profitability, cost and revenue. Therefore, the information related to monetary amounts is commonly used for evaluating financial performance, such as invoiced and paid amount (cf. Table 5.4, financial performance). Furthermore, such KPIs could be translated into cost and revenue. The translation depends on the role of the participants. For example, the paid monetary amount found in *Remittance advice message (REMADV)* could be translated to either cost, if the *NAD* (Name and address) or *FII* (Financial institution information) segment are qualified by the value “BY” (Buyer), or to revenue, if the *NAD* segment or *FII* segment are qualified by the value “SE” (Seller).

Table 5.4: Excerpt of inter-organizational KPIs identified from EDIFACT messages

KPI	Mapping to EDIFACT data / Literature Support
Satisfaction	
Ordered quantity[s,a,p]	Ordered quantity (Quantity in QTY qualified by value 21) from ORDERS, or INVOIC messages Support: [SK10; SG06]
Returned quantity[s,a,p]	Returned quantity (Quantity in QTY qualified by value 61) from INVOIC, RETANN, or INVRPT messages Support: [SG06; BS07]
*Customer retention rate	i) Message interchange sender or party identification (Party identification in NAD qualified by value BY) from ORDERS messages ii) Calculation: (Total customers − Number of new customers)/Total customers of pervious period Support: [HHS05; CGH09]
Financial performance	
Invoiced amount[s,a]	Invoiced amount (Monetary amount in MOA qualified by value 39 or 77) from INVOIC, or BANSTA messages
Paid amount[s,a]	Paid amount (Monetary amount in MOA qualified by value 11) from FINSTA, REMADV, or COACSU messages
*Revenue[s,a]	i) Payable, invoice line item, original, total payment and tax/duty amount (Monetary amount in MOA qualified by value 9, 77, 98, 139 and 161 respectively) from BANSTA, CREMUL, DEBMUL, COACSU or REMADV messages ii) The related business parties are seller, supplier, beneficiary and payee from NAD or FII qualified by value SE, SU, BE and PE respectively) Support: [Yus+04; HS05; BS00; CF05; MG05; SG06; HHS05; CGH09; KS03; Yeu08; GWI08; FHZ10; Li+06; RSK09; CP04]
Failure	
Number of application errors[s,a,p,c]	ERC in APERAK messages Support: [SMN09]
Reliability	
Lost goods quantity[s,a,p]	Lost goods (Quantity in QTY qualified by value 126) from the INVOIC message Support: [CQ03; SK10; SG06; Ara+07]
*On-time delivery[c,a,p]	i) Expected delivery date/time (Date/time in DTM qualified by value 10, 2, or 191) from ORDERS, DELFOR, DESADV, or DELJIT messages ii) Actual delivery date/time (Date/time in DTM qualified by value 11, 50, or 310) from DESADV, or RECADV messages iii) Calculation: COUNT(Deliveries arriving before or on the expected delivery date/time) Support: [WLN10; HSW04; JZ03; CQ03; AML10; CMM10; PSR01; Koh+12; NWL10; ZH09; SK10; PLC08; LNC02; SG06; Cha03; CGH09; BB10; BS07; CLS08; Yeu08; KT02; IGF04; Bag+05; SCW00; ZB07]
Adaptability/flexibility	
Quantity not available for despatch (fill rate)[s,a,p]	Quantity not available for despatch (Quantity in QTY qualified by value 255) from DESADV messages Support: [WLN10; HSW04; CQ03; OK03; SK10; LNC02; SG06; Cha03; Ara+07; KS03; Bag+05; ZB07; GWI08]
*Quantity ready for order[a,p]	i) Quantity on hand (Quantity in QTY qualified by the value 17) from QUOTES messages ii) Ordered quantity (Quantity in QTY qualified by the value 21) from ORDERS or QUOTES messages iii) Calculation: Quantity on hand − Ordered quantity Support: [PP10; OK03]
Customer Responsiveness	
Lead time[a]	Lead time (Date/time in DTM qualified by value 169) from DESADV messages Support: [WLN10; HSW04; JZ03; OK03; ZH09; SK10; SG06; SG06; Ara+07; BB10; BS07; SMN09; GPT01; CLS08; Yeu08; IGF04; Bag+05; SCW00]

*Time of order response[a]	i) Order date/time (Date/time in DTM qualified by 137) from ORDERS messages ii) Order response date/time (Date/time in DTM qualified by 137) from ORDRSP messages iii) Calculation: Order response date/time − Order date/time Support: [CQ03; BS00; AML10; ZH09; PLC08; SG06; KT02; CPL04; CP04]
Information Sharing	
Inventory report exchange[c]	INVRPT messages Support: [BS00; Wie+10; SK10; Zha02; ZB07]
Information Quality	
*Forecast accuracy[a,p]	i) Forecast/reserved quantity (Quantity in QTY qualified by value 247, or 248) from SLSFCT messages ii) Ordered quantity (Quantity in QTY qualified by value 21) from ORDERS messages iii) Calculation: Forecast/reserved quantity − Ordered quantity Support: [WLN10; PP10; CMM10; NWL10; SK10; SG06; BS07; GPT01; CWR08; IGF04; RGS11]
Credibility	
Disputed amount[s,a]	Disputed amount (Monetary amount in MOA qualified by value 257) from COMDIS message
Discount amount[s,a]	Discounted amount (Monetary amount in MOA qualified by value 52) from INVOIC, or QUOTES messages
*Contract violation (order quantity)[c]	Ordered quantity (Quantity in QTY qualified by value 21) from ORDERS messages ii) Minimum, maximum, or committed quantity (Quantity in QTY qualified by value 53, 54, or 66 respectively) from CNTCND messages iii) Calculation: COUNT if (Ordered quantity $<$ Minimum quantity or Committed quantity) or COUNT if (Ordered quantity $>$ Maximum quantity)
Loyalty	
Committed quantity[s,a,p]	Committed quantity (Quantity in QTY qualified by value 66) from CNTCND messages
Contract	
Contract exchange rate[c]	CNTCND messages
Customer Uncertainty	
Order change rate[c]	ORDCHG messages
*Changed order quantity[s,a,p]	i) Ordered quantity (Quantity in QTY qualified by value 21) from ORDCHG or ORDES messages ii) Calculation: Previous ordered quantity − Ordered quantity

Note 1: The superscript s,a,p,c on KPI names indicate applicable aggregation functions: sum, average, percentage, count.
Note 2: The superscript * denote the secondary KPIs.
Note 3: Segments and message types mentioned in this table are represented as abbreviation codes of EDIFACT release D10A. Full descriptions are provided in the following links:
`http://www.unece.org/trade/untdid/d10a/trsd/trsdi1.htm`
`http://www.unece.org/trade/untdid/d10a/timd/timdi1.htm`

Failure [CGP11] is one of the success factors in operational performance. The only information related to failure could be found in the *Application error and acknowledgment (APERAK)* message. Therefore, the KPIs identified from this data set reflect the failure at the application level, such as the number of application errors.

Reliability [HSW04] mostly focuses on delivery performance such as whether the delivery is on time, or how often the delivery is delayed, etc. Hence, the identified KPIs mainly focus on the delivery perspective including the completeness of goods/products delivered to customers (e.g. damaged/destroy goods, lost goods, etc.) and the number of on-time deliveries.

Adaptability/flexibility focuses on the ability of organizations to adapt themselves to changes or customer's requirements. In [JZ03] and [WE04], flexibility is the ability to deal with uncertainty, modification, or any variety of customer needs. Hence, we consider the readiness of goods/products that can be served for the customers orders as the KPIs reflecting this aspect. Such KPIs show the ability of organizations to handle the orders of customers.

Customer responsiveness tends to analyze how fast organizations respond to customer requests [HSW04; CPL04]. Requests and responses between business partners are usually found in the context of EDI messages since they require the exchange of information. As expected, we found a number of message types having these purposes (e.g., requesting and responding) including *Order (ORDERS)*, *Order response (ORDRSP)*, *Request for quote (REQOTE)*, and *Quote (QUOTES)* messages. Hence, we identify time of response for each pair of requesting message and responding message as one of the KPIs reflecting customer responsiveness. The other example is lead time which is the duration from the placement of an order until the goods are received. This can also reflect how fast an organization responds to its customers orders.

Information sharing tends to measure whether organizations and their business partners keep informing each other about changes or any information that affect their business [Li+06]. Therefore, the exchange of the information that is not related to business transactions such as sales forecast and inventory level can be considered to reflect information sharing between business partners. For instance, the number of inventory reports exchanged may reflect the success factor information sharing.

Information quality, in contrast to information sharing, focuses on the quality of the exchanged information such as accuracy, completeness and timeliness [LL06]. In other words, the information shared between business parties should be accurate, complete, and adequate for supporting decision making. This is difficult to determine from a single data element. However, by considering several data elements together we suggest that the accuracy can be determined by comparing forecast/plan information to the actual transactional information. For example, comparing sale forecast quantity to the actual order quantity. Moreover, the duration of delivery preparation (i.e. time from the delivery request notification until the requested delivery date) can also reflect whether the information is timeliness (i.e., is there enough time for the supplier to provide goods/products for an order).

Credibility [WLN10; AML10] reflects the belief of an organization towards its partners that they will perform the task as expected. Therefore, contract violation is suitable for measuring credibility. The disputed monetary amount possibly also indicates credibility of suppliers since disputed amount is usually raised by customers when they notice something unacceptable (e.g. invoiced amount, payment amount). In this case, suppliers may loose their credit. In contrast, allowance and discount could also help suppliers gain more credit.

Loyalty is determined in terms of the intention to continue doing business or buying product from an organization [GFC09]. Committed quantity can reflect this success factor, since it indicates the quantity of products that customers intend to buy or suppliers intend to provide. For example, large committed quantity can be interpreted that customer still tend to do a business with its supplier. In contrast, once the committed quantity is noticeably decreased, it might be a sign of detachment.

Contract [LLL09] is an agreement between partners which is defined in a formal way. Contract is considered as one mechanism for creating structural systems to which all partners must comply. The measurement of contract is quite straight forward by examining whether all agreements and obligations are formally defined and followed by all agreed parties. In the context of EDIFACT messages, there is *Contractual conditions message (CNTCND)* which provides the contractual information of business parties in order to enable an automation of transactions.

From this, the number of contracts, and the contract rate (frequency of contract generated) could reflect this success factor.

Customer uncertainty [LL06] is uncertainty from the customer's side. It refers to the unpredictable behavior of customers such as the change of orders and product requirements. Thereby, the number of order changes as well as changed ordered quantities could directly reflect customer uncertainty.

Evaluation of KPIs

The identified KPIs are evaluated through the support of literature. In particular, we assess those KPIs by answering the question of whether they are applied as indicators of performance and/or indicators of the success of IORs in the B2B context.

In doing so, we investigate measurement items (e.g., measures, and questions or scale items asked in surveys or interviews) of success factors found in the set of literature which is previously selected for the review presented in Chapter 4. If the KPIs are directly applied as an indicator, or mentioned as a determinant of success factors in a literature, the literature is considered as a support for those KPIs. For example, the study of measuring supply chain performance in [SG06], [BS07], [CGH09], and others explicitly mention as well as apply on-time delivery as a performance measure. Therefore, those studies are considered as supporting evidences for using the KPI on-time delivery. Hence, they are consequently added as a support for on-time delivery (cf. Table 5.4). Moreover, some studies describe the measurement of success factors through questions in a survey. Considering these questions, in case the literature contains questions related to the identified KPI, the literature is added as an additional support for the KPI. In the following an example is provided.

For instance, the study of supply chain practice and information sharing of Zhou and Benton Jr. [ZB07] used survey questions to collect data. One of the questions related to information sharing is "How often does your major customer electronically provide your firm with its information in inventory level or future demand forecasting?". This question is explicitly related to the frequency of exchanging inventory reports and sale forecasts. Therefore, the study of Zhou and Benton Jr. [ZB07] is added as a support for both the KPIs of inventory report exchange and sale forecast exchange (cf. Table 5.4).

According to the investigation, 19 out of the 53 primary KPIs and 15 out of the 18 secondary KPIs are confirmed that they have been applied as an indicator related to the success of IORs. Most of the secondary KPIs have strong support for their application in reality, whereas a few of primary KPIs are actually applied. The result shows that performance measurement in reality relies more on complex KPIs (i.e., secondary KPIs) derived from several aspects rather than primary KPIs which are calculated solely based on a single attribute. However, some primary KPIs are necessary for calculating complex KPIs. As shown in the result, a lot of primary KPIs considering financial performance (e.g., payable amount, final posted amount, invoiced amount, etc.) have no support from literature. However, they can be transformed or aggregated as revenue, profit, or cost. Moreover, the KPIs regarding credibility, loyalty, contract, and customer uncertainty have no support as well. These success factors are measured by using questions which are difficult or unable to be translated into quantifiable measures. Therefore, those KPIs are still unable to be claimed as an acceptable measure. Nevertheless, they can be considered as the first inspiration of quantifiable measurements for their corresponding success factors.

5.4 Summary

In this chapter we presented a method for identifying inter-organizational KPIs from business data in EDIFACT messages. The proposed method takes the frequencies of data elements into account, as well as the semantics of both, data elements and message types, in order to derive accurate results. We applied this method on a data set of real-world industry MIGs, presented a set of derived inter-organizational KPIs, as well as described guidelines for their calculation based on concrete EDIFACT data. Furthermore, we presented aggregations of these KPIs in order to define quantitative measurements for inter-organizational success factors. The work presented in this chapter has also been published in [Kra+13]. Moreover, we evaluated the identified KPIs by finding evidences of their application in reality from existing literature.

The resulting set of KPIs identified in this work answers the second research question: *"What are inter-organizational KPIs that can be derived from EDIFACT messages?"*. The KPIs, the inter-organizational success factors defined in the previous chapter, as well as their influencing relationships are stored in a knowledge base. This knowledge base serves as a basis for analysis tasks. In particular, this knowledge base is used as one of the components of our inter-organizational performance analysis framework. Mapping KPIs to actual data elements in EDIFACT messages allows automating the discovery and calculation of KPIs, and consequently, fosters automation of the performance evaluation. The detailed explanation on the application of this knowledge is provided in Chapter 6, Section 6.3.2.

CHAPTER 6

Inter-organizational Performance Analysis Framework

The objective of this research is not only to provide the quantitative measurement for IORs through inter-organizational KPIs, but also to lift performance measurements onto a strategic level. This has been introduced earlier as the third research question (cf. Chapter 1.1.3). In this chapter, we propose the framework for inter-organizational performance evaluation which applies quantitative KPIs as well as allows connecting those KPIs to organizations' strategies. Particularly, as introduced in Figure 1.6, the main idea of the framework is to connect business information concepts derived from EDI messages, KPIs, success factors, and business objectives for enabling the evaluation of IORs on a strategic level. The framework allows modeling and defining KPIs based on business information, and connecting these KPIs to business objectives through success factors. The detailed discussion of the proposed framework is provided in the following.

6.1 Motivation

Applying inter-organizational KPIs derived from EDI messages makes the evaluation of inter-organizational performance quantifiable, and in turn allows reflecting IORs tangibly. However, the KPIs derived from EDI messages are considered as performance measures at an operational level. To analyze and monitor business performance against business strategies or objectives, those KPIs must be linked to corresponding strategies. The alignment between KPIs and business strategies enables performance assessment against high-level business objectives. This helps managers or strategists to understand the impact of operational results on IORs.

To this end, we propose a framework for evaluating IORs by means of inter-organizational business performance analysis from EDI messages. On the one hand, the framework supports the identification of relevant KPIs from business information contained in EDI messages. Besides business information conveyed in EDI documents, the framework also takes the information of inter-organizational processes refer to as *choreography* into account. Such information includes event logs, and choreography models derived from the exchange of EDI messages. These sources of information are discovered by applying our business information extraction framework (cf. Chapter 3) as well as adapted process mining techniques [Eng+12a; Aal11; Aal+07] on EDI messages.

On the other hand, the framework allows defining business objectives following the Balanced Scorecard (BSC) method. It supports the alignment of identified KPIs and business objectives which enables the evaluation of inter-organizational business performance. The framework is realized as a plug-in for ProM [Don+05].

The main contribution of the work presented in this chapter is the integration of process mining techniques adapted for inter-organizational information exchange, i.e., a bottom-up approach, and the BSC method, i.e., a top-down approach. In particular, by deriving process models from the exchange log of EDI messages through process mining yields the actual process models used in real-world scenarios rather than the ones that are predefined. Therefore, deriving and calculating KPIs based on such mined models gives us the accurate results reflecting reality. While such bottom-up approach provides accurate performance results, the top-down BSC approach connects these performance results to strategies. In other words, it provides a way to interpret performance results as an achievement of organizations' strategies.

This chapter is organized as follows: Section 6.2 provides an overview on related work. In Section 6.3, the framework is described in detail including the architecture as well as the technical realization. A demonstration of the framework has been performed based on two case studies using real-world data which are presented separately in Chapter 7.

6.2 Related Work

Previously, we proposed the use of KPIs calculated from EDI data for the evaluation of IORs in order to improve quantifiability and explicitness. As presented in Chapter 4 and 5, we investigated appropriate success factors and ways of measuring them by conducting a literature review on factors influencing IORs as well as their measurement. Based on the success factors and their measurements found, we identified KPIs from EDI (UN/EDIFACT) messages and aggregated them into success factors.

Recently, the work of Engel et al. [Eng+12a] applied process mining techniques to discover process models from EDI messages for deriving KPIs reflecting business process performance. ProM provides several plug-ins supporting analysis based on low-level log data (e.g., ILP Miner [Wer+08], α-Miner [AWM04], performance analysis through process mining [Hor07]) as well as business data (e.g., data-aware process mining [LA13]). Results from process mining can also be applied for in-depth analysis of business processes for answering specific business-related questions. As in [ER14], they use the mined model of an inter-organizational purchasing process as well as related business information (e.g., requested and actual delivery dates, ordered quantities, etc.) for answering questions related to the operational performance regarding deliveries (e.g., "Which line items take the longest to deliver?", "Does the delivery time of line items vary depending on the delivery point?", etc.). These results can be used as KPIs for monitoring business performance since KPIs are defined based on information that actually appears in real business scenarios. However, a drawback of these bottom-up approaches is that they usually fall short of accurately reflecting business success on a strategic level.

Top-down approaches for performance analysis suggest to base the performance evaluation on business objectives and to translate these objectives into measurements which enables a performance evaluation on the business level. The BSC method [KN92] is a widely applied top-down measurement system [Eck06]. As discussed in state-of-the-art (cf. Chapter 2, Section 2.2), there are also several works on applying the BSC framework in inter-organizational contexts such as Supply Chain Management (SCM). For instance, Brewer et al. [BS00] and Bullinger et al. discuss the interrelationship between the BSC method and the SCM field and introduce approaches for supply chain performance analyzes based on BSCs. Kleijnen et al. [KS03] and Chia et al. [CGH09] study examples of KPIs commonly used for measuring supply chain performance following the BSC paradigm. However, top-down approaches are difficult to implement since business objectives and/or strategies are often too broadly defined and, hence, too ambiguous to identify appropriate KPIs. Therefore, best practice in the BSC framework suggests to align business strategy with KPIs through critical success factors [KN04]. However, success factors are still required to be translated into quantifiable KPIs in order to enable quantitative performance analysis.

In this work we introduce the performance analysis framework which integrates the process mining and the BSC approach. The framework is implemented by applying semantic technologies. In particular, we develop the BSC ontology for supporting the modeling of BSC elements and the automatic calculation of BSC models. Best to our knowledge, there are two related works on the BSC ontology. Bobillo et al. [Bob+09] proposed a fuzzy BSC ontology (fBSCO) supporting fuzzy BSC implementation. However, the fBSCO focuses on BSC perspectives and defines other BSC elements (e.g. business objectives, KPIs, etc.) as fuzzy variables which purposely supports fuzzy implementation. Furthermore, the relationships of each BSC elements are not explicitly provided in the fBSCO. Similarly, the ontology provided in [Nav+06] describes only the concept of BSC elements without any detail of their relationships and attributes. In contrast, our BSC ontology presented in this paper is designed for supporting full BSC implementation including the modeling of BSC elements including their relationships as well as the calculation of BSC models.

6.3 EDImine BSC Framework for Inter-organizational Performance Analysis

On the one hand, the bottom-up analysis of business (process) performance may fall short in accurately reflecting business success on a strategic level. One the other hand, top-down approaches tend to be difficult to implement. We suggest that the top-down and bottom-up approaches may complement each other. In particular, the bottom-up approach is utilized for defining the KPIs from real data which facilitates the calculation of KPIs in concrete use cases. Subsequently, the top-down approach allows aligning business objectives with KPIs which enables the evaluation of business performance of IORs on a strategic level. The framework proposed for analyzing business performance, namely the *EDImine BSC Framework*, integrates a bottom-up and a top-down approach. Figure 6.1 illustrates a high-level overview of the frame-

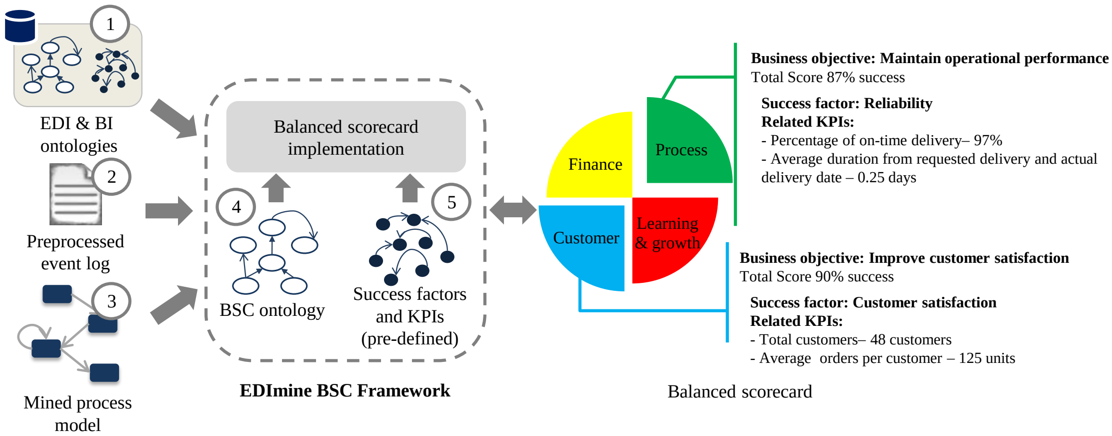


Figure 6.1: The EDImine BSC Framework

work. The framework is built upon (i) EDI & BI (Business Information) ontologies (cf. Fig. 6.1, Mark 1) [Eng+12b; Kra+12a], (ii) event logs (cf. Fig. 6.1, Mark 2), and (iii) process models (cf. Fig. 6.1, Mark 3). Figure 6.2 shows EDI data preprocessing resulting in those three artifacts. EDI messages as well as their contained values are stored and parsed as EDI ontologies (cf. Fig. 6.2, Mark 1) following the conceptualizing approach introduced by Engel et al. (cf. Chapter 2, Section 2.1.2). EDI ontologies are further extended with business information concepts by applying our information extraction approach presented in Chapter 3 (cf. Fig. 6.2, Mark 2). This yields EDI & BI ontologies. Furthermore, an even log is derived by using the EDI/event mapping and event correlation approach with the toolset presented in [Eng+13b; ER14]. Finally, process mining techniques such as *Heuristics Miner* algorithms [WAA06] are employed for discovering inter-organizational process models (cf. Fig. 6.2, Mark 4). Optionally, the derived event log can also be further preprocessed for filtering interested process instances, or cleaning undesired pro-

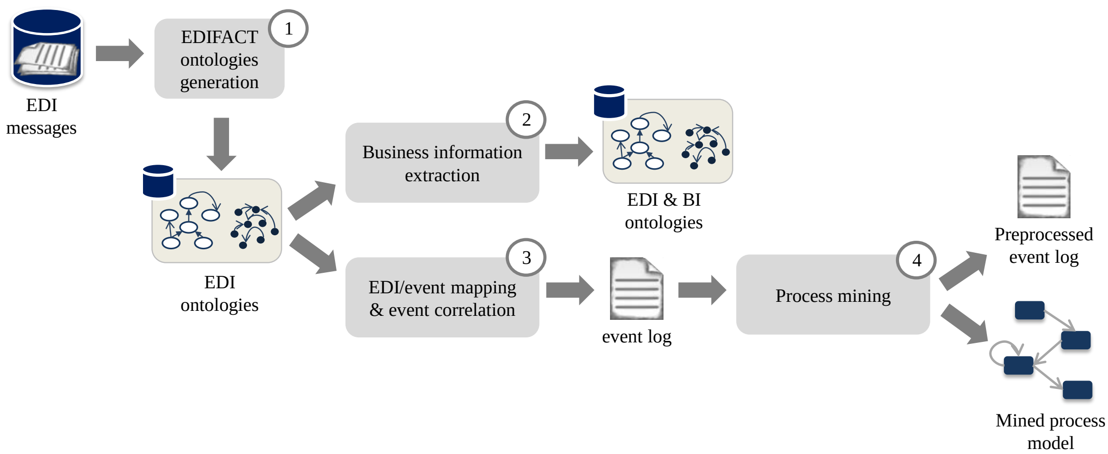


Figure 6.2: EDI data preprocessing

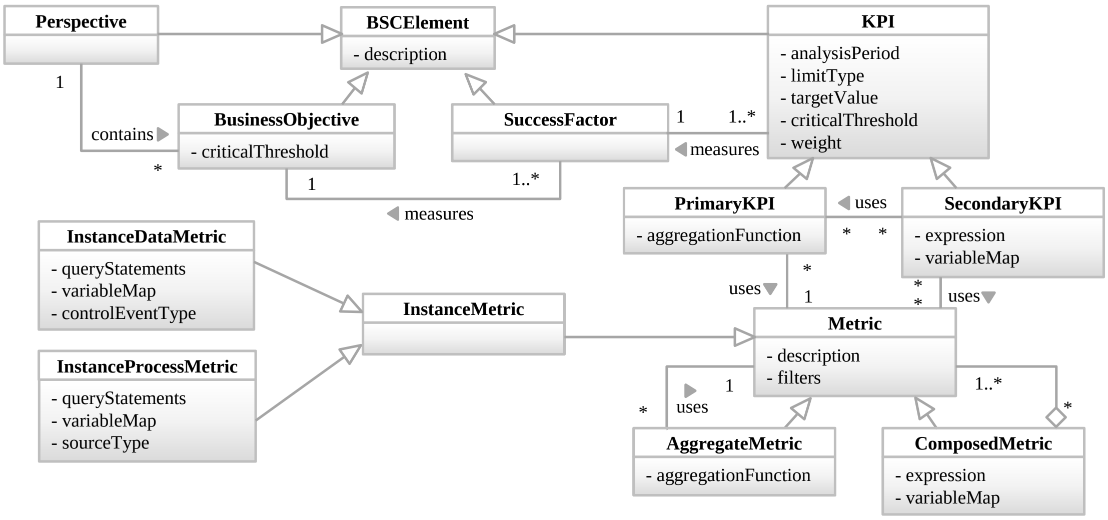


Figure 6.3: The balanced scorecard (BSC) ontology based on [WML08] as a UML class diagram

cess instances or events, etc. These preprocessed artifacts serve as data sources covering both, the business data and the process perspective, for supporting performance analysis.

