iae aix-en-provence
GRADUATE SCHOOL OF MANAGEMENT

TURUN KAUPPAKORKEAKOULU
Turku School of Economics

TILBURG UNIVERSITY

A COMMON INFORMATION MODEL FOR ENTERPRISE INTEGRATION

Master´s Thesis
in Management of
Information Technology

Author:
Marcos Ulises Tong Alvarez

Supervisors:
Dr. Eija Koskivaara
Dr. Paul Laifa
Dr. Anne Rutkowski

Thesis mentor:
Ph. D. Vilho Räisänen

1.6.2009
Espoo, Finland

Acknowledge

First of all, thanks are expressed to persons that made possible this thesis with their contributions:

- Thanks to the university thesis supervisor that supported the overall theoretical research process of the thesis and provided detailed feedback on the thesis report.
- Thanks to Nokia Siemens Networks for giving the opportunity to perform this collaborative research.
- Thanks to the thesis mentor that patiently guided the empirical research process and made significant contributions to improve the relevancy and quality of this thesis.
- Thanks to the Erasmus Mundus consortium of the International Master in Management of Information Technology (IMMIT) program, for providing the opportunity to study the program.
- Thanks to the IMMIT colleagues that during two years of students contributed to an excellent educational and multi-cultural experience.
- Thanks to the close friends and family for their support and encouragement.

Table of contents

List of figures

List of tables

List of abbreviations

ABE	Abstract Business Entity
BPEL	Business Process Execution Language
CIM	Common Data Model
CoPS	Complex Products and Services
DIS	Data Inventory Solution
DXSI	DataXtend Semantic Integrator
EA	Enterprise Architecture
EAI	Enterprise Application Integration
EI	Enterprise Integration
ERP	Enterprise Resource Planning
ESB	Enterprise Service Bus
eTOM	enhanced Telecommunications Operations Map
IM	Information Modeling
IS	Information Systems
IT	Information Technology
NGOSS	Next Generation Operation Support Systems
NSN	Nokia Siemens Networks
NPS	Network Provisioning Solution
OSS	Operations Support Systems
RSM	Rational Software Modeler
SOA	Service Oriented Architecture
SID	Shared Information / Data
Telecom	Telecommunications
TMF	Tele Management Forum
WSDL	Web Service Definition Language
XML	eXtended Markup Language
XSD	XML Schema Definition

1 INTRODUCTION

1.1 Research context

The present study combines two major areas of research, *information modeling* (IM) and *enterprise integration* (EI). IM is the process of creating and managing information models, for purposes such as representing, understanding and sharing the information created and used by an organization; this in order to guide the development of solutions, the improvement of operations and business, and the deployment of organizational best practices (Lankhorst, 2004; Lim, 1997). EI is mainly concerned with aligning the organization in terms of the strategy, business processes, information systems (IS) and technological resources across the layers and functions of an organization, in order to create competitive advantage (Gulledge, 2008).

The motivation of this research is based on the requirements of new business models for the provision of integrated solutions, including different EI capabilities. This is studied in relation to the use of IM for integrating the vertical and horizontal dimensions of an organization. Theory will be used to identify the relationship between IM and EI capabilities from the perspective of the dynamic business environments and the supporting IS. This is followed by an analysis of the IM integration's goals and challenges and the analysis on an approach for enterprise IM integration, the use of a single common information model (CIM) across the enterprise. Next, the CIM approach is carried out in practice in an integration solution use case at a telecommunications (telecom) company, Nokia Siemens Networks (NSN). Finally the results are assessed and presented in the form of a theoretical conceptual framework. The motivation and context of the research is expanded next, using the two main concepts: IM and EI.

The main motivation of this research is to contribute to the development of EI capabilities of companies, by studying the use of IM from an integral perspective across the enterprise. A positive relationship is expected to be found between the use of IM and the development of EI capabilities. However, there are challenges hindering the benefits of IM such as complex structures of large organizations, heterogeneity of information models and systems across the organization and lack of a standard and governed modeling practices. Efficient IM is a critical factor to overcome these challenges, to manage a consolidated business and technology strategy, and to create organizational structures that facilitate EI capabilities. Particular to this study is the evaluation of benefits of using a CIM across the enterprise levels and functions. CIM are used to horizontally integrate the information models used within a company and between industry participants, and to vertically integrate the end-to-end information needs of the value-chain operations (Tele Management Forum, 2009). The use of CIM is expected to improve the definition and reorganization of the enterprise internal and external assets with the purpose of reducing overlaps and redundancies (Petrie, 1992). This improvement is expected to result from CIM being able to help a company to improve the understanding of how the different business processes are enabled by applications and IS, to unify the enterprise information and to define the interfaces required to build reusable and reconfigurable services (Linthicum, 2003). In addition, a CIM is expected to improve the information transformation efforts required in application integration, thus making more efficient the integration of solutions in terms of time and cost. This situation becomes highly relevant when studied in the context of dynamic business environments as in the telecom industry.

EI is a critical factor for new business models which, unlike traditional business that are based purely on product manufacturing or services provision, are based on the combination of both products and services to provide value-added solutions to the customers. This phenomenon is more stressed for the high-cost, complex products and services (CoPS) such as airplanes or mobile phone networks. Thus, providers of CoPS require more EI capabilities than providers of consumer goods, like phone devices. While consumer goods integrated solutions are more or less standardized into package bundles, CoPS integrated solutions are tailored to the needs of each customer, which vary largely from simple to sophisticated buyers. Moreover, since CoPS have higher costs, providers strive for higher and recurring revenues from the provision of services during the entire life cycle of the products, thus a greater range of services per product and number of services per customer are normally present in CoPS (Davies, 2001). In this research, a study is conducted for the case of integration of solutions from the perspective of the downstream segment of the telecom industry, which is marked with a circle in the figure 1. Following is an explanation of the telecoms value-chain; other authors (Olsson, 2003) might have slightly different explanations but the main elements are normally similar.

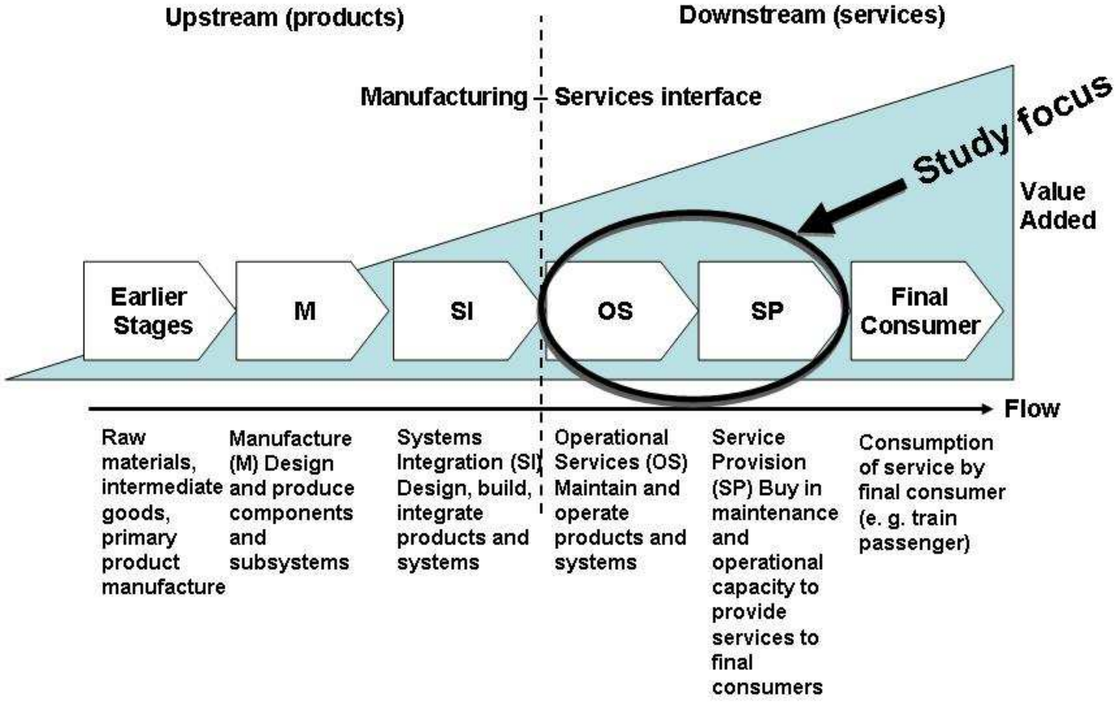

Figure 1. The value chain of CoPS (Davies, 2001).

A large range of participants are present in the telecoms value-chain. These include network equipment suppliers, network operators, media content providers, internet service providers, intermediaries, complementary service providers, storage service provides and service aggregators, to mention just a few. And as with other CoPS, business strategists have identified two segments of the telecom value-chain, upstream and downstream businesses (figure 1). The upstream segment includes firms based on the manufacturing of products and integrated services, for example parts manufacturers, network equipment providers and network system integrators. The downstream segment includes firms based on the provision of operational services and composed solutions, which are ultimately delivered to the final customer. Examples are network operators, internet service providers and complimentary service providers. Each of these two segments adds value in a different way:

> *"Upstream [markets] add value to the physical product through technology development and manufacture, understanding their customer's requirements, managing projects, and performing systems integration. Downstream [markets] add value by performing [..] service-based activities such as managing and maintaining system operations, customer care, advertising, billing, branding, marketing, and other service activities"* (Davies, 2001)

This distinction is relevant for this research to identify how different companies contribute in the integration of high-value solutions. As the value is added, going towards the final customer (see figure 1), the complexity of integrated solution increases and companies require more integration capabilities to reorganize their businesses into modular and flexible structures to provide integrated solutions (Braganza, 2002; Teece, 1997). In turn, companies' requirements for efficient integration of their supporting IS will tend to increase. In this study, a set of relevant EI capabilities are identified and linked to IS interoperability principles, followed by the study of using IM as a contributor to the development of such capabilities. The rest of this introductory chapter is intended to motivate and guide the reader to continue with the rest of this report. In the remaining of this chapter an overview of the current situation of the telecom industry is presented. This is followed by a definition of the more important concepts to set a common terminology base for the readers of this report. Next, the problem being research is described along with the purpose of this research. Then, the limitations that define the boundaries of this study are exposed. Finally, the chapter concludes with an overall description of the chapters included in this report.

1.2 Overview of the telecom industry business environment

Although initially limited by regulations, monopolies and the lack of infrastructure, the telecom industry has grown to become one of the largest contributors of the global economy. According to Loomis (1999) the average telecom share of gross domestic production (GDP), for the countries members of the organization for the economic co-operation and development (OECD), was expected to increase from 2.3% in 1997 to between 3.3% and 3.7% in 2006. The telecom industry has been enlarged with the entrance of new participants along with the expansion of the network infrastructure and market size, which now covers a large number of users, distributed in both urban and rural areas around the world. The telecom industry is also one of the most dynamics; there is a constant development of new technologies, services and business models. In consequence, telecom companies require the agility to react quickly and stay competitive. The following citation gives a good example of the dynamism of the telecom industry:

> "*'Anybody who tells you what the telecommunications industry will be like in five or 10 years is pulling your leg. Nobody knows. This industry has evolved so fast It's not even an evolution, It's a revolution' says David Bogaty, president & CEO of WorldNet Communications*" (Marino, 2009)

In addition, convergence of network technologies and services is a trend in the telecom industry. As opposed to having different services offered to customers via different channels, the tendency is that the different channels are being merged in a single network and the different technologies are transparent for the users. It is expected that network operators, including telephone, cable and ratio network companies will increasingly provide unified access and transmissions of data, voice and video. Similarly, the provision of network services also tends to converge. Phone, internet, media and entertainment services will increasingly converge as the use of internet through mobile devices increases. Two examples of this convergence is that service providers can query network traffic statistics from a market segment and specific network operators in order to create new marketing campaigns, and that service providers make available on demand video publicity for items related to internet search queries made by their customers. To provide these integrated solutions, telecoms operations and business strategy require special requirements in terms of IS integration.

Telecoms integrate a large number of IS to support their operations and the delivery of quality services. For this reason, the management of IS has been a major concern for telecoms, they require both quality and flexibility to be able to manage the market conditions and stay competitive (Loomis, 1999; Olsson, 2003). In addition, depending on the business strategy, telecoms might not only require the alignment of business units and the consolidation of internal systems, but also the ability to integrate external IS from suppliers, partners and competitors in order to provide new services or expand the existing ones (Davies, 2001). In addition, inherent to the combination of services, redundancies and inefficiencies might have an impact on the performance of integrated services. However, customers demand the same high levels of service quality than for individual services. This requires that integrations have to be done effectively to meet the performance and availability requirements of customers (Goodhue, Wybo and Kirsch, 1992). In consequence, investments in IS integration have become a significant part of the telecom expenses. It is agreed that there is a general need for telecoms to reduce the operational expenses (OPEX) for integrating and adapting new technologies and software. NSN estimates that 60% of operators' budget is spent in integration efforts (Nokia Siemens Networks website, 2009), and Radding (1999) mentions 50% for the same expenditure for information technology (IT) companies. A reason for these large costs is the lack of interoperability between IS coming from multiple vendors and developed in different platforms, like in the case of operating support systems (OSS) for the management of network elements. Interoperability is a key enabler factor for the integration of enterprise IS and solutions.

> *"Interoperability refers to the ability of a system (or process) to use information and/or functionality of another system (or process) by adhering to common standards [..] to achieve improved efficiency and effectiveness of internal process and system operations, timely procurements and fast product delivery, easy and instant access to all required job-relevant information by staff members, and enhanced reporting and monitoring facilities at the administrative level"*

(Vernadat, 2007)

As part of the interoperability of IS, data/information integration is a critical challenge. Integrating information/data models is error-prone and time-consuming. Normally, applications include proprietary information models that were made to meet the particular requirements of that application, and/or were influenced by the expertise and design preferences of the modeler. Consequently integrating heterogeneous information models might be highly time-consuming and costly.

> *"According to different estimates 15-30% of the engineering work in a capital investment project is spent in finding the data, discussing the meaning of the data and reentering the data into different systems"* (Paljakka, 2009)

Although data interoperability challenges are not an exception of the telecom industry, they are affected by the complexity, heterogeneity and performance requirements of the telecom technologies and networks. First, telecoms manufacturers need the flexibility to cope with a significant amount and complexity of information coming from the network equipment and their large amount of different components, equipment platforms, technologies and vendors. Second, system integrators and network operators need to integrate applications from different providers and partners, which normally use proprietary information models and heterogeneous interfaces. And third, network operators and service providers require the agility to handle the distribution of applications and the real-time nature of services. The following section introduces a number of concepts that are important to understand before continuing with the rest of this report (Olsson, 2003).

1.3 Main definitions and related concepts

This section contains definitions for the most important concepts that are used in the rest of this thesis report, and other related concepts, including the definition and purpose of information models, the definition and layers of enterprise architecture (EA), the definitions of integration and service and their relation to the EA layers, and the definition of service oriented architecture (SOA).

An *information model* is an explicit representation of concepts in a particular domain or set of domains. Concepts of the model correspond to either abstract entities, such as service or resource, or concrete entities, such as network equipment. There are different types of IM used for the development of IS. This study is particularly interested in entity-relationship models. References to information model in this document will refer to entity-relationship models. In entity-relationship models the concepts are represented as elements and connections between them are represented as relationships. An element is described with a name and a set of attributes that abstract its main characteristics. Relationships are also described by a name and attributes; however their most important feature is that they describe a connection between concepts. Normally, the elements are graphically represented as boxes or rectangles and the relationships as lines or arrows.

There are multiple uses of information models, however a fundamental use is to explicitly represent tacit knowledge that could be then stored, transferred and applied for a specific goal. This involves that what people knows in their heads is made explicit in models using certain notation, to allow other people or computers understand it (Nonaka, 1994). Furthermore, some of the main purposes of IM in an organization are to increase the organizational understanding, to improve the collaboration of multiple stakeholders and to document and deploy best practices across the organization. Since modeling an entire enterprise would produce a large model, that is difficult for people to use, models are normally focused on a part of the company information and they are tailored to meet the specific requirements from the perspective of a group of stakeholders. Thus, there are different types of information models depending on what is being modeled, their specific use and the perspective used to create the model. For instance, while a database model is used to represent the information stored in an IS, including attribute types and format; a high-level conceptual model might be used to represent how actors participate in a business use case. Furthermore, IM might be used not only for particular domains but also for creating the architecture for entire company (Petri, 1992)

> *"Enterprise architecture: a coherent whole of principles, methods, and models that are used in the design and realization of an enterprise's organizational structure, business processes, information systems, and infrastructure"* (Lankhorst, 2004)

The concept of EA is useful to situate the scope of this study. EA is the enterprise counterpart of the architecture of a building. EA is used to represent the elements of an entire company in terms of both the current situation of the company and a target situation, and to perform gap analysis between the two. There are EA frameworks that are used by companies as a blueprint to have a starting point in the creation of their architectural models. A classical framework for EA was created by Zachman (2009), based on different enterprise roles and their artifacts. Examples of these artifacts are product solution descriptions created by product owners, diagrams used by architects and designers and software programs created by developers.

> *"An artifact is a physical piece of information that is used or produced in a software development process, or by deployment and operation of a system. It is the representation, in the form of e.g. a file, of a data object or an application component, and can be assigned to (i.e., deployed on) a node"* (Lankhorst, 2004).

Zachman's layered framework has been extended in other ones such as archimate (Fredericks and Van der Weide, 2006) and the interop network of excellence framework (Chen, 2008). Based on the Zachman's framework and these two architectures, an EA matrix was created for the purposes of this study (see figure 2). This matrix will be useful to situate the scope of the research and to guide the reader through the different sections of this report. The vertical dimension has three layers corresponding to the ones in the archimate's framework: business, application and technology; and the horizontal dimension presents the static and dynamic aspects of corresponding to the information and functional aspects of enterprises (see figure 2). The layers describe the vertical hierarchy of the components of an enterprise and the aspects describe what information is used in the enterprise and how the information is used by the enterprise functions or processes. In particular, the scope of this study is focused on the information aspect at the business and application layers of this framework (see the circle inside the figure 2).

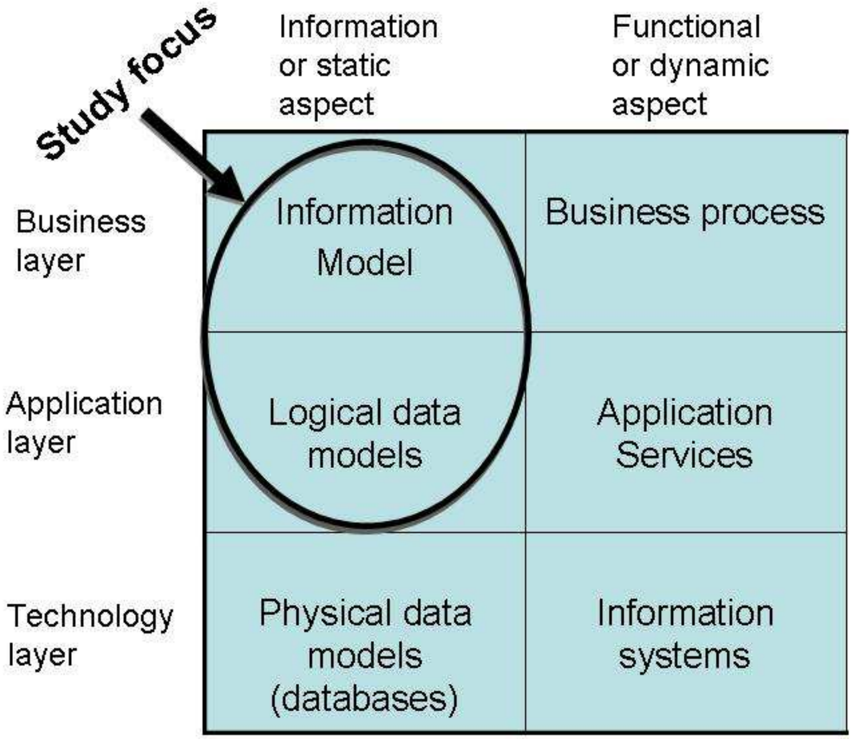

Figure 2. EA matrix of layers and aspects adapted from archimate (Chen, 2008; Fredericks and Van der Weide, 2006)

The descriptions of the three layers of the archimate's framework are as follows (Lankhorst, 2004):

1. The business layer offers products and services to external customers, which are realized in the organization by business processes (performed by business actors or roles).
2. The application layer supports the business layer with application services which are realized by (software) application components.

3. The technology layer offers infrastructural services (e.g., processing, storage, and communication services) needed to run applications, realized by computer and communication devices and system software.

EA is also useful to see an overall picture of the enterprise, where the different aspects and layers are related to each other. EA is a useful tool for performing integration within and across the enterprise (Chen, 2008; Gulledge, 2008; Presley, 2001). *Integration* is used here as the process of unifying, consolidating or coordinating two or more components into a larger unit with greater value than the combined parts. EI refers to the intra and inter organizational levels of integration to carry out more efficient operations, to leverage economies of scale and/or to react faster to the changing business environment. The following is the formal definition of enterprise integration that was used for this study:

> *"[EI] is about facilitating information, control, and material flows across organization boundaries by connecting all the necessary functions and heterogeneous functional entities (e.g., information systems, devices, applications, and people) in order to improve communication (data and information exchanges at system level), cooperation (interoperation at application level), and coordination (timely orchestration of process steps at business level) within this enterprise so that it behaves as an integrated whole"* (Vernadat, 1997).

EI is present at the three layers of EA:

1. *Business Process Integration* (BPI) refers to the alignment of the business processes, in a way that a company is able to control, re-engineer and made them interoperable.

   > *"A business process is a partially ordered sequence of steps executed to perform some enterprise goal"* (Vernadat, 2007).

2. *Enterprise application integration* (EAI) refers to the creation of aggregated services by integrating functions from different computer applications in an enterprise. Two of the main EAI activities are:
   - *Control and connectivity integration* defines how the information is exchanged between applications in terms of the communication session and data transmissions over the network.
   - *Data integration* refers to the syntactic and semantic aspects of the information that is exchanged between applications.

     "Data integration deals with aspects of integrating data within an enterprise. It deals with how data is modeled and the meaning of the data. It deals with normalization, validation, and integrity of data and what translations need to be applied to the data for exchange between applications within the enterprise or between the enterprise and outside

systems" (Smith, 2002)

3. *Physical or technology integration* is required at the network and platform layer to enable the communication between IS. It includes the integration of the various physical components, such as hardware equipment, network technologies, data storage devices, operative systems, packaged software and application servers. This level the interoperability is concerned with low level communication aspects such as networks channels and protocols.

Finally, *service* is a reusable, independent and defined piece of functionality or capability of a system. The purpose of a service is to decouple the consumer from the provider of information. In other words the external and internal views of the IS behavior are made independent (Lankhorst, 2004), thus facilitating the consumer requesting of services from the most convenient or optimal supplier. Service orientation refers to decomposing the organizational resources and structure into services. Services are reorganized independently and seamlessly, into aggregated services and integrated solutions to meet a particular goal. Service orientation is applied at the various enterprise layers; there are organizational services to provide functionality to customers, business services that are combined into business processes, application services to expose functionality of IS and other application to business services, data and network services that provide applications with access to the underlying technology (Erl, 2007). Service orientation is supported by methodologies such as service oriented architecture (SOA). The concept of SOA has been largely studied in previous literature (Erl, 2005; Erl, 2007; Gulledge and Deller, 2008; Josuttis, 2007; Papazoglou and Van der Heuvel, 2007; Rotem-Gal-Oz, 2009), in which SOA has been defined from different perspectives The next two SOA definitions were chosen to expose two of the main perspectives, business and application:

> *"SOA is an architectural concept that defines the use of services to support the requirements of consumers, while the service itself consumes as little resources as possible"* (Josuttis, 2007).
>
> *"SOA [is] an architectural style for building systems based on interacting coarse grained autonomous components called services. Each service expose processes and behavior through contracts, which are composed of messages at discoverable addresses called endpoints. Services' behavior is governed by policies which are set externally to the service itself"* (Rotem-Gal-Oz, 2009).

Using all the previous definitions, the next section continues with a description of the research problem and purpose, followed by the limitations of the study and the structure of this report.

