Antti Korpiniemi

Improving Data Management for Better Itemization Process

Towards the High Quality Data in Product Data Management

Helsinki Metropolia University of Applied Sciences

Master's Degree

Logistics

Master's Thesis

8th May 2016

Helsinki Metropolia University of Applied Sciences

The whole previous year has been an amazing journey which began in May 2015. It included many big events in my life, and also hard work for my thesis. First, it was challenging to find the correct focus for the topic, as there were many interesting possibilities. After discussion with my manager, I had focused on the current itemization processes in the case company and exploring methods that could be used to improve it. As I started my work on the Thesis, I was not sure how the interviewed people would receive my questions regarding itemization in the case company. I was surprised in a positive way that all my interviewees were very interested in my work and gave lots of valuable information that I was able to utilize and create a better overview of the company's itemization process. I would like to give my big thanks to all my interviewees and the stakeholders that I had discussions and feedbacks referring to my work. Especially, I want to express my gratitude to the data manager from the case company, who helped me a lot along my work on this thesis.

I would also like to warmly thank my school instructors, Dr. Juha Haimala and Dr. Thomas Rohweder for their helpful comments and guidance, which gave me new ideas how to continue with my work. I also want to give my big thanks to Zinaida Grabovskaia for the great feedback that she gave for my writings. And I want to thank also all my classmates as they generously shared ideas (and occasional laughs) which also guided me how to improve my work.

This has been an amazing year, but at the same time one of the hardest. As I went to the Logistic program's entrance exam, the day before our new family member was born. Thus, I had to work fulltime and take care of our newborn son, as my wife was also busy with her studies. The last few months were especially challenging and I used almost every evening and weekend to write my thesis to accomplish it. And I was lucky to manage! Therefore, I want to give my biggest thanks to my wife Maria for the endless support and caring for our child Aleksi, so I had time accomplish my thesis on time.

Now the approach for itemization improvement has been done and it is time to focus on new challenges, and give more time to my son.

Antti Korpiniemi
Helsinki, 8 May 2016

Author	Antti Korpiniemi
Title	Improving Data Management for Better Itemization Process: Towards the High Quality Data in Product Data Management
Number of Pages	82 pages + 9 appendices
Date	8th May 2016
Degree	Master of Engineering
Degree Programme	Logistics
Instructors	Juha Haimala, DSc (Tech), Principal Lecturer, Head of Industrial Management Zinaida Grabovskaia, PhL, Senior Lecturer

This Thesis focuses on improving data management for better itemization processes in the case company. Focus of the research is to identify and solve the root causes behind the problem why confirmed information in the product data management is not systematically received. This problem affects the processes where items are created and maintained and where there is a critical need for complete and high quality information.

The objective of this Thesis is thus to create an approach to improve the itemization process which would increase the quality of data in terms of component knowledge accuracy, reduce its delivery time, and would in the end increase competitiveness towards competitors. The study is conducted by, first, analyzing the current state of itemization process, combining the best existing knowledge, and creating an approach for information quality problem solving. This approach is then applied on an example case showing how to solving the data quality problem for the benefit of the case company.

The outcome of this Thesis is an approach for data quality problem solving. The approach consists of different methods based on: context, role, pattern, and aspect of data management that can be used together or individually to improve the received information quality, and a TDQM (the total data quality management) tool which can be used for continuous information quality improvement. The approach presents how to utilize the methods by showing the connections of high data quality dimensions between the methods.

The benefit of the research and created approach is that it can evoke questions and increase interest of the company to research their information quality and discover new methods how to improve the quality of data and the processes related to managing it. In the context of the case company, it could provide a more efficient itemization processes that would benefit the case organization and its stakeholders.

Keywords	Data Quality, Item, TDQM, Information Context, Itemization Process, Product Data Management, Information Quality, Data Quality Problem Solving

Contents

List of Figures

List of Tables

Acronyms

CM	Configuration Management
CSA	Current State Analysis
DQ	Data Quality
EDM	Engineering Data Management
ERP	Enterprise Resource Planning
GIT	Generic Import Tool
IoT	Internet of Things
IPM	Information Product Manager
IQ	Information Quality
IS	Information System
MDM	Master Data Management
OPAL	Operational Excellence
PCL	Product Lifecycle
PDM	Product Data Management
PLM	Product Lifecycle Management
PSP/IQ	Product and Service Performance model of Information Quality
RDO	Responsible Design Organization
RFQ	Request For Quotation
TDQM	Total Data Quality Management
WM	Workflow Management

1 Introduction

This thesis examines the existing itemization process as a part of a wider product data management and tries to find a new way to improve the current itemization process, so that it would tackle the root cause behind the data quality problem, and benefit internal customers in the case company.

1.1 Case Company Background

The case company for this research is a global company which provides technologies in two business area for mineral processing and metals, water and energy. It is one part of a consolidated group. Its history begins from 1910 when copper ore deposit was found in Eastern Finland and the company was founded for mining and metallurgical operations in the area. The original company went overseas in the 1970s and started the company's internationalization. The present company was organized as a legal consolidated group in October 2006 and changed its name in April 2007.

In minerals processing, the company provides design and delivers of mineral processing equipment, optimized processes, automation and control system as well as complete plants. The company provides sustainable solutions for metal processing, energy and water solutions and has a century of experience in developing these technologies and plants with lifecycle services. All products meet environmental discharge standards and reduce water and energy consumption. Majority of company's business sales relates to minerals and metals processing; as a consequence, the company's process model for value creating is focused to metal and mineral processing operations.

Presently, the company has sales and services in over 30 countries worldwide. Operations are divided to three regions: the Americas, EMEA (Europe, the Middle East and Africa) and APAC (Asia Pacific). Since the company has about 3500 direct and active suppliers, the company has prioritized supply development area since 2010. One key strategy of the company is to develop its supply chain management especially in the environmental-minded way. One major area of this is product lifecycle management (PLM), which in-

volves many different itemization processes. The itemization also plays a major role in the company strategy.

1.2 Key Concept of This Study

An item is a unique code which consist specified information to describe a tangible material or intangible service with information that consists of different attributes e.g. a product assembly, a work process or commercial component and presents it in a virtual environment. Component Engineering's *itemization process* includes either creation of new third party item or enrichment of existing legacy third party items. Documentation can vary between datasheets, order confirmation or purchase orders which will verify item information e.g. type, technical attributes, and manufacturer part number.

This work is primarily done by the Component Engineering team which is one part of digital platforms division in the organization. *Component Engineering team* in Product Data Management (PDM) department serves internal customers in global Outotec. Items are used globally by engineering, procurement, sales and service. For these operators, it is critically important to be able to find and recognize correct products for use everywhere around the world.

Information inside items will technically specify and describe every raw material, design and third party product that can be identified throughout their lifecycle, from creation to obsolete state. Unique codes also enable behavior control of different departments in organization for example sales, procurement and service. If the information is in good quality, it affects the company costs that could be reduced, it offers better inventory tracking and creating indicators for management, as well as makes quotations would more precise and delivered faster to customers and thereby increase competitiveness.

1.3 Business Challenge of This Study

Currently, the case company is developing its Product itemization process as part of the Product Lifecycle Management. Product itemization will have a major role to make supply chain more efficient in the following aspects: (a) increase component knowledge accuracy and thus reduce failures in projects, and (b) most importantly, reduce delivery time from

the customer order to a finished product delivery, and thus (c) increase cost competitiveness. The idea behind Product itemization is to define all company's own design products, raw materials and supplier commercial components with a unique item code that can be used globally to identify every single component. Figure 1 shows how product management is centered to cycle consist of different departments which use the items.