The framework consists of the BSC ontology (cf. Fig. 6.1, Mark 4) and a set of predefined success factors and KPIs (cf. Fig. 6.1, Mark 5). The BSC ontology provides a conceptual description of BSC elements including business objectives, success factors, and KPIs. The BSC ontology serves two purposes. First, it allows the modeling of definition and calculation of KPIs based on the aforementioned data sources. Second, it supports the alignment of KPIs and relevant business objectives according to the BSC method. Having a predefined set of success factors and KPIs available, allows to automatically identify available KPIs according to the business information available in the data sources.

In the following, the BSC ontology and the BSC calculation, including the method for calculating the scores of business objectives and KPIs, are described in detail. The knowledge base storing the set of predefined success factors and KPIs described in Chapter 4 and 5 is also provided along with ontological rules for automatic KPIs suggestion.

6.3.1 BSC Ontology

Figure 6.3 shows a simplified view of the BSC ontology represented as a UML class diagram. Please note that we represent the ontology as an UML class diagram because, in this section, it is necessary to describe BSC concepts together with their attributes. Attributes of concepts in the BSC ontology are designed based on the calculation method employed in the framework. In other words, these attributes are necessary for calculating the score of KPIs and business objectives. In our work, we do not calculate the score of success factors since success factors are only applied as a mediator to connect KPIs and their corresponding business objectives. In the following, we explain the BSC ontology along with the example provided in Figure 6.4.

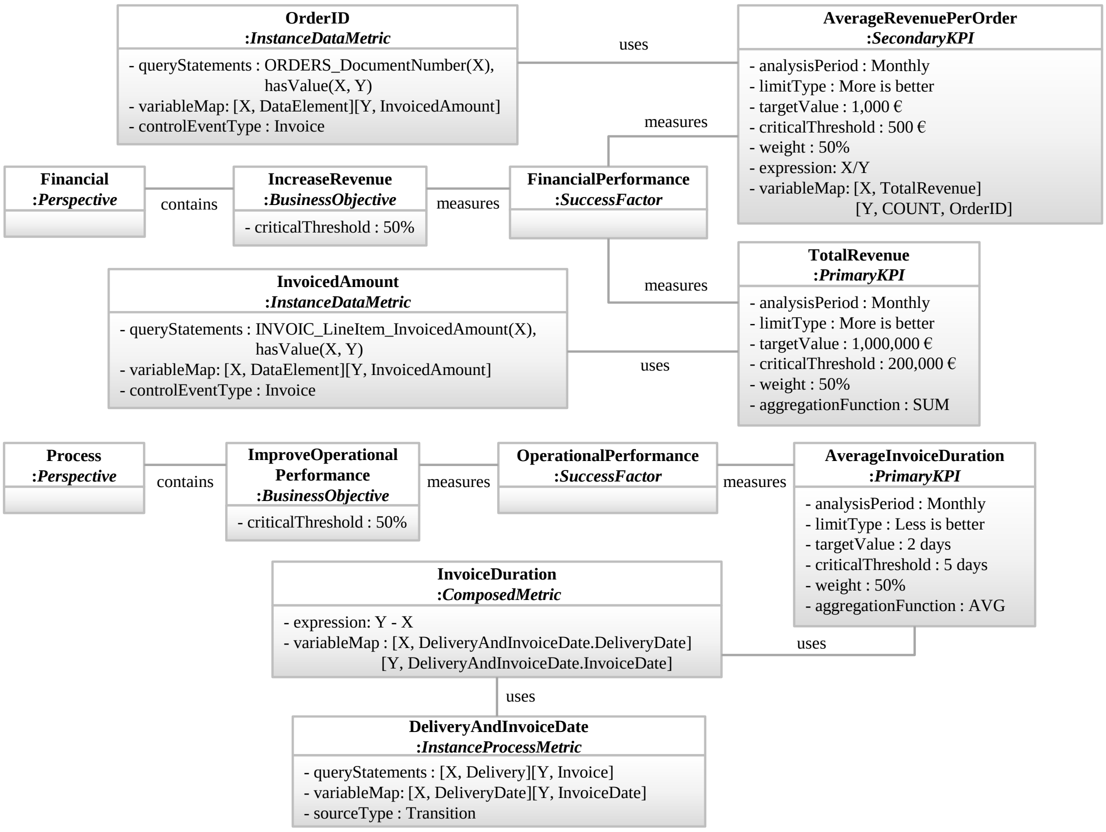


Figure 6.4: Example of an instance of the BSC ontology

In the BSC ontology, BSC elements are comprised of perspectives *(Perspective)* (e.g., finance, customer, process, learning and growth), business objectives *(BusinessObjective)*, success factors (*SuccessFactor*), and *KPI*s. A perspective contains related business objectives. A business objective can be measured by success factors which are measured by quantifiable KPIs. In other words, success factors are used as mediators to connect business objectives to KPIs. For example, Figure 6.4 describes a BSC model instance conforming to the presented BSC ontology. It illustrates a BSC model consisting of the `Financial` and `Process` perspective, the business objective `IncreaseRevenue` and `ImproveOperationalPerformance` , the success factor `FinancialPerformance` and `OperationalPerformance`, as well as the KPI `TotalRevenue`, `AverageRevenuePerOrder`, and `AverageInvoiceDuration`. In particular, the `Financial` perspective contains the business objective `IncreaseRevenue`. The business objective `IncreaseRevenue` is measured quantitatively by the KPI `TotalRevenue` and `AverageRevenuePerOrder` based on the success factor `FinancialPerformance`. The `Process` perspective contains the business objective `ImproveOperationalPerformance` which is measured based on the KPI `AverageInvoiceDuration` through the success factor `OperationalPerformance`. For the calculation of a business objective, we further define a critical threshold as an attribute of business objective as follows:

- **Critical threshold.** The critical threshold of a business objective is a percentage value that defines the lowest percentage of achievement that the business objective should have.

Furthermore, we categorize KPIs into two types: primary KPIs (*PrimaryKPI*) and secondary KPIs (*SecondaryKPI*). A primary KPI is a KPI that can be calculated by applying an aggregation function (i.e., sum, average, count, etc.) on a metric, whereas a secondary KPI can be calculated based on several metrics and other KPIs by following a calculation expression. Attributes of KPIs are defined as follows:

- **Target value.** The target value is required for considering whether the KPI is achieved. It is the desired value where the KPI should reach for. For example, the KPI `TotalRevenue`, provided in the example (cf. Fig. 6.4), has a target value as EUR 1,000,000. This means that the desired total revenue should reach EUR 1,000,000.

- **Limit type.** The limit type provides the desired direction of the KPI value. There are three limit types: "More is better", "Less is better", and "Two-side". The "More is better" indicates that actual values higher than the target value are preferred. The "Less is better" indicates that actual values lower than the target value are preferred. The "Two-side" indicates that actual values equal to the target value are preferred. For instance, the limit type of the KPI `TotalRevenue` (cf. Fig. 6.4) is "More is better" which implies that the total revenue more than EUR 1,000,000 (i.e., its target value) is preferred.

- **Critical threshold.** The critical threshold defined in this work is a relative value with respect to the target value. Depending on the limit type, it defines the boundary from the target value where the KPI becomes critical. For example, the critical threshold of `TotalRevenue` is defined as EUR 200,000 (cf. Fig. 6.4). According to its limit type which is "More is better", its critical threshold defines the lowest bound that the KPI should achieve. Therefore, if the actual value of `TotalRevenue` is less than EUR 800,000 (i.e., the revenue drops more than EUR 200,000 from the target value), the KPI is considered as critical. In contrast, the KPI `AverageInvoiceDuration` has the limit type "Less is better", the target value is 2 days and the critical threshold is 5 days (cf. Fig. 6.4). In this case, it indicates that the KPI becomes critical if its actual value is more than 7 (2 + 5) days.

- **Analysis period.** The analysis period defines a calculation period of the KPI such as weekly, daily, or monthly.

- **Weight.** The weight assigns a relative importance to the KPI. In our implementation, we define the weight of a KPI as a percentage value.

- **Aggregation function.** Primary KPIs apply aggregation functions (e.g., sum, average, count, etc.) for calculating their values. Therefore, primary KPIs have the additional attribute of an aggregation function in order to specify the function that will be applied on the related metrics. For instance, the KPI `Total Revenue` has an aggregation function "SUM" (cf. Fig. 6.4). This refers to the calculation of the KPI as a sum of its related metric `InvoicedAmount`.

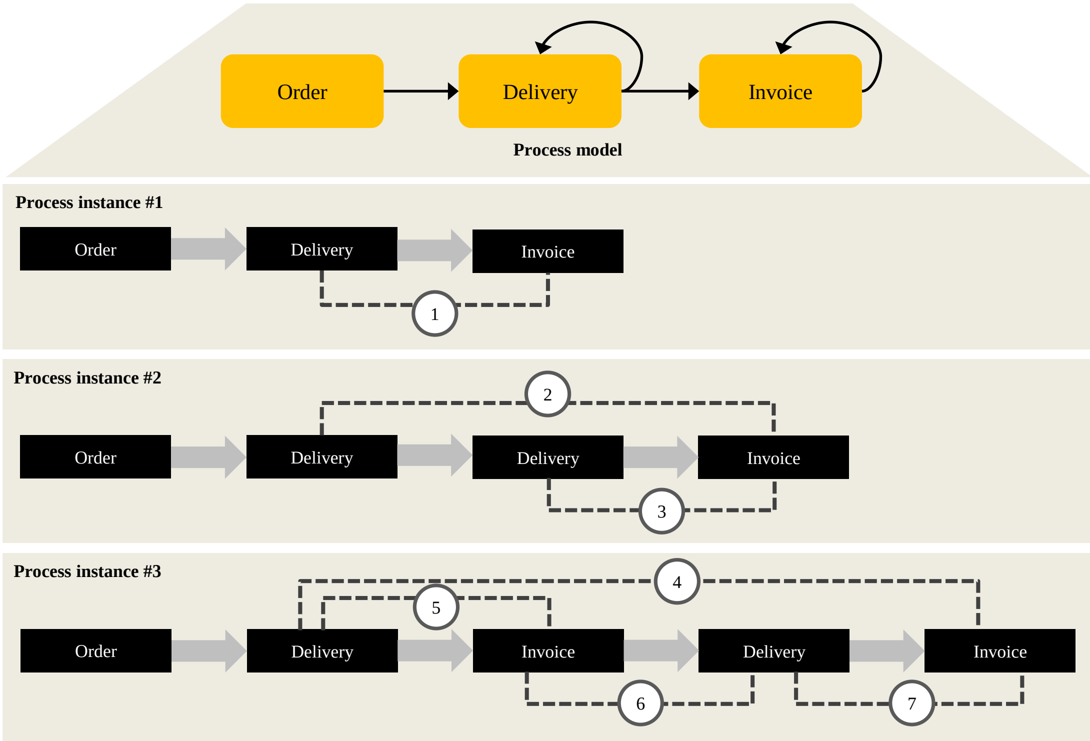


Figure 6.5: Example of process model and possible derived process instances

- **Expression.** For secondary KPIs which require the calculation based on algebraic expressions, the expression attribute is necessary for defining algebraic expressions over variables and constants. The KPI `AverageRevenuePerOrder` is an example of a secondary KPI (cf. Fig. 6.4). In the example, it is calculated based on the expression "X / Y". The binding of the variable "X" and "Y" is provided in the variable map which is explained in the following.

- **Variable map.** Variable map is an attribute for secondary KPIs. It defines the binding between variables (used in the expression) and values of the related metrics or KPIs. Let's consider the previous example of the KPI `AverageRevenuePerOrder`. As explained earlier, the KPI `AverageRevenuePerOrder` has a calculation expression as "X / Y". The variable map indicates that "X" is the value of the KPI `TotalRevenue` and "Y" is the total number of orders calculated by applying the "COUNT" function on the metric `OrderID`.

The other concept defined in the BSC ontology is *Metric*. The *Metric* is not a BSC element but it is required as a basis for calculating KPIs: metrics are calculated on individual events of a process instance, whereas KPIs aggregate one or more metrics over a specific period of time. Our employed ontology of metrics builds upon the work of Wetzstein et al. [WML08]. Metrics are divided into instance metrics *(InstanceMetric)*, aggregate metrics *(AggregateMetric)*, and composed metrics *(ComposedMetric)*. Instance metrics are based on query statements and can be further divided into instance data metrics *(InstanceDataMetric)* and instance process metrics *(InstanceProcessMetric)*.

Instance data metrics use queries on EDI & BI ontologies where raw data elements from EDI standards are conceptualized into generic BI concepts. The reader is referred to Chapter 3 for details on encoding EDI messages through BI concepts. These BI concepts are used in query statements of instance data metrics for querying data on a conceptual level (e.g., ordered quantity, invoiced amount, etc.). Instance data metrics may be employed primarily for metrics focusing on business performance that are calculated from business information in EDI messages (e.g., ordered quantity). Instance data metrics have the following attributes:

- **Query statement.** To retrieve the information from EDI & BI ontologies, a query statement must be defined. In our implementation, we define queries based on the Datalog syntax [MSS05]. As shown in an example provided in Figure 6.4, the instance data metric `InvoicedAmount` has the query as "INVOIC_LineItem_InvoicedAmount(X), hasValue(X, Y)". In particular, the first part is defined to query instances of invoiced amounts (i.e., "INVOIC_LineItem_InvoicedAmount(X)") and the second part is defined to query the value of those instances of invoiced amounts (i.e., "hasValue(X, Y)").

- **Variable map.** The variable map is used to bind variables in query statements to their representative names. For example, the variable map of the instance data metric `InvoicedAmount` indicates that the variable "X" refers to "DataElement" that contain the value of invoiced amount and the variable "Y" refers to "InvoicedAmount" (cf. Fig. 6.4).

- **Control event type.** Control event type specifies the type of event in a process model where the calculation mechanism should perform a query. For instance, let's assume that the example provided in Figure 6.4 is defined based on the process model depicted in Figure 6.5. In the example, the instance data metric `InvoicedAmount` has a control event type as "Invoice". This means, at querying time, the calculation mechanism should retrieve invoiced amounts that are only related to the events of type "Invoice" (cf. Fig. 6.5). Limiting the query for specific types of events can help improving the correctness of information retrieval as well as reducing search scope which in turn speed up querying tasks.

Instance process metrics use query statements that reference time-based values gathered from process models and event logs. Hence, instance process metrics may be used primarily for metrics focusing on process performance and are calculated from event data and process models, such as event sequence patterns or event timestamps (e.g., order date/time). In this case, in order to ensure the correctness of information retrieval, event sequence patterns have to be specified. Based on the specified patterns, the querying mechanism has to be able to perform querying tasks following the patterns. In doing so, the implementation of the querying mechanism of this framework leverages the concept of log replay [AAD12] to step through the event log and retrieve corresponding activity timestamps accurately. Instance process metrics consist of the following attributes:

- **Query statement.** The query statement of instance process metric consists of variable and event type. In particular, the only information that is retrieved for instance process metrics are timestamps of events. In order to query timestamps of events, the information of event type (which are of interest) and the related variable are required. For example, let's assume that the instance example provided in Figure 6.4 is defined based on the process model in Figure 6.5. In the example, the instance process metric `DeliveryAndInvoiceDate` has a query statement as "[X, Delivery][Y, Invoice]". Based on this query, at the calculation time, the mechanism will query all timestamps of "Delivery" events and timestamps of "Invoice" events as well as assign "X" and "Y" as a corresponding variable.

- **Variable map.** Similar to the variable map of instance data metrics, the variable map is used to define names for variables that appear in the query. For instance, as shown in the instance process metric `DeliveryAndInvoiceDate`, the variable "X" refers to "DeliveryDate" and the variable "Y" refers to "InvoiceDate".

- **Source type.** Source type is necessary for ensuring information retrieval based on the specified event sequence patterns. In our implementation, we allow two source types: "Event" or "Transition". If the source type is "Event", the mechanism will retrieve timestamps of all possible events that are defined in the query statement. If the source type is "Transition", the mechanism will retrieve timestamps of events that match the sequence order of event types defined in the corresponding query statement. For example, let's again assume that the instance example provided in Figure 6.4 is defined based on the process model in Figure 6.5. The instance process metric `DeliveryAndInvoiceDate` is defined for retrieving the timestamps of "Delivery" events and related "Invoice" events for further calculation of the invoice duration. However, according to the process model, several event sequences can be derived as shown in the example of process instances (cf. Fig. 6.5). In the case of process instances#1, it is obvious that the duration of invoice is the duration between the "Delivery" event and its consecutive "Invoice" event (cf. 6.5, Mark 1). However, in the process instance #2 and #3, querying becomes ambiguous since there are several possible pairs of the events "Delivery" and "Invoice". Therefore, it is required to specify an event sequence pattern in order to ensure the correctness of information retrieval. In the example, the instance process metric `DeliveryAndInvoiceDate` has a sauce type as "Transaction". This restricts the query mechanism to retrieve timestamps of "Delivery" events and timestamps of their consecutive "Invoice" events since the "Delivery" is firstly defined and followed by "Invoice" in the query statement (i.e., "[X, DeliveryEvent][Y, InvoiceEvent]"). Therefore, the timestamps of "Delivery" events and "Invoice" events of the pairs depicted in Figure 6.5, Mark 1, 3, 5, and 7 will be selected at calculation time.

Aggregate metrics aggregate values of instance metrics by using aggregation functions such as sum, average, count, etc. Composed metrics allow the use of algebraic expressions on several metrics in order to further aggregate metrics (e.g., duration between ordering and invoicing). The attributes belonging to aggregate metrics and composed metrics are similar to the attributes of KPIs:

- **Aggregation function.** The aggregation function is required for aggregate metrics. It defines the aggregation function (e.g., sum, average, count, etc.) that is applied on a related metric.

- **Expression.** Composed metrics requires an expression attribute for specifying the algebraic expression over constants and/or variables related to their corresponding metrics.

- **Variable map.** Furthermore, composed metrics have the attribute of variable map for indicating the binding between the variable and the values of the related metrics. For example, the composed metric `InvoiceDuration` maps the variable "X" to the "DeliveryDate" and the variable "Y" to the "InvoiceDate" of the metric `DeliveryAndInvoiceDate`.

6.3.2 Success Factors, KPIs, and Suggestion Rules

The above described ontology and the above described method for BSC calculation address the problems of calculating KPIs from heterogeneous EDI data schemas as well as aligning KPIs with business strategy as introduced in the previous chapter. For addressing the challenge of defining concrete KPIs for evaluating IORs from EDI messages, we conducted a systematic literature review on inter-organizational success factors as well as their related measurements. The selection of relevant studies was based on search criteria covering the topics of inter-organizational success factors, inter-organizational performance evaluation, and business partner selection. We considered only studies published in the period from 2000 to 2012. Using Google Scholar with these search criteria pointed us to 177 qualified published works. The reader is referred to Chapter 4 for details on the review. Based upon this literature review, we identified 56 success factors. We studied a sample of EDIFACT message type specifications in various releases (ranging from D96A to D10A) and identified sets of KPIs reflecting these success factors that can be calculated from information in such messages. Furthermore, we presented aggregations of these KPIs in order to define quantitative measurements for inter-organizational success factors. Details on the employed method for KPI identification can be also found in Chapter 5.

Knowledge Base of Success Factors and KPIs

The identified inter-organizational success factors and KPIs are stored as a knowledge base supporting automatic KPI identification and suggestion. In other words, those identified success factors, KPIs, as well as their relationships are stored as instances in the knowledge base conforming to the conceptual model shown in Figure 6.6, Mark 1. The conceptual model shows the modeling of the relationships between success factors (`SuccessFactor`), KPIs (`KPI`), and metrics (`Metric`). KPIs identified from EDIFACT messages relate to metrics that are required for their calculation through the `uses` property. Furthermore, they are associated with corresponding success factors by the object property named `measures`. In addition, success factors are also organized as a hierarchical structure and influence each other as explained earlier in Chapter 4, Section 4.4.1 and 4.4.2 with the properties namely `partOf` and `influences`.

An example of knowledge base at an instance level is provided in Figure 6.6, Mark 2. According to the example, `Satisfaction`, `FinancialPerformance`, and `Performance` are

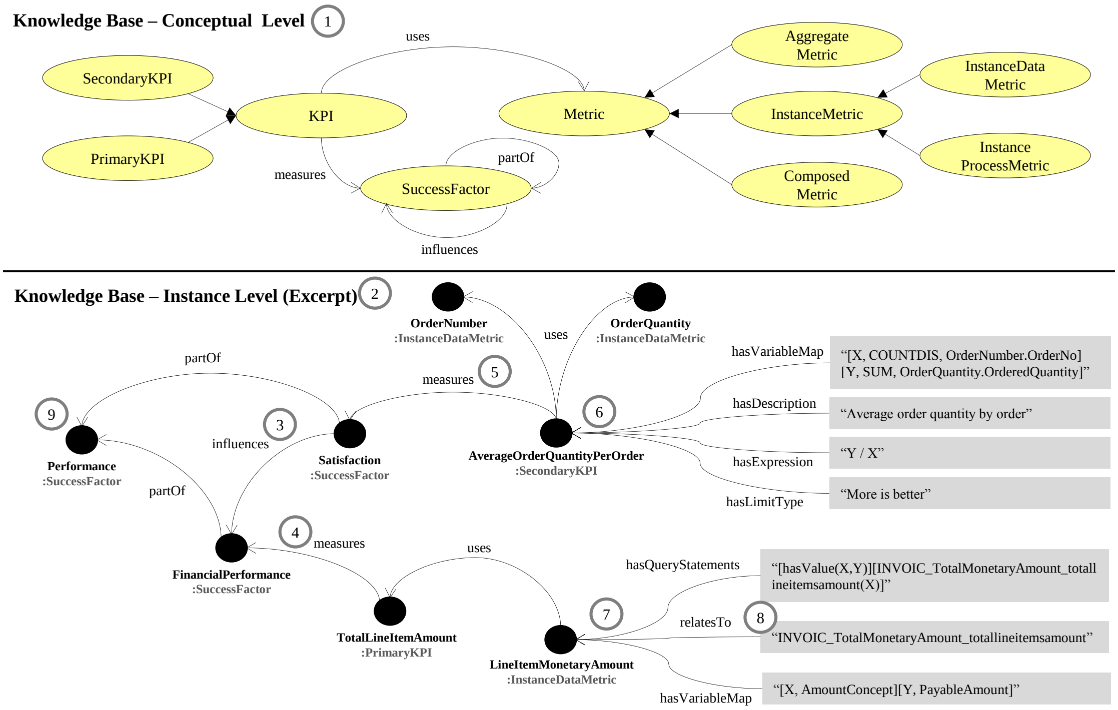


Figure 6.6: Predefined inter-organizational success factors and KPIs stored as a knowledge base

success factor. `Satisfaction` and `FinancialPerformance` are part of `Performance`. Furthermore, the success factor `Satisfaction` influences the achievement of `FinancialPerformance`. This relationship is expressed by `influences` property (cf. Fig. 6.6, Mark 3). The influencing relationships are derived by applying inference rules of inferring influencing relationships of success factors (cf. Chapter 4, Section 4.4.2) on the cause and effect model presented in Chapter 4, Section 4.4.2. Based on the result of inter-organizational KPI identificatin (cf. Chapter 5), the identified KPIs are assigned to their corresponding success factors. For example, the KPI `TotalLineItemAmount` measures the success factor `FinancialPerformance` and the KPI `AverageOrderQuantityPerOrder` measures the success factor `Satisfaction` (cf. Fig. 6.6, Mark 4 and 5). These KPIs are connected to metrics which are required for the calculation. Both KPIs and metrics have data properties containing necessary attribute values for their calculation (cf. Fig. 6.6, Mark 6 for example of KPI attributes and Mark 7 for example of metric attributes).