1.4 Research problem, purpose and limitations

In order to face the challenges of providing integrated solutions and to maximize the benefits of EI, an efficient management of the information generated and used organizations is required. A main EI challenge of is to cope with the associated costs of aligning the business, application and technology components of the EA, while procuring a restructuration that facilitates the reusability of services and agility to react to changes. This in turn is translated to IM challenges such as misalignment of information models used at the different layers of the organization, the effort associated with the development, maintenance and reusability of information models and the difficulties of integrating and maintaining heterogeneous, disperse and inconsistent data models used by the different IS (Robinson, 2008; Davies, 2001). One of the root causes of these issues is the lack of a *unified approach* to the creation and maintenance of information models. Such approach would include modeling methodologies, principles and frameworks used to meet the information requirements of the different stakeholders (Petri, 1992). Moreover, IM should supported by different industry standards and tools that have been developed facilitate the improvement of the design quality and value of models. Usually, these elements have been used in particular scoped to obtain individual gains in the modeling of specific aspects of a company. However, a holistic IM strategy is required to leverage the benefits of IM for EI.

Moreover, it is believed that standards and best practices, such as the use of industry reference models, could largely help companies to face their EI problems (TeleManagement Forum, 2009). For instance, by using a standard-based CIM as the enterprise unified model for both new developments and for integrating existing applications. However, it is not clear whether and under which conditions one single information model could support the EI challenges of a company, and what are the practical benefits and issues of using CIM. Furthermore, there is little theoretical background that could help managers and IT departments to understand the implications of using CIM, to make CIM implementation decisions to maximize its benefits. For this reason, the purpose of this study is to explore the business and IT relationships between IM and EI. The interest is to find how CIM could be applied to the information aspect of an EA (see figure 2) to help firms in the development of EI capabilities. The objective of this study is to use a combination of both theory and practice for:

- Establishing a link between EI and IM at the business and application layer using the concept of SOA.
- Creating a conceptual framework of the benefits and issues of using a CIM for creating EI capabilities.
- Showing for a case study firm the practical benefits, alternatives and issues of using a CIM in a SOA.

- Listing the strategic and tactical decisions that managers and IT teams should consider in the implementation of a CIM.

These objectives should be considered within the boundaries of the present research. This study was focused on the integration around the information aspect of EA at the business and application layers. A service oriented architecture and an EA framework was used for guiding the research, writing this report and organizing the findings. The research was limited in time and the empirical part of the research was conducted for the case of global company, NSN at the Espoo, Finland office. One particular industry reference model was used as base for the studying the practical benefits of using a CIM. A specific set of development standards and tools was used for studying a practical implementation of CIM. In addition, the study considered scenarios where the assets are already service oriented; the technology interoperability is expected to be addressed with web services and SOA tools. Moreover, this study was carried out in the context the telecom industry, a dynamic business environment of CoPS. Particularly, this research studied the integration of downstream solutions for the case of NSN.

1.5 Structure of this report

The rest of this report is structured in chapters as follows:

In chapter 2, the overall research strategy is outlined in terms of both methodology and research design. It starts by defining the process followed and the main research questions to be answered. Subsequently, the chapter presents an explanation of the action research method, which was found to be the most suitable method for the type of empirical study performed in this research. As part of the research design, this chapter describes the multiple types of data collected during the research process and the data analysis techniques. Lastly, this chapter presents the results validity and reliability considerations taken into account, in relation to the data collection and analysis.

In chapter 3, the topic of EI is explored in detail by making extensive reference to reliable literature and interviews with two selected experts in the field. The chapter starts by analyzing the objectives and challenges of EI relation to the provision of integrated solutions. The chapter continues by linking the development of both capabilities required for enterprise integration and principles of IS interoperability. Then the concept of EA is expanded, from the perspective of SOA, including the layers, structural elements, benefits and the challenges. Next, the chapter presents a discussion of the role of IM in the process of developing integrated solutions. Finally, the chapter ends with an analysis of modeling roles and the artifacts that are involved in the development cycle.

In chapter 4, aspects of IM are discussed in relation to the different layers of EI, and the use of a CIM is analyzed as an approach to facilitate EI integration capabilities. First, the integration, scoping and approaches of information models are presented as aspects to consider by managers in making IM integration decisions. Second, the chapter presents an explanation of the importance of data model integration, issues affecting it and their consequences. Third, the chapter exposes aspects and considerations of semantic integration. Fourth, the CIM approach is analyzed in terms of the expected benefits for IS interoperability and EI integration capabilities, along with the formulation of propositions for the building of a theoretical framework

In chapter 5, the report of the empirical research is presented. The chapter starts with an analysis the aspects to consider when selecting a CIM, and a telecom industry reference model is presented along with the reasons to use it, the TeleManagement Forum (TMF) shared information/data (SID) model. Next, a description of the NSN network fulfillment use case is presented. This is the context in which the empirical research was carried out, following a three-iteration action research process. Last in the chapter, the relevant findings of the three iterations are presented in a step-by-step basis.

In chapter 6, the results from the action research and interviews are analyzed using categorization. The chapter starts with a summary of the practical benefits, issues, conditions and alternatives of using CIM found during the action research. Next, the results from interviews are analyzed collectively by finding commonalities and relationships within categories. Subsequently, both the action research and the interviews are consolidated by triangulation of data sources. The triangulated results served to test the validity of the initial propositions presented in chapter 4. The chapter ends with the generalization assessment of benefits of CIM for IS interoperability and EI integration, which served to build the final conceptual framework.

And in chapter 7, the analyzed results are discussed in relation to the main research questions and contributions of the study are identified. Finally, a closing statement is made regarding the entire research, followed by research areas to be explored in the future.

2 RESEARCH STRATEGY

This chapter outlines the overall research strategy in terms of both methodology and research design. It starts by defining the process followed and the main research questions to be answered. Subsequently, the chapter presents an explanation of the action research method, which was found to be the most suitable method for the type of empirical study performed in this research. As part of the research design, this chapter describes the multiple types of data collected during the research process and the data analysis techniques. Lastly, this chapter presents the results validity and reliability considerations taken into account, in relation to the data collection and analysis.

2.1 Research process and questions

Based on the objectives of the research, literature review and a set of meetings with NSN stakeholders, areas of interest were identified for this research. Companies have identified the need for CIM and in industries like manufacturing, logistics and telecoms, industry reference standards available for companies to implement CIM (Gulledge, 2008; Reilly, 2008). In addition, a market exists for semantic integration tools that facilitate the implementation of CIM (Wilmes, 2008). However, there has been little scientific research to formally study CIM. Although companies claim benefits from CIM, research is needed on the factors that affect the leverage of such benefits. Moreover, it is interesting to study the practical implications of using CIM, particularly what conditions and alternatives exist in the implementation of CIM, and what issues could arise. Furthermore, this research has the purpose of contributing to the industry by serving as a guide for making managerial decisions regarding CIM. The following are the three main questions this research intends to answer:

1. How CIM can benefit the EI capabilities of telecoms?
2. What are the practical benefits, issues and alternatives of using CIM?
3. How using CIM affects managers' strategic and tactical decisions?

The theory building process proposed by Handfield (1998) serves as a guideline for the present research. The process consists of inductive and deductive reasoning phases, one after the other, using in each phase both theoretical and empirical research. Inductive reasoning consists of going from observations to theory, or from the specific to the general, through general propositions. Deductive reasoning is opposite to inductive, it goes from theory to observations, or from the general to the specific, through conjectures. Moreover, the approach consist of an eight-step cyclic process, as depicted in figure 3, where the left side of the circle corresponds to the inductive phase and the right side to the deductive phase. In addition, distinction made by Handfield

(1998) clarifies the nature of the activities during each step of the process. The upper side of the circle corresponds to theoretical activities while the lower side of the circle corresponds to the empirical research activities. Finally, it is noted by Handfield (2008) that research studies following this approach might include more than one loop to this cyclic process.

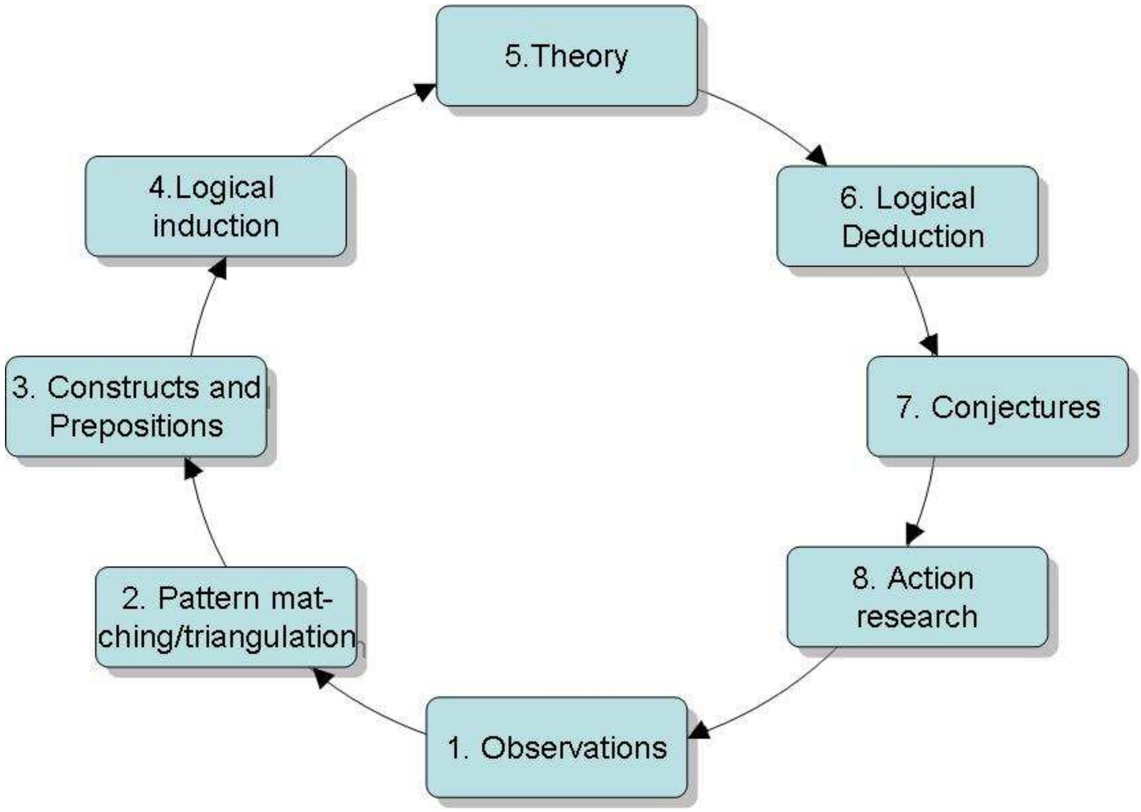

Figure 3. Overall research process (Handfield, 1998).

In this case study, one loop and a half were conducted in three phases: theory building, testing and refinement. These three phases are aligned with the main questions to be solved, as shown in figure 4. The first phase was inductive, following steps 1 to 5, the second phase was deductive, following steps 6, 7, 8 and 1, and the third phase was again inductive, following steps 2 to 5.

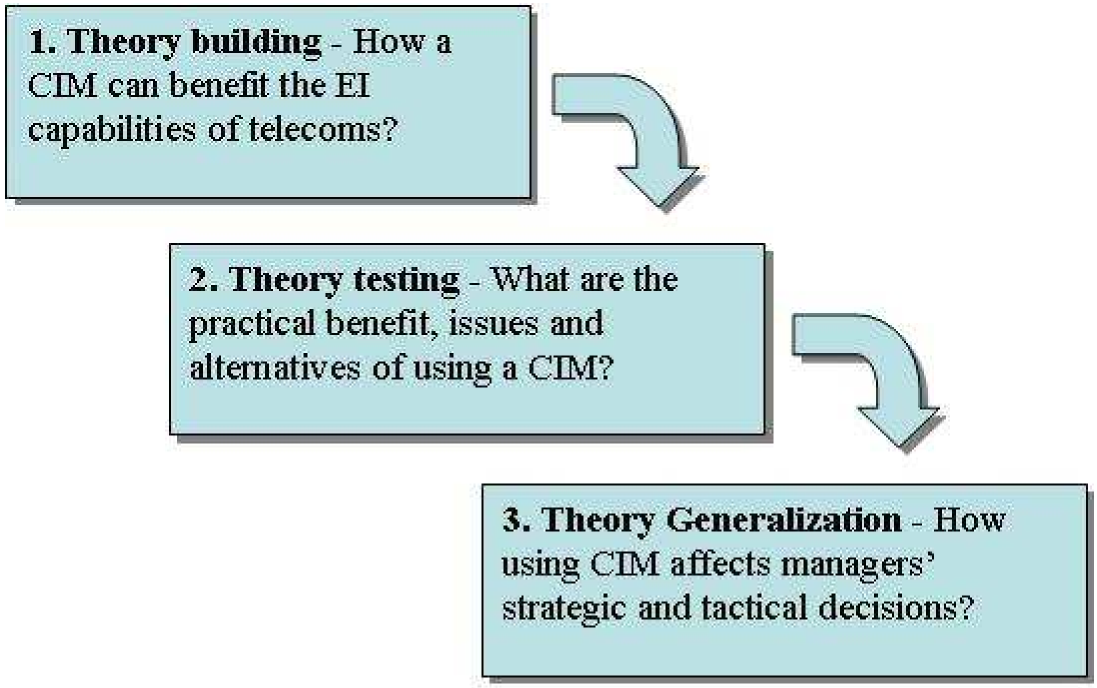

Figure 4. Research phases and questions

The overall research design was done following the recommendations of Yin (2002). The following paragraphs detail the process followed during each of the three phases, by making reference to the eight-step theory building process of figure 3:

In the first research phase inductive reasoning was performed to build an initial theory of EI and data integration challenges. The research started by compiling data from a number of sources, including journal articles, company documents from NSN, documents from standardization bodies and whitepapers from application integration vendors. In addition, observations from meetings with practitioners complemented these ideas from the practical perspective. Pattern matching was the second step and consisted in selecting and categorizing relevant ideas into the categories of two patterns. One pattern was a cause-context-phenomenon, corresponding to the causes of IS interoperability, under the context factors of the telecom company and their effect in the business. And the other pattern was action-condition-consequence, corresponding to the action of using CIM, given conditions that should hold true and the expected benefits were the consequence. In the third step of the research process, the result of pattern matching were rephrased as constructs and classified in three categories: cause, effect and context. In the fourth step, the constructs were complemented with discussions with practitioners and interviews with experts in the field. Finally, in step five the constructs were abstracted into propositions to build a theory. This first research phase concludes with the development of a conceptual framework to depict how a CIM might benefit the EI capabilities of telecoms. Finally, this framework is the base of the empirical research to operationalize the case study conducted in the second research phase.

In the second phase deduction reasoning was used to validate the empirical applicability of the initial framework. During this phase, an empirical action research was carried out following an iterative process, as described in section 2.2, at NSN with the collaboration of a research team. During each iteration, conjectures were made by the researcher to understand the outcome of each iteration. These conjectures were based on the conceptual framework, the opinions of the research team and, in the case of the second and third iteration, the lessons learned in previous iterations. The *action* performed by the researcher consisted of integrating information models using CIM for integrating applications. This action was performed under the scenario of a network fulfillment solution demo that is based on SOA and involves three IS. A detailed description of the scenario is presented in the chapter 5. Particularly, the SID model was used as a basis for the CIM. As a result of this second research phase, observations were obtained summarized and categorized in benefits, issues, conditions and alternatives of using CIM in practice.

Finally, in the third research phase, induction thinking was performed again to evaluate the generality of the findings. The process was similar to the one followed in the first phase, the difference is that generalizations were based on the findings of the action research and additional observations obtained from interviews with practitioners. All these observations were categorized in benefits, issues, conditions and alternatives of using CIM. Subsequently, conjectures were extrapolated back to propositions using triangulation of data sources and informants (Susman and Evered, 1978; Yin, 2002). At the same time, assessments were made to determine if each condition or alternative was a necessary and/or sufficient factor in the result obtained. The outcome of this third phase is a refined conceptual model that includes conditions affecting the realization of benefits of using CIM for developing EI and IS integration capabilities. In the following section, the action research methodology and process are detailed.

2.2 Action research

Action research was first introduced in 1946 by Kurt Lewin as an alternative to the traditional positivist science, such as the formalism, reconstructionism and pragmatism schools. In positivist, research knowledge is obtained from observations made by independent observers. The goal is to make predictions through explanations regarding future actions and behavior taken by actors. Moreover, the positivist research assumes that the context of study is structured logically, and predictions of taking an action could be made without taking the action, which is not the actual case in organizations. Conversely, action research does not require the context of the research to be logically structured. Given the way organizations are structured and the effect that persons play in modifying the organizational behavior, action research does not try to predict but to understand the effect of actions. In action research, the researcher acts on or in the system of study to not only understand but also to try to improve such system and generate critical knowledge of it. Action research considers the role of humans in creating artifacts to satisfy their particular needs. In action research a set of values are defined as the vision to reach and actions are directed towards achieving those values. Finally, the researcher and practitioners collaborate to learn from the consequences of trying alternatives in practice (Susman and Evered, 1978). Next, these characteristics are reviewed in the light of this research to explain the reason for choosing this method, the process followed and the participants involved.

Considering the characteristics of action research, it suits the study of organizational IM and EI. Telecom and IT dynamic environment makes the object of study a moving target, which is difficult to predict and control. The relevant values for the integration of systems and solutions were identified by performing practical evaluation of a CIM. Collaboration between researcher and practitioners allowed a clear understanding of the actual issues and implications of using a CIM. Moreover, since IM is an instance of humans defining artifacts for particular purposes; logical reasoning should be complemented with action (Baskerville, 1999).

Action Research answers the second research question as part of the second phase in the research process to test the initial theory developed in the first phase. The propositions of the first phase were taken to a more specific level, which could be measured during the action research. The process consisted of an iterative five-step process, carried out in three iterations. A detail description of each step is presented:

1. Diagnosing - First, areas of improvement were identified as application integration issues that could be solved by using a CIM.
2. Action planning - Various implementation alternatives were considered; depending on the options supported by the platforms and a selection was made based on the measures defined earlier during the second phase of the research process. During this step, conjectures were created to describe the expected improvements.
3. Action taking - The plan was executed. Any unforeseen decision taken or constraint was reported during this phase.
4. Evaluating - The results of the execution of the action were analyzed against the intended benefits described in action planning.
5. Specifying learning - In this step a decision was made to reject, modify or keep the conjectures as valid.

A number of participants were involved in this process, including the thesis worker, the university thesis supervisor, the thesis mentor and the research team leaded by a team manager. Their roles are as follow:

- The *thesis worker* performed the practical activities and requested assistance and advice from the research team. This person was also the main writer and owner of the action research report, included in this document.
- The university *thesis supervisor* tutored the work process through discussions, meetings and revisions of the report. Additionally, the issues were discussed in the research seminar meetings at the university.
- The *thesis mentor* guided the overall process and provided feedback through periodical meetings. This person made contributions and reviewed the action research report.
- The *research team manager* provided inputs in relation to the tactical goals and objectives, in addition to priorities. This in addition to facilitating resources and communication with other persons inside the organization.
- The *research team members* collaborated in the action taking and provided technical assistance, feedback and support for the action research; they also provided feedback for the evaluation of the lessons learned of the iterations.

2.3 Research data collection

Different data collection and reporting instruments were used for each of the research phases. They are summarized in table 1 for each of the three research phases and a detail description follows it.

Table 1. Research data collection and reporting design

Phase	Data collection	Data Reporting	Participants
1. **Theory building**	- Literature and documentation - Meetings with research team and thesis mentor - In-depth interviews with selected experts	- Categorized list of constructs and propositions - Meeting notes - Interview reports	- Thesis mentor - Research team manager - Three industry experts from TMF, NSN and Electricity Company
2. **Theory testing**	- Action taking and action evaluation	- Research log - Meeting notes - Customized technical reports	- Thesis mentor - Research team manager - Research team
3. **Theory generalization**	- Semi-structured interviews with different	- Interview reports	- Three architects, two from NSN and one from an

	practitioners		insurance company - Two NSN developers

During the first phase, a broad set of data sources was used to gather insights regarding the benefits of using a CIM. An articles database was created with categories such as EAI, semantic integration, SOA, research methodology, white papers and telecom. In addition, meeting notes were taken during the meetings held with the thesis mentor and the research team. In this phase, the communication between the thesis worker and research team was focused on clarifying concepts, goals and discussing what values the action should strive for. Furthermore, data for this phase is complemented with data collected from three interviews from industry experts. The three interviews were documented in written reports that were validated with the interviewee. The first interviewee, a TMF SID model administrator, was selected as a key person because of his important contributions to the development of the SID model and his experience as enterprise modeler, architect and teacher of the SID. The second interviewee is a data integration expert at an electricity company. This person was selected for the availability, experience and the current assignment carried out in that company. And the third interviewee is a technology and architecture manager at NSN. This person was chosen for his experience in IM and EA, and his work with the SID model.

For the second research phase, during the action research iterations, qualitative data were recorded in an electronic research log and technical reports. The research log maintained day-to-day findings of the action taken and decisions made. The technical reports included relevant information regarding the current procedure for data integration, the problems found, the ameliorations or proposals for improvements, the specific changes that are proposed and the evaluation of the current mode of action. The technical report information was categorized in four sections: business, application, information and tools. The business section refers to the managerial concerns, the application section to the IS view, the information section to the knowledge management perspective for both business and IS view, and the tools section the findings regarding the usage of the selected platforms and computer tools used during the action research. Additionally, as with the first phase, discussions held in meetings with the thesis mentor and the research team were documented with meeting notes.

Finally, for the third research phase, qualitative data was reported from semi-structured interviews with practitioners. These interviews were documented in written interview reports that were validated with the interviewee. First, the representative of the EA team from an insurance company was interviewed to confront the type of issues and alternatives of using a CIM as part of EA efforts. Second, two SOA practitioners from NSN were interviewed to contrast the practical findings and the conjectures of action research with their experiences. And third, two architects of the operations and business software unit at NSN were interviewed, as key informants of the current architectural situation in NSN and the applicability of the findings of action research.

2.4 Data analysis and validity

Data from a variety of sources was obtained in this research. During each of the three research phases, the data and analysis methods were chosen strategically for answering the three main questions of this research. Similarly, different data analysis techniques were chosen to match the purpose of each research phase. In addition, validity tests were applied to ensure the quality and reliability of the results (Yin, 2002; Voss 2002). In this section, the description of analysis techniques for each phase is provided, followed by the types of validity tests and their relationship to the analysis techniques.

Data from the first research phase was analyzed through pattern matching to understand the benefits of using a CIM in semantic and IS integration. Using a cause effect relationship, data coming from literature, documents and interviews was categorized in cause, phenomenon, condition, action/strategy, context and consequence. Data sources were included if they follow the pattern to the extent that they include information that fits in either in the phenomenon or the benefit category. For this analysis, the phenomenon was used to represent the semantic integration issues and the action/strategy represented the use of CIM. The collection of constructs for this pattern matching is presented in appendix 1. A criteria of strength was taken in order to include constructs, this was if a construct was supported by a scientific journal or by two different non-scientific sources. After completing at least four constructs for each category, constructs from different categories were combined to build propositions, presented in section 4.4. The propositions were the base for developing an initial conceptual framework of the effects of using CIM for semantic and ARE integration.

For the second phase, three iterations of empirical research were performed as described in section 2.2. The data collected during each iteration was reported in this document as a narrative of the five steps of the action research methodology. Subsequently, a narrative of the three phases was analyzed by identifying and categorizing constructs using categories and subcategories (see appendix 5). The main categories are benefits, issues, conditions and alternatives for using a CIM. Finally, a summary of the analysis is performed using the same categories.

For the third research phase, qualitative results obtained from interviews served to triangulate the results of action research. The interview reports were analyzed using the same categories than for action research results. Appendix 7 contains details about the interviewees and the interviews. These results were analyzed from two perspectives:

they served to validate the initial propositions constructed in the first phase and they served to generalize the results of the entire research.