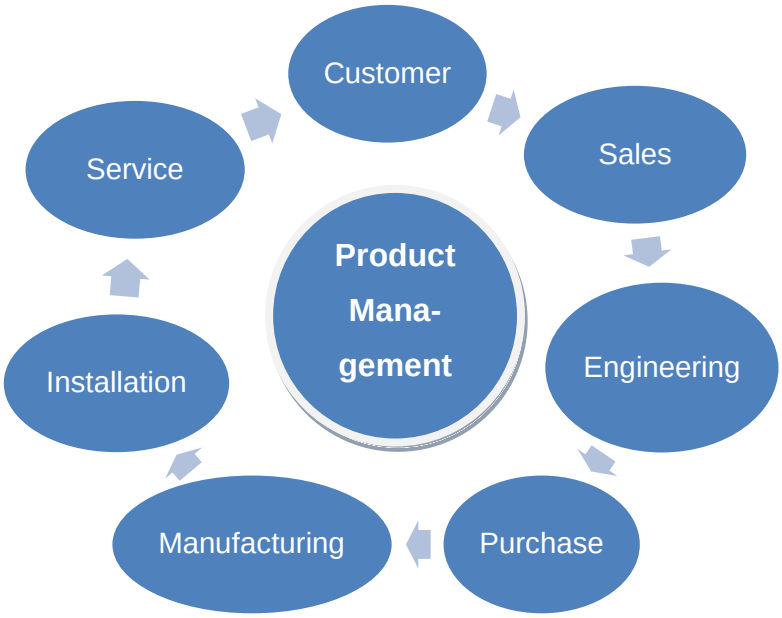


Figure 1. Item cycle in Product Management.

However, currently Component Engineering lacks *confirmed* information for itemizing the components in a complete format (i.e. creating and maintaining components). Current item information quality varies a lot and depends much on which party is requesting the items and who could provide the needed information. Therefore it is sometimes even impossible to recognize the real products described in the items, and the biggest challenge will be with the installed base data that is not utilized fully because of missing information.

Therefore, for the case company it is critically important to gather information from archives for item descriptions and implement critical information of the products into the data base and for every ones use in the case company. Properly defined information in the items would create more efficient process for item management (item creation and maintenance) which will eventually benefit the whole company especially the service business.

1.4 Objective, Outcome and Scope of the Study

The objective of this study is to create an approach to improve the itemization process for the case company. The eventual goal of this approach is to improve itemization processes by utilizing different methods that will view the itemization processes from the information quality point of view. The proposed approach will be done for use in product data management to improve the itemization processes in the case company.

The outcome of this study will contain proposal of an approach to improve the itemization process. The approach includes different methods which can be used to research item information quality. The research of this study will also clarify definition of information quality and its relation to context, explain the meaning of high data quality dimensions, and clarify why it is important to have quality information described in the items. The information supplements from different departments and data producers into item during its lifecycle. The quality of the information is thereby really important as it would help sharing information towards customers, internally and externally, and would help to use more advanced working tools. This study only presents the methods included to the approach and defines how to utilize them, if it would be implemented. The actual field test is not in the scope of this study and thereby left out.

The scope of this study concentrates on working with third party items, from commercial component perspective, and therefore, all documents linked to these items. The study will also give perspective of different departments which are involved in itemization processes. It will propose what could be changed and further developed with the help of better cooperation and enhanced internal visibility for producing better quality information for items.

This Thesis is written in 7 sections. The first section is the introduction to the Thesis. The second section describes the research methods. Section 3 includes the current state analysis. Section 4 presents discovers of existing literature and best knowledge that supports made improvements. Section 5 is about building new improvement. Section 6 will present validation results. Section 7 summarizes, presents observations and discusses about the conclusions of the Thesis.

2 Method and Material

This section discusses about the research design of this Thesis, describes the data collection and analysis, and the validity and reliability plan.

2.1 Research Approach

In this study, a research approach is qualitative case study which simplifies the examination of a phenomenon within its context utilizing multiple different data sources. Regarding to Yin (2003), a case study design should be considered when: (a) the focus of the study is to answer “how” and “why” questions; (b) you cannot manipulate the behavior of those involved in the study; (c) you want to cover contextual conditions because you believe they are relevant to the phenomenon under study; or (d) the boundaries are not clear between the phenomenon and context. (Baxter and Jack 2008: 545)

In this study, the research and the process improvement is executed for the organization and with help of different employees from the case company. The employees consist of people from different departments of the organization whose processes involve usage of items. This study answers *why* the itemization process should be improved in the case company, and *how* the improvement could be done and what would be the benefit of that improvement and to who. It is also important to understand the meaning and what are the contextual conditions that relates to existing phenomenon.

Commonly, if a case study includes a specific proposal, it increases the liability for the researcher to be able to frame the scope of the study and thereby increase the usefulness to complete the project. When the study contains a few specific proposals, the more likely it stays within the workable limits but when the study contains several proposals they all must have a distinct focus and purpose. The proposals may come from the literature, personal/professional experience, theories or generalizations based on experimental data. Data sources could include documentation, archival records, interviews, physical artifacts, direct conclusions, and participant-conclusions. (Baxter and Jack 2008: 551-554)

In this study, the idea is to search and discover solutions and utilize existing knowledge that would be available for Component Engineering to improve its current processes and

information knowledge which could be utilized in the itemization processes. Using this kind of research method helps to gain more understanding how data can be effectively managed.

2.2 Research Design

The research design of this study aims to create research approach to find solution and give proposal of improvement for the current Component Engineering itemization process. Figure 2 presents the research design of this study.

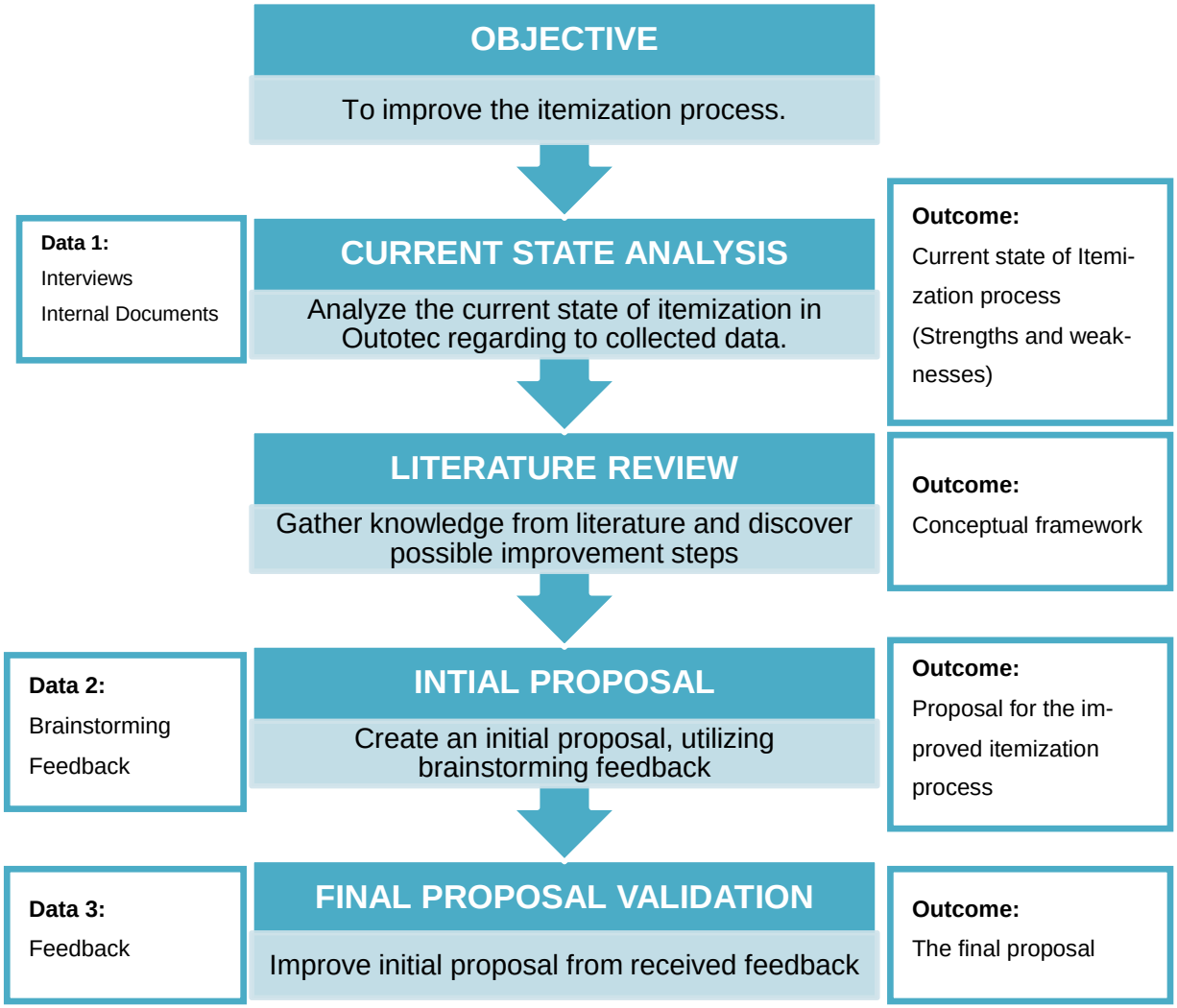


Figure 2. The research design.