KPI Suggestion Rules

Based on the knowledge base, the framework is able to suggest available KPIs for the selected success factors by (i) identifying available KPIs and (ii) applying suggestion rules. In order to identify available KPIs, the suggestion mechanism queries for KPIs which uses the available metrics. Metrics are considered available if they relates to BI concepts that appear in data source (i.e., EDI & BI ontologies). In particular, the mechanism searches for BI concepts contained

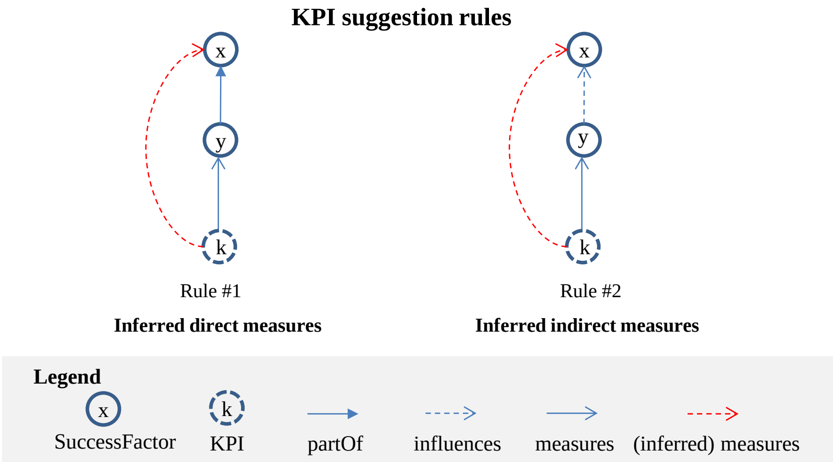


Figure 6.7: KPI suggestion rules

in data property named `relatesTo` (cf. Fig. 6.6, Mark 8) in data source. If all BI concepts related to a metric are available in data source, this metric is then available. Consequently, if all metrics used by a KPI are available, this KPI is also considered available. After available KPIs are identified, the mechanism applies suggestion rules. These rules are defined solely by the author. Based on the *partOf* and *influences* relationships between success factors, we define two suggestion rules as described in Figure 6.7. There are two type of suggestions: (i) direct measures, and (ii) indirect measures.

Direct measures are KPIs which are directly assigned to measures either a success factor or one of its sub-part. KPIs which are directly assigned to a success factor are automatically identified as a direct measure. Considering the example provide in Figure 6.6 at an instance level, the success factor `FinancialPerformance` has a `measures` relationship with the KPI `TotalLineItemAmount` (cf. Fig. 6.6, Mark 4). Therefore, the suggestion mechanism implies that the the KPI `TotalLineItemAmount` is a direct measure of the success factor `FinancialPerformance`. In the case of inferring direct measures from sub-part of success factors, the mechanism applies the inference rule #1.

Rule #1: If success factor y is a part of a success factor x and KPI k measures the success factor y, then it implies that KPI k directly measures the success factor x.

Moreover, a success factor can be influenced by other success factors. This can imply that if a success factor has an influence on a particular success factor, the achievement of this success factor will have an impact to the achievement of a particular one as well. Therefore, the measurement of a success factor influencing other success factors can indirectly indicate an achievement of others. These indirect measures are inferred by using suggestion rule #2.

Rule #2: If success factor y influences a success factor x and KPI k measures the success factor y, then it implies that KPI k indirectly measures the success factor x.

By reasoning over the knowledge base with these inference rules, the suggestion mechanism is able to suggest available KPIs for any selected success factors. For instance, to suggest direct measures for the success factor `Performance` (cf. Fig. 6.6, Mark 9), rule#1 is applied. In this case, the KPI `AverageOrderQuantityPerOrder` and `TotalLineItemAmount` will be suggested as direct measures for `Performance`. Since the success factors (i.e., the success factor `Satisfaction` and `FinancialPerformance`) that are measured by them are a part of the success factor `Performance`. In the case of indirect measures, the KPI `AverageOrderQuantityPerOrder` can be inferred as an indirect measure for the success factors `FinancialPerformance` by using rule#2 through the influencing relationship with the success factor `Satisfaction` (cf. Fig. 6.6, Mark 3).

6.3.3 BSC Calculation

The calculated result of a BSC model indicates the achievement of each business objective by means of scores. Best to our knowledge, there is no formal method for the calculation of a scorecard according to the BSC approach. In theory, the BSC is introduced as a conceptual framework for monitoring and analyzing business performance across several perspectives. In practice, the implementation of the BSC method as well as the calculation of a scorecard depends on organizations. Therefore, we rely on one of the existing BSC software tools. In particular, our calculation of scores applies existing methods of the ADOscore tool[1], which is summarized in the following. The score of a business objective is calculated as the weighted sum of the related KPIs' scores:

$Score_{BusinessObjective} = \sum_{i=1}^{n} Score_{KPI_i} \times Weight_{KPI_i}$ where n is a number of KPIs related to a business objective.

The calculation of a KPI depends on its target value, the *critical threshold*, and the limit type. As explained earlier, the target value is used for determining the success of KPIs by comparing it against the actual value whereas threshold is used for triggering a KPI's status. In our implementation, the status of KPIs is represented by traffic colors (i.e., green, yellow, and red). The color green, yellow, and red represent the status "Very good", "Not good", and "Critical" respectively. Therefore, we define two types of threshold for triggering the status: the Green-Yellow and the Yellow-Red threshold. The threshold Green-Yellow is used for triggering a KPI status from "Very good" to "Not good". The threshold Yellow-Red, i.e. the critical threshold, is used to trigger a status from "Not good" to "Critical". The calculation of the KPI score relies on the critical threshold being the lowest value that a KPI should have. The calculation of the KPI score is also influenced by the limit type which is divided into "More is better", "Less is better", and "Two-side". The calculation methods calculate the KPI score by means of percentage of achievement comparing to the target and the critical threshold. These methods are described in the following.

1. "More is better": The calculation of the KPI score is as follows[1]:

   $Score_{KPI} = \frac{actual-(target-criticalThreshold)}{target-(target-criticalThreshold)} \times 100$

[1] `http://www.boc-group.com/at/produkte/adoscore/` (visited March 01, 2014)

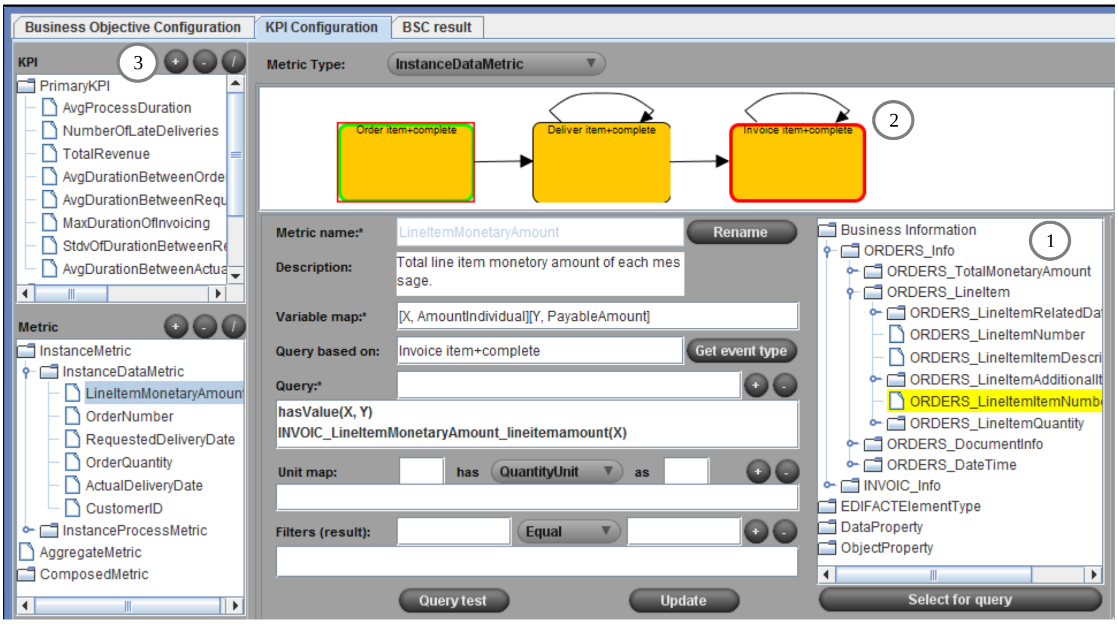


Figure 6.8: KPI and metric panel of the BSC EDImine plug-in for ProM 6

2. "Less is better": The calculation of the KPI score is as follows[1]:

   $$Score_{KPI} = \frac{actual-(target+criticalThreshold)}{target-(target+criticalThreshold)} \times 100$$

3. "Two-side": The calculation of the KPI score for the "Two-side" depends on the actual value. If the actual value is less than the target value then the calculation is performed as method 1, otherwise method 2 is applied.

6.3.4 The Implementation of the Framework

Based on EDI & BI ontologies, the event log derived from EDI messages, and the mined process model, the EDImine BSC Plug-in for ProM 6 allows the modeling of a BSC model composed of BSC elements and their relations. The BSC model is modeled as an OWL [AV04] ontology. We apply the OWL-API [HB11] for manipulating the BSC model, and the DReW system [XEH12] for reasoning and querying over the model and EDI & BI ontologies.

Figure 6.8 shows the KPI and metric configuration panel. The business information view (cf. Fig. 6.8, Mark 1) and process model view (cf. Fig. 6.8, Mark 2) are integrated for supporting the modeling task. For each activity type of a process model the related business information is displayed. Thereby, the information in both, the business data and process perspectives, are perceived coherently. At the same time, the plug-in also allows the modeling of KPIs (cf. Fig. 6.8, Mark 3) based on metrics as well as assigning other attributes (e.g., threshold, analysis period, weight, etc.) required for the BSC calculation.

In addition to the bottom-up identification of KPIs, the plug-in allows top-down modeling where business objectives are aligned with KPIs through success factors. In the business objective configuration panel BSC perspectives (cf. Fig. 6.9, Mark 1), business objectives (cf.

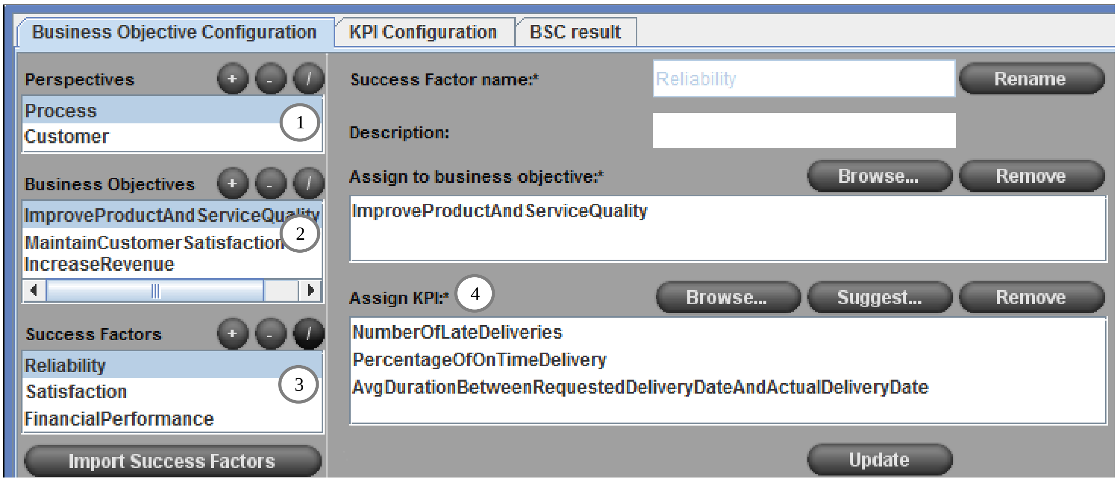


Figure 6.9: Business objective configuration panel of the BSC EDImine plug-in for ProM 6

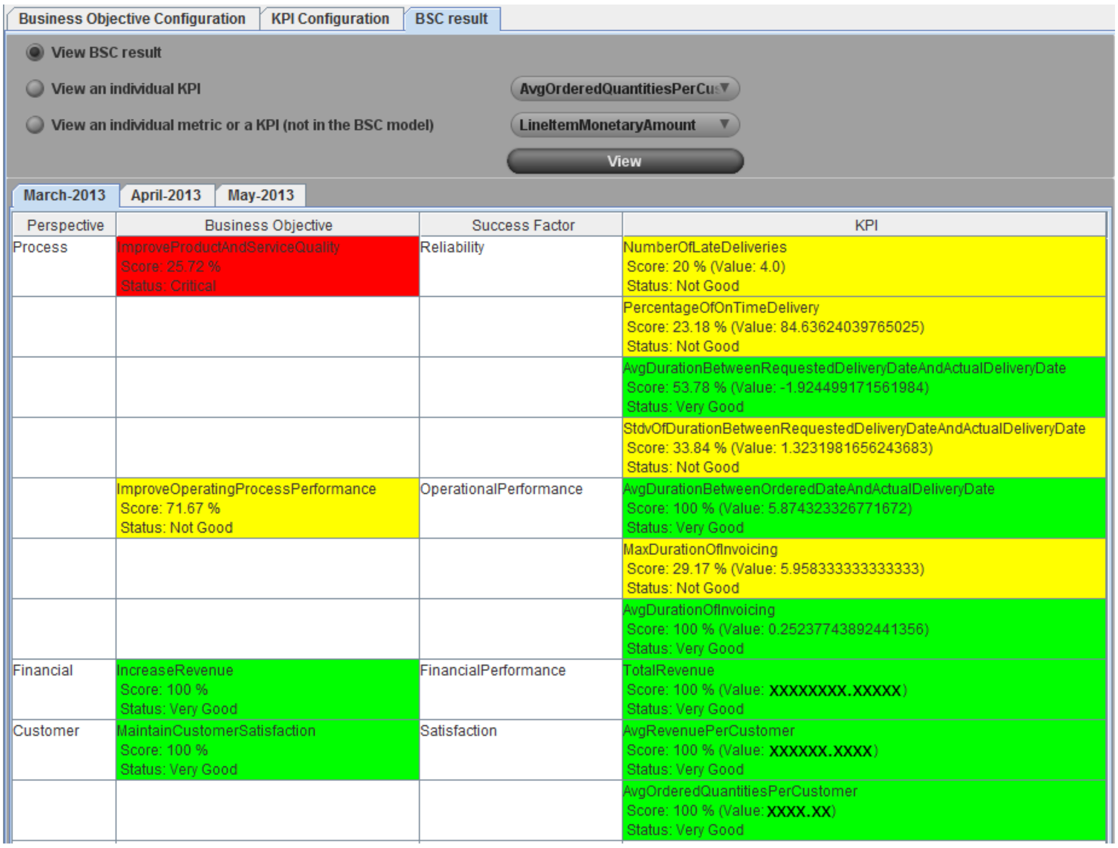


Figure 6.10: BSC result visualization panel of the BSC EDImine plug-in for ProM 6. Monetary figures and quantities of goods have been concealed for confidentiality purpose.

Fig. 6.9, Mark 2), and success factors (cf. Fig. 6.9, Mark 3) can be modeled together with their attributes (e.g., threshold, weight, etc.) and linked to each other as shown in Figure 6.9. KPIs which are modeled in the aforementioned KPI and metric configuration panel can be assigned to success factors (cf. Fig. 6.9, Mark 4) for enabling quantifiable measurement against business objectives.

Figure 6.10 illustrates a BSC result visualization calculated from a BSC model. The results calculated from a BSC model are visualized as BSC tables. Each table shows results belonging to each analysis period. The business objectives and KPIs are colored in traffic-light color code according to their target achievement status. Each business objective and KPI is visualized with a score (i.e., percentage of achievement comparing with target and threshold) and a status.

6.4 Summary

In this chapter, we presented the EDImine BSC Framework for evaluating inter-organizational performance based on EDI messages. The main contribution of this work is to leverage the advantages of both, bottom-up and top-down performance analysis approaches, by integrating process mining techniques and the BSC method in a single framework. The integration allows a bottom-up KPI definition and calculation and a top-down alignment of business objectives and KPIs. Furthermore, we realized the framework with semantic technologies which enables (i) the definition of KPIs on an abstract level regardless of the EDI transfer syntaxes in use, (ii) the modeling of BSC elements and their relations, and (iii) the automatic calculation of a BSC model. Our proposed BSC ontology used in the framework conceptualizes the complete set of BSC elements and metrics which are required for the conceptual modeling of BSC models and their corresponding calculation. This in turn supports the reusability of the BSC ontology.

The framework helps connecting the measurements on the operational level to the strategic level and, hence, it lifts the evaluation to a strategic level as well. Therefore, it solves the last research question: *"How can we lift the evaluation of business performance to the strategic level?"*. Consequently, the main research question about the evaluation of IORs can be addressed by using our framework which combines all of the previous contributed artifacts (i.e., the business information extraction approach, and the knowledge on inter-organizational success factors and KPIs). The framework has also been published in [Kra+]. In the following chapter, two case studies are presented where the EDImine BSC Framework has been applied for demonstrating the framework described.

CHAPTER 7

Case Studies

In the following, we present two case studies for evaluating inter-organizational performance by using the EDImine BSC Framework. The EDI data used in the first case study is collected from a beverage manufacturing company. In the second case study, the data used stems from a consumer goods manufacturing company. In this chapter, we provide an introduction to the implementation of the case studies as preliminaries. The detailed implementation of each case study is further discussed in the following sections.

7.1 Preliminaries

In the case studies, we analyze inter-organizational business performance of two companies based on samples of their EDI messages. The implementation of the case studies follows the steps depicted in Figure 7.1. Starting from the sample set of EDI messages collected from the companies, these messages are preprocessed for both (i) business-information-oriented preprocessing and (ii) process-oriented preprocessing. In business-information-oriented preprocessing, EDI messages are interpreted and stored in EDI ontologies and knowledge bases (cf. Fig. 7.1, Mark 1). Furthermore, the values contained in EDI ontologies are conceptualized into generic business information (BI) concepts by applying the business information extraction approach presented in Chapter 3 (cf. Fig. 7.1, Mark 2). In process-oriented preprocessing, the EDI data is mapped to process events which are further correlated into corresponding process instances (cf. Fig. 7.1, Mark 3). Those process instances as well as related events are stored as an event log. Depending on the intended analysis, the event log can be filtered for process instances which satisfy desired requirements. The event log is then used for mining process models by applying various process mining techniques (cf. Fig. 7.1, Mark 4). The preprocessed artifacts consisting of (i) EDI & BI ontologies and KB, (ii) preprocessed event logs, and (iii) mined process models are used as data sources for the performance analysis with the EDImine BSC Framework described in Chapter 6.

As mentioned earlier, the case studies presented in this chapter demonstrate the evaluation of inter-organizational performance of two companies by using our approach. The sample EDI

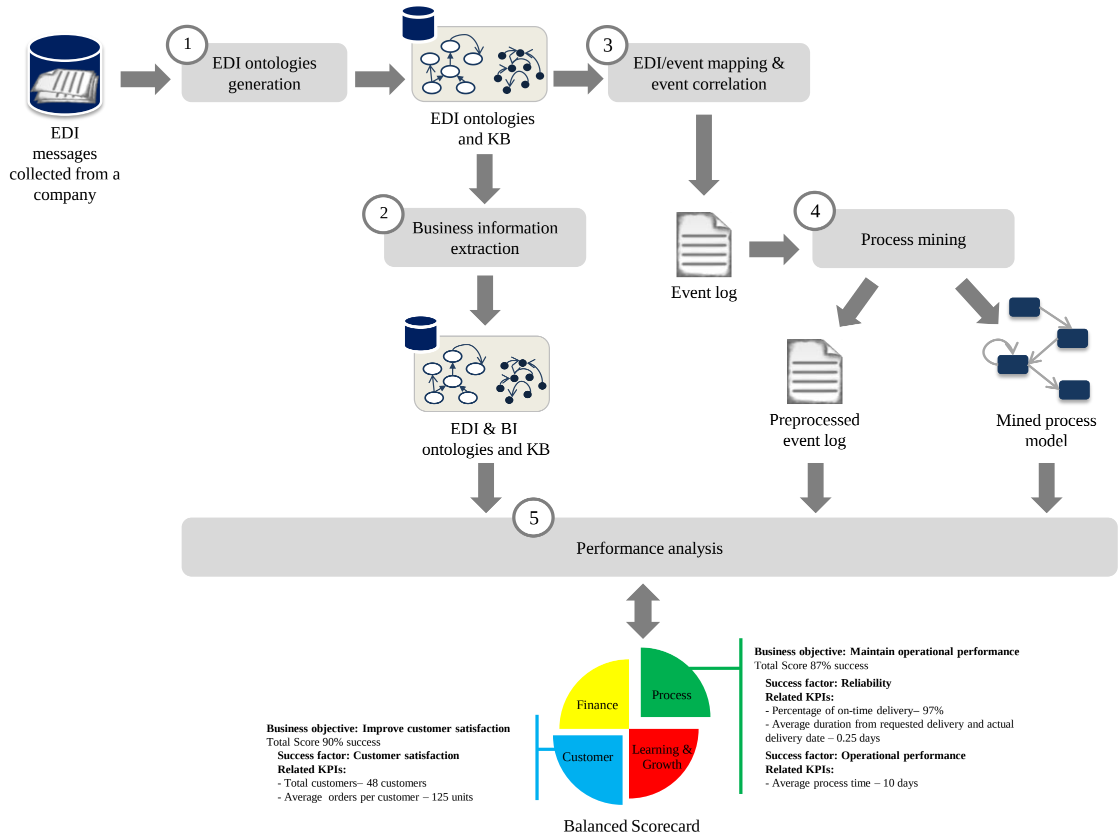


Figure 7.1: Implementation of the case studies

messages expressed in the UN/EDIFACT format are collected from two companies: (i) a beverage manufacturing company and (ii) a consumer good manufacturing company. The implementation of each case study follows the aforementioned steps shown in Figure 7.1. The case study of the beverage manufacturing company is presented in Section 7.2. Section 7.3 provides details on the second case study of a consumer good manufacturing company. The first case study focuses on performance analysis based on the processes reflecting the exchange of EDI messages (e.g. message flow). The second case study concentrates on the purchasing processes of individual line items. For the sake of confidentiality, in each case study we will further on refer to the beverage manufacturing company as *CompanyA* and to the consumer good manufacturing company as *CompanyB*. In addition, all monetary figures and quantities of goods have been multiplied by an undisclosed constant factor.

This chapter is structured as follows. Section 7.2 and Section 7.3 discuss the first and the second case study respectively. Each of them is self-contained. In particular, it provides the information of (i) the data set used in the case study, (ii) the data preprocessing performed according to the steps introduced in Figure 7.1, (iii) the balanced scorecard (BSC) implementation using the EDImine BSC Framework, and (iv) the results and discussion. Finally, this chapter is concluded in Section 7.4.

7.2 Case 1: Beverage Manufacturing Company

The first case study presented in this section is the evaluation of business performance of a beverage manufacturing company which is again referred as CompanyA. The data set contains sample EDI messages exchanged during purchasing processes of CompanyA, and their customers or business partners. Similar to the first case study, the process starts when a customer sends an order to CompanyA. In such an order, the customer usually specifies order details such as line items, ordered quantities, and requested delivery date. After goods have been shipped, CompanyA sends invoices for informing the customer about purchased goods, quantities, and prices.

In order to evaluate business performance of CompanyA on a strategic level, we implement the BSC method for evaluating the achievement of CompanyA's business objectives. Following the EDImine BSC Framework, we first preprocess sample EDI messages for deriving business information contained in the messages as well as an event log together with mined process model reflecting their purchasing processes. Consequently, we model a BSC for CompanyA by defining business objectives related to their mission and link them to success factors and KPIs derived from the discovered business and process information. The success factors and KPIs used in the case study are based on the knowledge from our conducted literature review about inter-organizational success factors and their measurements (cf. Chapter4). In particular they are selected from the *predefined knowledge base* of inter-organizational success factors and KPIs provided in the framework (cf. Chapter6, Section 6.3.2). The scores of the business objectives and KPIs, as defined in the following, are calculated based on the BSC model.

7.2.1 Data Set

The above described business process of CompanyA is supported by EDI messages that are interchanged between the IT systems of CompanyA and their customers. The data set consists of 282 received EDIFACT ORDERS (Purchase order) messages, and 427 sent INVOIC (Invoice) messages collected between August 6, 2012 and February 28, 2013 (dates refer to interchange timestamps). ORDERS messages were all encoded according to the D96A[1] EDIFACT release, while INVOIC messages were sent in D01B[2] releases of EDIFACT.

7.2.2 Data Preprocessing

Before analyzing business performance, the sample EDI data is preprocessed. As mentioned earlier, the framework allows identifying KPIs based on three artifacts derived from EDI data: (i) EDI & BI ontologies where EDI data is conceptualized and abstracted into business information concepts, (ii) an event log containing business processes together with related events, and (iii) an inter-organizational process model derived from such event logs. In the following, we describe the preprocessing of EDI data in detail.

EDI & BI Ontologies – Semantic Preprocessing. In the semantic preprocessing stage we parse sample EDI messages into EDIFACT ontologies [Eng+12b] where all data elements are

[1] `http://www.unece.org/trade/untdid/d96a/content.htm` (visited March 01, 2014)
[2] `http://www.unece.org/trade/untdid/d01b/content.htm` (visited March 01, 2014)

Table 7.1: EDI/event mappings used for the first case study

Activity	Event Attribute	Associated EDI Artifact				
		Message Type	**Segment Group**	**Segment**	**Composite Data Element**	**Data Element**
Receive order	*(Event trigger)*	ORDERS EDIFACT D96A	*(Message instance)*			
	time:timestamp		*(Message interchange timestamp)*			
	org:resource		*(Interchange sender)*			
	orderNumber		-	BGM	-	Document/message number (1004)
Send invoice	*(Event trigger)*	INVOIC EDIFACT D01B	*(Message instance)*			
	time:timestamp		*(Message interchange timestamp)*			
	org:resource		*(Interchange sender)*			
	orderNumber		1	RFF	Reference (C506)	Order number (purchase) (1154 [1153='ON'])

stored according to the structure of the related EDIFACT standard (cf. Fig. 7.1, Mark 1). The EDIminer toolset [Eng+13b] is used for generating EDIFACT ontologies and parsing all data elements from EDI messages into ontologies. Furthermore, we generate BI ontology using our proposed business information extraction approach based on a predefined mapping of BI concepts and actual EDI data elements (cf. Fig. 7.1, Mark 2). The resulting BI ontology contain generic BI concepts (e.g., ordered quantity, invoiced amount, etc.) that represent EDI data elements on a conceptual level regardless of a particular syntax.