Four main different validity tests used for case studies are leveraged in this research, construct, internal, external and reliability tests. First, construct validity is related to the correct identification of concepts and measures. This is addressed by using multiple sources of evidence and establishing a cause-effect chain of evidence. In addition, constructs validity is supported by revisions made by peer IM researchers and triangulation with interviews with key informants. Second, internal validity refers to the establishment of correct causal relationships. Meeting discussions with peers served to analyze the correctness of causal relationships learned during the empirical action research. Third, external validity is mainly concerned with the extent to which the results could be generalized. For this, triangulation was performed by interviewing people that is both internal and external to NSN. In addition, interviewees have different type of role in their organization, for instance, architects, developers and managers (Stuart, 2002; Yin, 2002). Fourth, to strengthen the reliability of the results, data was stored electronically and reviewed by peers and the actual informants, in the case of interviews. Lastly, the entire research process was performed as described in this chapter, allowing the possibility to repeat the same procedure, and data analysis was conducted by following the techniques presented in this section (Stuart, 2002; Voss, 2002).

3 ENTERPRISE INTEGRATION, ARCHITECTURE AND MODELING

This chapter discusses EI at the business and application layers of EA. First, the objectives and challenges of EI for the development of integrated solutions are discussed from the view of dynamic integration capabilities. Then the objectives and challenges are narrowed towards IS integration, and the development of IS interoperability principles is linked to such capabilities. Then the concept of EA is expanded, from the perspective of SOA, including the layers, structural elements, benefits and the challenges. Next, the chapter presents a discussion of the role of IM in the process of developing integrated solutions. Finally, the chapter ends with an analysis of modeling roles and the artifacts that are involved in the development cycle.

3.1 Objectives and challenges of enterprise integration

Driven by the provisioning of integrated solutions in an efficient and sustainable manner, providers of CoPS will pursue EI capabilities. These capabilities are driven by the achievement of tangible and intangible objectives in both horizontal and vertical EI. On the one hand, tangible objectives are for instance to increase of service quality and to decrease of unit cost and cycle time, and intangible benefits are for example to increase business process flexibility and the communication and cooperation among teams. On the other hand, vertical integration refers to the collaboration of the different organizational levels from top management down to tactical planning and operation levels. The purpose of vertical integration is to align the end-to-end business and IT strategy, thus enabling the integration of business process and the supporting IS. Horizontal integration refers to the collaboration and alignment of functional areas or business domains. The purpose of horizontal integration is to facilitate the timely exchange of information across organizational units in order to optimize the use of resources and improve the customer experience, among other goals. Consequently, companies moving towards the provision of integrated solutions need an integral strategy to achieve EI objectives (Lim, 1997).

Moreover, as providers of CoPS move downstream the value-chain, greater business agility is required to react timely and efficiently to the requirements of customers. Since value-added integrated solutions are directed towards the satisfaction of the changing customer specific requirements, dynamic coordination of the organizational units and resources is required, including organizational information, knowledge, technology, infrastructure and financial assets. And the development of *dynamic capabilities* will contribute to this objective. (Braganza, 2002). For this research, the dynamic capabilities identified by Teece (1997) were studied:

- *Coordination and integration* refers to both the internal and external activities of the firm, and how efficiently and effectively new technologies are integrated. This capability includes the gathering and processing of information for linking the customer experiences with the engineering design, and for coordinating the operations and supplier relationships (Teece, 1997).
- *Learning*, from the organizational perspective, is the ability to share and reuse knowledge generated inside and outside the organization. Learning allows to repeat best practices and solutions to problems, and to identify new opportunities. Learning requires the joint collaboration to understand the organizational logic and to find dysfunctional areas (Teece, 1997).
- *Reconfiguration and transformation* refers to the ability to reorganize the organizational internal and external assets, to transform the process and supporting systems in a way that minimizes the cost of changing (Teece, 1997).

However, developing dynamic integration capabilities in pursuing EI objectives presents a number of challenges. From the organizational perspective, generic EI challenges are communication issues implicit in political challenges, cultural differences and geographic dispersion of teams. There is also resistance to change and misfits between IS and organizational structures and practices. In addition, the provision of integrated solutions of CoPS pose extra challenges, for instance the responsiveness and availability of services is critical for the business, a failure in the systems could have large repercussions even for small periods of unavailability. Particularly, in industries with the dynamic characteristic of telecoms EI integration is a greater challenge. The telecoms' need to optimally reconfigure services requires the interoperability of IS made by multiple vendors, partners and even competitors. Furthermore, since the integration of IS play a crucial role in implementing EI initiatives, a link must be done between EI and IS integration in terms of both objectives and challenges. These two aspects are analyzed in the following section from the perspective of IS.

3.2 Objectives and challenges of information systems integration

In the present days, companies need to build partnerships with other participants of the business network. In order to achieve a greater customer satisfaction, economy of scale, competitive advantage and sustainability, companies should leverage information systems to facilitate communication with partners and its business integration. Moreover, the integration of information systems provide companies with the agility to offer integrated solutions in real-time and the automation, alignment, improvement of business processes. (Linthicum, 2003). Such agility poses the requirement for integrating systems in an efficient and effective manner. This includes both internal systems used either for a particular unit or for end-to-end operations, and external systems from customers, suppliers, distributors, banks, government, service providers and even competitors. Furthermore, different types of IS integration scenarios might be required for EI. Some examples are migrating IS to a new version, migrating to a new technology or framework, merging two different IS for a specific unit, consolidating IS from two business units, outsourcing IS to an external provider and consolidating IS from various parties to provide new services.

Important business parameters that are reflected in system integration requirements are the reduction of integration and maintenance costs, a smooth learning curve and a thorough change management. In consequence, companies should strive for system integration and seek for optimal ways to manage the integration process and reduce the so called 'integration tax' (Linthicum, 2003). Previous literature on EA proposes a set of principles for the development of interoperable components (Erl, 2007). It is believed that by following such principles, an organization could increase integration agility and reduce integration costs. (Erl, 2007; Gulledge, 2008; Malatras, 2008; Papazoglou and Van der Heuvel, 2007). In consequence, the development of IS interoperability principles should enable the development of EI capabilities, such as the ones defined in section 3.1. Thus there is a relationship between the IS interoperability principles, the IS integration tax and the EI dynamic capabilities (see figure 5).

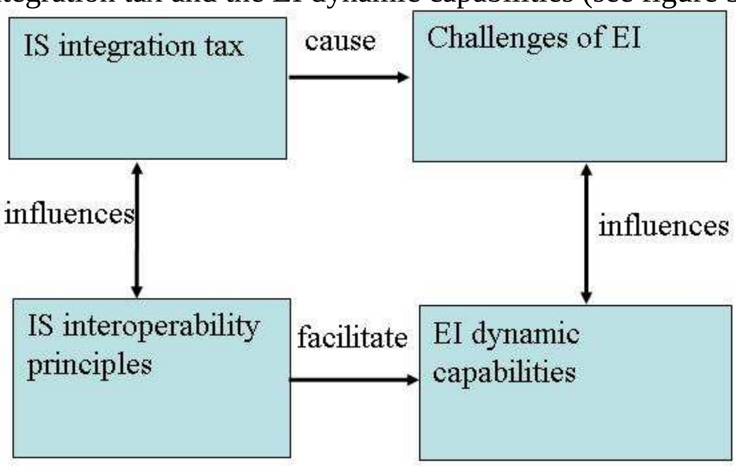

Figure 5. IS integration effects on EI.

The following *interoperability principles* were chosen as the most representative of the literature reviewed:

- *Using open standards* is desired because as they become more adopted by components, the communication naturally interoperable, reducing the need for transformations. An open standard is publicly available without need to pay for using it, additionally an open standard should not vendor specific (Erl, 2007).
- *Reusability* of components or services for more than one single purpose.

Having reusable components allow companies to leverage repeatedly the same logic and to save cost and time to market (Erl, 2007).

- *Loose coupling* refers to the degree of dependency between two components that are connected. A component is completely dependent when requires the existence of another component to function properly. In terms of IS, coupling refers to the relationship existing between a technical contract and the solution logic it is representing or the technology supporting it (Erl, 2007).

  > *"A [technical] contract is the complete specification of a service between a specific provider and a specific consumer"* (Erl, 2007).

- *Granularity* of components refers to the adequate level of detail, or abstraction, of the functionality provided by a system component (Erl, 2007). This principle is related to the structure of the components, which might follow patterns such as modular and/or hierarchical, among others.
- *Interpretability* of components is a capability for different persons to understand the functionality of a component. Additionally, interpretability favors the reusability of components and collaboration between stakeholders (Erl, 2007).

However, the complexity of IS integration accumulated from various types of integration scenarios is a major challenge. A number of factors influence the outcome of IS integration and might contribute to increasing implementation costs. Moreover, the recurrent appearance of these costs are know as *integration tax*, as an analogy to components of a balance statement that must be included in all cases. For example, IS integration projects might fail due to deficient IS quality, lack of IS support from the vendor, heterogeneous or incompatible systems, slowness or unreliability of networks used to share information, old systems not supporting the business needs, support of older versions stopped by vendors, lock-in effects associated to certain IS products, limited amount of control over participant IS, costly and time consuming evaluation of IS. (Chen, 2008; Hohpe and Wolf, 2003).

Finally, there is an effect between IS integration tax and the challenges of EI integration (see figure 5). The benefits of EI efforts should be analyzed against their associated costs and risks of integrating IS. This should be done by both business and technology stakeholders. Some of the IS decisions that commonly affect EI challenges are to select an IS vendor, to choose between internal and external IS development and management, and to choose between keeping an IS as-is, integrating an IS with others and replacing an IS. These types of decision are complex and they have been studied largely in previous literature (Bernroider and Stefan, 2001; Umar, 2009). In particular, this study is directed towards situations in which integrating IS was decided.

3.3 Service oriented architecture

One of the most promising approaches to achieve EI and IS interoperability is through SOA. SOA has been used as methodology for integrating distributed applications using platform independent standards and messaging technology. As described in the introduction of this report, SOA has both business and technology connotations. There is a large amount of previous literature studying both the high level business benefits of SOA and the low level technology details of SOA (Hale, 2006; Erl, 2005; Erl, 2007; Josuttis, 2007). In this report, the concept of SOA will serve as a bridge between EI and IM. In this section the main aspects of SOA are presented in relation to EI for the specific case of BPI and EAI, including challenges of using SOA, and IM is related to the development process of integrated IS in a SOA.

Before going into the details of SOA, it is important to discuss ‘services’ and how they relate to business processes. SOA is based on the concept of a service, a piece of functionality exposed by applications using a standards-based contract or interface. There are different types of services associated with several layers of a typical SOA. First, in the business layer, companies expose organizational services, based on well defined business processes that use application services. Second, in the application layer, more specific functions are exposed by applications services to support the execution of business processes. Third, in the technology layer, data and connectivity services are exposed by the different IS that form part of the company portfolio of IS. These are more specific services to access different types of assets, such as data and network access. These three layers of a SOA, depicted in figure 6, correspond to the three layers of the two-dimensional EA described in the introduction. Next, the structure of a SOA is discussed in relation to the integration of components at the three EA layers, from an IM perspective.

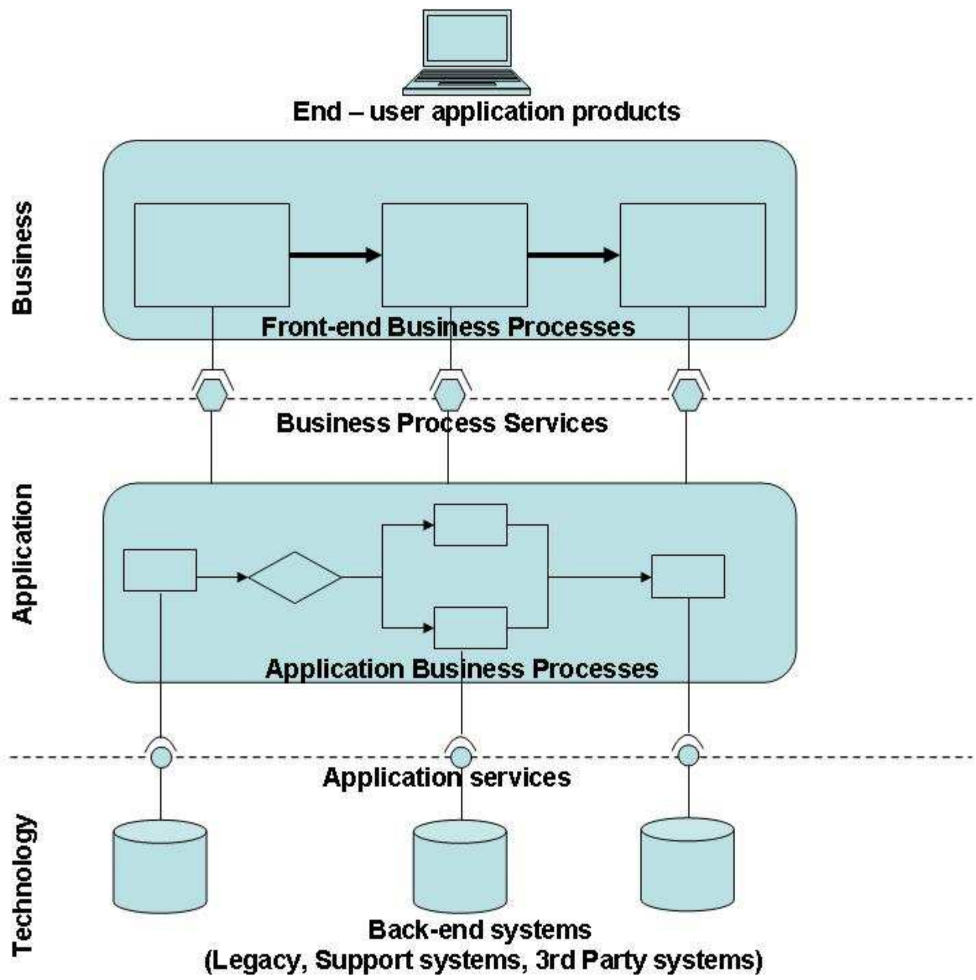

Figure 6. Example of SOA architectural layers.

Starting from the *business layer*, there are end-user products that support the actual use case requirements as part of a business integrated solution. These are for instance, in the case of telecoms, network management solutions for broadband provisioning, quality assurance, network traffic monitoring and billing reporting that are made available to end users through interfaces such as a web portal. These solutions are compound of *front-end business processes*, which represent flows of information between high-level concepts such as customer, network or bill. At this level there is no information regarding the supporting systems, but only the high-level steps followed in a use case. Front-end business processes are realized by business services that expose the functionality of applications. These business services are a bridge between the business and the application layer (see figure 6); they abstract the details of application functionality into a business function.

At the middle of the SOA architecture is the *application layer*, where business process services are realized by assembling application services into application business processes. These, as opposed to front-end business processes, are aware of the information system portfolio that is used to support the business and IT strategy.

Among others, applications might integrate legacy systems, acquired software packages, customer relationship management, enterprise resource planning and supply chain management systems. In a SOA, the functionality of these systems is normally exposed in the form of application web services (Erl, 2007).

And at the bottom of the figure 6 is the *technology layer,* where application services are realized by the back-end systems. In this way, existing IS assets could be leveraged in the creation of integrated solutions. For instance, a data inventory solution holding customer information might be exposed in the form of data services such as create customer or retrieve customer information. At this layer, implementation details, including the connectivity protocols and messaging standards are agreed between produces and consumers of services.

Next, three of the main structural elements of a SOA implementation are described along with their function (Erl, 2005; Zimmermann, 2009):

- *(Web) services.* Applications expose pieces of their functionality as reusable services. Web services are a technology for implementing application services using messages for communication requests and responses among applications. A service is exposed formally using a contract or interface, which roughly include operations and a data schema. The operations are like functions that might or might not have input and output parameters. The parameters are data elements that are defined in the web service data schema. In the case of web services, data schemas are defined using standard formats such as the unified modeling language (UML) and xml schema definition (XSD). And data schemas are included or imported to be used in the operations and web service definition, using standard languages like the web service definition language (WSDL). Finally, composite services aggregate functionality from various services.
- *Business processes.* Web services and composite services are used by business processes, which are meaningful for the business view. Business processes represent information workflows between web services to fulfill the needs of a use case. These workflows combine both sequential and conditional paths connecting business steps. Steps are normally requests to web services, case in which data element mappings are done to pass the information required by the web service parameters. Steps also represent human tasks. Business processes are implemented using a language like the business process execution language (BPEL), and these are stored and managed in a business process engine.
- *Enterprise service bus* (ESB). The main role of an ESB is to provide interoperability between different platforms. Among other functions, an ESB might include providing connectivity, data transformation and message management functions such as message validation, composition, routing and

transformation (Erl, 2005). Message validation includes the verification of the syntax and completeness of messages. In addition, business rules might be included to validate the semantics. Message composition might involve gathering data from different sources to create messages. Message transformation refers to making transformations of data format and mappings of data attributes from a source data model to the target data model. And message routing is the task of sending the message to the corresponding target system.

The SOA contributes to the development of IS interoperability principles. Although implementing SOA requires initial efforts and costs, by using SOA for EAI, companies eventually reduce interoperability issues of integrating applications. On the one hand, standards services are used to developed solutions, meeting the different requirements and incorporating heterogeneous applications, distributed among various parties, running on different platforms and having different data models. On the other hand, by using SOA the business processes are naturally decomposed into the underlying services, reducing the dependencies and increasing reusability. For more details on how SOA benefits EAI refer to previous literature (Erl, 2005; Gulledge, 2008).

> *"SOA addresses the requirements of loosely-coupled, standards-based and protocol independent distributed computing"* (Papazoglou and Van der Heuvel, 2007).

Furthermore, in the case of telecoms, previous research has contributed to the study of SOA methodology in telecoms EI. First, a model of SOA for service provisioning was proposed by Kim and Lim (2007). This architecture is compound of several layers that are defined in terms of services and is presented with the example of an order delivery process. Second, a framework for creating services utilizing telecom functionality is presented by Mittal (2008). Third, Duke (2005) described British Telecom's gateway approach for business-to-business (B2B) integration within telecoms. Third, the TMF has developed a suite of frameworks for the next generation operation support systems (NGOSS), which are useful to support a SOA strategy. The suite includes a business process framework, application architecture framework and a shared information/data model. In addition, NGOSS frameworks are complemented with standards and tools to facilitate the development of new systems and the integration of existing ones (Tele Management Forum, 2009). In the following section IM is discussed in relation EI and SOA, including the model elements and IM activities of SOA.

However, SOA still presents challenges regarding the integration and information management. Implementing SOA requires large efforts from the business, application and information perspectives. Misconceptions and lack of readiness and governance are among the major causes hindering the success of SOA implementations (Erl, 2005).

According to SOA maturity models, for instance the one proposed by the open management group (OMG), companies will require working on the consolidation of their information models to facilitate the different activities required in the development and maintenance of services (Ramakrishna, 2009). Furthermore, a critical challenge is the development of a unified data/information model for defining artifacts such as business requirements, service interfaces and data models. The following section explores aspects of this challenge in relation to the IM development cycle, including roles and artifacts.

3.4 Information modeling process for developing integrated solutions.

Information models are created at different times by stakeholders playing different roles, perspectives and information requirements. As a consequence there is normally a semantic or conceptual gap between the organizational levels and functional areas, which implies that even though people refer to the same domain, they use different vocabulary. For example, a sales executive, a marketing manager and an IT administrator might use the same concept of customer, however they would model represent it using different names or different attributes (Gulledge, 2008). The same situation is present in the development and management of IS, stakeholders with different roles and requirements produce or use different artifacts, specifically modeling documents (see figure 7). These are, for instance, business modelers, system designers and software developers. While business modelers work with high level artifacts such as strategic plans, business priorities and business models; system designers work with business processes workflows, system architectures and frameworks; and software developers work with database models and data flow diagrams. The artifacts and concepts used by these perspectives have differences in the meaning and the detail, thus affecting the way the stakeholders collaborate in the development of IS.

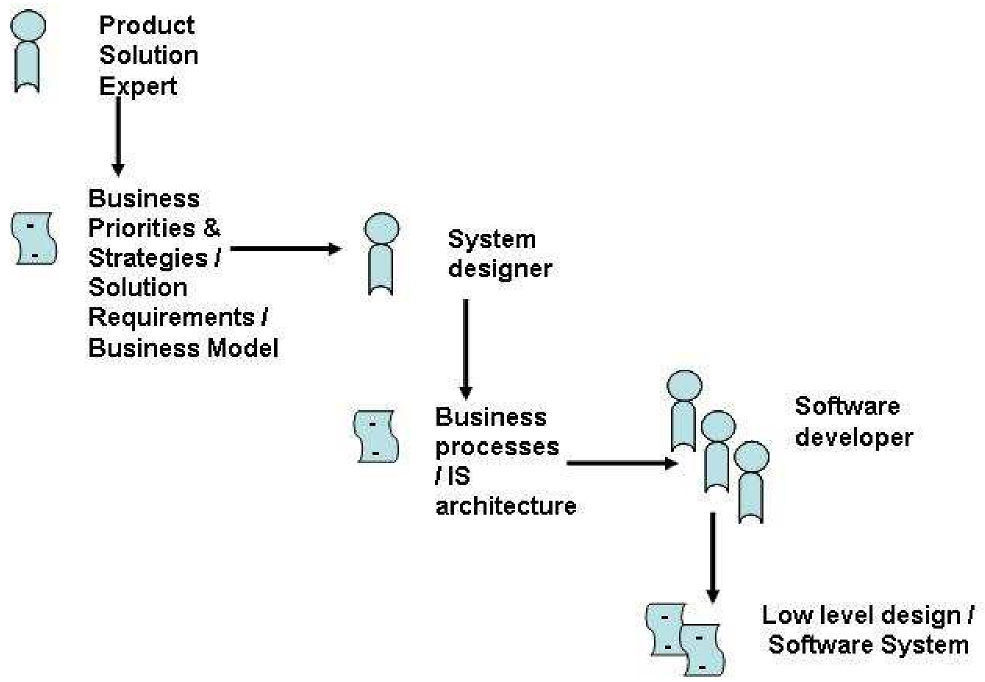

Figure 7. IM roles and artifacts in traditional development of IS.

Regarding the development of integrated IS, previous to SOA, the collaboration between business analyst and system designer was mainly limited to the exchange of artifacts (see figure 7). The business analyst was a domain expert, and it was in charge of analyzing the customer needs or business opportunities and writing down the requirements of new solutions or products. The business requirements handed to the system architect, which had limited business visibility but was an expert in the IS assets. The architect was in charge of interpreting the business requirements and roughly designing an automation IS to fulfill them (Erl, 2005). This high-level design was further detailed by software developers. As described in the next, this collaboration has been somehow enhanced with SOA.

With SOA, more modeling roles are normally present, which partitions the modeling activity or concerns in more segments than the traditional development (see figure 8). Given the layered structure of SOA model elements and artifacts, stakeholders should work together in the decomposition of business to system artifacts. This implies that business level artifacts should be realized by application layer artifacts and these by technology level artifacts. Two main approaches could be taken when implementing integrated solutions in SOA, a top-down approach consist of decomposing business process models into business services, composite application services, application processes and application services. This is also known as the “contract first” approach, and it is recommended as a best practice (Erl, 2005). However, a bottom-up approach is also possible when services are adequately defined and their composition/aggregation into composite services and business process is relatively simple. A mixed approach is also a normal scenario in the real world, where architects receive business solution descriptions and application service definitions, and they have to integrate the middle layers of services and business processes.