As can be seen in Figure 2, the research is divided into stages: (a) analyzing the current state of itemization process and discover its strengths and weaknesses (b) search for existing knowledge from literature for conceptual framework, (c) create initial proposal, and (d) validate the final proposal.

Stage 1, *current state analysis*. After identifying the business challenge, the actual problems in the current processes are analyzed how itemization is currently working. Current state analyze is conducted by interviewing managers from four different departments, one manager of each. Based on the interviews strengths and weaknesses of current itemization process is listed.

Stage 2, *literature review*. Existing models and knowledge for processes improvements are discovered from existing literature and utilized to form conceptual framework.

Stage 3, *initial proposal*. An initial proposal for the new process is formulated based on the data collection from interviews triggered by ideas from the conceptual framework. Several discussions with key manager and external consultants are held for planning and brain storming new ideas for the initial proposal.

Stage 4, *final proposal validation*. Potential improvements are done for initial proposal regarding feedback from key manager and few external consultants. From this feedback and discussion the improvements are suggested to the final proposal.

2.3 Data Collection and Analysis

In this study, the data is collected from several data sources: interviews and internal documents regarding to item concept and processes, and currently used tools. The data for this study is collected in three stages: (Data 1) a current state analysis, (Data 2) build an initial proposal model for new improvement and (Data 3) discussion of improvements for final proposal and validation from the management. Most of the data is collected through interviews and discussions. Table 1 presents detailed collection of the data for this study.

Table 1. Description of collected data 1, 2 and 3.

DATA	PURPOSE	DATA TYPE	DATA SOURCE	RESULT	ANALYSIS
Data 1	Identifying key S&W	1. Internal documents	1. Process map 2. Item Concept	- Description of current state of itemization process - paradigm from current process	Section 3, CSA
		2. Interviews with internal customers - Audio recording & memos	- Interviewee 1, Item Data Manager - Interviewee 2, Engineering Manager - Interviewee 3, Quotation Support Manager - Interviewee 4, Head of Automation & Instrumentations - Interviewee 5, Director - Supply & Distribution Network, Markets Unit/Spare & Wear parts	- How itemization is seen and items used in different departments	
Data 2	Building the proposal	1. Brainstorming sessions and feedback – field notes	1. Discussion with key manager 2. Discussion with customer(s)	- Gather best practice ideas to be implemented for new proposal	Section 5, Building the proposal
Data 3	Validating the proposal	1. Feedback from– audio records & field notes	1. Discussion with customer(s) 2. Discussion with key manager	- Comment round for final proposal and last changes	Section 6, Validation

As seen in Table 1, data collection was gathered for Data 1, 2 and 3. This data was collected with wide range and persons for interviews selected from different departments to get information how they see information affecting to their field of work in the company. This gives better understanding of current behavior in different transactions. The gathered data ranges from specialists to middle management level, by interviewing employees from different departments, meetings and discussions. Key data sources are discussed in more detail below.

Interviews, meetings

Interviews for Data 1 are based to general interview design. It consists of pre-constructed questions but they are used flexible to allow more questions or change questions depends what the interviewee responses and how the discussion develops. Table 2 below shows detailed information regarding the accomplished interviews and discussions.

Table 2. Data collection details.

	Informant	Data	Data type	Answers	Date	Duration	FN = Field notes TR = Tape recording
1	Item Data Manager	1 2 3	a) interview b) feedback c) feedback	Appendix 3 Appendix 8 Appendix 9	a) 5.2.2016 b) 22.4.2016 c) 26.4.2016	a) 66min b) 60min c) 60min	a) FN+TR b) FN+TR c) FN+R
2	Service Quotation Manager	1	a) interview	Appendix 4	a) 18.2.2016	a) 85min	a) FN+TR
3	Senior Product Engineer	1	a) interview	Appendix 5	a) 23.2.2016	a)105min	a) FN+TR
4	Head of Automation & Instrumentation	1	a) interview	Appendix 6	a) 8.3.2016	a) 75min	a) FN+TR
5	Director - Supply & Distribution Network, Markets Unit/Spare & Wear parts	1	a) interview	Appendix 7	a) 22.3.2016	a) 80min	a) FN+TR
6	Internal customer 1	3	a) feedback	Appendix 9	a) 26.4.2016	a) 60min	a) FN+TR
7	Internal customer 2	2 3	a) feedback b) feedback	Appendix 8 Appendix 9	a) 26.4.2016 b) 27.4.2016	a) 60min b) 30min	a) FN+TR b) FN+TR

As shown in Table 2, information referring to data collection. A total of five interviews are carried out and selected persons are working for the case company Outotec. Interviewees are selected from different areas of delivery chain to have perspective from the beginning and the end of the item lifecycle usage, as well as the connection to their itemization processes.

First interview was done with *Item data manager*, who has the best knowledge of current state of component engineering and a vision of the future targets and idea what needs to be done to achieve them. He was capable to give perspective from itemization concept and what were related to it and a general overview from company's whole product chain and how items are planned to use as part of product lifecycle management from designed product to delivered service solution.

Second interviewee, *a Quotation support manager* was interviewed to have perspective from spare parts business and quotations related issues. These answers were concentrated also to service business and clarify how business works in service department and how current tools and processes were working, and regarding to Request For Quotation (RFQ) procedures.

Thirdly, *a Senior product designer* from engineering gave answers how they are using items in their daily processes and what are positive and negative things that they have confront, after usage of items began in 2013. Interesting answer was also that what could be the possible bottle necks that are currently present or could be escalated. Solutions and opinion what Component Engineering is offering as a service to Engineering and what could be done better. Big question was that were there anything which could be improved.

Fourth interviewee, *a Head of Automation and Instrumentation* who was interviewed especially to have perspective from itemization and how its shown in instrument level which includes measuring equipments. Conversation also brought out knowledge how instrumentation uses items and what party actual should take responsible of certain type of items e.g. in third party item cases. Key question was how itemization process was currently supporting their operative work.

Fifth interviewee, *a Director of Service Spare Parts, Marketing and Sales* was selected to have best perspective from marketing and sales department. Idea was to have answers how spare part team uses items, how their process functions and what information was really needed for customer spare part delivery. Big question was above all how Component Engineering is supporting spare part service department from the itemization perspective and do they have useful information stored in the data base, which can be used for item information verification.

Pre-constructed questions consists eleven same questions for all interviewees regarding to itemization and use of items in Outotec. Expectation was that answers vary in certain questions and depending of the person who is interviewed. Question were considered carefully and left open enough in a way that answers can be vary without leading too much to any certain direction (see interview questions in Appendix 1). All interviews consisted field notes and tape recordings and each are transcribed in summarized answers.

In the proposal stage, Data 2 consisted of meetings with key manager and external consultants from the case company. Along with discovered literature for conceptual framework, the initial proposal was developed and presented in the fifth section of the study. In

the final phase Data 3 consisted of feedback and face-to-face discussion. Purpose was to have a feedback from initial proposal and analyze it by going trough comments from key manager and few external consultants in the case company and summarize the results and clarifying details for the final proposal.

All the data from interviews was documented by using field notes and tape recordings. Data 1,2 and 3 are summarized as field notes and written document and added as appendixes in this thesis work. Data was analyzed using content and thematic analysis to discover joint reflections that also can be shared with the case company.