Event Log – Event Mapping and Correlation. In order to generate an event log from the EDI data set, we start by defining a set of EDI/event mappings. The mapping follows the Message Flow Mining (MFM) method [ER14] focusing on generating event logs that reflect the message interchanges between business partners. Table 7.1 describes the MFM mapping used in this case study. According to the purchasing process of CompanyA, we define EDI/event mappings for *Receive order* and *Send invoice* as an activity. For the activity *Receive order*, we define a mapping that uses ORDERS message instances as event triggers and populate their timestamp attributes with the message interchange timestamp. Similarly, the activity *Send invoice* has a mapping that uses INVOIC message instances, and message interchange timestamp as event triggers and timestamp attributes respectively. The organizational resource (*org:resource*) associated with generated events is set to the interchange senders for all mappings. Furthermore, we add *orderNumber* as a common attribute to all mappings in order to allow the correlation of generated events to process instances.

Using the event mappings described, the data set under consideration yields a set of 709 events (282 *Receive order*, and 427 *Send invoice* events). As mentioned earlier, we intend to investigate the performance of the purchasing process composed of receiving orders and sending invoices events. We assume that *orderNumber*s are unique. Hence, we correlate events to process instances by grouping them according to *orderNumber* and filter the resulting process instances for cases which contain complete traces (i.e., having at least one activity instance of each activity type). In total, there are 282 complete cases consisting of 282 *Receive order*, 299 *Send invoice* events stored in an XES log.

Inter-organizational Process Model – Process Mining. Based on the event log from the event mapping and correlation preprocessing stage, we apply process mining techniques (i.e., the Heuristics Miner [WAA06]) for deriving an inter-organizational process model. Figure 7.2

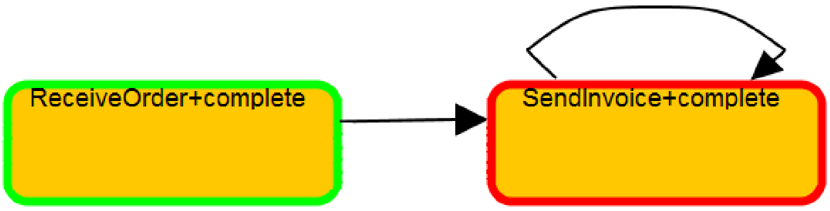


Figure 7.2: The mined process model of the first case study

illustrates the mined model of a message exchange during a purchasing process of CompanyA and their business partners.

7.2.3 Balanced Scorecard Implementation using the EDImine BSC Framework

For designing the BSC model for CompanyA we investigated their mission which is to become number one in their product line. To accomplish the mission, CompanyA may need to improve their business performance in all perspectives (i.e., finance, customer, process, and learning and growth) for raising their competitiveness. Therefore, we define "Increase revenue", "Increase customer satisfaction", and "Improve operational performance" as business objectives since they can support the success of the mission. Furthermore, we select appropriate success factors from our predefined set as shown in Table 7.2.

For each success factor, the framework can suggest the appropriate predefined KPIs from the available data (i.e., EDI & BI ontologies, the event log, and the mined process model derived from the sample of EDI messages). Moreover, the framework also allows modeling arbitrary user-defined KPIs. In this case, we define and select the available KPIs suggested from the framework and customize their attributes (e.g., weight, target value, thresholds, etc.) accordingly[3]. Since target and threshold values are confidential and hence concealed by CompanyA, we determine appropriate values based on historical data of CompanyA instead. The selected KPIs are shown in the BSC model described in Table 7.2. Furthermore, the calculations of KPIs are provided in Table 7.3.

Total revenue is selected to reflect financial performance because it indicates the status of the organization's earning. Since INVOIC messages contain information of the line item monetary amount, we can calculate this KPI by adding up all line item monetary amounts. We set the target value as 100,000. If it is less than 50,000, the KPI will be in critical state.

Average ordered quantities per customer is used to measure customer satisfaction. The reason is that if a customer is not satisfied, ordered quantities are expected to be reduced as a consequence. We expect that the ordered quantities reaches 3,750 per month. However, if it is less than half of the target value then it is considered critical.

Number of customers is the total number of customers in each month. Since the mission is to be the leading vendor in their product line, gaining more customers is necessary to expand

[3]In this case study, we specify critical thresholds as relative values with respect to target values in order to allow for the simple definition of thresholds for two-sided KPIs.

Table 7.2: The balanced scorecard for the first case study

Business Objective	**Success Factor**	**KPI**	**Weight (%)**	**Limit Type**	**Target (Critical thr.)**
Financial Perspective					
Increase revenue	Financial performance	Total revenue	100	More	100,000 (50,000)
Customer Perspective					
Increase customer satisfaction	Satisfaction	Average ordered quantities per customer	50	More	3,750 (1,875)
		Number of customers	50	More	6 (2)
Process Perspective					
Improve operational performance	Operational performance	Average duration between requested and actual delivery date	20	Two-side	0 (2)
		Average process time	30	Less	7 (14)
		Percentage of on-time deliveries	50	More	100 (20)

Analysis period of the BSC is set to "Monthly".
Critical thresholds of all business objectives are 50% (having a score less than 50% is critical).
Weight of related KPIs must be 100% in total for each business objective.

their market share. In this case, a unique sender of orders is considered as a new customer. Therefore, we calculate this KPI by counting unique ORDERS message senders (cf. Table 7.3).

Average duration between requested and actual delivery date aims to evaluate the operational performance since it can reflect the reliability of the delivery service. The expected duration is zero meaning the actual delivery date is exactly on the requested delivery date. The limit type of this KPI is set to *Two-side* since late or early deliveries are undesired. We set two days as the critical threshold which means that if the majority of the actual deliveries arrive two days later or earlier than the requested date, this KPI will become critical.

Average process time is the duration of the purchasing process. This KPI must be calculated based on process-oriented information. According to the mined model (cf. Fig. 7.2), it is the duration between when an ORDERS message is received and a last corresponding INVOIC message is sent. Since, according to the data, the process takes around two weeks, we set the target value to seven days and the critical threshold to 14 days. Therefore, process instances which take more than three weeks (21 days) are considered as critical cases. Furthermore, in

Table 7.3: KPI calculation of the first case study

KPI	**Calculation**
Total revenue	SUM(invoiced amount of line item in INVOIC)
Average ordered quantities per customer	SUM(ordered quantities of line item in ORDERS) / COUNTDIS(interchange sender in ORDERS) *Note: Counting distinct senders of ORDERS messages yields the total number of customers.*
Number of customers	COUNTDIS(interchange sender in ORDERS)
Average duration between requested and actual delivery date	AVG(actual delivery date in INVOIC − requested delivery date in ORDERS)
Average process time	AVG(timestamp of the last *Send invoice* event − timestamp of first *Send order* event)
Percentage of on-time deliveries	COUNT if (actual delivery date in INVOIC − requested delivery date in ORDERS) between 1 and -1 / COUNT(actual delivery date in INVOIC − requested delivery date in ORDERS) × 100

KPI calculation formulas are described as aggregation functions applied over sets of results calculated from algebraic expressions. These algebraic expressions are applied on each of the process instances which start in the given analysis period

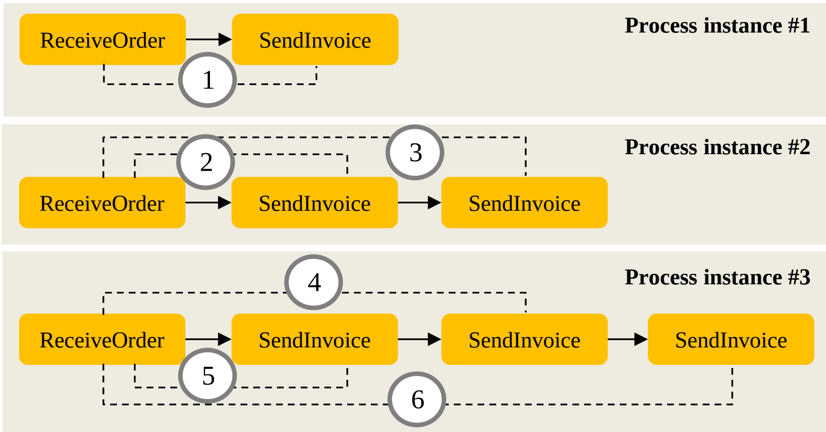


Figure 7.3: Examples of process instances possibly derived from the mined model (cf. Fig. 7.2)

order to ensure the correctness of information retrieval, the query mechanism has to be able to identify the first *Receive order* and the last *Send invoice*. In this case, process mining plays an important role to retrieve correct information against desired activity sequence patterns. Considering the examples of process instances provided in Figure 7.3, in the first process instance the calculation of this KPI is straight forward since there is one *Send invoice* activity. However, process instances #2 and #3 have multiple *Send invoice* activities which yield several possible pairs of *Receive order* and *Send invoice* (i.e., Fig. 7.3, Mark 2, and 3 for process instance #2 and Mark 4, 5, and 6 for process instance #3). Nevertheless, only the durations of *Receive order* and the last *Send invoice* (i.e., Fig. 7.3, Mark 1, 3, and 6) are of interest for calculating this KPI. In this case, by considering the mined process model (cf. Fig. 7.2) we define the activity sequence pattern such that in each process instance the timestamp of *Receive order* and the timestamp of *Send invoice* having no other subsequent activities will be retrieved. Based on this pattern, our query mechanism leverages the concept of *log replay* [AAD12] to step through the event log and retrieve corresponding activity timestamps accurately.

Percentage of on-time deliveries is the most weighted KPI, i.e., 50%, measuring the success factor "Operational performance". It indicates how well the delivery is performed by comparing the on-time deliveries against all deliveries. In this case, the deliveries arrived within the requested delivery date is considered on-time. Therefore, in the calculation we count the number of deliveries that arrived between one-day earlier or later the requested delivery date (i.e., between 1 and -1 day from the requested delivery date) as a number of on-time deliveries (cf. Table 7.3). The optimal case is all deliveries are on-time or 100% of on-time deliveries. If the percentage of on-time deliveries drops to 80%, it is considered as critical.

7.2.4 Results and Discussion

Based on the aforementioned BSC model and by using our developed prototype we calculate the achievement score of business objectives and KPIs according to the calculation methods explained in Section 6.3.3. Table 7.4 shows the calculated result of the BSC model. To avoid

Table 7.4: The balanced scorecard calculated from August 2012 to February 2013.

***Business Objective* & KPI**	**% of Achievement (Actual Value of KPI)**						
	Aug 2012	Sep 2012	Oct 2012	Nov 2012	Dec 2012	Jan 2013	Feb 2013
Financial perspective							
Increase revenue	73.36	68.98	68.26	100	100	100	94.28
Total revenue	73.36 (86,682)	68.98 (84,488)	68.26 (84,131)	100 (114,768)	100 (114,313)	100 (147,274)	94.28 (97,139)
Customer perspective							
Increase customer satisfaction	50	50	49.24*	96.29	63.96	87.73	51.38
Average ordered quantities per customer	100 (5,576)	100 (5,672)	48.48 (2,784)	92.59 (3,611)	77.92 (3,336)	75.47 (3,290)	52.76 (2,864)
Number of customers	0* (2)	0* (2)	50 (5)	100 (6)	50 (5)	100 (6)	50 (5)
Process perspective							
Improve operational performance	87.94	80.21	93.85	98.52	75.28	95.54	99.06
Average duration between requested and actual delivery date	100 (0)	95 (-0.1)	100 (0)	96.18 (-0.08)	94.37 (-0.11)	95.38 (-0.09)	96.42 (-0.07)
Average process time	59.79 (12.63)	65.14 (11.88)	79.51 (9.87)	97.62 (7.33)	59.23 (12.71)	88.21 (8.65)	99.27 (7.1)
Percentage of on-time deliveries	100 (100)	83.33 (96.67)	100 (100)	100 (100)	77.27 (95.45)	100 (100)	100 (100)

The business objectives and KPIs marked with * are in critical status.

the problem of over- and underestimation for both business objectives and KPIs we limit the maximum score and minimum score of achievement to 100% and 0%, respectively. For instance, “Total revenue” has the actual value of 114,768 in Nov 2012 (cf. Table 7.4). The related target, critical threshold, and limit type are 100,000, 50,000, and “More is better” (cf. Table 7.2). Hence, the score of this KPI is calculated by using method 1 (cf. Section 6.3.3). The resulting score is 129.54% (i.e., $\frac{114{,}768-(100{,}000-50{,}000)}{100{,}000-(100{,}000-50{,}000)} \times 100$) but due to the maximum limit it yields 100%.

The result reveals that CompanyA successfully achieved their business objectives in all perspectives. In customer perspective, the business objective “Increase customer satisfaction” has a significant improvement. Although the scores of the first three periods (i.e., Aug 2012, Sep 2012, and Oct 2012) are quite low, the score of the following periods are quite high due to the increasing number of customer in Oct 2012. The reason that the “Increase customer satisfaction” is in a critical status in Oct 2012 is that the average ordered quantities are considerably reduced (i.e., from 5,672 to 2,784). Consequently, the business objective “Increase revenue” has results following a similar trend. However, its statuses are still healthy since the total revenue is acceptable when comparing to target value and critical threshold. In addition, it also shows that CompanyA maintains their operational performance very well. Overall, the scores of the business objective “Improve operational performance” are quite steady. The delivery performance reflected by the KPI “Percentage of on-time deliveries” and “Average duration between requested and actual delivery date” is satisfactory. Also, the average process times throughout the periods are in an acceptable range (i.e., less than 21 days). Overall, the performance of CompanyA has improved. However, customer satisfaction in February 2013 significantly dropped according to the decrease in ordered quantities per customer. An investigation of this problem may be required to prevent drops in customer satisfaction which might in turn affect the performance in other perspectives.

However, in this case study, the KPIs in the learning and growth perspective related to employees and product development (e.g., number of new products, employee turnover rate, etc.) as well as some other KPIs which can directly reflect CompanyA's business objectives (e.g., profit, number of customer complaints, etc.) still could not be derived. This is because there is no such information available in EDI messages. Therefore, the information solely derived from EDI messages may be insufficient for reflecting performance in some perspectives. Moreover, some interested KPIs may not be able to be derived since there exists no required information in the available EDI data.

7.3 Case 2: Consumer Goods Manufacturing Company

In the second case study, we evaluate business performance of a consumer good manufacturing company which is further referred as CompanyB. In the preprocessing step, we start from a real-world sample of CompanyB's EDI interchange data and apply the Physical Activity Mining (PAM) method [ER14] for generating an event log reflecting the actual delivery process execution of CompanyB. Based on the preprocessed data and derived artifacts, we employ the EDImine BSC Framework in order to lift the gathered information to the strategic level and derive additional business intelligence of potential interest from the EDI data. Finally, we discussed our results with representatives of CompanyB.

In the following, we establish some basic facts and assumptions on CompanyB and its business processes that are relevant for the design of the case study. CompanyB declares its primary mission to be the provision of highest quality products and services. Moreover, since CompanyB delivers to a large number of individual supermarket branches, CompanyB's process of ordering, invoicing and delivery of goods to individual customers is of particular importance to the business' success and thus receives particular attention in this case study. This process starts when a customer sends an order to CompanyB. In such an order, the customer usually specifies a requested delivery date for the ordered goods. Subsequently, CompanyB despatches the goods. This is generally done in due time to meet the requested delivery date of the customer. If an order cannot be fulfilled at once, the ordered items may be shipped in partitions. After goods have been shipped, CompanyB sends invoices for the corresponding line items. Again, line items that were ordered in a single order may be scattered over different invoices.

7.3.1 Data Set

The above described business process of CompanyB is supported by EDI messages that are interchanged between the IT systems of CompanyB and its customers. The data set consists of 1389 received EDIFACT ORDERS (Purchase order) messages, 1289 sent DESADV (Despatch advice) and 1840 sent INVOIC (Invoice) messages collected between March 1, 2013 and June 5, 2013 (dates refer to interchange timestamps). ORDERS messages, which were received by CompanyB, were all encoded according to the D96A[4] EDIFACT release, while DESADV and INVOIC messages were sent both in D96A and D01B[5] releases of EDIFACT.

[4] `http://www.unece.org/trade/untdid/d96a/content.htm` (visited March 01, 2014)
[5] `http://www.unece.org/trade/untdid/d01b/content.htm` (visited March 01, 2014)

Table 7.5: EDI/event mappings used for the second case study

Activity	Event Attribute	Associated EDI Artifact				
		Message Type	Segment Group	Seg-ment	Composite Data Element	Data Element
Order item	*(Event trigger)*	ORDERS D96A	25	LIN	Item number identification (C212)	Item number (7140)
	time:timestamp		-	DTM	Date/time/period (C507)	Document/message date/time (2380 [2005='137'])
	org:resource		*(Interchange sender)*			
	itemID		25	LIN	Item number identification (C212)	Item number (7140)
	orderID		-	BGM	-	Document/message number (1004)
Deliver item	*(Event trigger)*	INVOIC (D96A/ D01B)	25 (D96A) 26 (D01B)	LIN	Item number identification (C212)	Item number (7140) (D96A) Item identifier (7140) (D01B)
	time:timestamp		-	DTM	Date/time/period (C507)	Delivery date/time, actual (2380 [2005='35'])
	org:resource		*(Interchange sender)*			
	itemID		25	LIN	Item number identification (C212)	Item number (7140) (D96A) Item identifier (7140) (D01B)
	orderID		1	RFF	Reference (C506)	Order number (purchase) (1154 [1153='ON'])
Invoice item	*(Event trigger)*	INVOIC (D96A/ D01B)	25 (D96A) 26 (D01B)	LIN	Item number identification (C212)	Item number (7140) (D96A) Item identifier (7140) (D01B)
	time:timestamp		-	DTM	Date/time/period (C507)	Document/message date/time (2380 [2005='137'])
	org:resource		*(Interchange sender)*			
	itemID		25	LIN	Item number identification (C212)	Item number (7140) (D96A) Item identifier (7140) (D01B)
	orderID		1	RFF	Reference (C506)	Order number (purchase) (1154 [1153='ON'])

7.3.2 Data Preprocessing

The preprocessing steps follows the process shown in Figure 7.1. Similar to the previous case study, we divided data preprocessing into three main steps: (i) semantic preprocessing of EDI & BI ontologies, (ii) event mapping and correlations, and (iii) mining the inter-organizational process model. In the following each step is elaborated on.

EDI & BI Ontologies – Semantic Preprocessing. We used the EDIminer toolset to parse the EDI messages into EDI ontologies and corresponding Message KBs. Furthermore, we generated BI ontology based on manually defined mappings of BI concepts to actual data elements of EDI messages. These mappings were defined such that semantically equivalent data elements of different EDIFACT standards releases were unified in common BI concepts and the hierarchical structure of these concepts reflects aggregations and/or compositions of these BI concepts.

Event Log – Event Mapping and Correlation. Similar to the previous case study, we defined a set of EDI/event mapping using the EDIminer toolset. The employed mapping definitions are shown in detail in Table 7.5. Because the data set under consideration contains messages based on both the D96A and D01B releases of EDIFACT, we define mappings for these two releases. However, since these two EDIFACT releases overlap in many cases, most mapped EDI artifacts are identical in both kinds of mappings; the cases in which the mappings differ are explicitly highlighted in Table 7.5. Since we use the EDIminer toolset for defining EDI/event mappings, these mappings are based on the above described ontological data model of EDIFACT messages and allow for direct access to qualified data elements. In case mappings are based on qualified data elements, these are shown as underlined labels in Table 7.5. For example, for EDIFACT release D96A, label *Document/message date/time* refers to the value of data element 2380 (Date/time/period) qualified by value '137' (code for "Document/message date/time") in data element 2005 (Date/time/period qualifier). This qualification relationship is specified in Table 7.5 as "(2380 [2005='137'])".

We consider the ordering, delivery and invoicing of goods as the crucial activities for our analysis since they are directly related to the performance of the delivery process. Hence, we define EDI/event mappings for *Order item*, *Deliver item* and *Invoice item* activities. Furthermore, since we intend to investigate delivery performance with regard to individual line items, we focus on the lifecycles of individual line items in the defined mappings as well. Consequently, we use individual line items in the EDI messages as event triggers for all of the three aforementioned activities.

Firstly, for the *Order item* activity we define a mapping that uses individual line items in ORDERS messages as event triggers and populate their timestamp attributes with the document dates of the messages (i.e., *Document/message date/time*).

Secondly, for the *Deliver item* activity, one may consider using individual line items in DESADV messages as event triggers. However, since the DESADV messages in our data set only contain document dates as well as estimated delivery dates, this would only allow us to generate events that reflect the shipment of goods or the estimated delivery of goods, respectively. However, in this case study we are rather interested in the actual deliveries of the goods at the customer's site. For this reason, we exploit that the INVOIC messages in our data set contain actual delivery dates for the invoiced line items and define a mapping for the *Deliver item* activity that uses individual line items in INVOIC messages as event trigger and corresponding values of Delivery date/time, actual as their timestamps. Consequently, we do not further consider the observed DESADV messages for our case study.

Thirdly, for the *Invoice item* activity we define a mapping that uses individual line items in INVOIC messages as event trigger and the invoice's document date (i.e., *Document/message date/time*) as a timestamp for the generated events.

Finally, we add common attributes to all three of the aforementioned mappings and map them to the corresponding EDIFACT data elements. We define the *itemID* and *orderID* event attributes in order to allow for subsequent correlation of generated events to process instances by means of *(itemID, orderID)* tuples. The organizational resource *(org:resource)* associated with generated events is set to the interchange senders from EDIFACT's UNB envelope for all mappings.

Using the above described event mappings, the data set under consideration yields a set of 52622 events (14026 *Order item*, 19318 *Deliver item* and 19318 *Invoice item* events). As mentioned earlier, we intend to investigate the performance of the delivery process from a line-item centric perspective in the context of individual orders. We assume that *orderID*s are generally unique. Hence, we correlate events to process instances by grouping them according to *(orderID, itemID)* tuples. This results in 21215 process instances (cases). We store the generated events and process instances in an XES log and use the LTL Checker Plug-In of the ProM process mining suite to filter the results for cases which contain complete traces (i.e., having at least one activity instance of all three defined activity types). This reduces the log size to 4751 compliant cases and 14779 events (4751 *Order item*, 5014 *Deliver item* and 5014 *Invoice item* events).

However, the above described reduced log is still afflicted with an apparent anomaly. Because the timestamps in the raw EDI data set are encoded with a precision of days only, in traces where delivery and invoicing was conducted on the same day, the timestamps for *Deliver item* events and *Invoice item* events are identical. This may lead to random order of these events in such traces in the event log. Notably, there are only 16 cases in which the invoicing date is strictly *before* the delivery date. However, in these cases the delivery date lies in a future (August 2013) that is beyond the time of writing this thesis (July 2013) and can therefore not specify an actual delivery date. These 16 cases all feature the same *orderID*. Hence, we assume that these 16 cases are part of an exceptional outlier and further assume that in the "real" business process delivery actually always occurs before invoicing. In other words, we assume that this anomaly results only from the way dates are encoded in our data set. Hence we further modify these logs programmatically so that *Deliver item* events always precede *Invoice item* in traces if (and only if) their timestamps are identical. Moreover, we remove the aforementioned 16 cases where the delivery dates lie in the future. The resulting log comprising 4735 traces serves as the basis for our subsequent analysis.

An analysis of the sender/receiver information in the EDIFACT interchange headers (i.e., the Global Location Numbers (GLNs) in the UNB segment) of the messages reveals that CompanyB receives ORDERS messages from 13 different customers and sends INVOIC messages to six different customers. Three of these customers are overlapping, i.e., they handle both orders and invoices electronically via EDI. These three customers account for 1574 ($\sim$33%) of the 4751 complete cases. The remaining 3177 ($\sim$67%) complete cases originate from a fourth customer where ORDERs are sent from a subsidiary company having a different GLN then the headquarters receiving the INVOICes. In this case study, we use the EDI data of four different customers for our subsequent analyzes.

Inter-organizational Process Model – Process Mining. The resulting event log is further mined for a process model. In this task, we applied the *Heuristics Miner* algorithms [WAA06] for discovering the process model. Figure 7.4 shows the mined process model as a *flexible model* representation.