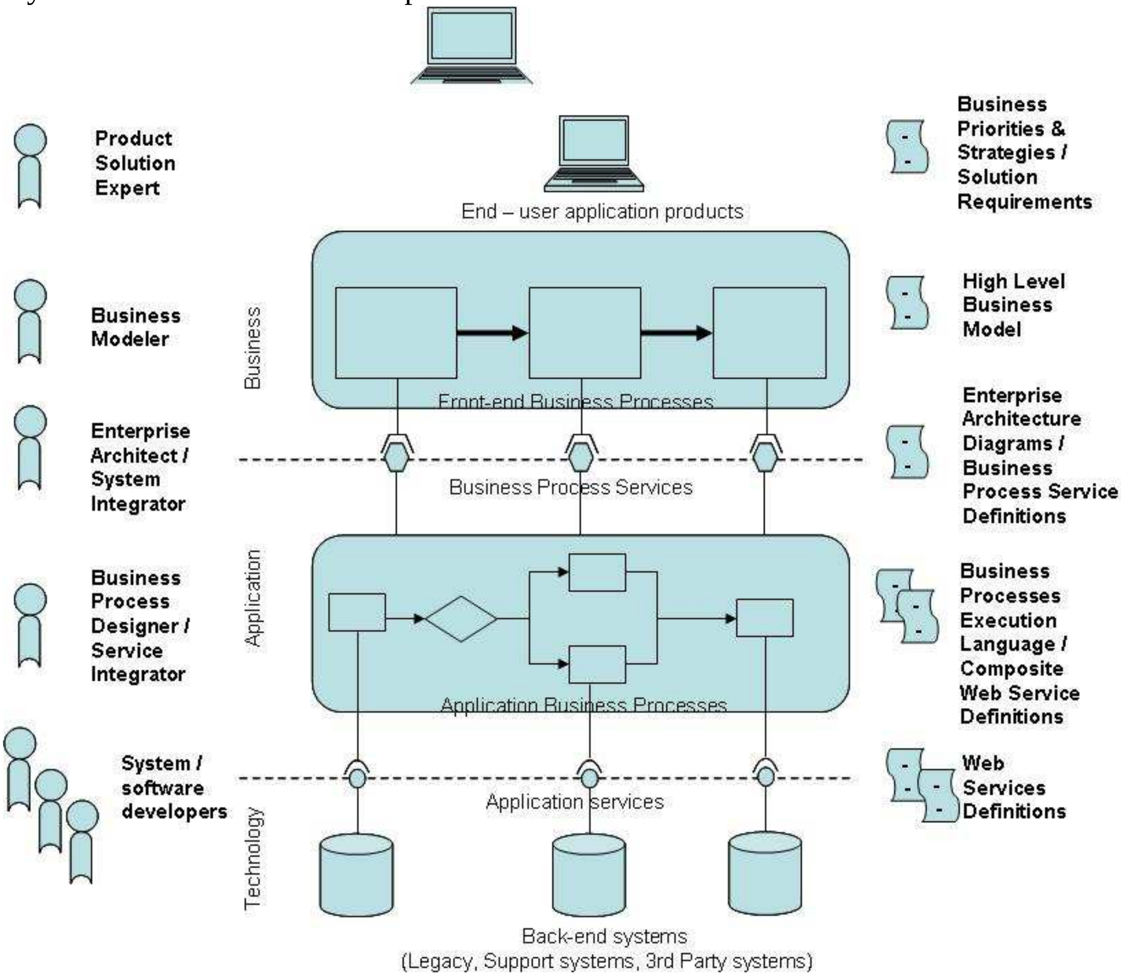

Figure 8. IM roles and artifacts in SOA development of IS

At this point, it is useful to describe the hierarchical structure of model elements of a SOA, from business processes to services and operations:

> *"Business Processes: A long running set of actions or activities performed with specific business goals in mind. Business processes typically encompass multiple service invocations. Examples of business processes are: initiate new employee, sell products or services, and fulfill order.*
>
> *Services: Represent logical groupings of operations. For example, if we view customer profiling as a service, then, lookup customer by telephone number, list customers by name and postal code, and save data for new customer represent the associated operations.*
>
> *Operations: Transactions that represent single logical units of work (LUWs). Execution of an operation will typically cause one or more*

> *persistent data records to be read, written, or modified. SOA operations are directly comparable to object-oriented (OO) methods. They have a specific, structured interface, and return structured responses. Just as for methods, the execution of a specific operation might involve invocation of additional operations"* (Zimmermann, 2009).

In the same way, interfaces and information flows can be defined hierarchically for modeling the interaction or information exchange between business processes and services. This decomposition is illustrated with the use of a telecom's IM scenario that was used for the empirical research, is described in section 5.2. The following chapter will presents topics related to integration of IM, from various perspectives of the development of IS.

4 INFORMATION MODELING INTEGRATION AND COMMON INFORMATION MODEL

This chapter discusses aspects of IM in relation to the different layers of EI, and the use of a CIM is analyzed as an approach to facilitate EI integration capabilities. First, the integration, scoping and approaches of information models are presented as aspects to be considered by managers in making IM integration decisions. Second, the chapter presents an explanation of the importance of data model integration, issues affecting it and their consequences. Third, the chapter exposes aspects and considerations of semantic integration. Lastly, the idea of CIM is analyzed in terms of the expected benefits for IS interoperability and EI integration capabilities.

4.1 Information models integration, scoping and approaches

IM might be used as an approach to perform enterprise or IS integration. The principle is to integrate models used by different parties, for instance people in different departments or different IS, and use the result to guide the integration of business processes and applications. Once models are integrated, there will be a common model that allows the different parties to understand each other. In particular for IS integration, IM is applied for multiple purposes, for example, to streamline the organization or process interactions, to predict of changes' effect, to integrate of services and to dynamically coordinate and control processes. In addition, IM might either focus on a particular domain or span multiple domains (Petrie, 1992; Lim, 1997). In this study, the focus is to use IM for the integration of services by helping to increase the interoperability of heterogeneous used to implement business processes. In the rest of this section, the various scoping levels of IM integration are discussed from three perspectives; this is followed by a summary of approaches to IM found in literature.

On the one hand, the scope of IM integration has at least three different perspectives. First, information models can be integrated at the *syntactic* and *semantic* levels. Syntactic integration considers the structure and format of models, while semantic integration is refers to the completeness and correctness of elements, in addition to the meaning and interpretability (Izza, 2006; Mirbel 1997). Second, scoping could be considered from the viewpoints of the EA layers: *business*, *application* and *technology*. At the business layer, IM is used to nail down the requirements of a new solution, which is further refined as various business processes and services, that when implemented would realize the requirements. At the application layer, a business process can be decomposed in a flow of steps that use different IS in the enterprise portfolio. And at the technology layer, data and technology services can be designed by linking the technology elements in the enterprise infrastructure. Multi-layer IM integration can be supported with frameworks like archimate that define relationships among EA layers (Lankhorst, 2004). And third, the scope of IM could be either *strategic* or *tactical*. This perspective refers to the alignment of the IM used in projects of an organization. Tactical integration will focus on clusters of projects where the models of a cluster are more related to each other than to the models in other clusters. Strategic integration is more ambitious, it is an attempt to span the integration to the entire organization (Robinson, 2008). These scoping perspectives are important for managers to understand the implications of IM integration and to take decisions based on that, for example the definition of a governance group.

> *"If the use [of SID] is strategic, some companies, especially big companies define a central group in charge. However it is not required, companies can opt for using the SID for tactical purposes. For example, within the same company, a group can use SID as a starting point to help developing applications and other group can use it for integrating applications"* (Reilly, 2009)

On the other hand, there are different approaches for at least four aspects of IM integration. First, models might be integrated *globally* or *pair wise*. Global integration is the integration of models to a central model in a hub structure. Pair wise integration, also called point-to-point integration, requires making individual integrations between pairs of models as needed. Second, models could be integrated by using either *translators* or *adapters*, also called wrappers. Translators are independent components that transform information between two data models, while adapters are build around an interface or data to make it conformant to a standard. Potentially, two or more translators could be chained to perform indirect transformations. And a translator might additionally perform *message routing*, which consist serving as an intermediate for the messages sent between systems. Using message routing approach for translators represents the implement of a data mediation layer between the middleware and the IS. Third, in relation to the IM integration of the EA layers, one might follow a top-down approach or a bottom-up approach. Top-down will start by linking business artifacts such as solution descriptions and requirement documents to business processes workflows, and these to business services, application business process scripts, composite application services and finally application services (see figure 8). The bottom-up approach will do the opposite. There is also the possibility to do a mixed approach. Finally, models could be integrated using a mater, unified or federated models. *Master model* is when all models are derived from a single reference model, thus all models naturally match. *Unified model* refers to the use of a standard model(s) to translate the information/data represented in different models. *Federated model* is a model governed by a central agent that maintains the model and decides what changes are made to it; this is normally implemented by some means of unification (Petrie, 1992; Fredericks and Van der Weide, 2006).

Table 2. Examples of managerial decisions to make for IM integration scoping

Scope alternatives	**Example of managerial decisions**
Syntactic or semantic	What tools exist to help in doing semantic integration? A common approach is to start by gathering together key stakeholders to define a common vocabulary.
Within an EA layer or multi-layer	Is this use case specific to an EA layer?
Strategic or tactical	Should CIM be deployed in the whole company or for used for specific needs on a project-by-project basis? For small and medium enterprise (SME) the first approach might be feasible, for large companies the second approach would be more realistic.

Table 3. Examples of managerial decisions to make for IM integration approaches

Integration approaches	**Example of managerial decisions**
Global or pair wise	Should all models be integrated to a central CIM or models should be integrated point-to-point as needed?
Translators or adapters	What is more convenient, a translator or an adapter? Translators could be more flexible but the adapters could be superior in performance
Master, unified or federated model	Are the service interfaces already defined or might CIM be used to re (define) them? Is it time/cost effective to use a CIM as a unified standard for all translations?

4.2 Data models integration

A data model is a logical representation of the concepts used under a particular domain and the relationships between those concepts. A computer data model is a computer understandable data model codified in a particular format that can be used to represent, store, search and retrieve the information from a data source (Allemang and Hendler, 2008). Data models integration is one of the most complex and critical issues when dealing with application integration, data integration could lead to projects failure and have severe consequences at the strategic level (Goodhue, Wybo and Kirsch, 1992). As the number of different applications increase, integrating them becomes more complex and sometimes unmanageable. If applications are not from the same vendor, which is typical in communication service providers' networks, each application might be using a different data model. Analogically, two different data models are like two different languages, for example Spanish and Chinese. Thus in order for a Spanish person and a Chinese person to communicate, a translation is required. Similarly, for integrating data models, it is required to define translations, commonly known as *mappings*, between the elements of one data model and the other. However, mappings could become complex and prone to errors, causing a considerable time required for data model integration. In turn, this situation affects directly the costs of IS supporting the provisioning of integrated solutions. In a case of Telstra, an Australian Telecom, Hamilton (1999) found that attempts to set up an integrated technology infrastructure failed, mainly due to data integration issues. Next, three of the main challenges of data model integration are discussed in detail.

A first challenge is the *diversity* of data models used by applications. Data models have been created at different times, by different people and to satisfy specific application needs. Thus different data models present disparity of formats, naming, data standards and procedures for creating and managing data (Zhao and Siau, 2007). In addition, there are several modeling paradigms for data models associated to different types of databases. Although relational databases are the most common, there are still hierarchical databases and more recently object oriented databases have been used. These paradigms imply a completely different way of modeling the data (McBrien and Poulovassilis, 1998). In consequence, there are *conflicts or inconsistency* in the way information is represented and interpreted from data models. Solving these conflicts is part of the *semantic* integration issue, which is detailed in a separate section. In consequence, the definition of semantics increase the complexity of projects and the time to market (Kobielus, 2007). The semantic integration problem is detailed in the next section.

A second challenge is the *instability* of data models. One of the main problems is the structural stability of data, in other words, the systems and their data models evolve continuously thus become a moving target of application integration. Even if the concept of customer is somehow standard among organizations it might not be fully stable at all times, for instance new details about the customer might be added. This is in part caused by the changing business needs and the rigidity of data models embedded in databases and in the logic of IS, causing slow rates of change and other problems (Hamilton, 1999). In addition, semantic instability is affected by the trend towards larger scopes of integration. This issue is summarized in the following citation:

> *"The integration problem is morphing from very simple to very complex, and even moves from a departmental problem to an enterprise-wise problem, and, ultimately, to a trading community problem"* (Linthicum, 2003)

This is particularly a problem in the telecom industry, where technology of hardware evolves at revolutionary steps and telecoms need to support both the existing core systems and the new coming systems (O'Reilly and Tushman, 2004).

Finally, a third challenge is the lack of integral *governance* for the management of data models. The topic of governance is large, going from the ownership of models to the use of best practices and the definition of policies. Aspects related to governance were out of the scope of this study. To mention a few, two main aspects of governance are the socio-political and the technical aspect. The socio-political aspect includes the stakeholder's requirements, expectations and practices. The technical aspect includes procedures, formats, platforms, standards and tools. In most cases, companies have not defined o are still in the process of defining governance groups and best practices. Moreover, the lack of regulatory agencies influences the creation of data models about the same domain, which might not only differ from but also contradict the other. As consequence, the data integration issue is worsening because if there is not common approach for modeling, the diversity and inconsistency is expected to be greater.

In the following section, the semantic aspect of data integration is introduced and the main challenges of semantic integration are outlined.

4.3 Semantic integration of models

> *"Semantic integration is the process that makes possible to [produce] a global [model] obtained from several [models], each of them describing the same reality in different views, in order to obtain the fullest view"* (Mirbel, 2007)

In order to integrate models, modelers have to consider several aspects of equivalence and consistency between the elements of the models involved. In finding such equivalence, it is necessary to take into account fundamental aspects of modeling for each of the models involved. These aspects include the *components* of models, the *representation* of concepts and reality by the model, the *interpretation* of the model by different users and the *purpose* of the model (Bergamaschi, 2001; McBrien and Poulovassilis, 1998; Mirbel, 2007). The components and representation of a model are somehow tangible and simpler aspects than the other interpretation and purpose aspects. Standards such as UML have been developed to facilitate the unification of model representations and computer tools are available, such as the one presented in appendix 3, for supporting the integration of the different components of models. Conversely, the interpretation and purpose of models are intangible and more complex tasks. This is due to the considerable work with humans and the lack supporting tools, thus more development is required. In this section, the equivalence is studied from the perspective of these four semantic aspects.

Semantic integration of models, from the *components* perspective, might follow the typical divide-and-conquer approach. This approach consist of dividing the original problem into smaller and manageable parts, finding the solution to these smaller parts and combining the smaller solutions into a solution to the original problem. This approach follows a top-down integration, however bottom-up approach, consisting of building islands of model integrations are initially formed and combined into larger integrated model subsets, until have a single complete model. Semantic integration of information models will normally consist of the definition of transformations or *mappings between two models.* These mappings represent translations of the elements of a source model to the equivalent elements of a target model. There is a possibility that some elements of a model do not have a corresponding element in the other model, thus only a partial mapping is possible.

Regarding model *representation*, mappings are based on equivalence of both the name and the concept represented by elements. Equivalence on name is based on dictionaries and other linguistic aspects related to the naming of the elements. Equivalence on concepts represented by the element refers to the level of detail and the level of abstraction of the elements. Level of detail indicates how much information of a concept is represented in the model element. For example, to represent customer information it might not be relevant to include their favorite color, but it is necessary to include their name. Level of abstraction indicates how general or broad a concept is. For example, while some models could represent a highly general concept like ‘resource’, other models could represent more specific concepts like ‘Physical Port’. Both the name and concept representation are linked to the other two semantic integration aspects, interpretation and purpose. In short, a model should use adequate names and represent the right concepts for the level and purpose of the users of the models.

An additional goal of semantic integration is that the *interpretation* of models should be equivalent. By applying mappings, a person or a computer should be able to take the data represented in either model and interpret it using the other model. However, since persons have different backgrounds and think differently, it is natural for different persons to have different interpretations of the same model elements. Semantic integration efforts might include meetings of different stakeholders to jointly agree on the definition of concepts such as ‘customer’ or ‘service’. This could be a highly time consuming task and there is no clear approach on ensuring a complete equivalence in the interpretation. As an alternative, a person could be designated to create a model for defining concepts and validating it with a group of stakeholders. Although that might be more efficient, semantic differences are still likely to be present, causing conflicts and potentially time-consuming communication for reducing such differences. The emphasis is then on how to reduce such difference in the most efficient way.

Furthermore, while doing semantic integration of models, it is important to step back and remember the main *purpose* of models in general. Models are supposed to help people to communicate, thus it is important to model the information in a way that the different people in an organization are able to understand. Models are also meant to explain the realty, by means of abstraction of details to an adequate level that is suitable for the users of the models, thus modelers should try to identify the important aspects that are relevant for a particular domains and stakeholders. Models are meant to serve as a blueprint and mediator, as a way to relate concepts seen from multiple viewpoints, thus it is critical to describe the information as accurately as possible in order to eliminate redundancies and inconsistencies between information models (Allemang and Hendler, 2008).

Finally, semantic integration in the context of telecoms has some particularities. A number of telecom standards also play an important role in the semantic integration of models; this because naming conflicts might be present for standard of the same technology And given both the number and complexity of the different types of network equipment, the models to describe the information are large. To provide an idea of this situation:

> *"The order of magnitude for the number of information elements that are handled by the information models in the company is around 1 million"* (NSN technology & architecture manager, 2009)

4.4 A common information model approach

Two main approaches could be taken for data integration, doing point-to-point data models integration and using CIM (Calvanese, 1998). The first one refers to the ad-hoc integration of data model as needed. In this case, for every new IS to integrate, a data integration is made for every communication link required between two IS. At most, if all IS have different data models then one transformation is required between each IS data model and another. Conversely, when using CIM each IS data model is integrated to the CIM, and indirect transformations might be made to enable the integration between two or more different data models.

CIM could be applied in the design and integration of business and application level models. At the business layer, it might be used to define as a common vocabulary and conceptual model that unifies the representation of information for the domains involved in the processes of an organization. At the application layer, CIM could be used to integrate data models to a unified model. In this way, CIM could provide a common language that could reduce the communication gap between teams and organizational levels. Moreover, communication inconsistencies between internal and external stakeholders could be reduced. Finally, using CIM could produce benefits such as increased understanding and learning, reduced integration time and costs and facility to reconfigure assets in the development of integrated solutions. Next, these benefits are expanded and abstracted in a set of prepositions and a conceptual framework. This conceptual framework was the base for the action research carried out at NSN, which is reported in chapter 5.

First, using CIM could facilitate the development of a single unified view of the data in a company, thus aligning the business understanding and business coordination. Findings of Francalanci and Morabito (2008) suggest that data integration contributes to the business performance of companies. They use the concept of *absorptive capacity*, as the ability of an organization to learn by identifying, assimilating and exploiting knowledge, and identify it as a medium for competitive advantage. For example, by having consistent information regarding the customer, stakeholders could increase their common understanding and achieve agreements on data models more easily than not having CIM. Reilly (2009) has observed this situation in exercises that have been made as part of SID courses. In one exercise, the same set of requirements was provided to two different teams that should use the SID as CIM, and the results the two teams produced were similar in both structure and names because they followed the same guidelines and used the same CIM. If this holds true in integration assignments, CIM could contribute to reduce integration efforts and at the same time increase the consistency/quality of customer services.

> *Proposition 1. Companies should increase their employees' coordination and understanding of the business by using a unified CIM. This is because using the CIM helps reducing communication gaps by providing a common vocabulary to unify the interpretation of data elements.*

Second, during application integration efforts, a considerable time is normally spent in learning the data models involved. By using CIM, companies save the time required to learn proprietary data models coming from multiple vendors and external software vendors (Reilly, 2009). And using CIM might facilitate the eventual reduction of application integration time and costs. For example, once two models are already integrated to the CIM, integrating a new application using the two data models would require simply a mapping between the new data model and the CIM. As opposed to requiring two mappings if no CIM was set in place. Studies have performed mathematical models to prove such statement. Weng (2007) concluded the minimum number of mappings for semantic interoperability among information sources is achieved when using the approach of a shared conceptual reference model.

> *Proposition 2. Companies that use CIM for integrating new applications could eventually reduce costs and time by reusing semantic integration assets: The mappings between data models and the CIM and the*

knowledge of the CIM are reused.

Third, using an industry standard as basis for CIM should facilitate the interoperability of IS. As the CIM adoption by industry participants increases, a common vocabulary would be shared among them, thus reducing the semantic differences present in their IS. This in turn should reduce the time and cost of integrating IS. In addition, since open standards are not vendor specific, the interfaces that are defined based on the standard will not be dependent on implementation or vendor details, thus reducing the coupling between IS of the same platform or vendor (Erl, 2007). Wilmes (2008) supports the idea that

> *Proposition 3. CIM based on open industry standards with sufficient adoption from the community would favor interoperability of IS by reducing dependencies on specific platforms and vendors and by standardizing the representation of data elements, thus eventually reducing semantic differences.*

Fourth, it is believed that CIM might also contribute to the stability of data models and facilitate the deployment of governance measures. CIM should provide a well structured model, provided that it is defined collaboratively, it should consider the different aspects of a business and arrange them in a way that separates the domains or concerns of the model. This in turn should facilitate the stability of the overall model, because changes could be made to specific domains without affecting others. And the control of the model should be facilitated by dividing the model in manageable subsets (Goodhue, Wybo and Kirsch, 1992).

> *Proposition 4. Companies that use CIM that covers the main concepts used in their existing processes, and in addition that model favors the interoperability principles, can develop capabilities of extensibility, reusability and flexibility to integrate new applications. This in turn increases the company's ability to reconfigure assets to create new integrated solutions.*

Finally, the previous propositions served as base for the building of a conceptual framework is presented in figure 9. In this framework the usage of CIM is defined as a factor that facilitates IS interoperability principles, which in turn reduce the integration tax. An additional relationship is defined from the usage of SID to the amelioration of the semantic integration issues, thus reducing the effect of these issues causing IS integration tax. With this model, the first phase of the research process is concluded. The next section describes the second phase of the research, where action research and induction thinking was used to validate the initial conceptual framework.

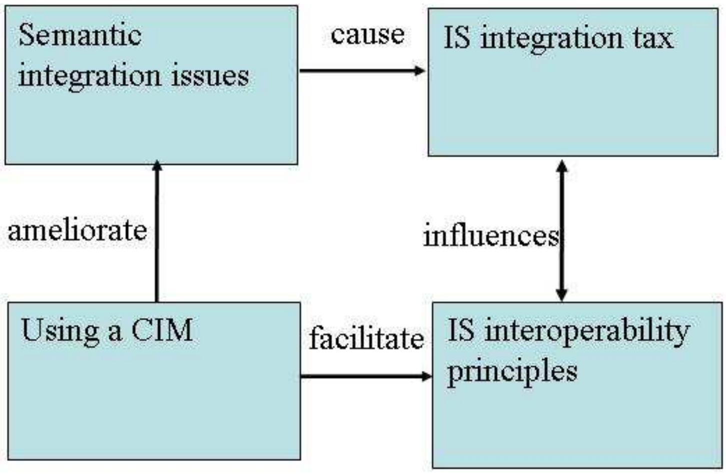

Figure 9. Conceptual framework of the benefits of using CIM.

5 COMMON INFORMATION MODEL IN PRACTICE

This chapter covers the practical aspects of using CIM for application integration. It starts by outlining two main aspects to consider in the selection of CIM and the reasons for using the TMF SID model as a basis for CIM. Next, the NSN case under which action research was carried is described. Finally, the report of action research is presented for each of the three iterations performed in the last three sections of the chapter.

5.1 Choosing a common information model

A company has two alternatives for using CIM, to develop its own model or to take an industry reference model as base and extend it for its needs. In the first case, there are two sub-alternatives, a model or group of models might be taken as base or a new CIM could be built from scratch. Developing the company model by its own has the advantage of being more adapted to the company situation, however that might be also a disadvantage when dealing with external 3rd party systems. In addition, developing a model by its own require normally larger efforts than reusing an industry reference model. In the second case, industry reference models are made with the purpose of providing a common base to facilitate the interoperation between enterprises and their IS for a given domain. This because they provide a more complete starting model, as opposed to start from scratch. And by using an industry standard, it provides the advantage that other participants are using the same model, facilitating the integration with their systems. Although work is required to learn the industry model and extend it to the company's requirements, it is agreed by authors and practitioners that using a reference model should eventually take less implementation and integration effort (Gulledge, 2008; Scheer and Hars, 1992) Finally, quality and adoption are crucial factors for the success of industry reference models (Reilly, 2008).

Some industry reference models have emerged in different industries, as a collaborative effort of companies to deal with the data integration issues. The supply chain operational reference (SCOR) model in manufacturing, the 3rd party logistics (3PL) model for logistics and the shared information/data (SID) model in telecoms (Gulledge, 2008; Tele Management Forum, 2009). Particularly, the TMF SID model was developed to meet the requirements of the telecom industry, while at the same time keep the standard compliance and flexibility (Reilly, 2008).