Tools used in the Itemization process

Many different systems which are defined as tools are used for itemization work. Some interview questions (Appendix 1) also relates to daily operated systems and how they were used in the case company. The purpose of these questions was to build common knowledge of what and how item information were used in these systems and which information especially is rigorous to achieve flawless operative workflow.

2.4 Validity and Reliability Plan

It is important for reader and researcher to understand how the validation and reliability of the study is ensured. The validity can be divided into four different types: construct validity; internal validity; external validity and reliability. Even though research is not a positivist form of case study it can still address similar concerns of rigor. Bottom line is that a study needs to include a discussion about aspects related to the validity and reliability of the done research (Näslund, Kale and Paulraj 2010: 338).

Three tactics could be used to increase *construct validity*. The first one is the use of multiple sources of evidence, secondary establish a chain of evidence and as thirdly to have the draft research reviewed by key informants (Yin 2003: 34-35).

In this study, challenges for qualitative study are to achieve proper internal validity as it does not measure anything concrete in numbers as it would be the case with quantitative study.

Internal validity can be seen as a particular strength of qualitative research in general, because the all gathered data itself could be sufficient to tell something regarding to the subject of the study (Quinton and Smallbone 2006: 128). According to Näslund, Kale and Paulraj (2010) to able perform relevant and rigorous research, researcher has to define relevant questions based on observation done in the case company and engage answers in a rigorous approach regarding to questions.

External validity relates to replication logic which would allow result of the study applied to other contexts or quite similar processes (Yin 2003: 36). These could potentially improve analysis and findings and as well as reduce problems with validity particularly with relation to authenticity and trustworthiness (Näslund, Kale and Paulraj 2010: 339).

In this study, to reduce biases and improve level of analysis, the research process includes of key people from the case organization and letting them elaborate and give critically feedback on the results. To present a more valid analysis, the study uses diagrams to provide readers a better understanding of the study. For this validation internal validity is achieved by interviewing employees from several different departments inside the case company who are closely related to this case study issue and write down field notes based to these conversations. These departments cover most of the business areas where itemization is used for. The interviews and feedbacks are seen as key points in this study to achieve a proper level of internal validity as explanation building can be done base to these answers.

Reliability is usually seen as a question that if this research would be repeated, could same results be achieved or if someone else conducted it (Quinton and Smallbone 2006: 129). This means that by documenting the used practices it would give possibility to replicate the methodology a second time (Yin 2003: 36).

In this study, reliability is improved by conducting a careful current state analysis. Reliability also influenced by the fact that the author works for the case company as a consultant as part of Product data management team and has been part of the team for few years and has relevant knowledge of current situation in the case company to be able to generate valid questions for interviewees. The questions is the same for every interviewee and the results are analyzed, summarized and logged as field notes, and the relevant documents are stored in order to retrieve and audit the research data later. The integrity and progress of the research project in general was also observed in regular status meetings with item data manager, where critical feedback can be utilized for the improvements of the research.

The validity and reliability of this study are again discussed in Section 7 with conclusions to verify the aspects discussed in this section. The next section discusses the current state of the itemization process in the case company to presenting the existing situation and identifying processes strengths and weaknesses for further presented improvement proposal.

3 Current State Analysis

This section reports the current state analysis of the itemization processes in the case company. The analysis aims to find results of strengths and weaknesses of the itemization process, and thereby form a focus for process improvement.

3.1 Overview of the CSA Procedure

Objective of this study is to improve itemization and current state analysis (CSA) section will support this by interviewing different departments and grow the knowledge of its capabilities to improve customer communication and co-operation with different departments which are useful for commonweal.

The targets for the CSA are to find answers to the following questions: (a) how component engineering's current model supports other departments itemization, (b) how items are used in other departments and their processes, (c) what are the bottlenecks of current itemization, (d) how the current guidelines, tools and support functions in different departments.

The CSA of this study consists of gathered information from interviews and information discovered from the existing documents of the current Library team's itemization process. As the interviewees for DATA 1 were from different department they had clearly their own perspective of things which was the wanted result for this study. Answers gave a good versatile picture how itemization is seen in different departments of the company in the current state. There was a big difference relating to answers, which depended in what department the interviewee was working in. As the question concerned itemization there was clearly less to answer what benefits itemization has brought or would bring in the future. It was interesting to notice how much variance there was between answers relating to working position, as it showed that for this case study the most valuable answers were received from people in higher position in the company.

The most valuable answer for Component Engineering point of view where received from *Service department*. There were many things in their processes that were unclear which

related to item information storage and way to manage it. Answers generated more questions relating to case study objective to improve itemization process.

Automation and instrumentation department have challenging situation regarding to ongoing itemization process in the company. The department does not use items in their daily work at all which also created a challenge for interview perspective. The challenge was accepted and questions had to be revised on the spot, to suit more of the situation without missing to objective where this study aims. A good overview was received of their department's current state and purchase processes and where the information is stored and why current itemization is seen as too complicated for their operative work.

Answer from *Product line* where positive as well as from others. They clearly had developed some standard processes to manage certain itemization procedures, like commercial component datasheet request for supplier, should be done and take it on use. They had good picture what benefits itemization has brought to project management; cost monitoring and reuse of items.

3.2 Current State of Itemization in the Case Company

The itemization process which includes creation and maintenance of the items in the case company is conducted by the Component Engineering department. The Component Engineering department manages third party item library in PDM, engineering data management (EDM), and ERP systems. Its core competences are item information quality level management, determination process management, user training, production of guidelines for item creation, and global third party itemization library management.

The current state of Component Engineering processes has been developing all the time from beginning of 2014 and for this study the CSA is presented as it is in January 2016. The implementation idea for itemization is a part of operational excellence (OPAL) process which was started at the end of 2011 and should have been completed at the end of 2014. However the schedule was only estimation and the process is still ongoing and should be ended in summer of 2016. The Product Life Cycle Management includes digitalizing the company products globally and creating items for every part that are included in the projects. At the beginning an actual Component Engineering was not supporting Engineering

as much as it could be. This was a big fault in the previous operational model which was defined first and which is considered to be changed. During few years Component Engineering department role has increased and other departments need for itemization support has been grown. To meet this demand a temporary branch; project itemization was also created to support Component Engineering and product lines for mass request to itemize projects for manufacturing purposes. Currently project itemization helps Library team with mass maintain requests and prepares and enriches items for item specialists for approval process. As this study concentrates to improve itemization processes, were Component Engineering plays major role, there has to be done research to discover how much actually Component Engineering department has influenced in other departments in Outotec.

The Component Engineering team consists of eight item specialists, library manager and item data manager. Item data manager also is responsible of the global item concept regarding to third party items. The Team is responsible for *standard* and *commercial item* library that contains all Outotec's commercial (third party) components used in designing products and service spare part sales. The team itemizes and maintains information that is important for the current and future service business in Outotec's core business areas. Itemization provides information and services for all employees inside the company who are using items in their daily work (in other words, Library team's customers). Table 3 presents the current Library team and its responsibility area for each member.

Table 3. Current Component Engineering's Library team roles and responsibility area.

Area / Role	Main responsible	Guideline responsible	Operative responsible	Secondary operative responsible
Raw & Bulk materials	1	1	2	10
Standard & Bulk Mechanical Components	1	1	2	10
Electrical & Bulk Electrical Components	4	4	3	
Mechanical & Bulk Components	5	5	6	2
EDM Admin	8	8	8	14
Duplicate Management	9	11	Antti Korpiniemi	12
Training	10	10	10	13
Concept	11	10	Antti Korpiniemi	12

As seen from Table 3, individuals are indicated only with a number and each number relates to one worker i.e. one worker can have several roles in different fields. Exceptionally author of this study is presented by name and has operative responsible of duplicate man-

agement and itemization concept. Areas are divided to seven different segments which consist of four item type areas, EDM library, duplicate item management, training and concept.

For different type of components there are specified specialists who will have responsibility to approve items into library. Itemized components can vary a lot for example by material, mechanical or electrical usability. In addition there is a difference that is the component standard product or is it customized someway by the manufacturer, and in which application it is used.

The Component Engineering's itemization work consists of *the Itemization process* (see the map below; Figure 3) which includes *a set of integrated tools.* The steps and tools are described in more detailed level in the next sub section.