7.3.3 Balanced Scorecard Implementation using the EDImine BSC Framework

For modeling and calculating the BSC, we define business objectives as well as corresponding success factors and KPIs and apply them on the input data. In order to evaluate business performance against business objectives by using the EDImine BSC Framework, we firstly define business objectives and related success factors to be used in the BSC for this case study. We considered CompanyB's primary mission statement – which is the provision of products

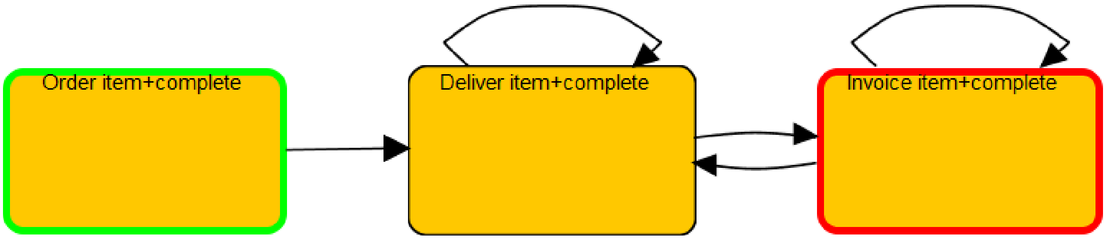


Figure 7.4: The mined process model of the second case study

and services of the highest quality – and translated it into business objectives which reflect this focus, as shown in Table 7.6. The “Improve product and service quality” business objective focuses on the quality of the manufactured goods as well as on related services such as delivery, after-sale services, etc. “Maintain customer satisfaction” reflects the organization’s intention to retain existing customers as well as to attract new customers as an indirect indicator of product and service quality. Furthermore, “Increase revenue” and “Improve operational performance” have been included as business objectives for CompanyB, as these represent typical goals of profit-oriented companies. Note that in real-world applications of the EDImine BSC Framework, business objectives may be derived from an already existing BSC of the company under analysis. As also shown in Table 7.6, for each of the business objectives we select the identified inter-organizational success factors from our knowledge base (cf. Chapter 6, Section 6.3.2) which relate to that objective. Based on the available data from CompanyB we identified concrete KPIs for measuring each success factor as shown in Table 7.6. The modeling of the KPI attributes such as critical thresholds relies on the information given from the company. Furthermore, the calculations of KPIs are provide in Table 7.7 In this case study, we consider business objectives having achievement scores less than 50% to be critical. The BSC is calculated monthly, hence, the score of business objectives and KPIs are calculated month by month.

Total revenue, as shown in Table 7.6, it is defined to reflect the success factor “Financial performance” We set the the target value of revenue to 600,000[6]. We consider a total revenue of less than 300,000 as critical.

Average revenue per customer and **Average ordered quantities per customer** are used to evaluate customer satisfaction. The target value of “Average revenue per customer” is one fourth of the target value of total revenue since CompanyB has four main customers as mentioned in the preprocessing step. Moreover, customer satisfaction is also reflected by ordered quantities which we model by means of a KPI “Average ordered quantities per customer”.

Number of late deliveries, **Percentage of on-time deliveries**, **Standard deviation of duration between requested delivery date and actual delivery date**, and **Average duration between requested delivery data and actual delivery date** are assigned as KPIs indicating the reliability of the company. In the process perspective, we focus on the performance of the delivery and invoicing processes. We define four KPIs related to delivery performance to reflect the success factor “Reliability”: “Number of late deliveries”, “Percentage of on-time deliveries”, “Standard deviation of duration between requested delivery date and actual delivery date”,

[6]We refrain from specifying units since all monetary figures and quantity figures have been altered in this thesis.

Table 7.6: The balanced scorecard for the second case study

Business Objective	**Success Factor**	**KPI**	**Weight (%)**	**Limit Type**	**Target (Critical thr.)**
Financial Perspective					
Increase revenue	Financial performance	Total revenue	100	More	600,000 (300,000)
Customer Perspective					
Maintain customer satisfaction	Satisfaction	Average revenue per customer	50	More	150,000 (60,000)
		Average ordered quantities per customer	50	More	18,000 (6,000)
Process Perspective					
Improve product and service quality	Reliability	Number of late deliveries	50	Less	0 (5)
		Percentage of on-time deliveries	30	More	100 (20)
		Standard deviation of duration between requested delivery date and actual delivery date	10	Two-side	0 (2)
		Average duration between requested delivery date and actual delivery date	10	Two-side	-1 (2)
Improve operational performance	Operational performance	Maximum duration of invoicing	50	Less	1 (7)
		Average duration of invoicing	50	Less	1 (2)

Analysis period of the BSC is set to “Monthly”.
Critical thresholds of all business objectives are 50% (having a score less than 50% is critical).
Weight of related KPIs must be 100% in total for each business objective.

and “Average duration between requested delivery data and actual delivery date”. “Number of late deliveries” can influence customer satisfaction and trust since late deliveries may harm the reputation of organizations. Since we want to emphasize the penalty on late deliveries, we give it a 50% weight which is half of the total score of the business objective “Improve product and service quality”. The optimal case is not to have any late deliveries, therefore we set the target value to zero and set the critical threshold[7] to five late deliveries. Similarly, the KPI “Percentage of on-time deliveries” reflects the reliability of CompanyB’s delivery service. The optimal case is to have 100% of on-time deliveries; less than 80% of on-time deliveries is considered critical. The KPI “Average duration between requested delivery date and actual delivery date” is also used to evaluate overall delivery performance. The duration between requested delivery date and actual delivery date should be as little as possible. We set the target value to -1 (i.e., delivery at most one day in advance) and the critical threshold to 2 days with the limit type as two-sided (i.e., more than three days early or one day late is considered critical).

Maximum duration of invoicing, and **Average duration of invoicing** are used as KPIs reflecting operational performance. For evaluating the operational performance, we focus on invoicing times and the duration between ordered date and actual delivery date. The KPI “Maximum duration of invoicing” is used to indicate the longest invoicing period after some delivery completed. We focus on the duration between delivery item event and its following invoice item event. In doing so, the calculation is implemented by subtracting timestamp of *Deliver-item* event from the corresponding timestamp of *Invoice-item* event. However, the querying mechanism of timestamps need to ensure the correctness of event sequence. Considering examples provided in Figure 7.5, there are several possibilities of event sequences according to the mined

[7] In this case study, we specify critical thresholds as relative values with respect to target values in order to allow for the simple definition of thresholds for two-sided KPIs.

Table 7.7: KPI calculation of the second case study

KPI	Calculation
Total revenue	SUM(invoiced amount of line item in INVOIC)
Average revenue per customer	SUM(invoiced amount of line item in INVOIC) / COUNTDIS(interchange sender in ORDERS) *Note: Counting distinct senders of ORDERS messages yields the total number of customers.*
Average ordered quantities per customer	SUM(ordered quantities of line item in ORDERS) / COUNTDIS(interchange sender in ORDERS)
Number of late deliveries	COUNT if (actual delivery date in INVOIC − requested delivery date in ORDERS) greater than or equal 1
Percentage of on-time deliveries	COUNT if (actual delivery date in INVOIC − requested delivery date in ORDERS) between 1 and -3 / COUNT(actual delivery date in INVOIC − requested delivery date in ORDERS) × 100
Standard deviation of duration between requested delivery date and actual delivery date	STDV(actual delivery date in INVOIC − requested delivery date in ORDERS)
Average duration between requested delivery date and actual delivery date	AVG(actual delivery date in INVOIC − requested delivery date in ORDERS)
Maximum duration of invoicing	MAX(timestamp of *Invoice item* event − timestamp of *Deliver-item event*)
Average duration of invoicing	AVG(timestamp of *Invoice item* event − timestamp of *Deliver-item event*)

KPI calculation formulas are described as aggregation functions applied over sets of results calculated from algebraic expressions. These algebraic expressions are applied on each of the process instances which start in the given analysis period

process model (cf. Fig. 7.4). In the case of process instance #1, it is obvious that the duration of invoicing is the duration between of *Deliver-item* event and its consecutive *Invoice-item* event (cf. Fig. 7.5, Mark 1). However, in the case of process instance #2 and #3, querying becomes ambiguous. In particular, there are two *Deliver-item* events followed by one *Invoice-item* event in process instance #2. This yields two possible pairs of *Deliver-item* event and *Invoice-item* event (i.e., Fig. 7.5, Mark 2 and 3). The confusion of acquiring the correct information becomes clearer in the example of process instance #3 where there are two *Deliver-item* events and each of them is followed by its corresponding *Invoice-item* event. This results in four possible pairs (i.e., Fig. 7.5, Mark 4, 5, 6 and 7). For calculating the duration of invoicing, we focus on time duration between *Deliver-item* event and its consecutive *Invoice-item* event. Since we want to measure time of invoice response after delivery finished. Therefore, the query is required to be limited to our interested pattern. As mentioned earlier, our framework leverages log replay concept for ensuring the correctness of information retrieval. By having a process model at hand, we can determine appropriate sequence patterns for querying any interested information. The log replay then enables querying mechanism to step through event log and to retrieve information correctly according to a specified pattern. Therefore, in our case of invoicing duration we specified that the query must follow the pattern having a *Deliver-item* event which consecutively followed by an *Invoice-item* event. Applying the query following this pattern on examples shown in Figure 7.5 the duration between event pair of 1, 3, 5, and 7 are retrieved. Normally, CompanyB's invoices should be issued 1-2 days after the delivery date. Hence, the target value is set to one day. However, invoicing later than one week is considered unusual. Hence, we set the critical threshold to seven days. In order to evaluate the overall performance of invoicing, the KPI "Average duration of invoicing" is applied. The query of invoicing duration of the previous mentioned KPI is also applied for this KPI. The majority of invoicing processes is expected to

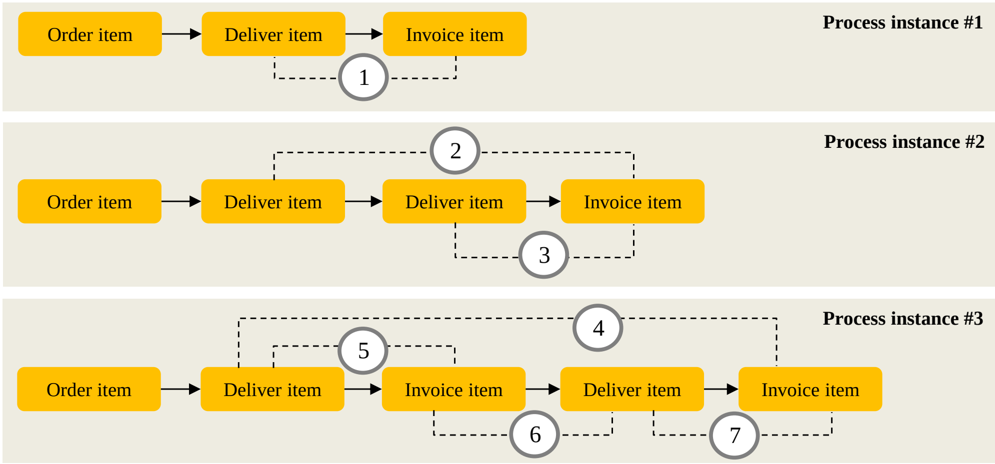


Figure 7.5: Examples of process instances possibly derived from the mined model (cf. Fig. 7.4)

last around 1-2 days. Therefore, the average duration of invoicing should be one day (i.e., one day after some delivery).

According to the above described BSC model and definition of KPIs, we calculate the achievement scores of each of the KPIs. In turn, the achievement scores of the business objectives can be calculated as the weighted sum of the related KPIs' scores. To avoid over- and underestimation, for both business objectives and KPIs we limit maximum and minimum scores to 100% and 0% respectively.

7.3.4 Results and Discussion

Since the EDI messages were collected between March 2013 and the beginning of June 2013, KPI scores and business objective scores were calculated for March 2013, April 2013 and May 2013. There are no results for the period of June 2013 because the EDI messages sent/received in this period belong to the process instances that start in the previous months (i.e., there are no *Order item* events in June). Table 7.8 shows the calculated BSC for these three months. Since in this case performance results are diverse. Therefore, we visualized the status of results by gray-scale. In particular, results with no highlight means they are still in a desired state, whereas results highlighted by light-gray and dark-gray means they are not good and critical respectively.

In the period of March 2013, CompanyB perfectly achieves its business objectives in both the financial and customer perspectives. The business objectives "Increase revenue" and "Maintain customer satisfaction" are successfully achieved with a score of 100% since all of their related KPIs score 100% as well. However, the KPIs of the process perspective exhibit less desirable scores. Delivery performance – reflecting the business objective "Improve product and service quality" – is much lower than targeted. There are four late deliveries in this month, which is only slightly below the critical threshold of five late deliveries per month. Similarly, the percentage of on-time deliveries and the standard deviation of duration between requested delivery date and actual delivery date are also achieved lower than the expectation. Although none of the KPIs

Table 7.8: The balanced scorecard calculated from March 2013 to May 2013

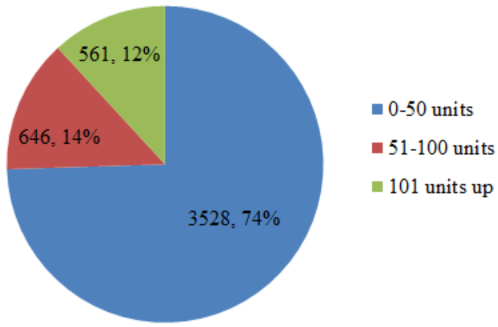
Business Objective / KPI	March 2013 Score (%)	March 2013 Actual Value	April 2013 Score (%)	April 2013 Actual Value	May 2013 Score (%)	May 2013 Actual Value
Financial perspective						
Increase revenue	100	n/a	58.61	n/a	2.74	n/a
Total revenue	100	682,088	58.61	475,832	2.74	308,209
Customer perspective						
Maintain customer satisfaction	100	n/a	36.58	n/a	0	n/a
Average revenue per customer	100	170,522	48.26	118,958	0	77,052
Average ordered quantities per customer	100	19,359	24.89	13,493	0	9,148
Process perspective						
Improve product and service quality	25.72	n/a	56.44	n/a	50.61	n/a
Number of late deliveries	20	4 times	100	0 times	80	1 times
Percentage of on-time deliveries	23.18	84.64%	0	78.45%	11.69	82.34%
Standard deviation of duration between requested and actual delivery date	33.84	1.32 days	28.89	1.42 days	28.54	1.43 days
Average duration between requested and actual delivery date	53.78	-1.92 days	35.52	-2.29 days	42.53	-2.15 days
Improve operational performance	64.59	n/a	57.15	n/a	57.15	n/a
Maximum duration of invoicing	29.17	5.96 days	14.29	7 days	14.29	7 days
Average duration of invoicing	100	0.25 days	100	0.28 days	100	0.27 days

The status of the results are visualized by gray-scale since the results are diverse. Results with no highlight means they satisfy the target values. Results highlighted with light-gray and dark-gray means their status are not good and critical respectively.

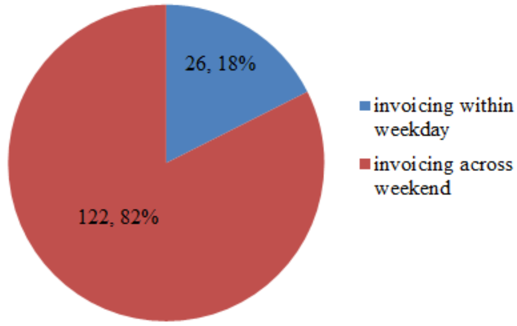
for "Improve product and service quality" is critical, the business objective itself is in a critical status since the overall achievement score is lower than 50%. However, the business objective "Improve operational performance", focusing on invoicing processes, is still acceptable.

In April and May 2013, the performance indicators of the financial and customer perspectives significantly drop. The total revenue decreases and its score drops approximately 50% each month. Similar to the financial perspective, the performance in the customer perspective drops as well. The "Maintain customer satisfaction" business objective is in a critical state. In particular, the score of the KPIs "Average revenue per customer" and "Average ordered quantities per customer" drop to 48.26% and 24.89%, respectively, in April and both severely drop to 0% in May. While the KPIs in the financial and customer perspectives keep declining, the KPIs of the process perspective remain steady.

In April, scores of the business objective"Improve product and service quality" seem to be slightly improved. The variation and the average duration between requested delivery date and actual delivery date are slightly worse than last month. In other words, deliveries are more unreliable. However, the KPI "Number of late deliveries" achieves a score of 100% since there are no late deliveries in this month. In contrast, the score of the KPI "Percentage of on-time deliveries" drops to 0%. This means that there are many early deliveries. Nevertheless, the very positive score of the KPI "Number of late deliveries" accounts for the overall acceptable score of the business objective "Improve product and service quality" of 56.44%. In the process perspective, CompanyB's performance in May is similar to that of April. There is only one late delivery which is reflected in a good status of the KPI "Number of late deliveries". In addition,



(a) Number of cases by ordered quantities



(b) Number of cases by invoice process time frame

Figure 7.6: Pie charts

the percentage of on-time deliveries as well as the average and variation of durations between requested delivery date and actual delivery date are also similar to the previous months.

We further investigated the cases of late deliveries as well as late invoices for deriving clues for such anomalies. In doing so, other analysis techniques can be applied for answering in-depth questions, such as "What are factors affecting delivery performance?", "How much does customer satisfaction depend on operational performance?", etc. (cf. [ER14]). According to the analysis of the cases having late deliveries, we found that their corresponding ordered quantities are quite large. In detail, we analyzed the number of cases according to ordered quantities. Figure 7.6a shows a pie chart describing the distribution of cases regarding ordered quantities. It is shown that the majority (74%) of all cases has ordered quantities up to 50. 14% of all cases have ordered quantities between 51-100, whereas the rest of the cases (12%) have ordered quantities above 100. By considering ordered quantities of five late-delivery cases, we found that two of them have ordered quantities of more than 100 (i.e., 460.8 and 194.4) and another two of them have ordered quantities between 51-100 (i.e., 64.8 and 97.2). These four cases have a high number of ordered quantities compared to the majority of all cases. Only one case of late deliveries has an ordered quantity of 43.2. However, this case refers to the line item contained in the same order of the other two late deliveries (the aforementioned cases having ordered quantities of 194.4 and 97.2). Therefore, the late delivery in this case may have been affected by large ordered quantities of the other two cases. Moreover, we investigated the cases having late invoices. There are 148 cases having invoices more than two days after deliveries. We analyzed the time frame between the activity of "Delivery item" and "Invoice item" of these cases by using the dotted chart analysis provided in ProM 6 [SA07]. Figure 7.7 shows the dotted chart of cases having late invoices. The analysis showed that 82.43% of late-invoice cases have time frames which included weekends, as depicted in Figure 7.6b. Therefore, we subtracted two days (i.e., Saturday and Sunday) from the duration between "Delivery item" and "Invoice item" of those cases in order to obtain the accurate total working days of invoicing. In total, we found 108 cases which took more than two working days for invoicing. Among these cases, 73% belong to cases of one particular customer.

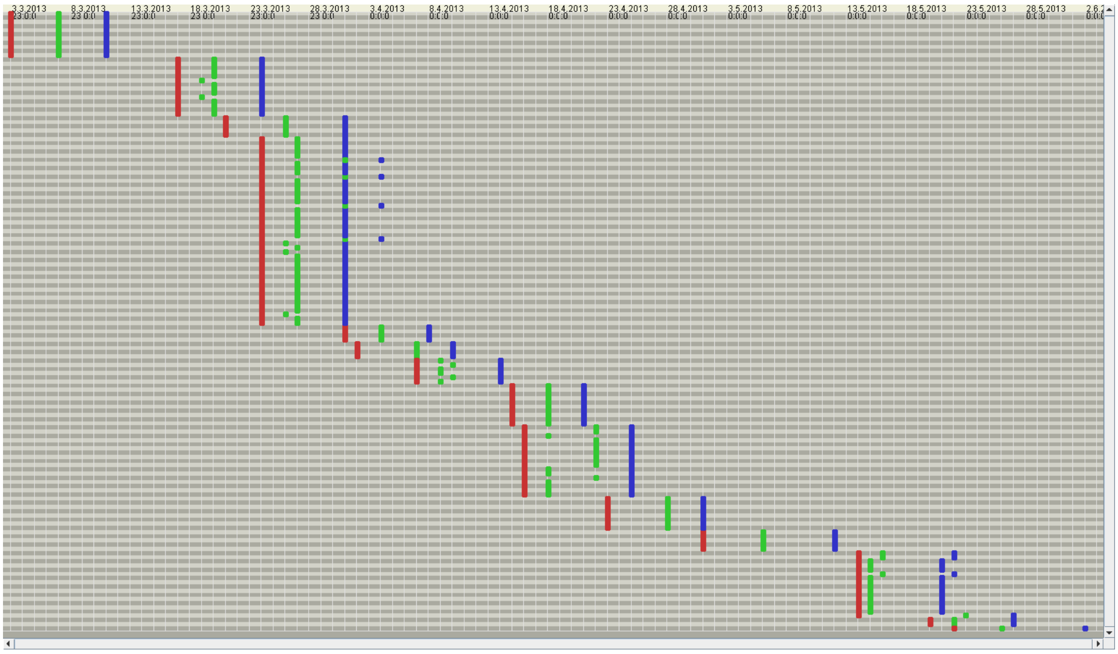


Figure 7.7: Dotted chart showing time frame of 148 cases of late invoices (Dots in red, green, and blue represent *Order-item*, *Deliver-item*, and *Invoice-item* event respectively.)

In summary, the scores of the business objectives in the financial and customer perspective keep falling in each of the months. The average of ordered quantities per customer dropped around 30% each month. Consequently, the total revenue also keeps declining. This might be the result of poor operational performance since all related business objectives score low. As a consequence this may reduce customer satisfaction which in turn leads to declining revenues. However, this cannot be concluded with certainty from the results since the analysis period of three months is too short. Nevertheless, the results suggest that CompanyB may investigate the underlying cause for the low scores of KPIs related to customer satisfaction as well as put additional efforts into the improvement of operations performance. From our in-depth investigations, three main insights can be derived. First, the analysis of late deliveries shows that ordered quantities may be the cause of the delays. Second, most of the late-invoice cases occurred during weekends. Finally, late invoices usually belong to the cases of one particular customer. According to our findings, the company should further analyze their manufacturing or delivery processes especially in the cases of large ordered quantities for finding the root cause of the poor delivery performance. Furthermore, the company should pay attention on their invoicing processes that are spanned over weekends in order to prevent the delays of invoices. Furthermore, they should inspect the reason of those late invoices towards the one particular customer that has been found as part of our investigation.

When comparing CompanyB's *real* revenue figures (i.e., as disclosed to us by a company representative) with the revenue figures from our results based on EDI data, it turns out that only a fraction (between 5% and 50%) of the actual revenue of the company gets reflected in our analysis. This implies that a significant portion of real-world business transactions of CompanyB is actually not reflected in the sample of EDI messages used in this case study. Furthermore, information in EDI messages is limited to certain kinds and, hence, some KPIs of interest may not be derived from EDI data at all. For instance, in this case study KPIs related to the learning and growth perspectives of BSCs (e.g., number of new products, employee turnover rate, etc.) as well as some KPIs which may directly reflect CompanyB's business objectives (e.g., profit, number of customer complaints, etc.) could not be derived.

7.4 Summary

In this chapter, we presented two case studies demonstrating the inter-organizational performance analysis with the EDImine BSC Framework (cf. Chapter 6). The first case study presented in this chapter has been published in [Kra+], while the second case study is currently under submission.

The case studies demonstrate that the framework supports an organization in evaluating inter-organizational performance based on EDI messages. With the integration of top-down and bottom-up performance analysis approaches (i.e., the BSC method and process mining), organizations are able to define their strategies and evaluate them quantitatively through the alignment of quantifiable KPIs. The case studies show that, on the one hand, the framework allows the companies modeling their balanced scorecards following a top-down approach. On the other hand, it allows the companies identifying, defining as well as calculating KPIs from the information derived using the bottom-up approach. By connecting those KPIs to the strategies defined in the scorecards, the companies can quantitatively assess their business performance against business objectives on a strategic level.

According to the case studies, we found that deriving business performance results solely from EDI data may not suffice for holistic business performance analysis. Furthermore, the case studies show that inter-organizational processes derived from EDI data are relatively simple in real-world setting. In other words, the result yielded from the presented framework covers only the analysis on transactions realized through EDI. However, in reality, inter-organizational business processes may be realized through other different approaches as well, such as paper-based documents, telephone, or emails. Moreover, one business transaction may rely on multiple approaches. For instance, a customer may place an order and receive an invoice as EDI messages, but the notifications about delivery or defects belonging to the same transaction may be communicated via emails or telephone. In order to derive more complex business processes and enable the performance analysis for the entire organization, every data source involved in business transactions have to be taken into account.

CHAPTER 8

Conclusion

In this thesis, we presented a framework for evaluating inter-organizational performance based on EDI messages which are exchanged in the course of IORs. In this chapter, a summary of this thesis is provided in Section 8.1 followed by a summary of the contributions in Section 8.2. Finally, the limitations of our work as well as an outlook on future works is provided in Section 8.3.

8.1 Summary of this Thesis

The goal of the research conducted in this thesis is to provide an approach for evaluating IORs from EDI messages. To accomplish the goal, we addressed the sub-problems of (i) extracting business information from EDI messages, (ii) identifying inter-organizational KPIs, and (iii) lifting performance evaluation to the strategic level. The ontological approach for business information extraction was developed for tackling the first problem. In addressing the second problem, we first identified inter-organizational success factors and studied their influencing relationships as well as their measurements by conducting a literature review. Then, we investigated industrial MIGs of UN/EDIFACT standards for identifying KPIs and aggregated those KPIs into corresponding success factors. For aggregating the KPIs we considered the measurements of the success factors found in the review. Built upon these artifacts (i.e., the business information extraction approach and the knowledge of inter-organizational success factors and KPIs), the framework for evaluating inter-organizational performance from EDI messages was developed. The framework addresses the third problem of lifting an evaluation to the strategic level. In the framework, we applied BSC method and process mining techniques for the top-down alignment of KPIs and business strategies as well as for the bottom-up definition and calculation of KPIs based on the information derived from EDI messages. The framework was demonstrated using two case studies. In the case studies, the performance evaluation was performed by using our proposed framework. The case studies showed the feasibility in both, practical and technical aspects.

8.2 Contributions of this Thesis

The research yields four main contributions which are not only applied in the scope of our work but are designed to be reused in future works. These contributions are summarized in the following.