> *"TM Forum is the world's leading industry association focused on improving business effectiveness for service providers and their suppliers. Serving the information, communications and entertainment*

> *industries, the Forum provides practical solutions, guidance and leadership to transform the way that digital services are created, delivered and charged. Members include the world's largest service providers, cable and network operators, software suppliers, equipment suppliers and systems integrators."* (Tele Management Forum, 2009)

The TMF SID model is the most widely used telecom industry CIM. The TMF SID presents the main features described previously for CIM. It is documented, structure, somehow complete, supported, adopted and accepted as standard. SID's documentation includes a complete definition of the domains, entities and relationships. The current version of SID included more than one thousand entities. SID's hierarchical structure favors flexibility and modularity (see appendix 2). And, the SID model covers domains from end-to-end processes from the various perspectives of the IS life cycle: business, systems and implementation. Such coverage allows telecoms working with end-to-end scenarios. In addition the SID model has been well adopted and supported. SID has been adopted widely and it has been accepted as an International Telecommunications Union's (ITU) recommendation. Lastly, the SID model is well supported by TMF and a large number of industry participants.

> *"[TMF] provides a wide range of information and support to help its members reduce the costs and risks associated with creating and delivering profitable services. These include industry research and benchmarks, technology roadmaps, best practice guidebooks, software standards and interfaces, as well as certified training, conferences and publications. The Forum also provides its member community with extensive marketing and networking opportunities, enabling business with new customers and partners."* (TeleManagement Forum, 2009).

Finally, the SID model has been successfully used in case studies such as British Telecom and Tektronix. First, one case is British Telecom's One IT project (Potter and Brady, 2005), in which the enhanced telecommunications operations map (eTOM) and SID were used to support transformation initiatives and system migration. The project turned to be a success from their perspective of British Telecom.

> *"The project delivers against both of Al-Noor Ramji's stated measures of success for One IT: reducing customer perceived cycle time and increasing the percentage of getting things right the first time. 'Paying attention to these metrics automatically leads to reduced cycle costs and, ultimately, increased customer satisfaction,' says Ramji"* (McCue, 2005).

In addition, a successful case is from Tektronix, a provider of measurement equipment that also delivers network management and assurance solutions. They have leveraged the SID model with the intention to unify the management of their software products and reduce the *integration tax* spent in application interoperation efforts. They extended the SID model to develop a detailed service assurance model. As a result, they were able to reduce costs of integrating other applications and the time to market for new service/resource key performance indicators (TeleManagement Forum, 2009). These two cases are only two of a number of case studies of successful implementation of the SID model, for which documentation is available at TMF for its members. For the practical exercise of this research, the SID model was chosen as a basis of CIM. The following sections describe the case study of this empirical research and results obtained.

5.2 Network fulfillment use case at Nokia Siemens Networks

NSN offers telecommunications infrastructure and services, in addition to operations and business support solutions. These solutions are based on products such as network equipment, software, maintenance and consulting services. On the one hand, the offerings include the management of operator software, the provisioning of NSN software and the integration of both. On the other hand, NSN also offers maintenance and consulting services, for which NSN units work closely with the operator teams to adjust the business, operations and infrastructure according to the changing needs of the operators. (Nokia Siemens Networks, 2009). Furthermore, as part of the solutions related to the network operators' software, NSN provides middleware software to simplify the integration and development of OSS. Research and development is carried out at NSN to investigate efficient methodologies and alternatives for the integration of OSS. Within these research efforts are the study of SOA and NGOSS frameworks in order to facilitate the integration of services and systems from network operators.

As part of the research efforts at NSN, a practical evaluation of data integration alternatives of using CIM was performed in the context of SOA. There was already available within NSN a demo implementation for a use case, which was utilized for this practical exercise. The use case demo consisted of a network fulfillment solution that supported the provision of network access and related services. Such solution includes a pre-integrated set of products, services and off-the-shelf modules, including network provisioning, inventory and order management products. Figure 10 presents a simplified representation of high-level architecture, corresponding to a customer order for new network access connection use case. An instance of this architecture could be the business scenario in which a customer asks a network operator's customer service agent for broadband internet access for home usage. To satisfy such customer demand, an agent might execute different actions in a customer relationship management (CRM) system. The agent might require querying the customer data stored in the network for information regarding the customer account and probably existing services and pending bills. Another business process could be activating the network provisioning service and issuing a bill for that particular service to the customer. These scenarios are supported products like the network fulfillment solution, which is exposed by means of a set of front-end business processes such as query customer data, activate service and issue bill.

In turn, the front-end business processes are realized by business services that are implemented by application business processes. The activation layer at figure 10 shows the simplified example of an activation business service, which was implemented by composing of steps such as checking the availability of a network service, activating the network provisioning service and logging the event. This particular example of business process would compose application services from three different systems. Notice that in this case, both the business and application services are exposed using web services. The difference is in the business services interfaces define high-level operations, meaningful for the business, and use a somehow abstract data model. Conversely, the application service interfaces define low-level operations, meaning more specific operations, which are system aware and use a more detailed data model.

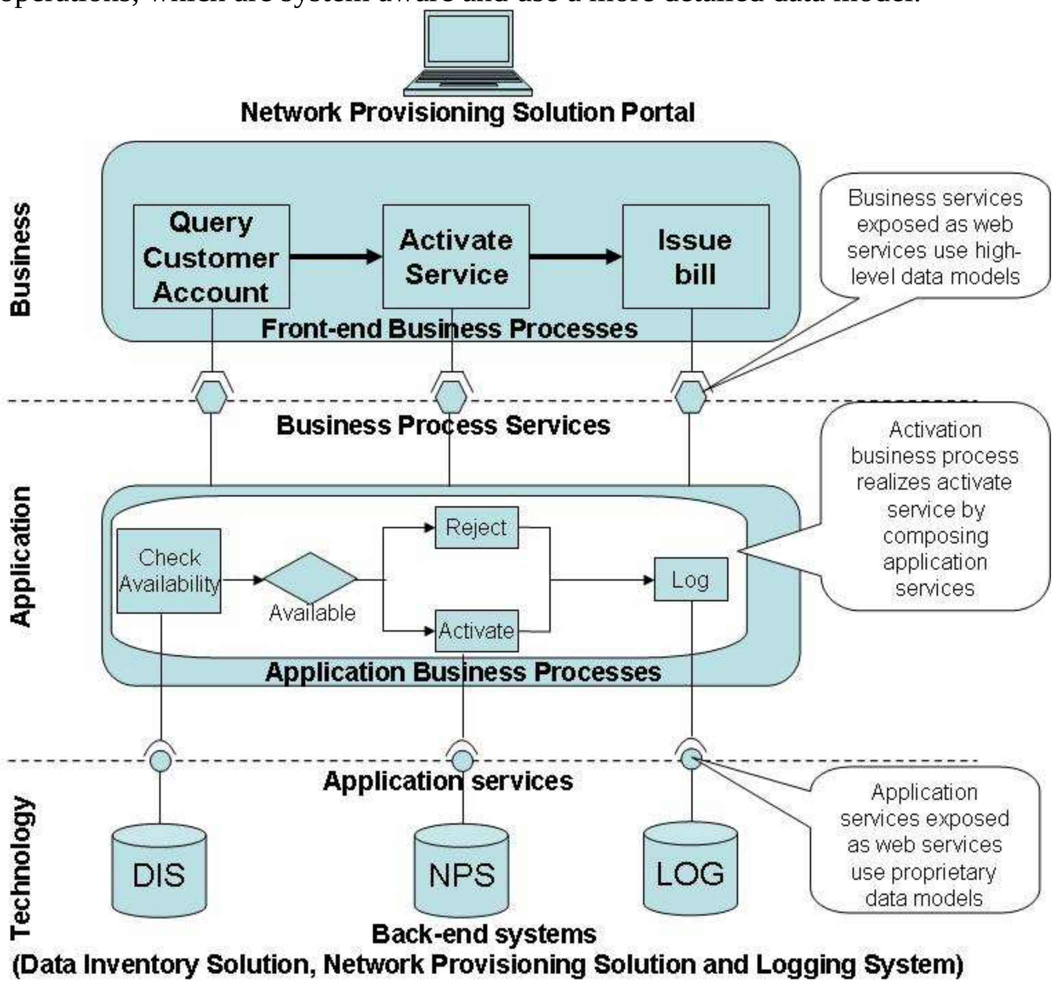

Figure 10. SOA architecture for the network provisioning use case

In terms of the application layer, three types of elements that participate in this use case are a web portal, an ESB and several OSS (see figure 10), such as a data inventory solution (DIS), a network provisioning solution (NPS) and a logging system (LOG), among other OSS. However, only the DIS was used for the purpose of this research mainly because of time limitations and the possibility to gather relevant results by simply including the DIS. Next, the involved systems and the flow of information are described. The portal plays the role of a business support system (BSS) and it is the first that participates in the use case from the business process viewpoint. The portal handles CRM operations relating to end-user orders coming from a browser, prepares and sends request messages to the ESB. The ESB acts as a traffic intermediary for all the messages sent between the applications. Particularly, the ESB platform used in this application architecture is also the host for the business process engine, which executes business workflows using business process execution language (BPEL). The ESB intercepts request messages from the portal, invokes the corresponding workflow or BPEL and, sends request messages to the DIS based on the process workflow. Similarly, based on the process logic, the ESB sends requests messages to the other two systems, NPS and LOG.

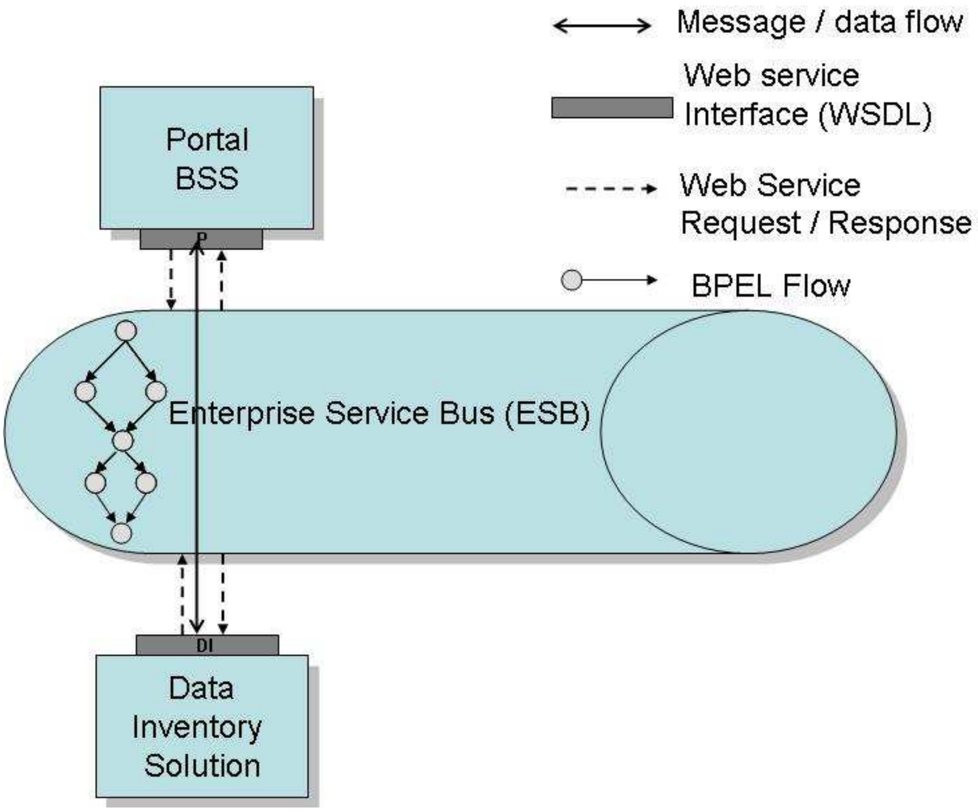

Figure 11. Network fulfillment use case architecture

To conclude with the description of the case study, the NSN specific context in which the empirical part of this research was done is described next in terms of models, tools and standards. First, the TMF SID model was chosen as basis for CIM due to its qualities, described in section 5.1. The DataXtend Semantic Integrator (DXSI) tool, from Progress Software, was selected for integrating data models. A detailed description of the features of this tool is found in appendix 3. And third, web services and the industry standards WSDL and XSD had been previously chosen for the definition of SOA interfaces and data model schemas respectively. Moreover, a number of preparation activities were performed before the action research started. First, a literature review of semantic integration served to identify existing methodologies to deal with semantic heterogeneities. Second, a period of three weeks was allowed for the thesis worker to learn the DXSI tool. And third, an overall review of the TMF SID documentation was performed. This concludes the description of the contextual aspects of the practical research. In the rest of this section an overview of the tasks performed as part of the action research, and the structure followed for reporting the action research results are presented.

Three iterations were performed by the thesis worker, in collaboration with the thesis mentor and the research team. During the first iteration, an attempt was made to use the SID model as CIM to perform transformations between the data models of portal and DIS. This was implemented by creating a *transformation application* that served a translator. In the second iteration, the SID model was extended to be used as CIM and the transformation application was used as a *mediation layer* or intermediate agent between the ESB and the DIS. In this iteration, the transformation application interface changed to resemble the one used by portal. Finally, in the third iteration, the message routing/transformation application was modified to expose CIM based interface, and an additional transformation application between portal and ESB was simulated. In this iteration, the ESB was dealing only with the CIM, and not with any other data model, which is known as *CIM realization*.

The next three sections comprise the report of the three iterations of action research. The results are organized following the sequence of the five action research steps: diagnosis, activity planning, activity taking, evaluating and specifying learning. First, in the diagnosing step, the data integration issues to be solved are presented. Next, the action planning step describes the alternatives that were considered and the decisions made. This step also includes the transition made from inductive to deductive thinking, through conjectures. Then, the details on the action taking step are presented. Following that are the results of evaluating against the expectations described in the action planning step. Finally, in the specifying learning section, the conjectures are refuted or accepted based on the results of the action taking.

5.3 Iteration 1: Data transformation application

The *diagnosis* step of the first iteration consisted of the analysis of the network provisioning solution scenario. A solution description document was provided by the research team and a number of meetings were held for transferring knowledge, from the research team to the thesis worker, about the use case architecture and current situation of the data integration action. Two main data integration issues were the learning curve of the data models involved and the large number of mappings required in the integrations. Each support system participating in this use case has its own proprietary data model, the portal data model was developed by the research team and the DIS data model was provided by a different team. Although a communication link existed, the learning curve on this data model was restricted by its documentation. Particularly, the version of the DIS and service interface that was available for this exercise, was initially developed for the purposes research and development, and it is not in commercial use, which is the reason for the documentation being neither formal nor complete. In addition, the manual mappings performed in ESB were identified as a time consuming task. As part of the business process flows, a number of steps required invoquing application services, each of which included a different data model, thus mapping between two or more data models where required for these steps.

Following this diagnosis, *action planning* included a review of semantic integration and application architectural approaches. The TMF's approach for semantic integration using the SID was taken as baseline (Reilly, 2007). This is a five-step general procedure for SID implementation, described by Reilly and Wilmes (2008). A number of scientific articles were found, containing different taxonomies and procedures for semantic integration of data. Examples of these taxonomies are the Weng (2007) taxonomy of major semantic mismatches for harmonization and the Rahm and Bernstein (2001) taxonomy of semantic matching approaches. In addition, notable work was found regarding semantic integration procedures by Mirbel (2007). As a result, the Reilly and Wilmes' (2008) procedure was adapted as described in appendix 4, and followed in this study.

In addition, the architectural alternatives for application data integration supported by the DXSI tool were discussed (see appendix 3). Specifically, a decision was made to choose transformation applications, as opposed to adapters (wrappers) around the applications to be integrated. The main reason was that transformation applications were found more flexible and simple to implement than adapters because programming an interface was not required. As a result, it was decided to create a transformation-only application, as shown in figure 12. With this approach, for every request message from portal to DIS, a call was first made from the ESB to the transformation application to convert the message from the portal data model to the DIS data model. Then the ESB would use the transformed request message to call DIS.

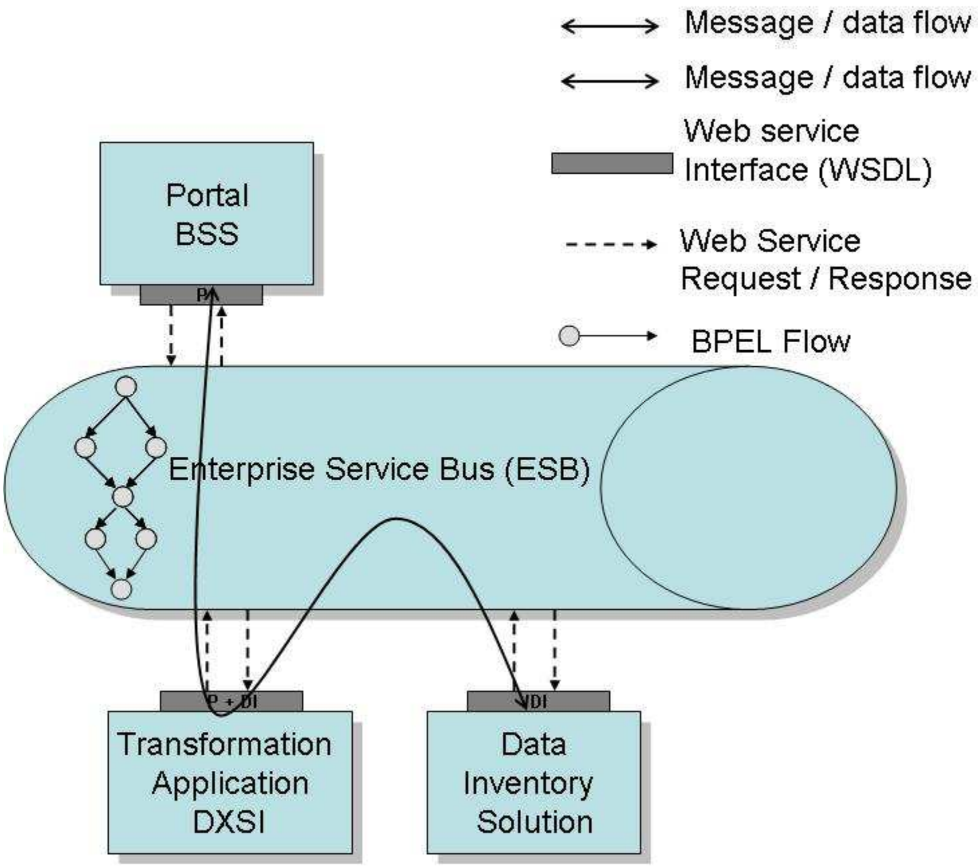

Figure 12. Transformation only application scenario

Next, given the characteristics of the SID (see appendix 2) and the architecture alternative used in this iteration, the following conjectures were derived from proposition 1:

> *Conjecture 1. Using the SID as CIM, it should be possible to relate the concepts described in the NSN network fulfillment business solution description document and to find a link to the technical data model specification (XSD files).*
>
> *Conjecture 2. Once familiar with the SID, it should be possible to map the elements in the portal and inventory solution data model to the SID.*

During the *action taking* phase, the procedure described in appendix 4 was followed. First, a high level analysis of the TMF SID model structure was required to get familiar with it. Then, the scoping of the integration was defined to consist of 14 abstract business entities (ABE) in 4 domains (see appendix 2). Second, the portal and the DIS data models were identified to be adequate for the purposes of this experiment, as they represent a complete end-to-end scenario. The NPS data model was not included because its level of detail does not correspond to the one in the other two data models, portal and DIS. Third, a detail analysis of the portal and DIS models was performed to understand the concepts of the different data models. The analysis was made using a combination of sources, including the business solution description document, the schema files (XSD) and the same research team. For the portal data model, containing simply 3 elements, the entire model was considered and for the DIS data model a subset of 18 elements out of 88 were considered. This subset corresponds to two of the main scenarios of the use case. In addition, an in-depth study of the SID documentation related to the selected ABE. For this, both the SID documentation and the DXSI tool were used to learn the concepts and their relationships.

Furthermore, using the procedure defined in the action planning, the portal and DIS data model entities and attributes were mapped to the SID model. First, the thesis worker received help from the thesis mentor to relate the entities to relate elements in the data models to the SID. Then design pattern's matching was used to link subsets of elements from the data models to the SID. For a definition of research pattern, refer to appendix 4. This was documented informally using an office document, enlisting the design pattern along with the related data elements. Second, the attributes of the data models were matched to SID attributes using a two-column table. Third, interface mappings were performed inside the DXSI tool, using the knowledge developed during the analysis. The three models (portal, SID and DIS) were imported into the tool and manual mappings were performed. First the entities and second the attributes of the data models were mapped to the SID using the graphical editor in DXSI. Finally, the same DXSI tool test environment was used to test the mappings with a sample request and test data.

In *evaluating* of the action taking, it was possible to relate the majority of concepts from the business solution description document to the SID model, and to navigate from any element to another using the DXSI tool. In addition, it was possible to relate business and technical elements coming from the solution description and the XSD files, respectively. As a result of the data element and design pattern mapping, the thesis worker's understanding of the portal and DIS concepts increased. In general, using the SID was useful as the base for CIM, as it serves as a bridge between concepts.

Conversely, it was found that even for small and medium data models with similar concepts to the ones in SID, there could be a considerable amount of semantic mismatches or conflicts of different type. First, few linguistic mismatches were present in the form of synonyms, for example 'Technical Service' and 'Resource Facing Service'. Second, although most elements of the portal model were already in the SID, some abstract elements in SID required to be extended in concrete elements in order to make them usable, for instance there was not a concrete entity for a 'Virtual Network'. Third, some elements that were present in the SID were missing attributes required for this network provisioning use case. As a result, extensions[1] to the SID model were required in order to map approximately 38 % of the element attributes in the DIS model. And for approximately one third of those attributes no corresponding data element was identified in the SID model. Since additional documentation or semantics was not included in the DIS data model, these element attributes were included in the CIM within any feasible data element. In this regard, it is important to notice that the SID model has been and will continue evolving towards a more complete industry reference model (Reilly, 2009), this with the help of TMF and its participants. Finally, the structure of the SID, which uses consistent design patterns, was perceived to be flexible for extending the SID with NSN elements and attributes for this use case.

Finally, *lessons learned* of this iteration include advantages and disadvantages of using the SID and of using a transformation-only application. First, it was possible to navigate the SID elements to relate business and technology/systems concepts. For this reason, conjecture 1 was retained as valid. Second, it was noticed that a large percentage of the DIS model attributes required extensions and some of the mappings required more domain knowledge. Thus mapping data models to the SID should be taken as an iterative process, it is not necessary to map the complete data model to elements in the SID the first time. For this reason, conjecture 2 should be modified to include a condition that the person doing the mappings should have sufficient command of the data models being mapped. Third, the SID model documentation, structure, design patterns and guidelines were perceived as useful. Apparently it is not common to find such benefits in a data model. It would be interesting to compare those aspects of SID with other models. Lastly, from the implementation perspective, the transformation only architecture had advantages and disadvantages. The main advantage is the separation of concerns, this because the transformation application generated by DXSI has a defined task of transforming data models. However, other integration tasks are still handled in the ESB, with the disadvantage of dealing with large data model, compound of both the source (portal) and target (DIS) data models. One last lesson is that not all models are at the same semantic level; this was the case of the NPS data model. The semantics of the information represented by the NPS data model is not represented as part of the model but it is present in the underlying data instances.

5.4 Iteration 2: Data mediation layer

In the *diagnosing* step of the second iteration, a number of problems were identified

[1] An extension is an adaptation to a model by adding and sometimes modifying a data model to meet the requirements of an application or solution.

regarding the integration interfaces and the transformation application generated with DXSI. A main problem was that the DXSI interfaces were highly complex. Since the interface data model was compound of both the portal and DIS data, its size and structure was not easy to visualize. Furthermore, the mappings done in DXSI to generate transformation applications were perceived not only unhelpful but also as double work. The reason is because data model mappings have to be defined and maintained not only in the ESB but also in the DXSI. Moreover, in the ESB an extra data model mapping was required for the new transformation application interface. In addition, this new interface was harder to map than the DIS interface, because it was compound of both the source (portal) and target (DIS) data models. Finally, in the first iteration the extensions to the SID model were made using DXSI, however the need for a modeling tool that could improve the visualization of the model was identified.

Regarding the use of the SID as CIM other issues were identified. Even with a delimited SID scope, adapting the SID for use case specific information needs is a time consuming task. Most of the integration time was for learning the SID and making decisions regarding the best way to extend the SID. This time was perceived to be greater than the one it would take for doing a point-to-point integration for this particular use case. One of the reasons was that time was required to learn a third model, the SID model, however this time is expected to be reduced in further integrations. Additionally, in terms of extensions to the SID, it was made clear the need for guidelines for the naming and structure of the added elements and the data format and version control of the extended model.