The base idea for the PDM Standard and Commercial library is to hold all common commercial and standard items used globally in Outotec. All items in the library are processed through quality check and are controlled by Component Engineering department. Library is divided into a hierarchical class structure with three stages and it is maintained in PDM system. Engineering class instructions and guidelines of item information requirements are maintained in master data management (MDM) system called EBX to provide one place to find and upkeep needed information for creation of the items.

The library is the master container for all third party items owned by Component Engineering and it will be referenced to all other locations and systems utilizing commercial and standard items. Library items should be identified primarily using original manufacturer and manufacturer code for commercial components and global standards for standard components to support global use of items. In special occasion item can be identified using supplier information and will be thereby defined as vendor item.

The processes for itemization are currently running and in daily use, but there are still many open questions how things should be done and what kind of guideline and process should be used to execute those actions. Itemization concept is developing all the time, and the main goal is to achieve level of confirmed information for items used in service

business. The lack of confirmed information also creates problems to item determination processes and change management, as the information in the item descriptions is the only way to define possible relation between different items.

3.2.1 Map of the Current Itemization Process

Component Engineering's item release process consists of two different channels which are producing tasks to team members. These tasks are taken care of by Component Engineering library team and project itemization. Itemization team has the same core guideline for both new item creation and maintenance processes for third party items.

In both cases the itemization is based on a customer request or proposal. *Requests* are for new item creation, maintain or replacement of items. *Proposals* are items which customers have created to PDM system manually or from EDM and promoted item to preview state of items lifecycle; this action will list item proposals to itemization team's backlog for future validation process. Figure 3 shows the map of the current itemization process.

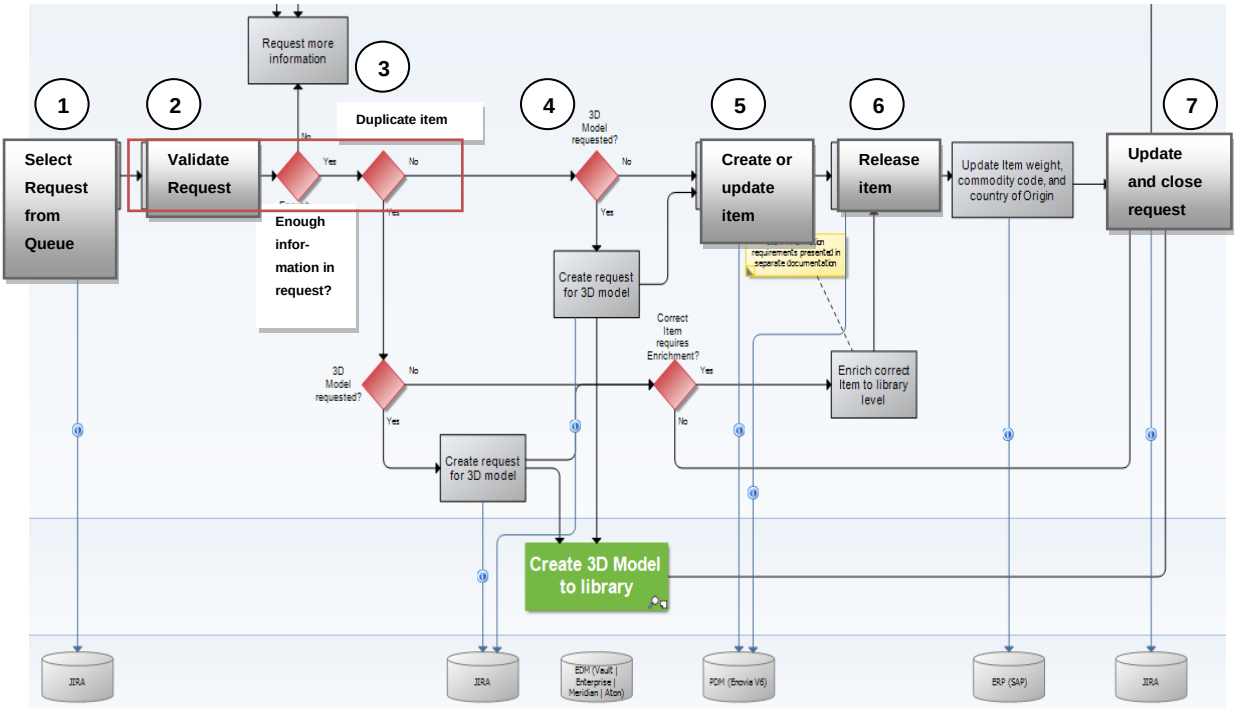


Figure 3. Map of the current itemization process (Data 1, Process chart 1).

As shown in Figure 3 above, the current itemization process consists of seven main steps. Step 1, using selection form users create item request which create two queues for itemization team to process. Requests are monitored from two different channels JIRA (a proprietary issue tracking software) and EBX backlog. To JIRA, requests are in ticket format and can contain creation, maintain and replace requests for different types of items and 3D models. The 3D model requests are always managed as sub-tasks and they are done by EDM Library specialist, which are clarified in step 4. These requests can be initiated by all users from Outotec, mainly designers, procurement and service. JIRA ticket requires mandatory information for created or maintained items and which will be used for item validation process and this means that requestor does not get the needed item number before the request has been completed in the new item creation case. Items which are created straight to Enovia and promoted to review state by users are monitored in EBX backlog. These are managed as proposals which mostly designers have created when they needed new item in the ongoing product structure building phase.

Next, in Step 2, an important request validation will be done in this stage. To validate request, item specialist will first check correctness and adequacy of item information e.g. manufacturer code against attached datasheet, which is mandatory attachment for commercial items. If item request does not meet the adequacy information an item specialist will make an additional information request which will be sent as an e-mail to original requester via JIRA. Additional information request has its own status in JIRA and item specialist will take request under work again after correct answer has been received. Figure 4, presents a simplified model of validation process for step 2. For new item creation and brown field item enrichment the procedure is little different. A new third party commercial item only needs a datasheet to verify the defined manufacturer code and after duplicate review it can be selected as a master item and release to library. In contrast the item maintain situation is different, as used item number already exists and it will need a verification of the contained information and this focuses verification of the manufacturer code with purchase or quotation order. Otherwise item cannot be recognized, in order to say that the existing legacy item really used to purchase the correct component that is defined in the description.

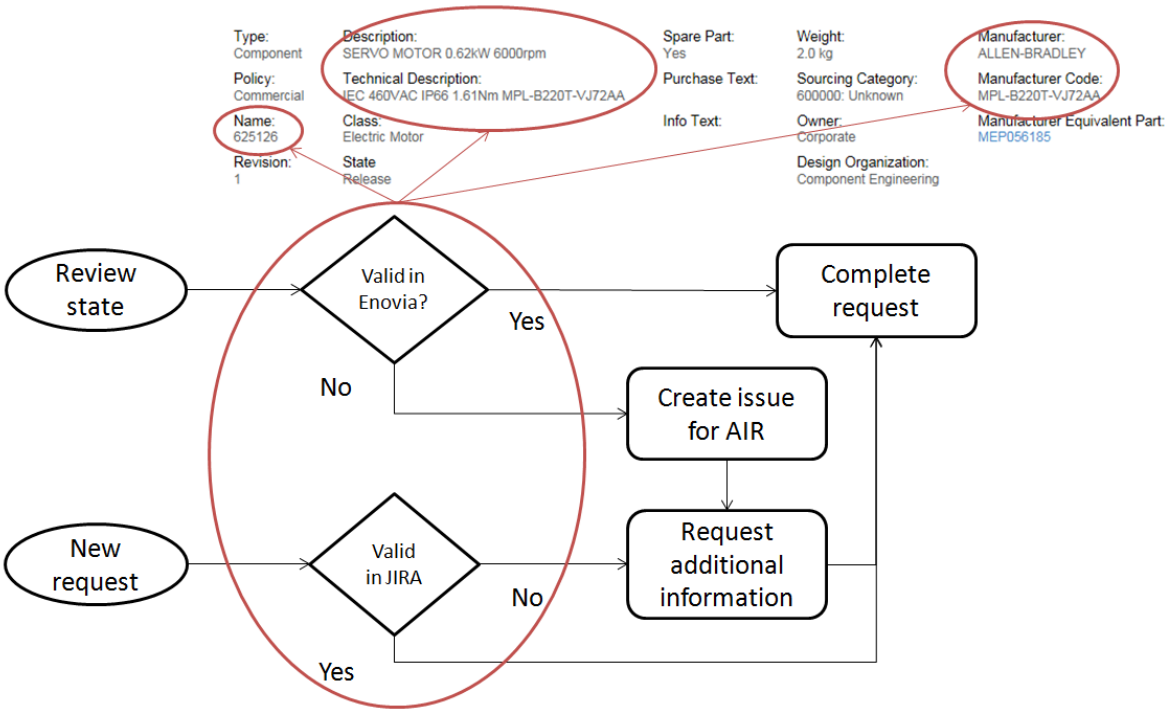


Figure 4. Item validation process to complete request (Data 1, Process chart 2).