An Approach for Business Information Extraction from EDI Messages. One of the cornerstones of the work presented is an ontological approach for business information extraction. The approach is built upon EDI ontologies where EDI data is preliminarily interpreted according to codes and qualifiers. Furthermore, the interpreted data is stored in ontologies as well [Eng+12b]. The business information extraction approach *extends* EDI ontologies by generating the Business Information (BI) ontologies on top of the existing EDI ontologies. The BI ontology contains generic BI concepts which are used to represent EDI data on a conceptual level. The generation of these BI concepts is based on user-defined mappings which are defined as ontological rules. Through the mappings, EDI data is classified into corresponding BI concepts, and hence they are conceptually represented as generic BI concepts.

Therefore, the approach supports automated interpretation of arbitrary EDI messages based on codes and qualifiers provided in their specifications. At the same time, the approach also provides flexibility through user-defined mappings which allows conceptualizing EDI data into generic BI concepts for any specific business domains.

Inter-organizational Success Factors and Influencing Model. In order to identify KPIs related to the success of IORs, we started by investigating inter-organizational success factors. The conducted review investigated success factors impacting IORs. In the review, we identified the success factors having an impact on the success of IORs as well as the influencing relationships between those success factors. Based on the review, we *created* a cause and effect model of inter-organizational success factors and *developed* inference rules deriving additional influencing relationships from the model. The model explains the minimum set of influencing relationships between success factors. The full set of influencing relationships can be obtained by applying the inference rules.

In the scope of this thesis, the cause and effect model is stored as a knowledge base that supports KPI suggestion against success factors. Furthermore, the knowledge of influencing relationships between inter-organizational success factors leads to the understanding of their impacts on the success of IORs. Such an understanding is required for strategic management, especially in inter-organizational settings. Having this knowledge at hand, strategists can identify as well as effectively define relevant success factors which are important for their organizations.

Inter-organizational KPIs and the Method of KPI Identification from EDI Messages. In addition studying inter-organizational success factors, we further investigated EDI messages in order to derive KPIs and aggregate these KPIs into appropriate success factors found in the aforementioned review. Therefore, a method for identifying KPIs from EDI messages was *created*. The method was applied on industrial MIGs which represent generalizations of EDI data used in specific industries. The idea of the method is to consider both, (i) frequency of occurrences of EDI data elements as well as (ii) semantics of EDI data elements and message types. The frequency of occurrences of EDI data elements implies the possibility of the data element appearing in the real-world scenarios, whereas considering the purpose of message types ensures the accuracy of the information. By applying this method, we *derived* a set of inter-organizational KPIs and grouped them in corresponding success factors.

The identified KPIs are added to the aforementioned knowledge base of success factors for supporting KPI suggestion. The set of identified KPIs and their calculation guidelines provide means of quantitative measurement for evaluating IORs. Furthermore, the method of KPI identifications developed in this thesis can be applied for deriving KPIs from any EDI standards, since it is independent of the syntax.

Inter-organizational Performance Analysis Framework. The EDImine BSC Framework presented in this thesis supports the inter-organizational business performance analysis from EDI messages. The framework integrates top-down and bottom-up performance analysis approaches. In detail, the framework allows top-down alignment of business objectives at a strategic level and KPIs at an operational level. At the same time, it also allows a bottom-up definition and calculation of KPIs based on the information derived from EDI messages. The information from EDI messages is derived using process mining techniques and other preprocessing methods (i.e., business information extraction). The BSC method allows us to connect KPIs to a strategic level, whereas process mining allows us to derive process models executed in reality. Using derived process models ensures that KPI definitions and calculations reflect the real-world context instead of merely relying on planned and/or a-priori models. According to the case studies presented in Chapter 7, the case studies show the feasibility of the framework in both, technical and practical aspects. Based on the case studies, the framework allows companies to evaluate their inter-organizational business performance quantitatively on a strategic level based on the information derived solely from reality (i.e., EDI messages and their exchange log). Consequently, this enables a quick assessment of the impact of business performance on IORs.

The integration of the BSC method and process mining leverages the benefits of both top-down and bottom-up approaches. Best to our knowledge, there exists no technical framework of performance analysis that provides such an integration. Hence, the framework *extends* state-of-the-art of BSC and process mining in the context of performance analysis.

Furthermore, the framework leverages semantic technologies for implementing the BSC method which enables the automation of BSC calculations as well as supports the definition of KPIs on a conceptual level. It is flexible in terms of the calculation since the ontology is independent from any particular transfer syntax (e.g. XML, EDIFACT, etc.) and for supporting different syntaxes *only* the calculation needs to be adapted at a technical level while the ontology always stays the same. Moreover, the application of semantic technology also provides the ability of inferring new knowledge through logical rules. As described in our work, we applied inference rules on the knowledge base of success factors and KPIs for automatically identifying available KPIs.

Moreover, the development of the contributions presented in this thesis is based on the combination of techniques from multiple *sub-disciplines* in the field of computer science. These include semantic and reasoning technologies, process mining, as well as management science. Hence, this work can be considered as interdisciplinary research where contributions are developed across these disciplines. Therefore, in addition to the aforementioned contributions, the value of this research is to leverage existing knowledge from different disciplines and apply it for tackling the problem beyond the context of those disciplines.

8.3 Limitations and Future Work

The scope of the proposed framework concentrates on the business performance analysis related to IORs. The analysis is based on information derived from EDI messages exchanged among business partners. In the following, the limitations of the framework as well as remaining challenges, beyond the scope of this thesis, are presented.

Integration of Different Data Sources. The case studies indicate that an analysis solely based on EDI messages covers only one part of an organization's performance. One reason is that the companies communicate and exchange business information with their business partners in several ways other than EDI. For instance, in the case study of a consumer good manufacturing company (cf. Chapter 7, Section 7.3) we found that only a small set of all transactions is realized by EDI. This implies that the remaining transactions are realized by some other approaches. These approaches include paper-based documents, fax, telephone, email, etc. Moreover, sometimes one complete transaction may be realized by different message exchange approaches for different tasks involved in one transaction. For example, a customer and a seller may exchange their orders and invoices using EDI messages but during the delivery period they may contact each other (e.g., enquiry about the deliveries, delivery problems report, etc.) via telephone or emails. Hence, in order to enable an analysis covering entire organization's performance, information from relevant data sources must be integrated for supporting an analysis task. This raises the challenge of an integration of heterogeneous data models, data representations, as well as data conceptualizations in different domains. One aspect of enabling such an integration would be the definition of one common data model for representing data from different heterogeneous sources/domains.

Intra- and Inter-organizational Processes. As mentioned earlier, in this thesis we focus on inter-organizational performance. The framework provides analysis capabilities based on inter-organizational processes and business information exchanged between business partners. However, investigating internal processes additionally may reveal important insights which might lead to a more effective (inter-organizational) business process management. In other words, the integration of inter- and intra-organizational processes provides an understanding of business activities within and outside organizational boundaries. This understanding can lead to an improvement of business processes. Furthermore, in case a problem arises an investigation on a complete and comprehensive view of business processes integrating both, inter- and intra-perspective, can help organizations tackling the problem properly. In particular, investigating a problem based on a complete view of processes, organizations can quickly identify the cause of the problem. Furthermore, organizations can forecast the impact of the problem on themselves and their business partners. The ability of problem prediction and planning is important for business collaboration, especially in the context of supply chains. For instance, a single problem spotted in a chain can affect business partners along the whole supply chain. Therefore, the integration between inter- and intra-organizational processes can be considered as a potential future research opportunity.

Qualitative Performance Analysis Approaches. While the presented framework provides a quantitative approach for measuring inter-organizational performance, qualitative measurements are beyond the scope of our framework. Nevertheless, according to our work on the identification of inter-organizational KPIs, we found that some success factors related to the success of IORs are difficult to be measured as numbers. In this case, the measurements through qualitative approaches are still necessary, such as interviews or surveys. Therefore, combining qualitative approaches with quantitative approaches can provide complete means for performance evaluation. However, the challenge of qualitative measurement is an ambiguity of measurement results since qualitative measures cannot be expressed as definite numbers. Hence, integrating qualitative and quantitative approaches becomes challenging. This poses the problem of interpreting qualitative results and appropriately integrating them in a quantitative approach.

Inter-organizational Success Factors and Influencing Model. In the work of identifying inter-organizational success factors, we identified success factors related to IORs as well as derived their influencing relationships. The influencing model, the so-called *cause and effect model*, explains the impacts of success factors on others. In this thesis, we focused only directed influencing relationships. The studies on the tendency of relationships (i.e., positive relationships or negative relationships) and the weighted impact on relationships for enhancing the influencing model remains for future research. The application of the model is not only limited to our proposed framework, but can also be used to provide an understanding of success factors and their impacts on IORs. This in turn supports strategic management in an inter-organizational context.

Case Studies. Case studies provided in this thesis only demonstrate the feasibility and applicability of the framework in both, technical and practical, aspects. However, they are not sufficient for evaluating the usability and quality of the framework since the companies did not participate in the case studies. Instead, they only provided us with sample EDI data. Although we communicated with the company in the second case study, they can only provide the basic information about their delivery approach, invoicing process, and estimated income. However, the information about their business strategies and KPIs of interest are difficult to obtain. This is because such information is confidential, cannot be disclosed to public, and our access to such companies is limited. Therefore, the evaluation of the presented framework remains for the future work.

Bibliography

[Aal11] Wil M.P. van der Aalst. *Process Mining: Discovery, Conformance and Enhancement of Business Processes*. Springer, 2011 (pages 5, 31–34, 92).

[AAD12] Wil M.P. van der Aalst, Arya Adriansyah, and Boudewijn F. van Dongen. "Replaying History on Process Models for Conformance Checking and Performance Analysis". In: *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery* 2.2 (2012), pp. 182–192 (pages 32, 34, 99, 115).

[AA05] Wil M.P. van der Aalst and Ana Karla Alves de Medeiros. "Process Mining and Security: Detecting Anomalous Process Executions and Checking Process Conformance". In: *the 2nd International Workshop on Security Issues with Petri Nets and other Computational Models (WISP 2004)*. Vol. 121. Elsevier, 2005, pp. 3–21 (page 32).

[AAW05] Wil M.P. van der Aalst, Ana Karla Alves de Medeiros, and Anton J.M.M. Weijters. "Genetic Process Mining". In: *Proceedings of the 26th International Conference on Applications and Theory of Petri Nets (ICATPN 2005)*. Ed. by Gianfranco Ciardo and Philippe Darondeau. Vol. 3536. Lecture Notes in Computer Science. Springer Berlin Heidelberg, 2005, pp. 48–69 (page 32).

[Aal+08] Wil M.P. van der Aalst, Marlon Dumas, Chun Ouyang, Anne Rozinat, and H.M.W.-(Eric) Verbeek. "Conformance Checking of Service Behavior". In: *ACM Transactions on Internet Technology* 8.3 (2008), 13:1–13:30 (page 32).

[Aal+07] Wil M.P. van der Aalst, Hajo A. Reijers, Anton J.M.M. Weijters, Boudewijn F. van Dongen, Ana Karla Alves de Medeiros, Minseok Song, and H.M.W. (Eric) Verbeek. "Business Process Mining: An Industrial Application". In: *Information Systems* 32.5 (5 2007), pp. 713–732 (pages 5, 92).

[AW04] Wil M.P. van der Aalst and Anton J.M.M. Weijters. "Process Mining: A Research Agenda". In: *Computers in Industry - Special issue: Process/workflow mining* 53.3 (2004), pp. 231–244 (page 32).

[AWM04] Wil M.P. van der Aalst, Anton J.M.M. Weijters, and Laura Maruster. "Workflow Mining: Discovering Process Models from Event Logs". In: *IEEE Transactions on Knowledge and Data Engineering* 16.9 (2004), pp. 1128–1142 (pages 15, 32, 92).

[Aal+12] Wil M.P. van der Aalst, Arya Adriansyah, Ana Karla Alves de Medeiros, Franco Arcieri, Thomas Baier, Tobias Blickle, Jagadeesh Chandra Bose Rantham Prabhakara, Peter van den Brand, Ronald Brandtjen, Joos Buijs, et al. "Process Mining Manifesto". In: *the 7th International Workshops on Business Process Intelligence (BPI 2011), Business Process Management Workshops*. Vol. 99. Lecture Notes in Business Information Processing. 2012, pp. 169–194 (page 34).

[AML10] Eamonn Ambrose, Donna Marshall, and Daniel Lynch. "Buyer Supplier Perspectives on Supply Chain Relationships". In: *International Journal of Operations & Production Management* 30.12 (2010), pp. 1269–1290 (pages 59–61, 63, 65, 78, 86–88, 154, 155).

[AV04] Grigoris Antoniou and Frank Van Harmelen. "Web Ontology Language: OWL". In: *Handbook on Ontologies*. Springer, 2004, pp. 67–92 (page 105).

[Ara+07] Lusine H. Aramyan, Alfons G.J.M. Oude Lansink, Jack G.A.J. van der Vorst, and Olaf van Kooten. "Performance Measurement in Agri-food Supply Chains: A Case Study". In: *Supply Chain Management: An International Journal* 12.4 (2007), pp. 304–315 (pages 59, 86, 152–155).

[Ash+09] Bahar Ashnai, Maria Smirnova, Sergei Kouchtch, Qionglei Yu, Bradley R. Barnes, and Peter Naudé. "Assessing Relationship Quality in Four Business-to-Business Markets". In: *Marketing Intelligence & Planning* 27.1 (2009), pp. 86–102 (pages 59, 60, 63, 154).

[Baa+03] Franz Baader, Diego Calvanese, Deborah McGuinness, Daniele Nardi, and Peter Patel-Schneider, eds. *The Description Logic Handbook: Theory, Implementation, and Applications*. Cambridge University Press, 2003 (page 48).

[Bag+05] Prabir K. Bagchi, Byoung Chun Ha, Tage Skjoett-Larsen, and Lars Boege Soerensen. "Supply Chain Integration: A European Survey". In: *The International Journal of Logistics Management* 16.2 (2005), pp. 275–294 (pages 86, 152–155).

[BDO06] Alistair Barros, Marlon Dumas, and Phillipa Oaks. "Standards for Web Service Choreography and Orchestration: Status and Perspectives". In: *International Workshop on Web Service Choreography and Orchestration for Business Process Management (WSCOBPM 2005), Business Process Management Workshops*. Ed. by C. Bussler and A. Haller. Vol. 3812. Lecture Notes in Computer Science. Springer, 2006, pp. 61–74 (page 1).

[BHM07] Domenico Beneventano, Sabina El Haoum, and Daniele Montanari. "Mapping of heterogeneous schemata, business structures, and terminologies". In: *the 18th International Workshop on Database and Expert Systems Applications (DEXA 2007)*. IEEE Computer Society, 2007, pp. 412–418 (pages 23, 37).

[Ber94] John Berge. *The EDIFACT Standards*. Blackwell Publishers, Inc., 1994 (pages 1, 20, 35).

[BS07] Rajat Bhagwat and Milind Kumar Sharma. "Performance Measurement of Supply Chain Management: A Balanced Scorecard Approach". In: *Computers & Industrial Engineering* 53.1 (2007), pp. 43–62 (pages 29, 86, 87, 89, 151, 153–155).

[BB10] Barbara Bigliardi and Eleonora Bottani. “Performance Measurement in the Food Supply Chain: A Balanced Scorecard Approach”. In: *Facilities* 28.5 (2010), pp. 249–260 (pages 86, 151, 153–155).

[Bob+09] Fernando Bobillo, Miguel Delgado, Juan Gómez-Romero, and Enrique López. “A Semantic Fuzzy Expert System for a Fuzzy Balanced Scorecard”. In: *Expert Systems with Applications* 36.1 (2009), pp. 423–433 (page 93).

[Bre+07] Pearl Brereton, Barbara A. Kitchenham, David Budgen, Mark Turner, and Mohamed Khalil. “Lessons from Applying the Systematic Literature Review Process within the Software Engineering Domain”. In: *Journal of Systems and Software* 80.4 (2007), pp. 571–583 (pages 52, 53).

[BS00] Peter C. Brewer and Thomas W. Speh. “Using the Balanced Scorecard to Measure Supply Chain Performance”. In: *Journal of Business Logistics* 21.1 (2000), pp. 75–93 (pages 15, 30, 86, 87, 93, 153–155).

[BKV02] Hans-Jörg Bullinger, Michael Kühner, and Antonius Van Hoof. “Analysing supply chain performance using a balanced measurement method”. In: *International Journal of Production Research* 40.15 (2002), pp. 3533–3543. eprint: `http://www.tandfonline.com/doi/pdf/10.1080/00207540210161669` (page 29).

[BSV13] Andrea Burattin, Alessandro Sperduti, and Marco Veluscek. “Business Models Enhancement through Discovery of Roles”. In: *IEEE Symposium on Computational Intelligence and Data Mining (CIDM 2013)*. 2013, pp. 103–110 (page 33).

[CP08] Jesús J. Cambra-Fierro and Yolanda Polo-Redondo. “Creating Satisfaction in the Demand-Supply Chain: The Buyers’ Perspective”. In: *Supply Chain Management: An International Journal* 13.3 (2008), pp. 211–224 (pages 59, 60, 62, 63, 85).

[CP02] Amelia S. Carr and John N. Pearson. “The Impact of Purchasing and Supplier Involvement on Strategic Purchasing and Its Impact on Firm’s Performance”. In: *International Journal of Operations & Production Management* 22.9 (2002), pp. 1032–1053 (page 58).

[Cas08] Mary Casey. “Partnership - Success Factors of Interorganizational Relationships”. In: *Journal of Nursing Management* 16.1 (2008), pp. 72–83 (pages 13, 78).

[Cha03] Felix T.S. Chan. “Performance Measurement in a Supply Chain”. In: *The International Journal of Advanced Manufacturing Technology* 21.7 (2003), pp. 534–548 (pages 59, 65, 86, 153, 154).

[CQ03] Felix T.S. Chan and H.J. Qi. “An Innovative Performance Measurement Method for Supply Chain Management”. In: *Supply Chain Management: An International Journal* 8.3 (2003), pp. 209–223 (pages 86, 87, 152–155).

[CP04] Injazz J. Chen and Antony Paulraj. “Towards a Theory of Supply Chain Management: The Constructs and Measurements”. In: *Journal of Operations Management* 22.2 (2004), pp. 119–150 (pages 58, 59, 63, 65, 86, 87, 154, 155).

[CPL04] Injazz J. Chen, Antony Paulraj, and Augustine A. Lado. “Strategic Purchasing, Supply Management, and Firm Performance”. In: *Journal of Operations Management* 22.5 (2004), pp. 505–523 (pages 58, 59, 62, 63, 66, 87, 88, 154, 155).

[Che11] Jao-Hong Cheng. “Inter-organizational Relationships and Information Sharing in Supply Chains”. In: *International Journal of Information Management* 31.4 (2011), pp. 374–384 (pages 13, 57, 59, 60, 63, 64, 78, 85).

[CS12] Jao-Hong Cheng and Jiuh-Biing Sheu. “Inter-organizational Relationships and Strategy Quality in Green Supply Chains - Moderated by Opportunistic Behavior and Dysfunctional Conflict”. In: *Industrial Marketing Management* 41.4 (2012), pp. 563–572 (pages 58–60, 66).

[CMM10] Mee-Shew Cheung, Matthew B. Myers, and John T. Mentzer. “Does Relationship Learning Lead to Relationship Value? A Cross-national Supply Chain Investigation”. In: *Journal of Operations Management* 28.6 (2010), pp. 472–487 (pages 60, 64–67, 86, 87, 154, 155).

[CGH09] Adrien Chia, Mark Goh, and Sin-Hoon Hum. “Performance Measurement in Supply Chain Entities: Balanced Scorecard Perspective”. In: *Benchmarking: An International Journal* 16.5 (2009), pp. 605–620 (pages 15, 29, 86, 89, 93, 154).

[Cho+11] Kavita Choudhary, Umasanker Pandey, Malay Kumar Nayak, and Durgesh Kumar Mishra. “Electronic Data Interchange: A Review”. In: *Proceedings of the 3rd International Conference on Computational Intelligence, Communication Systems and Networks (CICSyN 2011)*. July 2011, pp. 323–327 (page 1).

[CWR08] Marloes J.T. Claassen, Arjan J. van Weele, and Erik M. van Raaij. “Performance Outcomes and Success Factors of Vendor Managed Inventory (VMI)”. In: *Supply Chain Management: An International Journal* 13.6 (2008), pp. 406–414 (pages 87, 154, 155).

[CE93] Martha C. Cooper and Lisa M. Ellram. “Characteristics of Supply Chain Management and the Implications for Purchasing and Logistics Strategy”. In: *The International Journal of Logistics Management* 4.2 (1993), pp. 13–24 (page 58).

[CF05] Daniel Corsten and Jan Felde. “Exploring the Performance Effects of Key-Supplier Collaboration: An Empirical Investigation into Swiss Buyer-supplier Relationships”. In: *International Journal of Physical Distribution & Logistics Management* 35.6 (2005), pp. 445–461 (pages 65, 86, 154).

[CGP11] Daniel Corsten, Thomas Gruen, and Marion Peyinghaus. “The Effects of Supplier-to-Buyer Identification on Operational Performance - An Empirical Investigation of Inter-organizational Identification in Automotive Relationships”. In: *Journal of Operations Management* 29.6 (2011), pp. 549–560 (pages 59, 60, 63, 65, 66, 78, 85, 87, 154).

[CLS08] Paul D. Cousins, Benn Lawson, and Brian Squire. “Performance Measurement in Strategic Buyer-supplier Relationships: The Mediating Role of Socialization Mechanisms”. In: *International Journal of Operations & Production Management* 28.3 (2008), pp. 238–258 (pages 63, 86, 153–155).

[Dau+06] Patricia J. Daugherty, R. Glenn Richey, Anthony S. Roath, Soonhong Min, Haozhe Chen, Aaron D. Arndt, and Stefan E. Genchev. "Is Collaboration Paying Off for Firms?" In: *Business Horizons* 49.1 (2006), pp. 61–70 (page 1).

[Dav99] Fred R. David. *Strategic Management: Concepts and Cases*. 13th. Prentice Hall, 1999 (page 28).

[DW09] Jim Dawson and John Wainwright. "Processing EDI Code Pairs". In: *Pro Mapping in BizTalk Server 2009*. Apress, 2009, pp. 317–333 (pages 23, 37).

[Don+05] Boudewijn F. van Dongen, Ana Karla Alves de Medeiros, H.M.W. (Eric) Verbeek, Anton J.M.M. Weijters, and Wil M.P. van der Aalst. "The ProM Framework: A New Era in Process Mining Tool Support". In: *Proceedings of the 26th International Conference on Applications and Theory of Petri Nets (ICATPN 2005)*. Ed. by Gianfranco Ciardo and Philippe Darondeau. Vol. 3536. Lecture Notes in Computer Science. Springer Berlin Heidelberg, 2005, pp. 444–454 (page 92).

[Dou96] William R. Doucette. "The Influence of Relational Norms and Trust on Customer Satisfaction in Interfirm Exchange Relationships". In: *Journal of Consumer Satisfaction, Dissatisfaction and Complaning Behavior* 9 (1996), pp. 95–103 (page 62).

[Duf+12] Rachel Duffy, Andrew Fearne, Sue Hornibrook, Karise Hutchinson, and Andrea Reid. "Engaging Suppliers in CRM: The Role of Justice in Buyer–Supplier Relationships". In: *International Journal of Information Management* 33.1 (2012), pp. 20–27 (pages 13, 63, 65, 78).

[Eck06] Wayne W. Eckerson. *Performance Dashboards*. John Wiley & Sons, Inc., Hoboken, New Jersey., 2006 (pages 15, 28, 30, 51, 93).

[Emm90] Margaret A. Emmelhainz. *Electronic Data Interchange: A Total Management Guide*. New York, NY, USA: Van Nostrand Reinhold Co., 1990 (page 19).

[Eng+12a] Robert Engel, Wil M.P. van der Aalst, Marco Zapletal, Christian Pichler, and Hannes Werthner. "Mining Inter-organizational Business Process Models from EDI Messages: A Case Study from the Automotive Sector". In: *Proceedings of the 24th International Conference on Advanced Information Systems Engineering (CAiSE 2012)*. Springer, 2012, pp. 222–237 (pages 2, 12, 92).

[Eng+13a] Robert Engel, Worarat Krathu, Christian Pichler, Marco Zapletal, and Hannes Werthner. "Towards EDI-Based Business Activity Monitoring". In: *Workshop on Methodologies for Robustness Injection into Business Processes (MRI-BP 2013), the 17th IEEE International EDOC Conference (EDOC 2013)*. IEEE, 2013 (page 82).

[Eng+11] Robert Engel, Worarat Krathu, Marco Zapletal, Christian Pichler, Wil M.P. van der Aalst, and Hannes Werthner. "Process Mining for Electronic Data Interchange". In: *Proceedings of the 12th International Conference on Electronic Commerce and Web Technologies (EC-Web 2011)*. Vol. 85. Lecture Notes in Business Information Processing. Springer, 2011, pp. 77–88 (pages 2, 5, 12).

[Eng+12b] Robert Engel, Christian Pichler, Marco Zapletal, Worarat Krathu, and Hannes Werthner. “From Encoded EDIFACT Messages to Business Concepts Using Semantic Annotations”. In: *Proceedings of the 14th IEEE International Conference on Commerce and Enterprise Computing (CEC 2012)*. IEEE, 2012, pp. 17–25 (pages 23–27, 37, 38, 40, 48, 94, 111, 130).

[ER14] Robert Engel and Jagadeesh Chandra Bose Rantham Prabhakara. “A Case Study on Analyzing Inter-organizational Business Processes from EDI Messages using Physical Activity Mining”. In: *Proceedings of the 47th Hawaii International Conference on System Sciences (HICSS 2014)*. IEEE, 2014 (pages 92, 94, 112, 117, 126).