In the *action planning* the rational software modeler (RSM) tool was chosen for making extensions to the SID model. This step included also a change in the integration interfaces and a different architectural alternative. As opposed to merging the portal and DIS data models, a simpler data model based on the portal data model was used for the integration interface of the DXSI transformation application. This was combined with a change in the architectural approach from a transformation only to a message forwarding transformation application, depicted in Figure 13. With this new approach, the DXSI transformation application had two responsibilities, besides transforming data models; it was in charge of forwarding the messages to the DIS. And as opposed to the transformation-only application, with the message forwarding approach the ESB did not have to deal with the DIS data model.

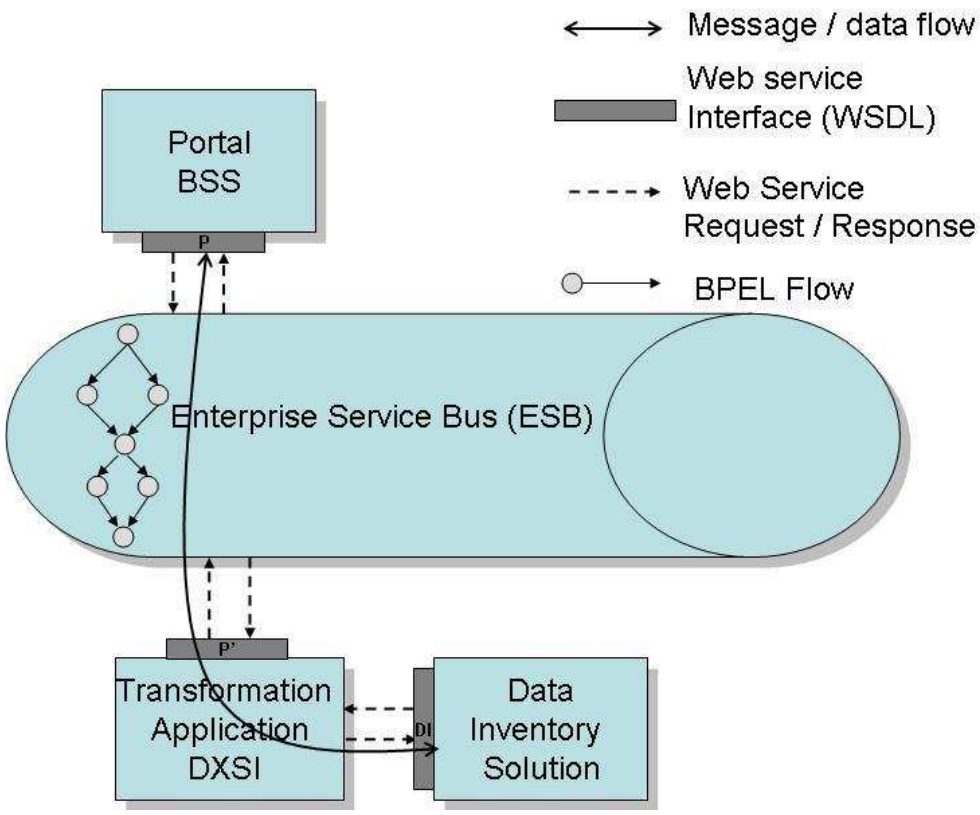

Figure 13. Data mediation layer scenario

Next, given the characteristics of the SID (see appendix 2) and the architecture alternative used in this iteration, the following conjectures were derived from proposition 2:

> *Conjecture 3. After making a first set of extensions to the SID concepts for a particular use case, further extension to SID for that use case should take less time than it took making the first extensions.*
>
> *Conjecture 4. The effort required in to integrate the participant applications in the ESB should be perceivably reduced by using a message forwarding transformation application compared to point-to-point integration.*

During the *action taking* step, additional adjustments to the extended SID were made using the RSM tool. This was done to take into account the suggestions made by the thesis mentor and the guidelines from Reilly (2007). The adjustments took one day, as opposed to three days spent for the extensions in the first iteration. These adjustments increased the level of detail in the extended SID, with the intention to improve the organization of the information coming from DIS. The RSM tool was used to model the extensions, the resulting model was exported from RSM and imported into DXSI. The mappings in DXSI for the adjusted model did not represent a significant effort in terms of semantic matching. The mappings from the first iterations were reused and the new mappings of this iteration were simply added to those. For the portal interface, the mappings for the new transformation interface were recreated but they were highly similar to the ones in the previous transformation interface. Additional configuration and logic was necessary inside the DXSI tool to direct the messages coming from the ESB to the corresponding operation in the DIS interface. Finally, the new interface provided by the transformation application did not represent a reduction of effort, from the application implementation perspective, because mappings were still needed inside the ESB to map the new transformation application interface. No reduction of effort was perceived by the person making the applications integration.

Finally, a *lesson learned* from this second iteration was that some time savings were perceived in terms of the SID model learning curve. Moreover, a navigability test of low level elements was performed to ensure that, starting from an element in the extended SID, it was possible to go to any other element by following the relationships of intermediate elements. For these reasons, conjecture 3 was kept as valid. However, even if the interfaces to be integrated are similar, in names and structure, if they are not the same, manual mappings have to be done in ESB. There was not a significant benefit perceived from the new architectural approach, thus conjecture 4 was rejected. Finally, the need for governance was identified as helpful to reduce the semantic integration effort, mainly guideless and standards for performing SID extensions.

5.5 Iteration 3: Common information model realization

In the third iteration, it was *diagnosed* that the command of the SID model and the involved data models was greater. All attributes of the DIS data model were mapped to a corresponding concept in the extended SID model. However, the main issue was that integration interfaces were not CIM compliant. Compliance to CIM is desirable because CIM based interfaces would be more reusable once CIM is increasingly adopted. In addition, the integration effort in the ESB was not improved by the introduction of the message forwarding transformation application.

During the *action planning* step, it was decided to try a SID compliant CIM for integrating the systems inside the ESB (see figure 14). In this new approach, CIM would be used inside the ESB. That is the case of the extended SID interface, marked in the figure 14 as ‘NSN x SID’. But, since the CIM model is large, a manageable subset would be extracted, containing simply the required elements for this use case. Two DXSI transformation applications would be involved: one for transforming data models and forwarding messages between the portal and the ESB, and a second one between the ESB and the DIS. The integration interfaces would be redefined to use the CIM as data model.

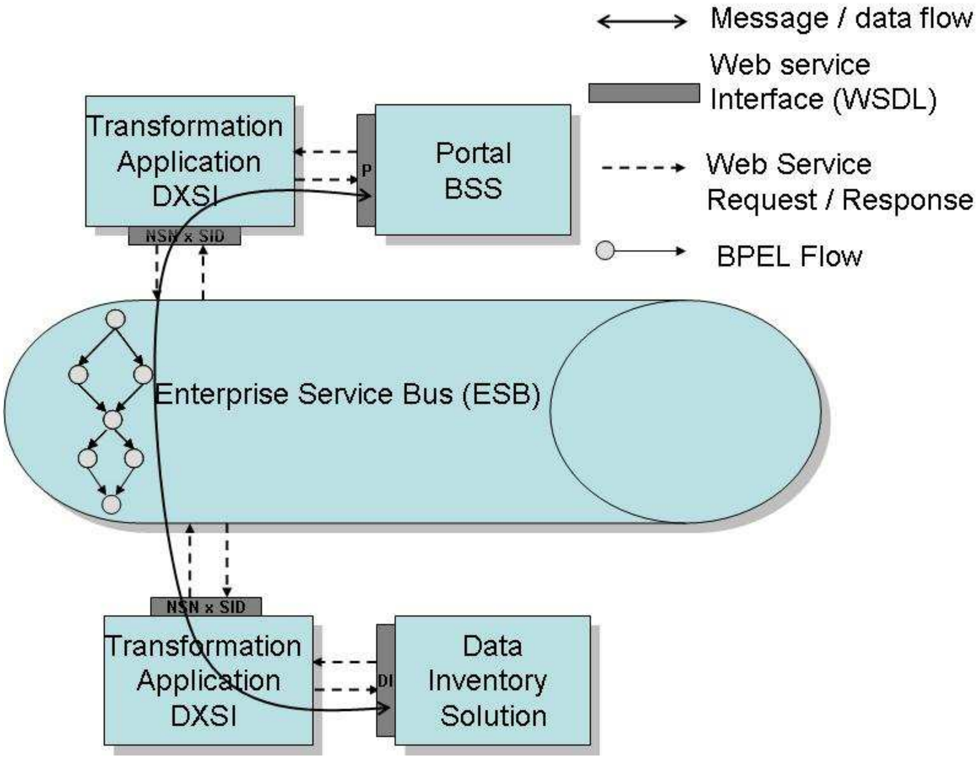

Figure 14. CIM realization scenario

Next, given the characteristics of the SID (see appendix 2) and the architecture alternative used in this iteration, the following conjectures were derived from proposition 3:

> *Conjecture 5. Using a SID compliant integration interfaces will reduce the number of mappings needed in the ESB.*
>
> *Conjecture 6. Using SID compliant integration interfaces will improve and unify the understanding of the use case data.*

During the *action taking* step, a manageable subset of the CIM model was extracted to be used for the integration interfaces. Two interfaces were required, one for each of the transformation applications. The CIM subset model contained approximately 90 elements, from which most of them were SID abstract elements, and some concrete NSN elements that inherited from abstract elements for the purpose of instantiation. During the integration there were implementation issues related to the management of abstract elements. Apparently although standards such as WSDL and XSD are well adopted, there are few aspects that are not handled equally by all development platforms. The web services interoperability (WS-I) organization (2009) has worked in the development of a set of recommendations for WSDL and XSD, in an attempt to manage such inconsistencies, however interoperability problems might still be present among development platforms as of now. As a workaround, a modification to the CIM was made to allow mapping the two transformation interfaces in the ESB. Finally, for the specific tool chain used in this exercise, there were format conversion issues. The SID model was extracted from the TMF online resources as UML format and imported into RSM, extended as XSD format and imported into DXSI. As a result, syntactic and semantic errors were present.

As a result of this third iteration, the effort in the ESB was perceivably reduced. The number of mappings was reduced by using CIM, the main reason being that mappings are done at a high level, i.e. entity and data model level as opposed to attribute level. For this reason, conjecture 5 was kept as valid. However, the integration interface became unnecessarily large and difficult to read. A large number of optional elements were present in the integration interface; most of them were not needed for in this particular use case. Additionally, from the ESB perspective, it was difficult to understand and to work with the new schema, it was too large and the element names did not have the adequate level of detail, in other words they were abstract elements that were not meaningful for this specific use case. This situation represents an issue regarding the separation of tasks among modeling roles, since the interface does not defines the actual data that is needed for a particular service, the system integrator (ESB) depends on the data integrator (DXSI) to perform actual data mapping for the ESB. For these reasons, conjecture 6 had to be complemented and retested to include an adequate level of detail and sufficient command of the model elements involved in the use case.

With this third iteration, the action research phase of the research was completed. A number of lessons learned were obtained from the practical exercise. The next chapter presents a summary of the results of these three iterations and analyzes them in triangulation with the interviews applied to practitioners as part of the third research phase.

6 RESULTS DISCUSSION AND ASSESMENT

As part of the third research phase, a number of interviews were applied to industry practitioners. The goal was to have a source of comparison between the results of action research and the experiences of other people from both NSN and other industry companies. Performing this comparison was particularly useful for addressing the external validity of the research results. In the next sections, a summary of the practical benefits, issues, conditions and alternatives of using CIM found during the action research is presented. Next, the results from interviews are analyzed collectively by finding commonalities and relationships within categories. Subsequently, both the action research and the interviews are consolidated by triangulation of data sources. The triangulated results served to test the validity of the initial propositions presented in chapter 4. Finally, in this chapter is the generalization assessment of benefits of CIM for IS interoperability and EI integration, which served to build the final conceptual framework.

6.1 Action research results summary

At the end of the three action research iterations, an analysis was made to summarize and categorize the findings. Four categories were selected to represent the benefits, issues, conditions and alternatives of using CIM. These categories are described in table 8 of appendix 5. The table 4 summarizes the results of this analysis:

Table 4. Lessons learned from action research.

Category	**Lessons learned**
Benefit	- SID allows to link business and system concepts. - Documentation and structure of the SID was perceived useful. - Working with the same model (the SID) reuses learning (learning curve). - Using CIM reduces the integration work; this is because the same model is being used in both sides of the mappings.
Issues	- A large amount of extensions are needed even for small models. - The way standards such as WSDL and XSD are handled by the tools might not be consistent, causing interoperability issues. - Using different tools for doing modeling, mapping and integrating applications might require model format conversions. These conversions might have conflicts and errors. - By using CIM as data model for service interfaces, it is likely to

	produce large and complex interfaces, mainly because references to parent elements include a large number of unnecessary elements. As a result, the interfaces are difficult to read and to work with.
Alternatives	- Using a transformation only application requires more application integration work in the ESB, two data model mappings are required instead of one. The advantage of it is the separates concerns; the transformation task is completely separated. - Using a message forwarding transformation application does not save integration effort, because even if the interfaces to be integrated are similar, ESB mappings are required. The advantage is that it helps to separate dependencies between the ESB and the applications.
Conditions	- The information models to be integrated should describe the semantics of the information being modeled; otherwise, additional documentation is required to perform the semantic integration. - The semantic integration tool should support the integration needs in terms of aspects such as the type of model, the technologies and the activity or procedure for semantic integration. - Governance measures should be placed for managing the way extensions and mappings are made. - To reduce the effort in mappings in ESB, all interfaces should use the same model. - The persons doing the semantic integration should be able to understand the semantics of data models to integrate.

6.2 Interview results

In addition to the interviews of the first research phase, in the third phase, five more persons were interviewed to support the general applicability of the action research results. Details about the interviewees and the interviews were included as part of appendix 7. For each interview, electronic interview reports were generated and validated by the interviewees. Next, the data was summarized and organized using categories and subcategories (see table 8 and 9 of appendix 5 for a description of each). The categories were benefits, issues, conditions and alternatives of using a CIM, and they were codified with the letters B, I, C and A respectively. The subcategories were enterprise, IS and modeling aspects of integration, and they were codified with the letters E, S and M respectively.

As a result of the categorization, the table 5 presents the interview results grouped by interviewees. A cross reference was appended to each entry in this table, which is a code comprised of the category code, optionally the subcategory code and a sequential number for the entries of the same categorical group within each interview. It is important to notice that the intention is not to use the results the interviews as statistical representation of a population, but to provide external validity and to understand the relationships between concepts within and across categories.

Table 5. Categorized interview results

Interviewee Codes	**Statement id and description**	**Category Code**
NSA	With the SID, the knowledge and lessons learned from other people is reused. For instance, design patterns are a form of reuse of solutions to known problems.	BE1, BS1
	At the lowest level, when dealing with the data model you might modify the structure to increase performance.	A1
	Care should be taken to keep the semantics.	C1
	For low level modeling roles, it is still necessary to know higher levels of SID for the domain being modeled. This in order not to break the semantics of the SID.	C2
NBA	The documentation of SID provides savings in training, logistics and other areas of IM management.	BE1
	Top-down approach is useful to understand the business.	BE2
	Using SID represents an industry alignment, which means you could speak the same ‘language’ than the industry.	BE3, BM1
	Separating different modeling roles is useful to separate concerns.	BE4
	SID separates domains, and thus concerns cleverly. If a person works with a SID domain well knowledge of the domain, but does not need to know other domains.	BM2
	By following a top-down approach for the implementation of NGOSS frameworks (eTOM and SID), at some point when you reach lower levels the resemblance with the standards disappears.	IS1
	An issue is that children classes inherit a number of unnecessary attributes.	IM2
	SID is large, complex and with deep structure.	IM3
	An alternative is to make a simpler model and link them to the complete model.	A1
	A transformation language could allow making	A2

	references from simpler models to the complete model	
	Governance should handle both services (SOA) and information (CIM) together.	C1
	Separating concerns is how you are able to see the implications of each modeling role.	C2
NSD	Using SID could reduce impacts of changes; the level of detail is organized so that high-level processes do not need to deal with technical issues	BM1
	Using SID could reduce impacts of changes also when changing a sub process, other sub process are not impacted.	BS1
	We realized that the same structure was present in different systems to manage the same data, thus the need for CIM was noticed.	IM1
	It was not clear where the border between information and data model is.	IM2
	When we started working with the data inventory schema, we realized that its size would be a problem.	IM3
	In our case, we do not have role separation but process separation. Organizations manage different structures, in CIM separation of roles and artifacts should consider that.	C1
	The semantic integration effort depends on the starting point; we normally had manageable data models. Some elements were mapped easily and other ones were simply mirrored to the database elements.	C2
ESI	With CIM companies would simply need to define their adapter in order to be able to exchange information with other companies.	BS1
	Implementing this project is a long term investment. If CIM does not receive enough support, there is the risk of lack of funding.	IE1
	Modelers should focus on the important aspects first; too much detail might be an issue at the beginning.	IM1
	Data integration is an iterative process. You can start simple, with a dictionary and mapping static data	A1
	The greater the number of participants the greater the benefit of the information exchange framework	C1
	Careful mapping needed	C2

	Good design is needed	C3
	Timing is very important, if someone is too late, it delays the entire design process. An alternative is to set up penalties for these cases	C4
	The initial goal is to gather the information of the companies and create a common terminology	C5
IEA	Currently we have been doing point-to-point integration but the goal is to use the industry reference model as common model for all transformations.	BS1
	An available insurance industry standard architecture is somehow complete, which includes enough detail.	BM1
	Sometimes you have to do bottom-up because legacy systems still need to be supported.	I2
	An appropriate governance model is required to be defined as early as possible.	C1
	The intention is to use the modeling suite of products of our provider to handle integrally business processes, information and software design.	C2
	We have done EA for a particular business unit and market segment.	A1
	EA should be carried out step by step in pilot projects.	A2

The rest of this section presents the analysis of the interview results. The analysis consisted in finding relationships among the interview results within each category:

For the *benefits* of using CIM, there is apparently a relationship between the structure of CIM, the knowledge reusability and the effort required during integrations. On the one hand, if the CIM is structured in a way that separates the concerns of the various business domains, persons could focus on a particular domain and reuse the learning from one integration project to another. And on the other hand, the same separation of concerns should favor the loose coupling of applications, thus providing more flexibility to react to changes and reducing the impact of changes.

In addition, connections were found in the *conditions* necessary to implement CIM. Aspects such as governance, modeling roles and tooling are all related and should be considered together. There are tools available for modeling business processes, information models, system architectures and software design. Each particular modeling role should have the appropriate tool, and the chain of tools should support the collaboration of the different rules participating in the entire development life cycle. Finally, governance mechanisms should be set in place for harmonize the practices used by the different roles in the development of the different artifacts. As a result, it should be possible to maintain a somehow standard way of working, which is conditional to leverage the benefits of using CIM.

Lastly, *issues* and *alternatives* of using CIM are related because different issues might be present or not, depending on the alternatives taken. On one hand, depending on the integration approach, different issues are found. If a top-down approach is taken for the definition of processes and services, at some point there is an issue of linking the underlying IS to meet the needs of the services that realize the functionality of business processes. If a bottom-up approach is used because legacy systems have to be integrated, then the result might end up being a functionality that is difficult to reuse for different purposes. On the other hand, modeling decisions and alternative tools will impact the issues present during implementation of CIM. The way design patterns are used could benefit the consistency and flexibility of SID extensions, or might affect the mapping effort and the semantics of concepts, in case the reality does not match exactly a pattern. Finally, the set of tools could arise issues in terms of format conversions and facilitate or difficult the mapping activity.

6.3 Triangulation of results and theory validation

In order to leverage the most learning of the different types of results obtained in this research, *triangulation* of data sources was made between the action research and interview results. Similar to the analysis of the individual data sources, the results from action research and interviews were consolidated and organized using the categories defined in appendix 5. For the purposes of triangulation, each of the three iterations and the five interviews with practitioners were considered as separate cases (see appendix 6). The cases were codified using a three-character code, for the iterations they were IT1, IT2 and IT3; and for the interviews the interviewee code was used. The triangulated results are presented in table 6, grouped by categories and making cross-reference to the supporting data entries from the action research iterations and interviews (see appendix 5). Next, the triangulated results were confronted with the propositions used during the first research phase to validate and complement them (see section 4.4).

The elements of each proposition were analyzed by performing two types of tests. The first test was that for each condition of the proposition, there should be enough support from empirical data. The second test was to qualify each causal relationship between the conditions and the effects of the proposition, based on the theory of causality. This theory states that in a cause-effect relationship, causes might be *necessary* and/or *sufficient* to obtain the effect. The case in which a cause is both necessary and sufficient indicates equivalence, in other words the effect will always be present when the cause is present (Mahoney, 2008).

Table 6. Triangulation between action research and interviews results

Category	Idea	Support (case-category)
Benefit	CIM provides a common language across different stakeholders and across companies.	NBA-BE3, ESI-C5
	SID is useful to understand both business and system view, and to close the gap between them	IT1-BE1, NBA-BE3, NBA-BM1, IEA-BM1
	Model documentation of SID can be useful to obtain benefits, they represent a reuse of knowledge	IT1-C1, NBA-BE1
	SID allows reusing model learning / knowledge, for instance with design patterns	IT1-C2, NSA-BE1
	Compliance with an industry standard is a desired to 'speak the same language' than the rest of the industry	IT3-IM2, NBA-BE3, NBA-BM1
	When using the same SID, integration times or costs should be reduced.	IT2-BE2, NBA-BE1
	The number of data model mappings is reduced with CIM	IT3-BS1, IEA-BS1
	Extensions are reused when using CIM	IT2-C1, IT2-C3
	Using CIM design patterns helps to make extensions and to keep consistency.	IT1-C3, NSA-BS1
	Using CIM allows to easily integrate models	IT3-BS1, ESI-BS1
	Using CIM might reduce impact of changes if its structure separates concerns clearly, when changing a sub process does not affect other sub processes. Separation of concerns helps to identify the implications of the different modeling roles.	NBA-BM2, NSD-BS1, NBA-C2
	Also, having the translation function clearly defined will reduce impacts of changes (translators or adapters)	IT1-BS1, NSD-BS1
Condition	Any person involved in the integration, should be familiar with the domain of the information model that is concerned.	ITE1-IE1, NSA-C2
	The suite of modeling tools should support the overall IS development life cycle and the collaboration of the different modeling roles. In other words, adequate integration of tools is a requirement.	IT1-C5, IT2-IM1, IEA-C2
	Governance of models should be defined for the	IT1-C7, IT2-C3,

	modeling practices and it should be in parallel for both services and information, and integrally for all the involved modeling roles and artifacts. In order words integration of governance is also a condition.	MBA-C1, IEA-C1
	All interfaces should use the same model in order to leverage benefits. As more and more participants use the CIM, benefits would be greater.	IT3-C1, ESI-C1
	Mappings should be done carefully to ensure quality	NSA-C2, ESI-C2, NSA-C3
	Adequate level of detail should be present in the CIM or part of the CIM used for integrations. Using CIM might reduce impact of changes, if its structure defines clearly the level of detail, the high-level issues do not need to deal with the technical issues.	IT3-IM1, IT3-IS1, NSA-A1, NSD-BM1, IEA-BM1
Issue	Interface composition might be an issue with large CIM, for example the SID model, which have a complex and deep hierarchy of elements. Unnecessary elements had to be included in iteration 3 because of element hierarchy and references, thus the size of interfaces exploded.	IT3-IM1
	Sometimes it is necessary to do a bottom-up approach because legacy systems need to be used.	IEA-I2
Alternative	A simplified model should be used in defining service interfaces, transformation languages and tools should be explored to find a suitable alterative that could both keep compliance to standards and allow the creation of flexible and understandable interfaces	NBA-A1, NBA-A2, IT3-A4
	Implementing CIM for enterprise semantic integration can be done progressively, for instance with pilot projects	IEA-A1, IEA-A2
	Using a subset of the CIM is easier to visualize	IT3-BM1

The result of the validation of the four initial propositions from section 4.1 will be presented next along with generalization assessments:

> *Proposition 1. Companies should increase their employees' coordination and understanding of the business by using a unified CIM. This is because using the CIM helps reducing communication gaps by providing a common vocabulary to unify the interpretation of data elements.*

Proposition 1 includes two conditions, reduction in the communication gap reduction and a common vocabulary. Regarding the communication gap three conditions were found. First, the fact that SID covers and links together the business and system concepts was found a necessary condition [IT1-BE1, NBA-BE3, NBA-BM1, IEA-BM1]. Second, it was a necessary condition for the information/data models to be defined at the same abstraction level than the SID, not meeting this condition was the cause for not being able to map the NPS to the SID in iteration 1 of action research [IT1-IM5, IT3-IM1]. And third, it was a necessary condition for the person making the data models integration to have knowledge of the data models, including the SID domains required in the use case. To meet this last condition, two sub-factors were found to be sufficient, but not necessary: If the modeler making the integration had sufficient command of data models involved [IT1-IE1, NSA-C2] or sufficient documentation was available to understand the semantics of the model elements [IT1-IM1], it would eventually be possible to connect the business and system concepts.