As shown in Figure 4, third party item proposals coming from PDM are also validated by checking the correctness of information defined to item description field, datasheet and manufacturer part number. If information is valid a duplicate check will be done for proposed item and it will be released to library if it is unique. This action will be described in step 3.

Next, in Step 3, if the information described in the item can be verified, a duplicate check for the item will be done in the next step. This means that item information e.g. manufacturer and manufacture code or standard (e.g. ISO or DIN) and material standard will be used to search duplicate items among all master data that exists in the EBX system. If duplicate item or items will be found, a more specific comparison has to be done and will include master item selection procedure which will remain only one item to be used. This refers only for maintain item requests. If the new proposed item is a duplicate, it will be obsolete in PDM and JIRA ticket will be rejected and indicated to use correct item which already exists in the system. Before answering to the request item specialist will confirm that existing item is in the Standard and Commercial library and meet's required quality

level. If item needs to be maintained, this will be done by going through normal enrichment procedure. If item is unique in JIRA item specialist takes items under work and does next steps to enable that items can be released to PDM systems third party component library, this will be described more accurately in steps 5 and 6.

Next, in Step 4, after request validation and duplicate check, if request has contained also request of 3D-model creation it will be taken under work by EDM library specialist who will manage sub-task of the parent request. Before it is possible to complete parent request in JIRA all sub-tasks have to be completed. The specialist role depending of the request is to maintain 3D-model in Solid Works or Inventor and updated it in a way that it can be used in production. Although the main task is to add item number in EDM to existing 3D model or its configuration to match correct item in PDM library.

Next, in Step 5 and 6, item specialist will create new item or maintain existing legacy item by following item guidelines. In the new item creation all information will be described as defined in the guidelines, the datasheet will be connected with document object and manufacturer and manufacturer code are added by using special object which is connected to item. The process includes selection for valid engineering class which is selected subject to what the product or component is and is it mechanical or electrical device. Item type and policy are also critical to be defined correctly because different types are affecting how item can be used in PDM and EDM systems as the unit of measure will change. Item policy will define its lifecycle behavior and mandatory information fields and it is different for standard, commercial and bulk items. After these selections item will finally be attached to global plants and released. When item is at release state the information will transferred to ERP system via integration from PDM.

Next, in Step 7 item specialist closes ticket in JIRA, reply to user via e-mail that item is released to library and ready for use, and then defines used time for processed ticket. If request would have been invalid in step 2 ticket would have been rejected and reason of rejection would have been sent to requester via e-mail.

The project itemization executes also same processes as library team from step 1 to 5, but they are not able to promote items in release state. They will only enrich items to stage where library item specialist will do the final check and release item to production which means that after release item is ready to be used in design structures and purchases in ERP. The library team can also assign tasks from received tickets for project itemization to maintain and create items.

Request for items comes mainly from product line designers, engineering and spare part service. Minority of the request comes from pricing, instrumentation and manufacturing sites. Communication between customers and library team is managed via JIRA, e-mail or in some cases using online conversation. Mainly e-mail is preferred through JIRA because of its ability to store history of made actions and information change towards customers.

3.2.2 Definition of Process for Item Type and Policy

For operative processes library team and all the other departments using third party and design items always has to use specific criteria to define correct type and policy for items. This is really important as with help of the guideline item can be defined for correct unit of measure, engineering class, and responsible design organization (RDO) which relates to policy definition that affect’s on items behaviour also in PDM, EDM and ERP systems and allows different transactions and information usability for items.

Currently guideline does not give exact definition to all cases which leads to situation where there is possibility that it’s difficult to define correct responsibility organization for the item because responsibility is divided to either Component Engineering or Product line who owns the design product. Figure 5 shows the current itemization process steps for library item type and policy definition.

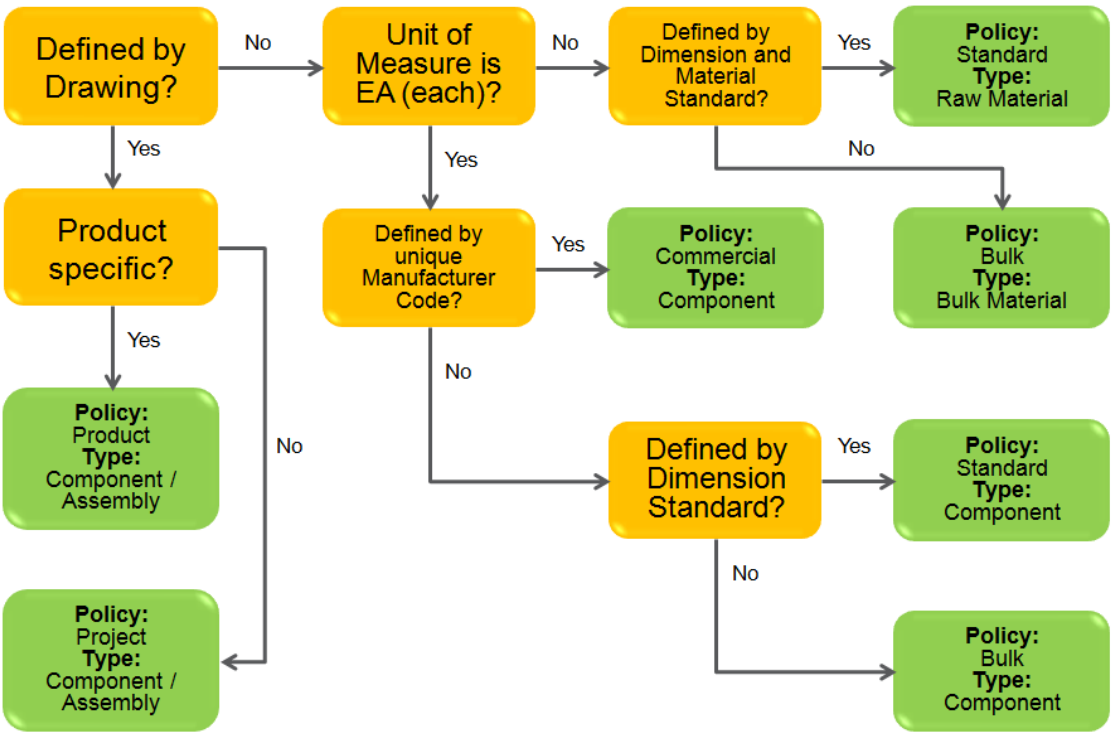


Figure 5. Defined steps for library item type and policy selection (Data 1, Process chart 3).

As show in Figure 5, items can be defined to certain type and policy by described information and documents that are included them. These definitions are done and checked for all items during either creation or maintain process. The type: component, assembly or raw material, specifies the basic purpose of the item in the item structure. The item policy: commercial, standard or bulk, defines information requirements for the item, and the behavior of the item regarding to characteristics of the item e.g. the amount of item lifecycle stages (Data 1, Item Concept). They have important role regarding to item lifecycle which gives overview of item's creation, production use and obsolescence process.

3.2.3 Overview of the Tools Used in the Current Itemization Process

Used soft ware solutions for itemization are important for functional daily operations. For Component Engineering the main tools are product data management software (Enovia), master data management (MDM) EBX which is used as item master, issue and project tracking software JIRA and enterprise resource planning (SAP) for business and ware-

house management. Figure 6 shows systems which are currently used in itemization processes in the case company.