[Eng+13b] Robert Engel, Jagadeesh Chandra Bose Rantham Prabhakara, Christian Pichler, Marco Zapletal, and Hannes Werthner. “EDIminer: A Toolset for Process Mining from EDI Messages”. In: *CAiSE’13 Forum, the 25th International Conference on Advanced Information Systems Engineering (CAiSE 2013)*. CEUR-WS.org, vol. 998, 2013, pp. 146–153 (pages 94, 112).

[FKK10] Kamel A. Fantazy, Vinod Kumar, and Uma Kumar. “Supply Management Practices and Performance in the Canadian Hospitality Industry”. In: *International Journal of Hospitality Management* 29.4 (2010), pp. 685–693 (pages 58, 59, 62, 63, 66, 154).

[FHZ10] Barbara B. Flynn, Baofeng Huo, and Xiande Zhao. “The Impact of Supply Chain Integration on Performance: A Contingency and Configuration Approach”. In: *Journal of Operations Management* 28.1 (2010), pp. 58–71 (pages 59, 63, 86, 154).

[FB06] Doug Foxvog and Christoph Bussler. “Ontologizing EDI Semantics”. In: *the 1st International Workshop on Ontologizing Industrial Standards (OIS 2006), Advances in Conceptual Modeling - Theory and Practice*. Vol. 4231. Lecture Notes in Computer Science. Springer Berlin Heidelberg, 2006, pp. 301–311 (pages 23, 27).

[FB05] Douglas Foxvog and Christoph Bussler. “Ontologizing EDI: First Steps and Initial Experience”. In: *the International Workshop on Data Engineering Issues in E-Commerce (DEEC 2005)*. IEEE Computer Society, 2005, pp. 49–58 (pages 23, 27).

[GFC09] Irene Gil-Saura, Marta Frasquet-Deltoro, and Amparo Cervera-Taulet. “The Value of B2B Relationships”. In: *Industrial Management & Data Systems* 109.5 (2009), pp. 593–609 (pages 60, 65, 88).

[GWI08] Kenneth W. Green Jr, Dwayne Whitten, and R. Anthony Inman. “The Impact of Logistics Performance on Organizational Performance in a Supply Chain Context”. In: *Supply Chain Management: An International Journal* 13.4 (2008), pp. 317–327 (pages 59, 66, 86, 153, 154).

[GPT01] Angappa Gunasekaran, Christoph Patel, and Ercan Tirtiroglu. “Performance Measures and Metrics in a Supply Chain Environment”. In: *International Journal of Operations & Production Management* 21.1 (2001), pp. 71–87 (pages 86, 87, 151, 153–155).

[HCV09] Kim Sundtoft Hald, Carlos Cordón, and Thomas E. Vollmann. “Towards an Understanding of Attraction in Buyer–Supplier Relationships”. In: *Industrial Marketing Management* 38.8 (2009), pp. 960–970 (pages 61, 65).

[Han+09] Robert Handfield, Kenneth Petersen, Paul Cousins, and Benn Lawson. “An Organizational Entrepreneurship Model of Supply Management Integration and Performance Outcomes”. In: *International Journal of Operations & Production Management* 29.2 (2009), pp. 100–126 (page 154).

[HJ92] Jan B. Heide and George John. “Do Norms Matter in Marketing Relationships?” In: *The Journal of Marketing* 56.2 (1992), pp. 32–44 (page 62).

[HS05] Kevin B. Hendricks and Vinod R. Singhal. “Association Between Supply Chain Glitches and Operating Performance”. In: *Management Science* 51.5 (2005), pp. 695–711 (pages 86, 154).

[HRS10] Miguel Hernández-Espallardo, Augusto Rodríguez-Orejuela, and Manuel Sánchez-Pérez. “Inter-organizational Governance, Learning and Performance in Supply Chains”. In: *Supply Chain Management: An International Journal* 15.2 (2010), pp. 101–114 (page 59).

[HHS05] Aref A. Hervani, Marilyn M. Helms, and Joseph Sarkis. “Performance Measurement for Green Supply Chain Management”. In: *Benchmarking: An International Journal* 12.4 (2005), pp. 330–353 (pages 86, 154).

[Hev11] Alan R. Hevner. *Design Science Research in Information Systems, talk in the Trends in E-Commerce Lecture at the Vienna University of Technology*. 2011. URL: `http://www.ec.tuwien.ac.at/files/u326/Vienna_2011_DSR_Presentation_hevner.pdf` (pages 10, 11).

[Hev+04] Alan R. Hevner, Salvatore T. March, Jinsoo Park, and Sudha Ram. “Design Science in Information Systems Research”. In: *Management Information Systems Quarterly* 28.1 (2004), pp. 75–105 (page 7).

[HF89] Ned C. Hill and Daniel M. Ferguson. “Electronic Data Interchange: A Definition and Perspective”. In: *EDI Forum: The Journal of Electronic Data Interchange* 1.1 (1989), pp. 5–12 (pages 19, 20).

[Hor07] Peter Hornix. “Performance Analysis of Business Processes through Process Mining”. MA thesis. Department of Mathematics and Computer Science, Eindhoven University of Technology, Eindhoven, The Netherlands, 2007 (pages 15, 34, 92).

[HB11] Matthew Horridge and Sean Bechhofer. “The OWL API: A Java API for OWL Ontologies”. In: *Semantic Web* 2.1 (2011), pp. 11–21 (page 105).

[Hsu+08] Chin-Chun Hsu, Vijay R. Kannan, Keah-Choon Tan, and G. Keong Leong. “Information Sharing, Buyer-Supplier Relationships, and Firm Performance: A Multiregion Analysis”. In: *International Journal of Physical Distribution & Logistics Management* 38.4 (2008), pp. 296–310 (pages 59, 78, 85).

[Hu+10] Shanshan Hu, Qing Ye, Wei Chi, and Barbara B. Flynn. “Supply Chain Structure, Inventory Turnover, and Financial Performance: Evidence from Manufacturing Companies in China”. In: *Proceedings of the 21th Annual Conference of the Production and Operations Management Society (POMS 2010)*. 2010 (page 59).

[HSW04] Samuel H. Huan, Sunil K. Sheoran, and Ge Wang. “A Review and Analysis of Supply Chain Operations Reference (SCOR) Model”. In: *Supply Chain Management: An International Journal* 9.1 (2004), pp. 23–29 (pages 59, 60, 86–88, 153–155).

[IGF04] Karthik N.S. Iyer, Richard Germain, and Gary L. Frankwick. “Supply Chain B2B E-Commerce and Time-Based Delivery Performance”. In: *International Journal of Physical Distribution & Logistics Management* 34.8 (2004), pp. 645–661 (pages 86, 87, 153–155).

[Jan11] Gerrit K. Janssens. “Electronic Data Interchange: From its Birth to its New Role in Logistics Information Systems”. In: *International Journal on Information Technologies and Security* 3.3 (2011), pp. 45–56 (page 1).

[JBT07] Christian C. Johnson, Irv Belman, and John Thompson. *Balanced Scorecard for State-Owned Enterprises: Driving Performance and Corporate Governance*. Asian Development Bank, 2007 (page 30).

[Joh+04] David A. Johnston, David M. McCutcheon, F.Ian Stuart, and Hazel Kerwood. “Effects of Supplier Trust on Performance of Cooperative Supplier Relationships”. In: *Journal of Operations Management* 22.1 (2004), pp. 23–38 (pages 13, 59–61, 63, 78, 154).

[JZ03] Patrik Jonsson and Mosad Zineldin. “Achieving High Satisfaction in Supplier-Dealer Working Relationships”. In: *Supply Chain Management: An International Journal* 8.3 (2003), pp. 224–240 (pages 59, 60, 62–65, 86, 87, 153–155).

[KD10] Yildiray Kabak and Asuman Dogac. “A survey and analysis of electronic business document standards”. In: *ACM Computing Surveys* 42.3 (2010), 11:1–11:31 (page 1).

[KT02] Vijay R. Kannan and Keah Choon Tan. “Supplier SSelection and Assessment: Their Impact on Business Performance”. In: *Journal of Supply Chain Management* 38.3 (2002), pp. 11–21 (pages 86, 87, 154, 155).

[KN92] Robert S. Kaplan and David P. Norton. “The Balanced Scorecard - Measures That Drive Performance”. In: *Harvard Business Review* January-February (1992), pp. 71–79 (pages 15, 28, 93).

[KN96] Robert S. Kaplan and David P. Norton. “Using the Balanced Scorecard as a Strategic Management System”. In: *Harvard Business Review* 74.1 (1996), pp. 75–85 (page 30).

[KN04] Robert S. Kaplan and David P. Norton. *Focusing Your Organization on Strategy-with the Balanced Scorecard*. Harvard Business School Publishing, 2004 (pages 28, 30, 51, 93).

[KW05] Weiling Ke and Kwok Kee Wei. "Critical Factors Affecting the Firm to Share Knowledge with Trading Partners: A Comparative Exploratory Case Study". In: *Proceedings of the 7th International Conference on Electronic Commerce (ICEC 2005)*. ACM, 2005, pp. 177–183 (page 66).

[Kim+10] Kyung Kyu Kim, Seung-Hoon Park, Sung Yul Ryoo, and Sung Kook Park. "Inter-organizational Cooperation in Buyer–Supplier Relationships: Both Perspectives". In: *Journal of Business Research* 63.8 (2010), pp. 863–869 (pages 61, 64–66).

[KS03] Jack P.C. Kleijnen and Martin T. Smits. "Performance Metrics in Supply Chain Management". In: *Journal of the Operational Research Society* 54.5 (2003), pp. 507–514 (pages 15, 29, 86, 93, 153, 154).

[Koh+12] Marko Kohtamäki, Jukka Vesalainen, Stephan Henneberg, Peter Naudé, and Marc J. Ventresca. "Enabling Relationship Structures and Relationship Performance Improvement: The Moderating Role of Relational Capital". In: *Industrial Marketing Management* 41.8 (2012), pp. 1298–1309 (pages 59, 61, 66, 86, 154).

[KMD02] Masaaki Kotabe, Xavier Martin, and Hiroshi Domoto. "Gaining From Vertical Partnerships: Knowledge Transfer, Relationship Duration, and Supplier Performance Improvement in the US and Japanese Automotive Industries". In: *Strategic Management Journal* 24.4 (2002), pp. 293–316 (page 58).

[Kra13] Worarat Krathu. "Review of Success Factors in Inter-organizational Relationships: A Cause and Effect Model". unpublished; under submission. 2013 (pages 18, 75).

[Kra+13] Worarat Krathu, Robert Engel, Christian Pichler, Marco Zapletal, and Hannes Werthner. "Identifying Inter-organizational Key Performance Indicators from EDIFACT Messages". In: *Proceedings of the 15th IEEE Conference on Business Informatics (CBI 2013)*. IEEE, 2013, pp. 276–283 (pages 14, 18, 90).

[Kra+12a] Worarat Krathu, Christian Pichler, Robert Engel, Marco Zapletal, and Hannes Werthner. "Semantic Interpretation of UN/EDIFACT Messages for Evaluating Inter-organizational Relationships". In: *Proceedings of the 5th International Conference on Advances in Information Technology (IAIT 2012)*. Vol. 344. Communications in Computer and Information Science. Springer, 2012, pp. 81–93 (pages 13, 17, 49, 94).

[Kra+] Worarat Krathu, Christian Pichler, Robert Engel, Marco Zapletal, Hannes Werthner, and Christian Huemer. "A Framework for Inter-organizational Performance Analysis from EDI Messages". In: *Proceedings of the 16th IEEE Conference on Business Informatics (CBI 2014), To appear* (pages 16, 107, 128).

[Kra+12b] Worarat Krathu, Christian Pichler, Marco Zapletal, and Hannes Werthner. "Semantic Inter-organizational Performance Analysis using the Balanced Scorecard Methodology". In: *Proceedings of the 35th Jubilee International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO 2012)*. 2012 (pages 18, 35).

[KHT07] Daniel R. Krause, Robert B. Handfield, and Beverly B. Tyler. "The Relationships between Supplier Development, Commitment, Social Capital Accumulation and Performance Improvement". In: *Journal of Operations Management* 25.2 (2007), pp. 528–545 (pages 13, 59, 63, 78).

[LNC02] Kee-hung Lai, Eric W.T. Ngai, and T.C.E. Cheng. "Measures for Evaluating Supply Chain Performance in Transport Logistics". In: *Transportation Research Part E: Logistics and Transportation Review* 38.6 (2002), pp. 439–456 (pages 86, 153, 154).

[LKS07] Chang Won Lee, Ik-Whan G. Kwon, and Dennis Severance. "Relationship Between Supply Chain Performance and Degree of Linkage among Supplier, Internal Integration, and Customer". In: *Supply Chain Management: An International Journal* 12.6 (2007), pp. 444–452 (page 63).

[Leo+06] Nicola Leone, Gerald Pfeifer, Wolfgang Faber, Thomas Eiter, Georg Gottlob, Simona Perri, and Francesco Scarcello. "The DLV System for Knowledge Representation and Reasoning". In: *ACM Transactions on Computational Logic* 7.3 (2006), pp. 499–562 (page 70).

[LA13] Massimiliano de Leoni and Wil M.P. van der Aalst. "Data-aware Process Mining: Discovering Decisions in Processes Using Alignments". In: *Proceedings of the 28th Annual ACM Symposium on Applied Computing (SAC 2013)*. ACM, 2013, pp. 1454–1461 (pages 15, 34, 92).

[LL06] Suhong Li and Binshan Lin. "Accessing Information Sharing and Information Quality in Supply Chain Management". In: *Decision Support Systems* 42.3 (2006), pp. 1641–1656 (pages 13, 58, 61, 63–65, 67, 78, 88, 89).

[Li+06] Suhong Li, Bhanu Ragu-Nathan, T.S. Ragu-Nathan, and S. Subba Rao. "The Impact of Supply Chain Management Practices on Competitive Advantage and Organizational Performance". In: *Omega* 34.2 (2006), pp. 107–124 (pages 58, 59, 63, 66, 67, 86, 88, 154).

[Liu+12] Yi Liu, Ying Huang, Yadong Luo, and Yang Zhao. "How Does Justice Matter in Achieving Buyer-Supplier Relationship Performance?" In: *Journal of Operations Management* 30.5 (2012), pp. 355–367 (pages 59, 65, 66).

[LLL09] Yi Liu, Yadong Luo, and Ting Liu. "Governing Buyer–Supplier Relationships through Transactional and Relational Mechanisms: Evidence from China". In: *Journal of Operations Management* 27.4 (2009), pp. 294–309 (pages 59, 62, 66, 88).

[MB00] Michael Maloni and W.C. Benton Jr. "Power Influences in the Supply Chain". In: *Journal of Business Logistics* 21.1 (2000), pp. 49–74 (pages 63, 65).

[Mea13] Timothy Meaklim. "Considering Game Theory to Improve Leadership in Partnership Working within the UK Public Services". In: *International Journal of Leadership in Public Services* 9.1/2 (2013), pp. 22–31 (page 74).

[MG05] Mary J. Meixell and Vidyaranya B. Gargeya. “Global Supply Chain Design: A Literature Review and Critique”. In: *Transportation Research Part E: Logistics and Transportation Review* 41.6 (2005), pp. 531–550 (pages 86, 154).

[MFN96] Jakki J. Mohr, Robert J. Fisher, and John R. Nevin. “Collaborative Communication in Interfirm Relationships: Moderating Effects of Integration and Control”. In: *The Journal of Marketing* 60.3 (1996), pp. 103–115 (page 58).

[Möl10] Klaus Möller. “Partner Selection, Partner Behavior, and Business Network Performance: An Empirical Study on German Business Networks”. In: *Journal of Accounting & Organizational Change* 6.1 (2010), pp. 27–51 (pages 59, 65, 67).

[Moo11] Chul Woo Moon. “The Influence of Credibility on Satisfaction with Supplier Performance in the Inter-Organizational Relationship”. In: *International Business & Economics Research Journal* 10.9 (2011), pp. 47–58 (pages 59, 61, 64).

[MSC11] Hee-Cheol Moon, Sang-Ryul Shim, and Pin Cho. “Issues in the International Standards of Electronic Documents for Global e-Trade”. In: *Journal of Information and Management (Japan Society for Information and Management)* 31.4 (2 2011), pp. 168–176 (page 1).

[MH94] Robert M. Morgan and Shelby D. Hunt. “The Commitment-Trust Theory of Relationship Marketing”. In: *The Journal of Marketing* 58.3 (1994), pp. 20–38 (page 58).

[MSS05] Boris Motik, Ulrike Sattler, and Rudi Studer. “Query Answering for OWL-DL with Rules”. In: *Journal of Web Semantics: Science, Services and Agents on the World Wide Web* 3.1 (2005), pp. 41–60 (pages 42, 43, 48, 99).

[MSP11] Tore Mysen, Göran Svensson, and Janice M. Payan. “Causes and Outcomes of Satisfaction in Business Relationships”. In: *Marketing Intelligence & Planning* 29.2 (2011), pp. 123–140 (pages 65–67).

[Nav+06] Rene F. Navarro-Hernandez, Alonso Perez-Soltero, Gerardo Sanchez-Schmitz, Mario Barcelo-Valenzuela, and Cesar E. Rose. “An Ontological Model to Support the Implementation of Balance Scorecard in the Organizations”. In: *Proceedings of the 10th Word Multiconference on Systemics, Cybernetics and Informatics (WMSCI 2006)*. 2006, pp. 324–328 (page 93).

[NM07] John von Neumann and Oskar Morgenstern. *Theory of Games and Economic Behavior (60th Anniversary Commemorative Edition)*. Princeton University Press, 2007 (page 74).

[Nur08] Juha-Miikka Nurmilaakso. “EDI, XML and E-Business Frameworks: A Survey”. In: *Computers in Industry* 59.4 (2008), pp. 370–379 (pages 1, 20).

[NWL10] Gilbert N. Nyaga, Judith M. Whipple, and Daniel F. Lynch. “Examining Supply Chain Relationships: Do Buyer and Supplier Perspectives on Collaborative Relationships Differ?” In: *Journal of Operations Management* 28.2 (2010), pp. 101–114 (pages 86, 87, 154, 155).

[OK03] Andreas Otto and Herbert Kotzab. “Does Supply Chain Management Really Pay? Six Perspectives to Measure the Performance of Managing a Supply Chain”. In: *European Journal of Operational Research* 144.2 (2003), pp. 306–320 (pages 59, 86, 152, 153, 155).

[PP10] Dimitris Papakiriakopoulos and Katerina Pramatari. “Collaborative Performance Measurement in Supply Chain”. In: *Industrial Management & Data Systems* 110.9 (2010), pp. 1297–1318 (pages 86, 87, 155).

[PLC08] Antony Paulraj, Augustine A. Lado, and Injazz J. Chen. “Inter-organizational Communication as a Relational Competency: Antecedents and Performance Outcomes in Collaborative Buyer–Supplier Relationships”. In: *Journal of Operations Management* 26.1 (2008), pp. 45–64 (pages 58, 59, 62, 63, 86, 87, 154, 155).

[PKS96] Walter W. Powell, Kenneth W. Koput, and LaurelL. Smith-Doerr. “Interorganizational Collaboration and the Locus of Innovation: Networks of Learning in Biotechnology”. In: *Administrative Science Quarterly* 41.1 (1996), pp. 116–145 (page 58).

[PSR01] Damien J. Power, Amrik S. Sohal, and Shams-Ur Rahman. “Critical Success Factors in Agile Supply Chain Management - An Empirical Study”. In: *International Journal of Physical Distribution & Logistics Management* 31.4 (2001), pp. 247–265 (pages 86, 154).

[PC10] Daniel Prajogo and Mesbahuddin Chowdhury. “The Relationship between Supplier Management and Firm’s Performance: A Multidimensional Perspective”. In: *Proceedings of the 8th International Conference on Supply Chain Management and Information Systems (SCMIS 2010)*. IEEE. 2010, pp. 1–7 (pages 58–60, 63).

[PS08] Keith G. Provan and Jörg Sydow. “Evaluating Inter-organizational Relationships”. In: *The Oxford Handbook of Inter-Organizational Relations* (2008), pp. 691–718 (pages 2, 6, 13).

[RGS11] Usha Ramanathan, Angappa Gunasekaran, and Nachiappan Subramanian. “Supply Chain Collaboration Performance Metrics: A Conceptual Framework”. In: *Benchmarking: An International Journal* 18.6 (2011), pp. 856–872 (pages 87, 155).

[Rat00] Pauline Ratnasingam. “The Influence of Power on Trading Partner Trust in Electronic Commerce”. In: *Internet Research* 10.1 (2000), pp. 56–63 (page 63).

[RA08] Anne Rozinat and Wil M.P. van der Aalst. “Conformance Checking of Processes Based on Monitoring Real Behavior”. In: *Information Systems* 33.1 (2008), pp. 64–95 (page 32).

[RRG00] Ricky Ryssel, Thomas Ritter, and Hans Georg Gemünden. “Trust, Commitment and Value-creation in Inter-organizational Customer-supplier Relationships”. In: *Proceedings of the 16 th IMP Conference (IMP 2000)*. 2000 (pages 61, 65, 66).

[RSK09] Il Ryu, SoonHu So, and Chulmo Koo. “The Role of Partnership in Supply Chain Performance”. In: *Industrial Management & Data Systems* 109.4 (2009), pp. 496–514 (pages 59, 61–65, 86, 154).

[SMN09] Murali Sambasivan, Zainal Abidin Mohamed, and Tamizarasu Nandan. "Performance Measures and Metrics for E-Supply Chains". In: *Journal of Enterprise Information Management* 22.3 (2009), pp. 346–360 (pages 86, 152, 153, 155).

[Sau+04] Carol Saunders, Yu Andy Wu, Yuzhu Li, and Shawn Weisfeld. "Interorganizational Trust in B2B Relationships". In: *Proceedings of the 6th International Conference on Electronic Commerce (ICEC 2004)*. ACM, 2004, pp. 272–279 (pages 13, 61, 78).

[Sep08] Risto Seppänen. "Trust in Inter-organizational Relationships". PhD thesis. School of Business, Lappeenranta University of Technology, Lappeenranta, Finland, 2008 (page 59).

[SBS07] Risto Seppänen, Kirsimarja Blomqvist, and Sanna Sundqvist. "Measuring Inter-organizational Trust - A Critical Review of the Empirical Research in 1990–2003". In: *Industrial Marketing Management* 36.2 (2007), pp. 249–265 (pages 13, 61, 78).

[SG06] Craig Shepherd and Hannes Günter. "Measuring Supply Chain Performance: Current Research and Future Directions". In: *International Journal of Productivity and Performance Management* 55.3 (2006), pp. 242–258 (pages 86, 87, 89, 151–155).

[SCW00] Hojung Shin, David A. Collier, and Darryl D. Wilson. "Supply Management Orientation and Supplier/Buyer Performance". In: *Journal of Operations Management* 18.3 (2000), pp. 317–333 (pages 59, 66, 86, 153–155).

[SS05] Togar M. Simatupang and Ramaswami Sridharan. "The Collaboration Index: A Measure for Supply Chain Collaboration". In: *International Journal of Physical Distribution & Logistics Management* 35.1 (2005), pp. 44–62 (page 153).

[Sin+11] Doordarshi Singh, Jaspreet Singh Oberoi, Inderpreet Singh Ahuja, and Fatehgarh Sahib. "A Survey of Literature of Conceptual Frameworks Assessing Supply Chain Flexibility". In: *International Journal of Applied Engineering Research* 2.1 (2011), pp. 172–182 (page 60).

[SPP10] Mirko Sokovic, D Pavletic, and Karmen Kern Pipan. "Quality improvement methodologies–PDCA cycle, RADAR matrix, DMAIC and DFSS". In: *Journal of Achievements in Materials and Manufacturing Engineering* 43.1 (2010), pp. 476–483 (page 2).

[SA07] Minseok Song and Wil M.P. van der Aalst. "Supporting Process Mining by Showing Events at a Glance". In: *the 17th Annual Workshop on Information Technologies and Systems (WITS 2007)*. 2007, pp. 139–145 (page 126).

[SK10] Gunjan Soni and Rambabu Kodali. "Internal Benchmarking for Assessment of Supply Chain Performance". In: *Benchmarking: An International Journal* 17.1 (2010), pp. 44–76 (pages 59, 86, 87, 151–155).

[Son11] Peter Sonntagbauer. "E-Business, Emerging Trends in the European Union". In: *Proceedings of the 2nd International Conference on ICT Innovations 2010*. Vol. 83. Communications in Computer and Information Science, pp. 40-50. Springer, 2011 (pages 1, 20).

[SQ03] Christine Storerl and Mohammed Quaddus. “Preliminary Evaluation of Inter-organizational Information Systems and Relationships”. In: *Proceedings of the 7th Pacific Asia Conference on Information Systems (PACIS 2003)*. 2003 (pages 13, 78).

[The+01] Marinos Themistocleous, Zahir Irani, Robert M. O’Keefe, and Ray Paul. “ERP Problems and Application Integration Issues: An Empirical Survey”. In: *Proceedings of the 34th Annual Hawaii International Conference on System Sciences (HICSS 2001)*. Vol. 9. IEEE Computer Society, 2001, p. 9045 (page 1).

[Tim01] Paul Timmers. *Electronic Commerce: Strategies and Models for Business-to-Business Trading*. Management/Information Systems / Wiley. Wiley, 2001 (page 19).

[TN99] Alberto de Toni and Guido Nassimbeni. “Buyer-supplier Operational Practices, Sourcing Policies and Plant Performances: Results of an Empirical Research”. In: *International Journal of Production Research* 37.3 (1999), pp. 597–619 (page 58).