Furthermore, proposition 1 indicates that using CIM can facilitate the understanding across business and technology domains by having a common vocabulary [NBA-BE3, ESI-C5]. A clear necessary condition for this is that the model is commonly used among the participants [IT3-C1, ESI-C1], and in this case no additional condition was found, thus the benefit should hold true in all cases. Lastly, all conditions for proposition 1 were not found to be specific to SID, they might be generalized to other CIM.

> *Proposition 2. Companies that use CIM for integrating new applications could eventually reduce costs and time by reusing semantic integration assets: The mappings between data models and the CIM and the knowledge of the CIM are reused.*
>
> .

Support was found for proposition 2 thus using CIM should allow some reduction of integration time and costs [IT2-BE2, NBA-BE1]. A necessary condition for such reduction was to reuse some asset such as documentation [IT1-C1, NBA-BE1], knowledge of the models (or parts of the models) [IT1-C2, NSA-BE1], the mappings [IT3-BS1, IEA-BS1] or the extensions [IT2-C1, IT2-C3]. No additional condition was found to affect this relationship, thus the presence of reuse should reduce effort.

> *Proposition 3. CIM based on open industry standards with sufficient adoption from the community would favor interoperability of IS by reducing dependencies on specific platforms and vendors and by standardizing the representation of data elements, thus eventually reducing semantic differences.*

Support from proposition 3 was found [IT3-BS1, ESI-BS1]. A sufficient condition for that is the standardization of models used by different participants [IT3-IM2, NBA-

BE3, NBA-BM1] and interfaces that are involved in an integration [IT3-C1, ESI-C1].

> *Proposition 4. Companies that use CIM that covers the main concepts used in their existing processes, and in addition that model favors the interoperability principles, can develop capabilities of extensibility, reusability and flexibility to integrate new applications. This in turn increases the company's ability to reconfigure assets to create new integrated solutions.*

Support for proposition 4 was found [IT3-BS1, ESI-BS1] with the necessary condition that CIM had clear separation of concerns or domains [IT3-IM1, IT3-IS1, NSA-A1, NSD-BM1, IEA-BM1]. An additional necessary condition is the adequate level of detail in the CIM or part of the CIM that is used to do the integration [NBA-BM2, NSD-BS1, NBA-C2]. Lastly, a 'helpful' condition that was not found sufficient or necessary is the use of SID design patterns; however it was clearly supported by multiple cases [IT1-C3, NSA-BS1]. Future research might be interested in researching design patterns in more detail.

Finally, these propositions were abstracted in relation to the development of IS interoperability principles and EI dynamic capabilities. These abstractions served to refine the initial conceptual framework of benefits of using CIM (see figure 9), which was combined with the model of objectives and challenges of IS integration and EI (see figure 5). The result was the final theory presented in the next section.

6.4 Theory refinement

This section describes how using CIM facilitates the development of IS interoperability principles and EI dynamic capabilities.

First, the *learning* capability is facilitated by the reuse of knowledge and the learning curve, which are built when using CIM continuously, extensively and consistently among the different modeling roles during the entire development cycle of IS. Knowledge of the CIM, its extensions and mappings might be reused by using CIM, and by using an industry standard model as CIM, part of the knowledge of the people that created the standard is reused. An example is the documentation and the design patterns in a model like SID that represents solutions to design problems that are normally faced. Moreover, as stakeholders continue working with the same model, they built knowledge on the model, which is reused from project to project. Models like SID that have a clear separation of domains or concerns should help the different stakeholders to focus on the part of a model that is of their interest, thus allows them to learn those domains more fluently than having to learn an entire enterprise model.

Finally, CIM facilitates also the transferring, and thus the reuse of knowledge built within the organization. When using a common language for the different domains of end-to-end business and operation, people communicate ideas in a way other people are able to interpret them accurately.

However, there are conditions for developing learning capability by using CIM, which when not present could become issues. On the one hand, CIM might become an issue if it is not approached correctly. For instance, since CIM connects a large amount of concepts from the different enterprise domains, a person might attempt to learn the entire model at once, which would be a burden in a use case that only involves a small set of concepts. On the other hand, a condition for leveraging CIM learning is that command of the CIM is required for at least the domains involved in the use case. A person should be able to understand the semantics of the part of the model the person is working with. In consequence, significant initial time and costs might be spent in learning CIM.

Second, the *coordination and integration* capabilities are also favored by the CIM separation of concerns. This characteristic of CIM is useful to separate the modeling tasks in a way that avoids redundancies. Reusability of model transformations is also a feature of CIM that benefits the integration capability. When using CIM for integrating more than three systems, the number of mappings between data models is fewer than doing point-to-point integrations. And by using an intermediate model to perform all data model transformations, it should become easier to decouple the mapping tasks without integration effort. As opposed to having two persons that own two different systems participating in the mapping between the data models of the two systems, each person could focus on mapping the data model one IS to the CIM model. Moreover, the coordination of teams could be facilitated by using CIM because the interpretability of the modeling artifacts becomes more unified or standard. This should also be reflected in the efficiency of meetings and joint efforts related to the definition of concepts used in models. Finally, the use of an industry reference standard could facilitate the coordination of internal and external parties that are required in normal development for instance for requirement gathering.

Nevertheless, there are also issues and conditions affecting the developing of coordination and integration capabilities by using CIM. One of the main issues of using CIM in SOA is the implementation of service interfaces is the level of detail of the information elements. When using CIM as a basis for defining service interfaces' schemas, element and attribute names defined at a high level of detail might not be appropriate or meaningful to low level modeling of data In addition, a number of attributes that are inherited from high level abstract elements cause the low level elements to be large and complex. A second issue is the tool support for the coordination of the different modeling roles. The set of tools used for implementing

CIM should allow the integration of artifacts from a tool to another, without having syntactic issues or altering the model semantics. And a third condition is the governance required for defining standard measures and principles to be used in the different modeling tasks by the various modeling roles. Governance might become an issue if for instance the definition of services and information elements are not handled together, causing coordination issues among roles and their artifacts.

And third, the *reconfiguration and transformation* capabilities might be improved when using CIM. In this case, the contribution of CIM depends largely on the model design. The structure of CIM that use design patterns and design principles facilitates the flexibility for performing integration efforts. In addition, if CIM is structured adequately, it would contribute to having the right level of detail in the different modeling artifacts. This in turn provides the ability to organize the services modularly and reconfigure them easily when changes when needed.

Furthermore, taking the base of both conceptual models, IS integration effects on EI (see figure 5) and benefits of using CIM (see figure 9), a final theoretical framework was build to depict the general implications of using CIM for EI and IS integration (see figure 15). Two new elements are added to the framework, semantic integration issues and the use of CIM. Two links were added as well. A link is drawn between the use of CIM and the IS interoperability principles. From the results, a conclusion is that using CIM could increase the reusability, compliance to standards, loose coupling and interpretability of service interfaces. And a second link is between the use of CIM and the semantic integration issues. CIM might facilitate the reuse and transferring of knowledge regarding the semantics of information model entities. In addition, a well documented model could improve the interpretation of elements, and a well structured model might facilitate the separation of domains, concerns and purposes of a model, thus reducing the number of conflicts and inconsistencies. The arrangement of the concepts in this model form two horizontal cause-effect chains that go from information aspects in the left side, through IS or application aspects in the center, to enterprise wise aspects in the right side.

Finally, two contextual boxes were added to the model to represent the factors that affect the cause-effect chain. The upper gray box indicates the telecom context elements and encloses the semantic integration issues, the IS integration tax and the challenges of EI. This context represents the challenges of dynamic environment of telecoms for the provision of integrated solutions, the heterogeneity of technologies, platforms and systems from multiple parties that normally participate in integrations; and the semantic integration issues underlying the complexity of the business and operations support systems. The lower gray box indicates the contextual elements of using CIM, including the conditions, issues and alternatives. This context represents the elements that should be considered in the implementation of CIM to leverage its benefits. For instance the tool chain, the modeling roles and the governance aspects, all of them should also be integrated when implementing CIM.

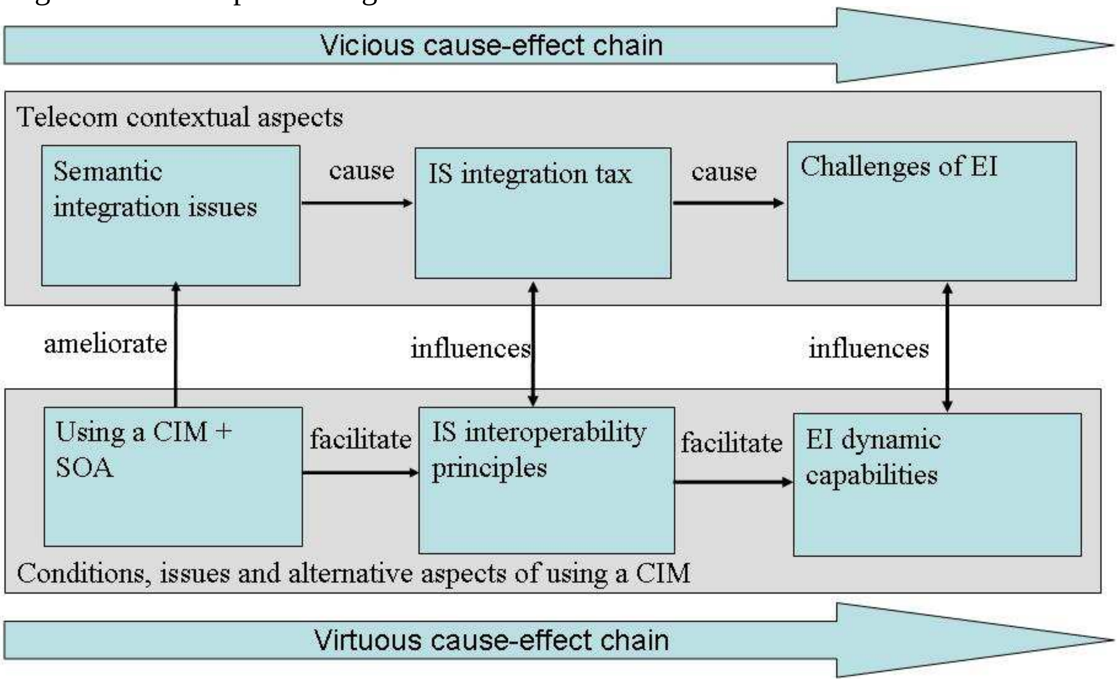

Figure 15. Virtuous effect of CIM effects in enterprise integration.

7 CONCLUSIONS

7.1 Closing statement

Unlike traditional business that based purely on product manufacturing or services provision, new business models are based on the combination of both products and services to provide value-added solutions to the customers. Companies that are moving towards the provision of these integrated solutions would need a larger set of capabilities geared towards the integration of the intra and inter organizational functional entities, resources and IS. These integration capabilities are also required to be dynamic, in the sense that companies should be able to react to changing business environments by reconfiguring their structural elements efficiently and rapidly. This is highly important for the case of telecom operators, where integrating solutions is highly costly and complex, mainly affected by interoperability challenges of integrating services from multiple partners and heterogeneous IS. Specifically, a main root issue is the heterogeneity of information used by the different stakeholders.

This issue is reflected in the disparity of the different information models used in the development of IS. In order to face these challenges, developing an integrated and service-oriented enterprise architecture will contribute to the creation of efficient service structure and to align the business and information technology strategy and infrastructure. And industry reference information models have been developed for unifying both the intra and inter organizational interoperation of IS. This research studied the use of a telecom industry reference model as CIM for the case of a NSN. The focus of the research was the benefits, issues and alternatives of using CIM for developing dynamic integration capabilities.

Theoretical research was done to create a conceptual framework that explains the cause-effect relationship between aspects of information modeling, IS interoperability and enterprise integration. The concept of dynamic capabilities served to define the type of requirements needed for the provision of integrated solutions. Three main capabilities were studied: organizational learning, coordination / integration, and reconfiguration / transformation. In addition, reference was made to IS interoperability principles as enablers of those three dynamic capabilities. Particularly, this studied included the compliance to open standards, reusability and loose coupling and the interpretability of components. Finally, EI and IS integration was analyzed from the perspective of IM. Objectives and challenges of information and data integration were described and the use of CIM was explored as an alternative to improve the semantic integration efforts and facilitate IS interoperability principles.

Furthermore, an empirical research was performed to validate the framework and to learn the practical implications and alternatives of using CIM. The telecom industry reference model from TMF, the SID model, was used CIM for integrating data models as part of application integration required in a service oriented use case. This empirical research was done in three progressive iterations, using action research methodology. Finally, interviews with practitioners from multiple industries served to confront the action research results and to refine the initial conceptual framework. Results from both action research and interviews were triangulated using categories such as benefits, issues, conditions and alternatives for using CIM.

As a result of this study, a cause-effect framework was built to describe the relationship between information modeling, IS integration and EI, including: a vicious cause-effect chain referring to the challenges and issues present during the integration efforts, which is affected by the context of the telecom industry. A virtuous cause-effect chain referring to the benefits of using CIM, for developing application interoperability principles and dynamic EI capabilities, which is affected by a context of conditions, issues and alternatives. The theory resulting from this study contributes to both the industry and the scientific community.

Industry community, including managers and practitioners might benefit from this study in a number of ways. The refined theory provides managers with an overall overview of the connection between IM and EI from perspectives of both semantic integration issues and benefits of using CIM. Based on the three EI dynamic capabilities studied in this research, managers could identify benefits of using CIM in terms of business agility. This would be useful in making managerial decisions regarding for aligning business and IT strategies at both the intra and inter organizational level, for instance, by coordinating the different stakeholders regarding the alternatives for information/data models integration and the deployment of best practices through governance policies. At the tactical level, managers should consider restructuring organizational resources in a way that favors the reusability of knowledge, documentation and data mappings of CIM. Moreover, the results of this study might be useful for industry practitioners to understand the practical implications of using a CIM. The empirical part of this research explored a number of CIM implementation alternatives and a number of lessons learn could be leveraged by industry practitioners. This includes the benefits and issues obtained from using CIM for integrating data models in a SOA application integration demo. In addition, the present study provides insights regarding the conditions that should be considered in order facilitate the benefits of CIM and the issues that might be present in implementing CIM.

Finally, the present study might be used by the scientific community as a base for further studies of CIM and EI. The next section will present some areas of interest for further research and suggestions to expand the present theory.

7.2 Further research

This section will outline areas of expansion of this research and other research opportunities that will complement the results of this study. The present study is a basis for the study of CIM in the context of EI in the context of the telecom industry. The main focus of this research was the IM aspects of an EA.

Moreover, expansion for this research is expected to come from different angles. First, dynamic EI capabilities could be studied for cases in which CIM is used in combination with a common business process framework, like the eTOM framework from TMF. Additional enterprise and IS integration benefits are expected from using them in parallel. Particularly the combination of eTOM and SID has the advantage that both use the same terminology and a group of the TMF community is currently working to map the two frameworks. Second, this research might be expanded to study the adoption of CIM by teams coming from multiple companies and/or situated at distant geographical locations. Third, the present research might be extended the study the use of CIM for providing integration solutions that embrace the different participants of a value chain, including manufacturers, integrators, operators and service providers.

Nevertheless, the conceptual framework build during this research is expected to be complemented and made more detailed, by focusing on specific aspects of it. Further research might focus on the degree in which particular aspects of CIM favor specific interoperability principles or the degree to which the interoperability principles favor specific dynamic EI capabilities, such as the ones used in this research or other capabilities. Additionally, further studies might focus on alternatives for performing semantic mapping of data models with CIM, including topics such as automation of mappings and tooling support. Finally, governance aspects of CIM should be developed in terms of what aspects need to be governed for implementing CIM, who should govern and how to enforce the decisions made.

8 REFERENCES

Allemang, Dean –Hendler, Jim (2008) Semantic web for the working ontologist. Morgan Kaufman, ISBN 978-0-12-373556-0

Malatras, et. al. (2008) A service-oriented architecture for building services integration. Journal of Facilities Management, Vol. 6, No: 2, 132-151

Baskerville, Richard L. (1999) Investigating information systems with action research. Communications of the Association for Information Systems, Vol. 2, No: 19

Bergamaschi, et. al. (2001) Semantic integration of heterogeneous information sources. Data & Knowledge Engineering, Vol. 36, No: 3, 215-249

Bernroider Edward - Koch Stefan (2001) ERP selection process in midsized and large organizations. Business Process Management Journal, Vol. 7, No: 3, 251-257.

Borgatti, Steve (2009) Introduction to grounded theory. <http://www.analytictech.com/mb870/introtoGT.htm>, retrieved 26.05.2009

Braganza, Ashley (2002) Enterprise integration: creating competitive capabilities. Integrated management systems, Vol. 13, No: 8, 562-572

Calvanese, et. al. (1998) Information integration: Conceptual modeling and reasoning support. Proc. of the 6th Int. Conf. on Cooperative Information Systems (CoopIS-98), 1998, 280-291.

Chen, e. al. (2008) Architectures for enterprise integration and interoperability: Past, present and future. Computers in Industry, Vol. 59, No: 7, 647-659

Cummins, Fred (2002) Enterprise Integration. Wiley & Sons. ISBN 978-0-471-40010-3.

Davies, et. al. (2001) Integrated solutions: the new economy between manufacturing and services. SPRU, University of Sussex, Falmer, Brighton, 1-43 <http://innovationwiki.brighton.ac.uk/images/2/2a/Integrated_Solutions_The_New_Economy.pdf >, retrieved 04.05.2009

Dittberner Associates (2005) Dittberner's Examination of Nokia's OSS middleware proposal.<http://www.zdnet.com.au/whitepaper/0,2000063328,22083387p-16001429q,00.htm>, retrieved 07.05.2009

Duke, et. al. (2005) Towards B2B integration in telecommunications with semantic web services, Springer-verlag berlin heidelberg, 710-724

Erl, Thomas (2005). Service-oriented architecture: concepts, technology, and design. Prentice-Hall, ISBN 0-13-185858-0.

Erl, Thomas (2007) SOA principles of service design. Prentice Hall, ISBN-13: 9780132344821

Francalanci, C. – Morabito V. (2008) IS integration and business performance: the mediation effect of organizational absorptive capacity in SMEs. Journal of Information Technology, Vol. 23, No: 4, 297-312

Fredericks, Paul – Van der Weide, Theo (2006) Information modeling: The process and the required competencies of its participants. Data & Knowledge Engineering, Vol. 58, No: 1, 4-20

Goodhue, Dale – Wybo, Michael – Kirsch, Laurie (1992) The impact of data integration on the costs and benefits of information systems, MIS Quarterly, Vol. 16, No: 3, pp. 300.

Gulledge, Thomas (2008) Architecture-driven Enterprise Integration. International Journal of Management and Enterprise Development, Vol. 5, No: 3, 265-307

Gulledge, Thomas – Deller, Greg (2008) Service-oriented concepts: bridging between managers and technologists. Industrial Management & Data, Vol. 109, No:1, 5-15

Hale, John (2006) Enterprise data integration: business boon or budget breaker? SOA World Magazine, October 13, 2006.

Hamilton, Douglas (1999) Linking strategic information systems concepts to practice: systems integration at the portfolio level. Journal of Information Technology, Vol. 14, No: 1, 69-82

Handfield, Robert (1998) The scientific theory-building process: a primer using the case of TQM. Journal of Operations Management, Vol. 16, No: 4, 321-339

Zimmermann, et. al. (2009) Elements of service-oriented analysis and design: An interdisciplinary modeling approach for SOA projects. IBM Technical Library. <http://www.ibm.com/developerworks/webservices/library/ws-soad1/>, retrieved 18.05.2009

Izza, et. al. (2006) Dealing with semantic application integration within large and dynamic enterprises. International Journal of Cooperative Information Systems, Vol. 15, No: 4, 507–534

Josuttis, Nicolai (2007) SOA in practice: The art of distributed system design. O’Reilly, August, 2007. ISBN 978-0-596-52955-0

Kim J.W. - Lim K.J. (2007) An approach to service-oriented architecture using web service and BPM in the telecom-OSS domain. Internet Research, Emerald Group Publishing Limited, Vol. 17, No: 1, 99-107

Lam, Wing (2005) Investigating success factors in enterprise application integration: a case-driven analysis. European Journal of Information Systems, Vol. 14, No: 2, 175-187

Lankhorst, Marc (2004) Enterprise architecture modeling—the issue of integration. Advanced Engineering Informatics, Vol. 18, No: 4, 205-216

Lim, et. al. (1997) Enterprise modeling and integration: a taxonomy of seven key aspects. Computers in industry, Vol. 34, No: 3, 339-359

Linthicum, David (2003) Next generation application integration: from simple information to web services. Addison-Wesley Information Technology Series. ISBN 0-201-84456-7

Loomis, David (1999) The future of the telecommunications industry. Springer, ISBN: 978-0-7923-8667-4, 195-197

Mahoney, James (2008) Towards a unified theory of causality. Comparative political studies, Vol. 41; No: 4-5, 412-436

Marino, John (2009) 'Convergence' will drive telecom industry in the future. Caribbean Business, Special Edition, 66-67

McBrien, Peter – Poulovassilis, Alejandra (1998) A formalisation of semantic schema integration. Information Systems, Vol. 13, No: 5, 307-334

McCue, Andy (2005) BT CIO Al-Noor Ramji: Inside BT's 'One IT' strategy, the offshoring issue 90-day "hothouses". Silicon.com, CIO interviews <http://www.silicon.com/cio-interviews/0,3800015327,39153745,00.htm> retrieved 26.05.2009

Mirbel, Isabel (1997) Semantic integration of conceptual schemas. Data & Management Engineering, Vol. 21, No: 2, 183-195

Mittal, et. al. (2008) SewNet - A framework for creating services utilizing telecom functionality. WWW2008, Refereed Track: XML and Web Data - XML II, 875 – 844 <http://www2008.org/papers/pdf/p875-mittal.pdf> retrieved 07.05.2009

Nokia Siemens Networks, 2009. Company web site. <http://www.nokiasiemensnetworks.com>, retrieved 07.05.2009

NSN technology & architecture manager (2009) Interview. Topic: Benefits of using the SID. Interview held by phone on 15.04.2009 during 60 minutes

Nonaka, Ikujiro (1994). A dynamic theory of organizational knowledge creation. Institute of Business Research, Organization Science, Vol. 5, No:1, 14-37

Olsson, Anders (2003) Understanding changing telecommunications. John Wiley and Sons, ISBN 0470868511, 9780470868515

O'Reilly, Charles – Tushman, Michael (2004) The ambidiextrous organization. Harvard Business Review, Vol. 82, No: 4, 74-81

Paljakka, Matti (2009). SEFRAM service framework for process industry information exchange, Tekes research. <https://www.automaatioseura.fi/confprog/downloadfile_public.php?conference=16&filename=16-14056.pdf>, retrieved 07.05.2009

Papazoglou, Mike – Van der Heuvel, Willem-Jan (2007) Service oriented architectures: approaches, technologies and research issues. The VLDB Journal, Vol. 16, No: 3, 389-415

Petrie, Charles (1992) Enterprise integration modeling. Scientific and Engineering Computing Series, the MIT Press, ISBN 0-262-66080-6

Potter, Morgan – Brady, Alison (2005) Strategic changes and frameworks affecting future OSS/BSS in the telecommunications industry. BT Technology Journal, Vol. 23, No: 3, 30-44

Presley, et. al. (2001) Engineering the virtual enterprise: An architecture-driven modeling approach. The International Journal of Flexible Manufacturing Systems, Vol. 13, No: 2, 145-162

Radding, A. (1999) Return on Investment: Measuring technology's dividends, <http://findarticles.com/p/articles/mi_m0FOX/is_19_4/ai_60369921/>, retrieved 07.05.2009

Rahm, Erhard – Bernstein, Philip (2001) A survey of approaches to automatic schema matching. The International Journal on Very Large Data Bases, Vol. 10, No: 4, 334–350

Ramakrishna, Raju (2009) An evaluation of SOA frameworks. Leading edge forum. <http://www.csc.com/aboutus/leadingedgeforum/knowledgelibrary/uploads/SOA%20Frameworks%20-%20RRaju.pdf>, retrieved 27.05.2009

Reilly, John (2007) Getting started with the SID: A SID modeler’s guide. ISBN 978-0-9794281-0-4

Reilly, John (2009) Interview with TeleManagement Forum distinguished member. Topic: Benefits and usage of the SID model. Interview held by phone on 04.03.2009 for 90 minutes

Reilly, John – Creaner, Martin (2005) NGOSS distilled: The essential guide to the next generation telecoms management. TeleManagement Forum. ISBN 978-0-9551087-0-9

Reilly, John – Wilmes, John (2008) Application integration using the SID. Published by TM Forum. ISBN 978-9-9794281-1-1.