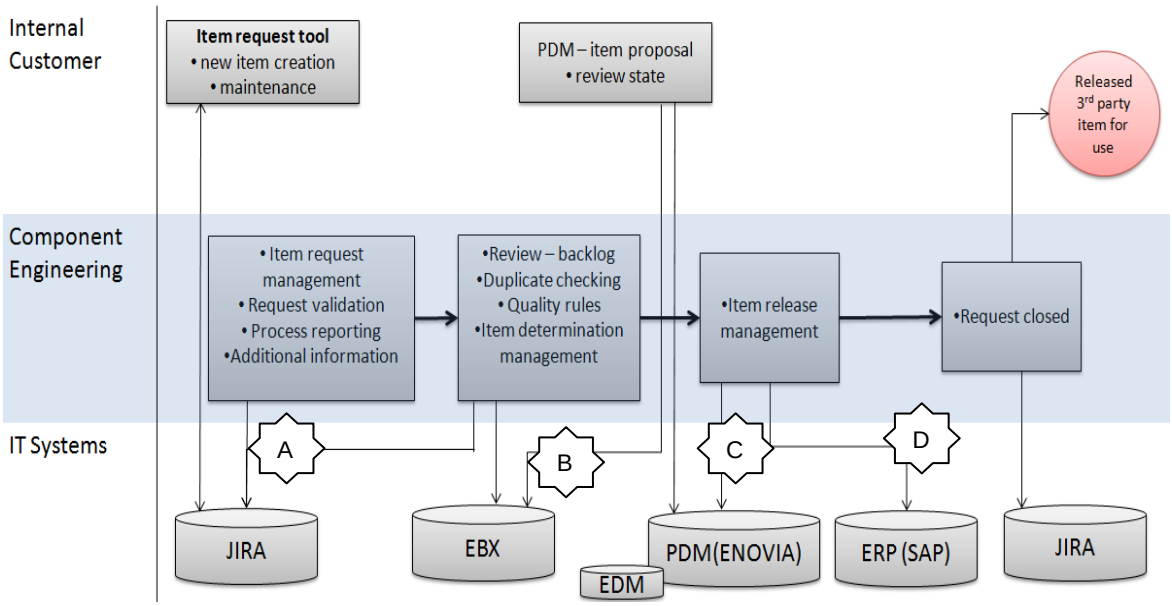


Figure 6. IT systems used for daily practical itemization work.

As shown in Figure 6, different systems are needed to manage different steps of the itemization process.

a. JIRA as the tool for Steps 1 to 4 and 7 is used for issues management. Internal customers use an item request tool to create request which are managed in JIRA. The system offers capability to track working time, assign tickets and communication between item specialist and customer via email. Validation is done regarding the information that is included in the ticket.

b. EBX as the tool for Steps 2 and 3 is the most efficient tool for validation and item determination management. Duplicate checks for items are made with EBX which automatically search possible matches for item data base according to defined manufacturer code, standard or material in the item description fields. As a result from EBX match item can be taken to enrichment process and released to production. EBX is used as item master as it contains master data for the items used in

Enovia, ERP and old legacy systems. All relevant data for itemizing work will be uploaded to EBX where it can be structured in a way that is will be useful and add information knowledge of the items. With help of this system item specialist are able to find possible duplicate items, control item library classification guidelines, process and follow item determination, select master item and manage harmonization work done for the items.

c. ENOVIA is the PDM system which basically includes all actions and lifecycle regarding items which are related to designed products. Standard and commercial items library is managed from Enovia. The system is the master tool for step 5 and 6 will be used for operative work. As a master system all item creation and maintain are done in Enovia. From Enovia item information is transferred via integration to ERP (SAP). A generic import tool (GIT) is used for mass item requests and updates in Enovia and it is the only tool for this purpose excluding administrator tools. The whole item lifecycle can be managed in Enovia and currently third party item has five different lifecycle stages: Created, Review, Approved, Release and Obsolete. Each of these stages will support the item handling process. Enovia has integration with EDM which is used to 3D modelling. All new products are designed in EDM and structure is uploaded to Enovia integration. For these actions both PDM and EDM systems have to have same item codes for third party items existing to work properly.

d. ERP (SAP) as the tool for Steps 6 and 7. When item is processed in Enovia and promoted to release state it will be transferred to SAP. Important item information like commodity code, country of origin, lead time, prices, old material number and local purchase text for service business are added straight to SAP by service department. When item is available in SAP it can be used for quotation requests, purchasing and deliveries. Most of the transactions are from service and projects business. Some of the existing order confirmations also exist in SAP in PDF format. They are attached to purchase orders that are linked to item used in specific projects.

3.2.4 Current Performance of Itemization Process

Only rough estimates of time handling and expenses can be given from current processes. Currently there is not established way to measure itemization work and present accurate costs of itemization. It is unknown how much an item costs to case company and it also varies between different item policies.

As there are two different channels for Component Engineering itemization processes. The current request handling process is working fine without any major conflicts. Ideal request handling time for single request varies between 15 minutes to few hours depending of the requested item and available information. Single item handling time from work backlog is different than JIRA request handling time as item proposal is processed and actual item exists already and it only should be validated and maintained is necessary.

The standard and commercial library has been reclassified recently and engineering classification has been simplified for component by reducing the amount of engineering subclasses. Legacy classification had over thousand different engineering classes and new one only under three hundred. This has been a good improvement and has enhanced the process of correct item engineering class defining and reduced time used for engineering class selection by item specialists.

3.3 Bottlenecks in the Current Itemization Process

Current itemization process has its bottle necks which should be improved at some point to enhance item usability in the case company.

First of all, to receive confirmed information for the commercial components the process demand certain amount of documentation e.g. datasheet, purchase order, quotation order or order confirmation which could be used for information verification and help to connect them to some context.

Second, item lifecycle has shortage of specific state which would be helpful for product- and spare part management. E.g. at the moment there is no controlled ramp down for items and therefore existing replacement determination process is difficult to manage.

Thirdly, lack of training and inefficient use of instructions creates misunderstandings which are also decrease current performance and slows down response to requests as the additional information is needed. Most of the misunderstandings come along from wrong working method, unclear email messages and lacks of required information.

Fourthly, the current storage for guidelines and instructions are managed in EBX, Item management Wiki and Operational Model Handbook (company's own intra sites), network hard drive and personal hard drives. A problem occurs from too many channels were the item guideline is stored in different systems. Upkeep of all these places is really challenging and responsibility of the update is also uncertain. Figure 7 shows the most critical bottlenecks spot where improvement should be made in the itemization process.