[Ver+11] H.M.W. (Eric) Verbeek, Joos C.A.M. Buijs, Boudewijn F. van Dongen, and Wil M.P. van der Aalst. “XES, XESame, and ProM 6”. In: *Information Systems Evolution (CAiSE Forum 2010, Selected Extended Papers), the 22nd International Conference on Advanced Information Systems Engineering (CAiSE 2010)*. Ed. by W. M. P. van der Aalst, J. Mylopoulos, N. M. Sadeh, M. J. Shaw, C. Szyperski, P. Soffer, and E. Proper. Vol. 72. Lecture Notes in Business Information Processing. Springer, 2011, pp. 60–75 (pages 5, 15).

[VM06] Ann Vereecke and Steve Muylle. “Performance Improvement Through Supply Chain Collaboration in Europe”. In: *International Journal of Operations & Production Management* 26.11 (2006), pp. 1176–1198 (page 1).

[Vij10] Leo R. Vijayasarathy. “Supply Integration: An Investigation of Its Multi-dimensionality and Relational Antecedents”. In: *International Journal of Production Economics* 124.2 (2010), pp. 489–505 (page 65).

[VGS07] Ken Vollmer, Mike Gilpin, and Jacqueline Stone. *B2B Integration Trends: Message Formats*. Forrester Research, 2007 (pages 1, 20, 35).

[WAA06] Anton J.M.M. Weijters, Wil M.P. van der Aalst, and Ana Karla Alves de Medeiros. *Process Mining with HeuristicsMiner-algorithm*. BETA Working Paper Series WP 166. Eindhoven University of Technology, Eindhoven, 2006 (pages 32, 94, 112, 120).

[Wer+08] Jan Martijn E.M. van der Werf, Boudewijn F. van Dongen, Cor A.J. Hurkens, and Alexander Serebrenik. “Process Discovery Using Integer Linear Programming”. In: *Proceedings of the 29th International Conference on Application and Theory of Petri Nets and Other Models of Concurrency (PETRI NETS 2008)*. Vol. 5062. Lecture Notes in Computer Science. Springer, 2008, pp. 368–387 (pages 15, 92).

[WML08] Branimir Wetzstein, Zhilei Ma, and Frank Leymann. "Towards Measuring Key Performance Indicators of Semantic Business Processes". In: *Proceedings of the 11th International Conference on Business Information Systems (BIS 2008)*. Ed. by Witold Abramowicz and Dieter Fensel. Vol. 7. Lecture Notes in Business Information Processing. Springer Berlin Heidelberg, 2008, pp. 227–238 (pages 95, 98).

[WLN10] Judith M. Whipple, Daniel F. Lynch, and Gilbert N. Nyaga. "A Buyer's Perspective on Collaborative Versus Transactional Relationships". In: *Industrial Marketing Management* 39.3 (2010), pp. 507–518 (pages 59–61, 65, 66, 86–88, 153–155).

[Wie+10] Frank Wiengarten, Paul Humphreys, Guangming Cao, Brian Fynes, and Alan McKittrick. "Collaborative Supply Chain Practices and Performance: Exploring The Key Role of Information Quality". In: *Supply Chain Management: An International Journal* 15.6 (2010), pp. 463–473 (pages 59, 63, 87, 153).

[WE04] Ka-shing Woo and Christine T. Ennew. "Business-to-Business Relationship Quality: An IMP Interaction-Based Conceptualization and Measurement". In: *European Journal of Marketing* 38.9/10 (2004), pp. 1252–1271 (pages 60, 63, 87).

[XEH12] Guohui Xiao, Thomas Eiter, and Stijn Heymans. "The DReW System for Nonmonotonic DL-Programs". In: *Proceedings of the Joint Conference of the 6th Chinese Semantic Web Symposium and the First Chinese Web Science Conference (SWWS 2012)*. Ed. by Juanzi Li, Guilin Qi, Dongyan Zhao, Wolfgang Nejdl, and Hai-Tao Zheng. Springer Proceedings in Complexity. Shenzhen City, China, Nov. 2012, pp. 383–390 (pages 47, 70, 105).

[Yeu08] Andy C.L. Yeung. "Strategic Supply Management, Quality Initiatives, and Organizational Performance". In: *Journal of Operations Management* 26.4 (2008), pp. 490–502 (pages 59, 86, 153–155).

[Yus+04] Yahaya Yusuf, Angappa Gunasekaran, Ezekiel Olu Adeleye, and K. Sivayoganathan. "Agile Supply Chain Capabilities: Determinants of Competitive Objectives". In: *European Journal of Operational Research* 159.2 (2004), pp. 379–392 (pages 86, 154).

[ZH06] Akbar Zaheer and Jared Harris. "Interorganizational Trust". In: *Handbook of Strategic Alliances*. Ed. by Oded Shenkar and Jeffrey J. Reuer. Sage Publications, 2006. Chap. 10, pp. 169–197 (pages 13, 61, 78).

[Zha02] Fang Zhao. "Performance Measures for Inter-organisational Partnerships". In: *Proceedings of the 7th International Conference on ISO 9000 and TQM (7-ICIT 2002)*. 2002 (pages 59, 62, 65, 87, 153, 154).

[ZB07] Honggeng Zhou and W.C. Benton Jr. "Supply Chain Practice and Information Sharing". In: *Journal of Operations Management* 25.6 (2007), pp. 1348–1365 (pages 59, 66, 78, 86, 87, 89, 153, 154).

[ZH09] Gerrit Willem Ziggers and Jörg Henseler. "Inter-firm Network Capability: How It Affects Buyer-supplier Performance". In: *British Food Journal* 111.8 (2009), pp. 794–810 (pages 58, 59, 86, 87, 152–155).

APPENDIX **A**

Inter-organizational KPIs

In the following, the complete sets of inter-organizational KPIs identified by applying the KPI identification method discussed in Chapter 5 are provided. These KPIs are presented as (i) a set of primary KPIs which are calculated by directly applying aggregation functions on business information found in EDIFACT messages and (ii) a set of secondary KPIs which are calculated based on primary KPIs or require complex calculation by using algebraic expressions. Furthermore, they are grouped into corresponding success factors found in the review presented in Chapter 4. A set of primary KPIs and secondary KPIs are presented in Table A.1 and Table A.2 accordingly.

Table A.1: Primary inter-organizational KPIs identified from EDIFACT messages

KPI	Mapping to EDIFACT data / Literature Support
Satisfaction	
Ordered quantity[s,a,p]	Ordered quantity (Quantity in QTY qualified by value 21) from ORDERS, INVOICE, ORDCHG, RECADV, or RETANN messages Support: [SK10; SG06]
Number of orders placed[c]	ORDERS messages
Returned quantity[s,a,p]	Returned quantity (Quantity in QTY qualified by value 61) from INVOIC, RETANN, or INVRPT messages Support: [SG06; BS07]
Received and accepted quantity[s,a,p]	Received and accepted quantity (Quantity in QTY qualified by value 194) from INVOIC, RECADV, RETANN, or RETINS messages
Received but not accepted quantity[s,a,p]	Received but not accepted quantity (Quantity in QTY qualified by value 195, or 196) from INVOIC, RECADV, RETANN, or RETINS messages Support: [SG06; BB10; GPT01]
Returned by consumer quantity[s,a,p]	Returned by consumer quantity (Quantity in QTY qualified by value 210) from INVRPT, or RETANN messages
Rejected return quantity[s,a,p]	Rejected return quantity (Quantity in QTY qualified by value 269) from RETINS messages Support: [SG06; BB10; BS07; GPT01]
Financial performance	
Payable amount[s,a]	Payable amount (Monetary amount in MOA qualified by value 9) from BANSTA, CREMUL, DEBMUL, FINSTA, or PAYMUL messages
Final posted amount[s,a]	Final posted amount (Monetary amount in MOA qualified by value 60) from BANSTA, or FINSTA messages

Original amount[s,a]	Original amount (Monetary amount in MOA qualified by value 98) from COACSU, BANSTA, CREMUL, DEBMUL, FINSTA, REMADV, or INVOIC messages
Invoiced amount[s,a]	Invoiced amount (Monetary amount in MOA qualified by value 39 or 77) from BANSTA, FINSTA, REMADV, IFCSUM, or COACSU messages
Paid amount[s,a]	Paid amount (Monetary amount in MOA qualified by value 11) from FINSTA, REMADV, or COACSU messages
Total payment amount[s,a]	Total payment amount (Monetary amount in MOA qualified by value 139) from FINSTA, REMADV, or COACSU messages
Remitted amount[s,a]	Remitted amount (Monetary amount in MOA qualified by value 12) from CREMUL, DEBMUL, FINSTA, PAYMUL, REMADV, or COACSU messages
Total credit amount[s,a,p]	Total credit amount (Monetary amount in MOA qualified by value 346) from FINSTA messages
Total debit amount[s,a]	Total debit amount (Monetary amount in MOA qualified by value 347) from FINSTA messages
Credit note amount[s,a]	Credit note amount (Monetary amount in MOA qualified by value 210) from CREMUL, DEBMUL, or FINSTA messages
Debit note amount[s,a]	Debit note amount (Monetary amount in MOA qualified by value 211) from FINSTA messages
Received amount[s,a]	Received amount (Monetary amount in MOA qualified by value 119) from BANSTA, or FINSTA messages
Other transport charges[s,a]	Other transport charges (Monetary amount in MOA qualified by value 104) from IFCSUM messages Support: [CQ03; OK03; ZH09; SG06; Bag+05]
Charge summary total[s,a]	Charge summary total amount (Monetary amount in MOA qualified by value 24) from IFCSUM messages
Freight charge[s,a]	Freight charge amount (Monetary amount in MOA qualified by value 64) from IFTMAN messages Support: [CQ03; OK03; ZH09; SG06; Bag+05]
Other cost[s,a]	Other cost (Monetary amount in MOA qualified by value 160) from IFTMAN messages
Loading and handling cost[s,a]	Loading and handling cost (Monetary amount in MOA qualified by value 81) from IFCSUM messages Support: [CQ03; OK03; ZH09; SG06; Bag+05]
Tax/duty amount[s,a]	Tax/duty amount (Monetary amount in MOA qualified by value 55, 124, 150, or 161) from INVOICE, TAXCON, CUSDEC, FINSTA, or IFCSUM messages
Charge amount[s,a]	Charge amount (Monetary amount in MOA qualified by value 8, 23, or 259) from CREMUL, DEBMUL, FINSTA, IFCSUM, or COACSU messages
Adjusted amount[s,a]	Adjusted amount (Monetary amount in MOA qualified by value 5, 165) from REMADV, PAYMUL, CREMUL, DEBMUL, FINSTA, or COACSU messages
Invoice quantity[s,a,p]	Invoiced quantity (Quantity in QTY qualified by value 47) from INVOIC, MSCONS, or RECADV messages
Ordered quantity[s,a,p]	Ordered quantity (Quantity in QTY qualified by value 21) from ORDERS, INVOICE, ORDCHG, RECADV, or RETANN messages
Number of payment cancellation[c]	FINCAN messages
Failure	
Number of application errors[s,a,p,c]	ERC in APERAK messages Support: [SMN09]
Number of normal acknowledgment[s,a,p,c]	APERAK messages (without ERC)
Reliability	
Short shipped quantity[s,a,p]	Short shipped quantity (Quantity in QTY qualified by value 119) from DESADV, INVOIC, RECADV, or RETANN messages Support: [CQ03; SK10; SG06; Ara+07]
Damaged goods quantity[s,a,p]	Damaged goods quantity (Quantity in QTY qualified by value 124) from INVOIC, RECADV, or RETANN messages Support: [CQ03; SK10; SG06; Ara+07]
Over shipped quantity[s,a,p]	Over shipped quantity (Quantity in QTY qualified by value 121) from INVOIC, RECADV, or RETANN messages Support: [CQ03; SK10; SG06; Ara+07]

Lost goods quantity[s,a,p]	Lost goods (Quantity in QTY qualified by value 126) from the INVOIC message Support: [CQ03; SK10; SG06; Ara+07]
Destroyed quantity[s,a,p]	Destroyed quantity (Quantity in QTY qualified by value 65) from RECADV, or RETANN messages Support: [CQ03; SK10; SG06; Ara+07]
Outstanding quantity[s,a,p]	Outstanding quantity (Quantity in QTY qualified by value 73) from RECADV, or RETANN messages
Adaptability/flexibility	
Quantity not available for despatch (fill rate)[s,a,p]	Quantity not available for despatch (Quantity in QTY qualified by value 255) from DESADV messages Support: [WLN10; HSW04; CQ03; OK03; SK10; LNC02; SG06; Cha03; Ara+07; KS03; Bag+05; ZB07; GWI08]
Quantity on hands (fill rate)[s,a,p]	Quantity on hands (Quantity in QTY qualified by value 17) from QUOTES messages Support: [WLN10; HSW04; CQ03; OK03; SK10; LNC02; SG06; Cha03; Ara+07; KS03; Bag+05; ZB07; SS05; GWI08]
Back order quantity[s,a,p]	Back order quantity (Quantity in QTY qualified by value 83) from INVOIC, RECADV, or RETANN messages Support: [SG06; Ara+07]
Customer Responsiveness	
Lead time[a]	Lead time (Date/time in DTM qualified by value 169) from DESADV messages Support: [WLN10; HSW04; JZ03; OK03; ZH09; SK10; SG06; SG06; Ara+07; BB10; BS07; SMN09; GPT01; CLS08; Yeu08; IGF04; Bag+05; SCW00]
Information Sharing	
Inventory report exchange[c]	INVRPT messages Support: [BS00; Wie+10; SK10; Zha02; ZB07]
Sale forecast exchange[c]	SLSFCT messages Support: [Wie+10; Zha02; ZB07]
Credibility	
Allowance amount[s,a]	Allowance amount (Monetary amount in MOA qualified by value 8, or 204) from COACSU, CREMUL, DEBMUL, FINSTA, ORDCHG, or QUOTES messages
Outstanding amount[s,a]	Outstanding amount (Monetary amount in MOA qualified by value 263) from COACSU, or FINSTA messages
Disputed amount[s,a]	Disputed amount (Monetary amount in MOA qualified by value 257) from COMDIS message
Discount amount[s,a]	Discounted amount (Monetary amount in MOA qualified by value 52) from FINSTA, REMADV, INVOIC, PAYMUL, or QUOTES messages
Allowance quantity[s,a,p]	Allowance quantity (Monetary amount in MOA qualified by value 130) from DESADV, INVOIC, QUOTES, RECADV, or RETANN messages
Free goods quantity[s,a,p]	Free goods quantity (Monetary amount in MOA qualified by value 192) from DESADV, or QUOTES messages
Loyalty	
Committed quantity[s,a,p]	Committed quantity (Quantity in QTY qualified by value 66) from DELFOR, SLSFCT, or CNTCND messages
Contract	
Contract exchange rate[c]	CNTCND messages
Customer Uncertainty	
Order change rate[c]	ORDCHG messages

Note 1: The superscript [s,a,p,c] on KPI names indicate applicable aggregation functions: sum, average, percentage, count.
Note 2: In this table, message types are represented as code only (e.g., *ORDERS* corresponds to Purchase order messages). The full description of segments and message types in EDIFACT release D10A is provided in http://www.unece.org/trade/untdid/d10a/trsd/trsdi1.htm and http://www.unece.org/trade/untdid/d10a/timd/timdi1.htm respectively.

Table A.2: Secondary inter-organizational KPIs identified from EDIFACT messages

KPI	Mapping to EDIFACT data / Literature Support
Satisfaction	
Customer retention rate	i) Message interchange sender or party identification (Party identification in NAD qualified by value BY) from ORDERS messages ii) Calculation: (Total customers − Number of new customers)/Total customers of previous period Support: [HHS05; CGH09]
Order quantity per customer	i) Message interchange sender or party identification (Party identification in NAD qualified by value BY) from ORDERS messages ii) Ordered quantity (Quantity in QTY qualified by value 21) from ORDERS iii) Calculation: (Ordered quantity / Total sender or party) Support: [BS00]
Revenue per customer	i)Message interchange sender or party identification (Party identification in NAD qualified by value BY) from INVOIC messages ii) Invoiced amount (Monetary amount in MOA qualified by value 39 or 77) from INVOIC messages iii) Calculation: (Invoiced amount / Total sender or party) Support: [BS00]
Financial performance	
Revenue[s,a]	i) Payable, invoice line item, original, total payment and tax/duty amount (Monetary amount in MOA qualified by value 9, 77, 98, 139 and 161 respectively) from BANSTA, CREMUL, DEBMUL, COACSU or REMADV messages ii) The related business parties are beneficiary, beneficiary's bank, party to receive refund, payee, receiving financial institution, party to receive commercial invoice remittance, seller's financial institution, receiver of cheque, selller, seller's agent/representation or supplier from NAD or FII qualified by value BE, BF, DCX, PE, RB, RE, RH, RV, SE, SR, or SU respectively) Support: [Yus+04; HS05; BS00; CF05; MG05; SG06; HHS05; CGH09; KS03; Yeu08; GWI08; FHZ10; Li+06; RSK09; CP04]
Cost[s,a]	i) Payable, invoice line item, original, total payment and tax/duty amount (Monetary amount in MOA qualified by value 9, 77, 98, 139 and 161 respectively) from BANSTA, CREMUL, DEBMUL, COACSU or REMADV messages ii) The related business parties are buyer, buyer's agent/representative, Debtor, Invoicee, Payor, Payer, Paying financial institution or, Cheque order from NAD or FII qualified by value BY, AB, HX, IV, PL, PR, PB, or PE respectively) Support: [WLN10; JZ03; CQ03; Joh+04; MG05; PLC08; LNC02; SG06; Cha03; Ara+07; GPT01; CWR08; SCW00; CGP11; RSK09; CP04]
Delivery cost[s,a]	i) Calculation: SUM(Primary KPI of (the transport charges, charge summary total, freight charge, other cost, loading and handling cost, charge, invoiced amount) Support: [WLN10; CQ03; ZH09; LNC02; SG06; BB10; Bag+05]
Profit[s,a]	i) Calculation: Secondary KPI of Revenue − Secondary KPI of Cost and/or Delivery cost Support: [WLN10; Yus+04; Han+09; Ash+09; HS05; BS00; Joh+04; NWL10; MG05; ZH09; SK10; SG06; Ara+07; CGH09; BS07; GPT01; Zha02; CPL04; Yeu08; FKK10; GWI08; FHZ10; Li+06; CP04]
Reliability	
On-time delivery[c,a,p]	i) Expected delivery date/time (Date/time in DTM qualified by value 10, 2, or 191) from ORDERS, DELFOR, DESADV, or DELJIT messages ii) Actual delivery date/time (Date/time in DTM qualified by value 11, 50, or 310) from DESADV, or RECADV messages iii) Calculation: COUNT(Deliveries arriving before or on the expected delivery date/time) Support: [WLN10; HSW04; JZ03; CQ03; AML10; CMM10; PSR01; Koh+12; NWL10; ZH09; SK10; PLC08; LNC02; SG06; Cha03; CGH09; BB10; BS07; CLS08; Yeu08; KT02; IGF04; Bag+05; SCW00; ZB07]
Delay delivery[c,a,p]	i) Expected delivery date/time (Date/time in DTM qualified by value 10, 2, or 191) from ORDERS, DELFOR, DESADV, or DELJIT messages ii) Actual delivery date/time (Date/time in DTM qualified by value 11, 50, or 310) from DESADV, or RECADV messages iii) Calculation: COUNT(Deliveries arriving after the expected delivery date/time) Support: [HSW04; JZ03; CQ03; AML10; CMM10; LNC02; SG06; Cha03; BB10; BS07; KS03; Yeu08]
Adaptability/flexibility	

Quantity ready for order[a,p]	i) Quantity on hand (Quantity in QTY qualified by the value 17) from QUOTES messages ii) Ordered quantity (Quantity in QTY qualified by the value 21) from ORDERS or QUOTES messages iii) Calculation: Quantity on hand − Ordered quantity Support: [PP10; OK03]
Customer Responsiveness	
Time of order response[a]	i) Order date/time (Date/time in DTM qualified by 137) from ORDERS messages ii) Order response date/time (Date/time in DTM qualified by 137) from ORDRSP messages iii) Calculation: Order response date/time − Order date/time Support: [CQ03; BS00; AML10; ZH09; PLC08; SG06; KT02; CPL04; CP04]
Time of order query response[a]	i) Order query date/time (Date/time in DTM qualified by 137) from OSTENQ messages ii) Order report date/time (Date/time in DTM qualified by 137) from OSTRPT messages iii) Calculation: Order report date/time − Order query date/time Support: [CQ03; AML10; SG06; BB10; BS07; GPT01; KT02]
Time of request for quote response[a]	i) Request for quote date/time (Date/time in DTM qualified by 137) from REQOTE messages ii) Quote date/time (Date/time in DTM qualified by 137) from QUOTES messages iii) Calculation: Request for quote date/time − Quote date/time Support: [CQ03; AML10; SG06; KT02]
Lead time[a]	i) Order date/time (Date/time in DTM qualified by 137) from ORDERS messages ii) Good receipt date/time (Date/time in DTM qualified by 310 or 50) from DESADV or RECADV messages iii) Calculation: Good receipt date/time − Order date/time Support: [WLN10; HSW04; JZ03; OK03; ZH09; SK10; SG06; SG06; Ara+07; BB10; BS07; SMN09; GPT01; CLS08; Yeu08; IGF04; Bag+05; SCW00]
Information Quality	
Delivery preparation time[a]	i) Order or delivery plan issue date/time (Date/time in DTM qualified by 137) from ORDERS, DELFOR or DELJIT messages ii) Requested delivery date/time (Date/time in DTM qualified by 10, 2 or 191) from DELFOR, DESADV, DELJIT, ORDERS, or ORDCHG messages iii) Calculation: Order or delivery plan issue date/time − Requested delivery date/time
Forecast accuracy[a,p]	i) Forecast/reserved quantity (Quantity in QTY qualified by value 247, or 248) from SLSFCT messages ii) Ordered quantity (Quantity in QTY qualified by value 21) from ORDERS messages iii) Calculation: Forecast/reserved quantity − Ordered quantity Support: [WLN10; PP10; CMM10; NWL10; SK10; SG06; BS07; GPT01; CWR08; IGF04; RGS11]
Credibility	
Contract violation (order quantity)[c]	Ordered quantity (Quantity in QTY qualified by value 21) from ORDERS messages ii) Minimum, maximum, or committed quantity (Quantity in QTY qualified by value 53, 54, or 66 respectively) from CNTCND messages iii) Calculation: COUNT if (Ordered quantity $<$ Minimum quantity or Committed quantity) or COUNT if (Ordered quantity $>$ Maximum quantity)
Customer Uncertainty	
Changed order quantity[s,a,p]	i) Ordered quantity (Quantity in QTY qualified by value 21) from ORDCHG or ORDERS messages ii) Calculation: Previous ordered quantity − Ordered quantity

Note 1: The superscript [s,a,p,c] on KPI names indicate applicable aggregation functions: sum, average, percentage, count.

Note 2: In this table, message types are represented as code only (e.g., *ORDERS* corresponds to Purchase order messages). The full description of segments and message types in EDIFACT release D10A is provided in http://www.unece.org/trade/untdid/d10a/trsd/trsdi1.htm and http://www.unece.org/trade/untdid/d10a/timd/timdi1.htm respectively.

APPENDIX B

List of Abbreviations

Abbreviations of Terms

ANSI X12	American National Standards Institute X12
B2B	Business-to-Business
BI	Business Information
BSC	Balanced Scorecard
CDE	Composite Data Element
DE	Data Element
EDI	Electronic Data Interchange
GLNs	Global Location Numbers
IORs	Inter-organizational Relationships
KB	Knowledge Base
KPI	Key Performance Indicator
MFM	Message Flow Mining
MIGs	Message Implementation Guidelines
OWL	Web Ontology Language
PAM	Physical Activity Mining
PDF	Portable Document Format
SCM	Supply Chain Management
UN/EDIFACT	The United Nations Electronic Data Interchange for Administration, Commerce and Transport
WSML	Web Service Modeling Language
XES	eXtensible Event Stream
XML	Extensible Markup Language

Abbreviations of EDIFACT Message Types

APERAK	Application error and acknowledgement message
BANSTA	Banking status message
CNTCND	Contractual conditions message
COACSU	Commercial account summary message
COMDIS	Commercial dispute message
CREMUL	Multiple credit advice message
CUSDEC	Customs declaration message
DEBMUL	Multiple debit advice message
DELFOR	Delivery schedule message
DELJIT	Delivery just in time message
DESADV	Despatch advice message
DIRDEB	Direct debit message
FINCAN	Financial cancellation message
FINSTA	Financial statement of an account message
IFCSUM	Forwarding and consolidation summary message
IFTMAN	Arrival notice message
INVOIC	Invoice message
INVRPT	Inventory report message
MSCONS	Metered services consumption report message
ORDCHG	Purchase order change request message
ORDERS	Purchase order message
ORDRSP	Purchase order response message
OSTENQ	Order status enquiry message
OSTRPT	Order status report message
PAYMUL	Multiple payment order message
QUOTES	Quote message
RECADV	Receiving advice message
REMADV	Remittance advice message
REQOTE	Request for quote message
RETANN	Announcement for returns message
RETINS	Instruction for returns message
SLSFCT	Sales forecast message
TAXCON	Tax control message

Abbreviations of EDIFACT Segments

BGM	Beginning of message
DTM	Date/time/period

ERC	Application error information
FII	Financial institution information
LIN	Line item
MOA	Monetary amount
NAD	Name and address
QTY	Quantity
RFF	Reference
UNB	Interchange header