Robinson, Greg (2008) Using a common information model for enterprise integration. Enterprise Integration, Utility Automation, January 2008, 56-67

Rotem-Gal-Oz, Aaron (2009). What is SOA anyway? *<http://www.rgoarchitects.com/Files/SOADefined.pdf>, retrieved 05.05.2009*

Scheer, August-Wilhelm - Hars, Alexander (1992) Extending data modeling to cover the whole enterprise. Association for Computing Machinery. Communications of the ACM, Vol. 35, No: 9, 166-173

Smith, et. al. (2002) Enterprise integration. News@SEI, Vol. 5, No: 4, <http://www.sei.cmu.edu/news-at-sei/columns/the_architect/2002/4q02/architect-4q02.htm>, retrieved 10.05.2009

Stuart, et. al. (2002) Effective case research in operations management: a process perspective. Journal of Operations Management, Vol. 20, No: 5, 419-433

Susman, Gerald – Evered, Roger (1978) An assessment of the scientific merits of action research. Administrative Science Quarterly, Vol. 23, 582-603.

Teece, David – Pisano, Gary – Shuen, Amy (1997) Dynamic capabilities and strategic management. Strategic Management Journal, Vol. 18, No: 7, 509-533

TeleManagement Forum (2008) Shared Information/Data Model. TeleManagement Forum, GB222 Release 7.5, Version Number 7.6.

TeleManagement Forum (2009) Introduction to TM Forum. Tele Management Forum. <http://www.tmforum.org/IntroductiontoTMForum/5749/home.html>, retrieved 09.04.2009

TeleManagement Forum (2009) Leveraging the SID Model to Develop a Unified Model for Service and Resource Assurance Application. Tektronix case study. <http://www.tmforum.org/CaseStudies/LeveragingtheSIDModel/37478/article.html>, retrieved 26.05.2009

Umar, Ajmad – Zordan, Adalberto (2009). Reengineering for service oriented architectures: A strategic decision model for integration versus migration. The Journal of Systems and Software, Vol. 82, No: 3, 448-462

Vernadat, F. B. (2007) Interoperable enterprise systems: Principles, concepts, and methods. Annual Reviews in Control. 2007, Vol. 31, 137-145

Voss, Chris, et. al. (2002) Case Research in Operations Management. International Journal of Operations and Production Management, Vol. 22, No: 2, 195-219

Web Services Interoperability Organization (2009). About WS-I. <http://www.ws-i.org/about/Default.aspx>, retrieved 25.05.2009

Wilmes, John (2008) Data integration: defining the common model, Global Telecoms Business CEO and CFO Guide to OSS/BSS: March/April 2008, 72-73 <www.progress.com/progress/dataxtend/docs/article_defining_common_model.pdf>, retrieved 25.05.2009

Wittgreffe, J. P. – Dames, M. P. (2005) From desktop to data centre - addressing the OSS challenges in the delivery of network-centric ICT services. BT Technology Journal, Vol. 23, No: 3, 65

Weng, et. al. (2007) User-centered semantic harmonization: A case study. Journal of Biomedical Informatics, Vol. 40, No: 3, 353–364

Yin, Robert K. (2002) Case study research. design and methods. Sage Publications, Vol. 5, ISBN 0-7619-2553-8

Zachman, Jonh (2009) The Zachman Framework™: the official concise definition. <http://www.zachmaninternational.com/index.php/the-zachman-framework>, retrieved 13.04.2009

Zhao, Luyin – Siau, Keng (2007) Information mediation using metamodels: an approach using xml and common warehouse metamodel. Journal of Database Management, Vol. 18, No: 3, 69-82

Appendix 1 – Constructs for EI and data integration challenges

The following table contains the constructs generated for the first phase of this research. Each construct consist of a statement and a list of references that include the statement. The constructs were organized according to the categories proposed by Borgatti (2009).

Table 7. Constructs used for pattern matching in the first research phase.

Category	Constructs
cause	Data heterogeneity of formats (Zhao and Siau, 2007; Cummins, 2002). Data volatility, changes of data over time (Hamilton, 1999; Cummins, 2002). Data interoperability, same data modeled very differently (Dittberner Associates, 2005; Wilmes, 2008). Communication gap between business and IT people (Cummins, 2002).
context	Source of uncertainty defines the desirability of integration, which is more suitable for interdependence between units than for complexity of task or unstable environment (Goodhue, Wybo and Kirsch, 1992). Telecoms have two types of uncertainty: future architecture and integration of multi-vendor and multi-technology OSS (Dittberner Associates, 2005). Mergers and acquisitions landscape of the telecoms industry and the increased requirement to partner with external organizations to deliver new services such as content (Wilmes, 2008). Adapting to change as OSS/BSS interfaces change and systems are added and removed over time (Petri, 2008). "The primary challenge lies not in the ICT infrastructure itself, but in the increasing demands placed upon the OSS to deliver holistically across the range of ICT operations" (Wittgreffe and Dames, 2005).
phenomenon	Integration tax, semantic integration is time consuming and complex (Hamilton, 1999; Dittberner Associates, 2005). Data integration is a main cause of failure of IT portfolio alignment. These issues are causing OSS transformation programs becoming unsuccessful (Hamilton, 1999; Wilmes, 2008). Changes in business processes require a long and tedious implementation that involves many people, considerable time, and usually trial-and-error solutions (Cummins, 2002).

action	Approaches point-to-point vs. conceptual model (Calvanese, 2008). Minimalist vs. strategic (Robinson, 2008). Types of model integration: master, unified and federated (Lim, 1997). Architecture driven integration of IS (Hamilton, 1999). SOA to hide complexity of underlying telecom protocols (Mittal, 2008). Solutions of information interoperability include the middleware-based interoperability and the mediation-based interoperability (Zhao and Siau, 2007). A standard data model as common model offers a reduced learning curve. Industry/enterprise standards should be applied (Wilmes, 2008; Cummins, 2002). Stable integration standards (Robinson, 2008). An enterprise data model is very important (Lam, 2005). An enterprise should leverage industry efforts and extend or complement them with local standards where necessary (not in all cases, especially if data exist within a single application or function). (Cummins, 2002). "Unless SOA can be aligned to business processes, it will be viewed as a risky proposition that adds to costs without directly addressing business needs" (Gulledge and Deller, 2008).
condition	Cost control, use of tools (Hale, 2006). Software tools could help in facing integration challenges (Petrie, 1992). Vendor support / partnership (Robinson, 2008). Contribution of stakeholders, willingness to change/improve (Robinson, 2008). Depends on interdependence of units, need for flexibility and difficulty of working with integrated data (Goodhue, Wybo and Kirsch, 1992). Absorptive capacity of the organization should also increase (Francalanci and Morabito, 2008). SOA and IS integration should be process driven (Gulledge and Deller, 2008). Only using a Common Data Model and ensuring data interoperability can maximize the potential of SOA (Wilmes, 2008) Implementing a CDM will require the continual evolution of the common model and the business requirements placed upon it

	(Wilmes, 2008). “There must be an agreement on data. If an organization cannot agree on standardized master data to load [..], then executives should be suspicious of activities that require the challenging type of governance that SOA implementation requires” (Gulledge and Deller, 2008). Customizations and extensions are always needed (Wilmes, 2008) Impact analysis and change management that are essential for governance and maintenance (Wilmes, 2008) It is necessary for business functions to be performed in a reasonably consistent fashion (Cummins, 2002)
consequences	Flexibility / transparency to deal with diversity of data and sources (Izza, 2006; Robinson, 2008) Unify business / data consistency (Hale, 2006; Robinson, 2008) Cost reduction of implementation and maintenance (Robinson, 2008) Business agility and risk reduction (Robinson, 2008) SID-based data services enable more rapid integration of OSS/BSS and adaptation to change as system interfaces are modified and as systems are added and removed (Petrie, 2008)

Appendix 2 – TeleManagement Forum's shared information/data model structure

The SID model is an object-oriented model that covers the domains of the entire end-to-end network operators' business. The SID is organized in three views, the business, the system and the implementation view.

First, the business view is further composed domains, which are areas of concern for a business or layer of the value-chain. There are eight domains in SID: market/sales, product, customer, service, resource, supplier/partner, enterprise and common business. In turn, domains are composed of a number of abstract business entities (ABE), which are high-level concepts that are of interest for the business. For example, within the customer domain there are ABEs such as customer, customer order and customer bill. Finally, ABEs are composed of business entities, which are tangible or intangible concepts that are represented by a name and a set of attributes. These business entities might have relationships with other business entities to represent connections between them.

Second, the system view is also decomposed in domains, abstract system entities (ASE) and system entities. The difference is that system entities extend the business entities, from an object-oriented perspective, which implies that a system entity will have the same properties than the corresponding business entity plus additional system related details, such as the attribute type. The system view is available in various formats, such as UML, XSD and RSM. Finally, the implementation view is the data model representation of the system view (Reilly and Creaner, 2005).

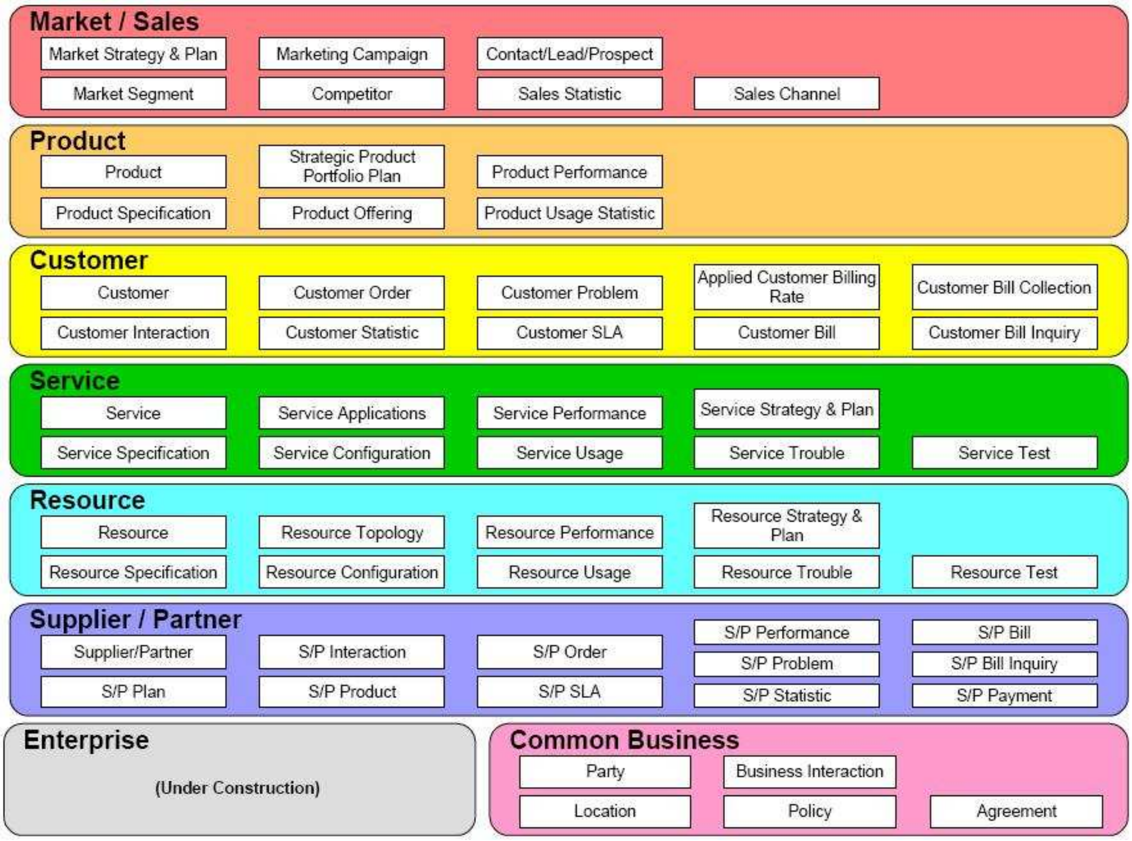

Figure 16. Shared Information/Data (SID) domains and level 1 ABEs (TeleManagement Forum, 2008)

Appendix 3 – Progress DataXtend semantic integrator

The tool is used to integrate various data models to CIM and generate a transformation modules or applications. Within the tool it is possible to import, visualize, navigate and map data models, among other functions. Next is an unofficial list of the features of this tool:

- Various importing formats include XSD, WSDL and UML. Visualization allows, among others, a hierarchical view and an entity-relationship view.
- Navigation is a useful feature of this tool, in which a user is able to go from one entity to the related entities simply by double click.
- Mapping is done either graphically, by drag and drop, or with wizards provided by the tool. As part of the mappings, the tool allows the definition of 'computed' attributes and other business rules, including transformations of formats and routing of mappings based on conditions.
- A test environment is embedded in the tool to be used in the development cycle to track mappings and detect mapping errors.
- An impact analysis feature allows managing both the data models and the mapping files.
- Lastly, the deployment function includes multiple platform options for the generated transformation application or module, which performs mappings and acts as a translator or adapter is obtained from this tool.

Finally, various architectural alternatives are supported for the generated application, such as a transformation only application and a mediation (message routing) application. In the former, the generated application is limited to transforming information from a data model to another. The latter option is used to generate a mediation application that serves as an interlocutor, in other words all communication is done with the transformation application as opposed to directly contact any OSS.

Appendix 4 – Procedure for semantic integration

The following procedure is an adaptation of the Reilly and Wilmes' (2008) procedure for application integration using the SID. Additional details and steps were added from literature on semantic integration approaches (Mirbel, 2007; Rahm and Bernstein, 2001; Weng, 2007).

1. Define the scope of the project by selecting SID domains, ABEs and business entities that should be considered. For a definition of domain and ABE refer to the appendix 2.
2. Identify and analyze data models involved in the integration, to identify the ABEs that will be mapped for each data model. With this, the scope of the project is confirmed or refined.
3. Perform structural mapping of data models. Using SID design patterns described in Reilly (2007), a rough structural mapping is performed for a particular set of elements. Then a more detailed element level mapping is performed using linguistic matching, for instance, based on the element name and description.

    > *"A design pattern describes a common problem and provides a corresponding solution"* (Erl, 2007)

4. Attribute mapping of data models. Linguistic matching based on name and data type. Two types of matching were identified:
    - Simple Attributes - simple one to one mapping. Includes the conversion of types but not format conversion or formulas.
    - Complex Attributes - complex one to one mapping, requiring formulas for validation or format conversions, or one to many mapping.
5. Relationship mapping is the last step of the process. It consists of validating the element relationships in relation to the SID model. This is done by navigating through the elements and looking at the cardinality of the relationships. Cardinality of a relationship is the minimum and maximum number of instances of a model element that might participate in such relationship.

Appendix 5 – Categories and subcategories for the analysis of results

The following categories were chosen as representative elements of a cause-effect relationship of the use of CIM in enterprise and application integration. Two types of effects, or dependent variables, are the benefits (B) and issues (I) caused by using CIM. The other two categories are causes, or dependent variables, which correspond to conditions (C) that should be present and alternatives (A) that might be taken in the implementation of CIM. Furthermore, the benefits and issues categories are decomposed one level down into subcategories to indicate the EA aspect they are related to, from enterprise (E), IS (S) or modeling (M). The codes corresponding to the subcategories are appended to the codes of their parent category.

Table 8. Categories for the analysis of results

Code	**Category**	**Type of variable**	**Description**
B	Benefit	Dependent	Valuable aspect, feature or condition
I	Issue	Dependent	Aspect hindering a benefit or causing inefficiencies
C	Condition	Independent	Aspect needed to leverage a benefit
A	Alternative	Independent	An action or decision aspect.

Table 9. Subcategories for dependent variables in the analysis of results

Code	**Subcategory**	**Description**
BE	EI dynamic capability	Integration, coordination, learning, reconfiguration or transformation capabilities.
BS	IS Interoperability principles	Open industry standards, reusability, loose coupling, granularity and interpretability.
BM	Model properties of CIM	Documentation, structure, design patterns, adoption, support from a community, among other desired features.
IE	EI challenges	Organizational issues, resistance to change, incompatibility of vendors, complexity and heterogeneity of IS, large number of participants (stakeholders), among other enterprise level factors
IS	Integration tax	Time and cost (real or perceived) of integration effort.
IM	Semantic integration issues	Heterogeneity, volatility, inconsistency or lack of governance

Appendix 6 – Summary of action research results

The following tables summarize the results obtained from the three action research iterations. The three iterations were coded as IT1, IT2 and IT3, corresponding to the iterations 1 to 3 respectively. Each entry contains a statement about the action research extracted from the report. Entries are grouped in the main categories and labeled sequentially using the corresponding subcategory code, from the ones presented in appendix 5.

Table 10. Action research results of first iteration.

Category-Subcategory	**IT1 – Data transformation application**
Benefit	BE1. Ability to relate business concepts were mapped to data elements. BS1. Translation function is clearly separated from other mediation tasks
Issue	IE1. Learning curve for both CIM and proprietary models IM1. Incomplete documentation of DIS data model 2 IS1. The DXSI transformation service had a large and complex interface, containing both the portal and DIS data models. IM2. Since the entire SID model was used, it was hard to visualize navigate the model. IM3. Different types and several semantic mismatches to handle, it became difficult. IS2. The mappings done in DXSI were not perceived as useful, because additional mappings were required for the same data model in ESB IS3. The mappings done in ESB were doubled, two mappings rather than one was needed to integrate the DIS. IM4. Mapping the SID was time-consuming, mainly due to learning the SID and making decisions about how to extend the SID. But the time is expected to be reduced in further integrations. IM5. NPS data model was found to be at a different level of abstraction, thus difficult to do semantic mapping.

[2] The reason for the documentation not being complete is that the version of DIS used for this practical exercise was not a commercial version

Alternative	A1. A semantic translator for a tactical demo at the application layer, using SID as unified CIM for a global translation. Taken entire SID without modifications. A2. Documented manual semantic mappings in a table format using office software. A3. A simpler to manage subset of DIS data model elements used for this first iteration. A3. Extended SID using the DXSI tool.
Condition	C1. SID documentation is useful to perform semantic mappings in DXSI. C2. SID design patterns are useful to understand concepts of both SID and DIS models, and to do extensions. C3. SID structure and design patterns were perceived to be flexible for extending the SID for application specific data elements. C4. Knowing the type of semantic mismatches helped to identify them. C5. DXSI navigability was useful. C6. Previous knowledge of the DXSI tool and the overall structure of SID model. C7. It was identified the need for guidelines regarding the elements, data format and version control of SID extensions.

Table 11. Action research results of second iteration.

Category	Iteration 2 - CIM-Based mediation layer
Alternative	A1. Same as first iteration but performing global translation with an extended SID model as CIM. A2. Some TMF guidelines and design patterns were used in this iteration. A3. The SID was further extended, in addition to extensions from the first iteration. A4. The RSM tool was used for modeling and implementing the extensions to the SID model. A5. A change in the application architecture was to use a message-forwarding transformation application, to form a data mediation layer between the ESB and the DIS. A6. A simpler data model, based on the one from portal, was used for the integration interface of the DXSI transformation application.
Benefit	BM1. It was possible to navigate between low-level and high-level concepts. BE1 The command of the portal, SID and DIS models increased.

	BE2. After making a first set of extensions to the SID model, making further extensions was perceived to take less time.
Issue	IM1. There were format conversion errors when exporting from RSM to import into DXSI. IS1. The total time taken to make extensions to the SID, including the first and second iterations was perceived to be greater than the one spent for point-to-point integrations. IS2. Additional logic was needed inside DXSI in order to send the right service request message; the issue is that the logic is now in two places, thus affecting maintainability. IM2. The integration interfaces were not SID compliant, reusability was perceived to be lower than if they were SID compliant.
Condition	C1. New extensions were built in addition to the previous extensions from iteration 1. C2. Mappings from iteration 1 were reused. C3. Governance regarding SID extensions was identified as needed.

Table 12. Action research results of third iteration.

Category	**Iteration 3 - CIM realization**
Alternative	A1. Same as second iteration but using a unified and federated CIM, a subset of CIM was extracted as data model. A2. All interfaces handled by ESB were based on the CIM. A3. A subset of the CIM was used for semantic integrations in DSXI. A4. Further studies could try using defining a simplified CIM by ‘flattening’ the CIM, which consist of pushing the attributes of abstract entities into their children entities. As a result, the flattened CIM model would contain just the required concrete classes.
Benefit	BM1. It was simpler to visualize and navigate the CIM subset. BS1. The effort in the ESB was perceivably reduced because the number of mappings was reduced.

Category	Iteration 3 - CIM realization
Issue	IM1. The integration interfaces became large and complex to manage, because the CIM subset was included as part of the interfaces. Although smaller than the entire SID, the CIM subset was difficult to visualize and manage. IM2. It was confusing to determine what SID compliant is. Apparently it is not the same to be CIM compliant that SID compliant. IS1. The CIM subset included a large number of optional entities, mostly abstract classes with optional attributes that were not used in this case. This was an issue when using the CIM subset as interface because a dependency was created between the system integrator and the information modeler to know what concrete classes to use and whether attributes were required or not.
Condition	C1. A main condition was that the same model is used for all interfaces integrated in the ESB.

Appendix 7 – Interview results data

The following tables present the data obtained from the interviews with experts and practitioners. The first two tables correspond to the interviewee and interview details. The last table presents the interview results with cross reference to the interviewee code and a corresponding categorization code, in appendix 5 are the codes for categories and subcategories. The code for the interview entries were built by appending the category code (B, I, A or C), optionally a subcategory code (E, S or M) and a sequential number within each the categorization group of each interview.

Table 13. Interviewees

Code	**Company**	**Position**	**Background on the topic**
NSA	NSN	System architect	Started working as system architect 10 years ago and then changed to research and development systems architect, when started working with the NGOSS frameworks, including the SID. Also contributed in the definition of concepts for in a TMF initiative for prototype reference architecture.
NBA	NSN	Business architect	Worked 5 years as a consultant, including the roles of system integration tester and developer, has worked as business architect with extensive use of eTOM from 2005 to 2008 and with use of SID since April 2008.
NSD	NSN	Two SOA developers	One has worked for 15 years at NSN, working first in the development of radio network simulators and for the last three years in the research of SOA, including the development of demos. The other one has worked in the development of those demos.
ESI	Electricity company	System Integration developer	Worked for 20 years in the automation field, including the definition of relational and object-oriented databases for the process automation industry. Has worked for 5 years in an industry initiative with other companies, in the definition of a data integration framework and the development of integration tools used by those companies.
IEA	Insurance company	Enterprise architect	Has worked for 12 years in the field of EA, currently works as part of an EA group and in development projects. Currently the head of strategy a unit: process & technology.

Table 14. Interview details

Inter-viewee Code	**Topic**	**Duration (minutes)**	**Channel**	**Date**
NSA	Enterprise architects at NSN. Topic: Benefits, issues and alternatives of using the SID.	60	In person	15.05.2009
NBA	Enterprise architects at NSN. Topic: Benefits, issues and alternatives of using the SID.	90	Phone	21.04.2009
NSD	Implementations issues in relation to data integration.	60	In person	18.05.2009
ESI	Data integration in the "Service framework for process industry information exchange" project.	60	Phone	15.04.2009
IEA	Enterprise integration modeling: issues and strategies	90	Phone	19.05.2009