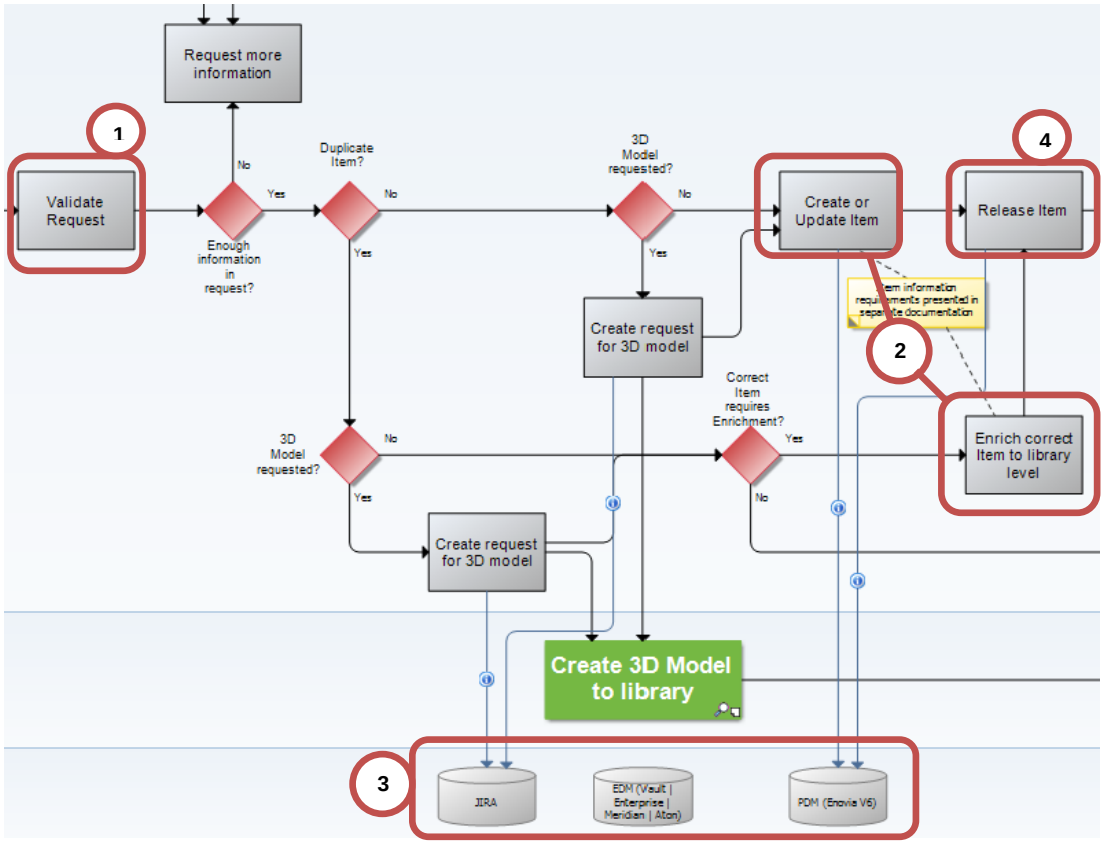


Figure 7. Map of bottlenecks in current itemization process (Data 1, Process chart 4).

As seen from Figure 7, the main challenges are appearing mainly from processes but also used systems have a lot to be improved. Following challenges are described regarding five steps in Figure 7.

Challenges related to Step 1

The most vital step is validation of proposed or requested item because it demands knowledge and accuracy to make correct observation regarding to request or proposed third party item. If there is lack of correct information and it is not noticed a possibility to create or maintain duplicate item is highly possible. This problem occurs especially when component has clear manufacturer code but the code actually contains values that should be specified separately and most likely are order specific. To support validation process and decision making, the information contained in the item description and documentation must be available and correct to support the decision making i.e. information has to have certain level of quality.

Regarding to interview answers, guidelines and instructions are not complete and contain steps that are not defined. At the moment tools are distributed to places where they are hard to find. Still is seen, that there is lack of information in instructions and they are not clear and change management should be therefore improved. Continuous training would be seen as a key to help to perform daily practices in more effective way and support co-operation between departments.

Challenges related to Step 2

After validation and regarding to section 3.2.2, for PDM and EDM purposes items type and policy are needed to be defined correctly and this definition is quite challenging time to time. By selecting specific type the responsibility of the item cannot be Component Engineering instead it will moved under product line responsibility and this means that item is not a part for standard and commercial library. Challenge is that item type and policy should be considered from engineering perspective which defines also items UoM (Unit of measure) that will affect its behavior in PDM and EDM systems. Manufacturer part number has major role defining correct policy as the existing code has to be unambiguous to define component as commercial item. After the actual enrichment or creation process is

done a challenge is to find proper engineering class from the standard and commercial library for the item. There can be situation when component's function is hard to define to enable the selection of correct engineering class for the item.

Interviewees also saw that there is a need to have new item policies to support item management in a way that it would help to classify items easier and define correct responsibility organization for the item and it would clarify what items actually would be necessary to classify in the Standard and Commercial library.

Challenges related to Step 3

Current systems are not containing all needed information in one system which would be used for item creation, maintaining and decision making of master items in duplicate item cases. Integrations are missing between different systems and many data tables in EBX are updated manually by uploading exported excel to system. Also information is still stored to several different systems with limited access to it. Systems are used only by certain departments and legacy data is still stored to legacy systems instead of migrating or storing them to one place where they could be available for everyone how would have benefit to access that information. This challenge relates also to stored item guidelines and instructions which are placed in many different places for different people to use. This causes challenges to update those documents and reduces their accountability.

Most of the interviewees answered that Enovia key user concept was seen as good way to share information effectively between employees. As there should be key user for every system it is not working for Enovia key users with current setup because selected persons do not have needed time to do the practical work or they are not trained enough.

Challenges related to Step 4

After the release the items are available in ERP which they can be used for quotations, purchase and delivery purposes. As standard and commercial library contains mainly spare parts they should always be connected to some design structure e.g. product. This means that all items should be able to be connected to some context which relates use of the items in certain sites, plants, products and equipments. The challenge is that at least all

legacy (brown field) items should always be connected to some context. In principle this should be possible in every case when legacy item's information is verified because the information and connection should be founded from the existing documents.

There was not so clear visibility from respondents what comes to Component Engineering supporting for different departments in the organization. This occurs mainly from different way of using items among product lines and lack of standardized design procedures. Even thought productization along with standardization was core things to develop in OPAL which was referred in section 1.2, the actual deployment has not taken its place in practice as it should have been.

Table 4. Summary of identified problems.

Area	Data validation	Library management	Tools for item enrichment	Guidelines / Instructions	Master data management	Cost and Metrics
Object	-Component proposals from PDM -JIRA requests -Global price list items -Brown field items -Preferred by supply items -eCatalogue -Stock level per item -Legacy systems data -Transactions	-Vendor item classification -Obsolete item classification -Class definitions -Amount of different class	-System reliability -Continuous work -Used in processes	-Presented in common place -Share point -Item Management Wiki -Training videos/ presentations	-Management (schedule, documentation) of existing and new scopes for maintain processes -Determination and obsolete process -Prioritization of scopes -Item information verification -Data Governance	-Time tracker for JIRA -> for cost estimation -Average amount of items managed in certain time
Component	-Datasheets from suppliers/manufacturers -Purchase orders -Quotation orders -Documentation -SAP change history	-Library for Vendor items	-Generic import tool update for mass item updates -Need for more integrated systems	-Key User concept	-Data information management process -Duplicate recognition -Item lifecycle -Item policy -Connection to context	-Lead time (process time)

Table 4 above shows how problem areas were divided roughly to six areas and what object where relate to those areas and what specific components possible causes the challenges in the current itemization process. Based to interviews and own observations most of the challenges appear from data management and data validation. These are closely connected together as they have many process relating dependencies between them. More of these problem areas will be clarified in section 3.7.

3.4 Other Challenges Revealed by Departments

Automation and instrumentation department have challenging situation regarding to ongoing itemization process in the company. This department does not use items in their daily work at all which also created a challenge for interview perspective. The challenge was accepted and questions had to be revised to suit more of the situation without missing to objective where this study aims. A good overview was received of their department's current state and purchase processes and where the information is stored and why current itemization is seen as too complicated for their operative work.

Answer from product line where positive as well as from others. They clearly had developed standard process how certain itemization process, like commercial component datasheet request for supplier, should be done and take it on use. They had good picture what benefits itemization has brought to project management, cost monitoring and reuse of items.

Question will be that how existing legacy data will be managed properly and what should be taken account when old information is wanted to bring into use. Is all that information really necessary or should something just leave as it was and create new instead. For Component Engineering this question is relevant and current position in the organization model should be reconsidered as if there could be more influence to components used in the case company as it was explained in earlier section 3.1.

Based to interviews answers a summary was made for easier comparison of the results (Appendix 2.). As the question related mostly to itemization the work connection was relevant first question and almost all interviewees used them daily. The daily item usage

pointed to designing, quoting, delivering and invoicing which covers most of the important departments in the case company.

Commonly used system for all departments was ERP (SAP) which was used for item purchases and presenting item information. Especially service unit were using eCatalogue and old legacy systems for brown field item enrichment work which was only way to connect item to certain project context. New discoveries were that JIRA was used for quoting purposes and Maximo for workflow.

One of the key questions was to know what item description field and information was important. Sales description was important because it is description which Outotec's will be visible to customer and it should not contain too detailed information of the actual component. From commercial items perspective the manufacturer and manufacturer part number are important but technical description has also major importance for internal use. Referring to received answers the key discoveries where that most of the departments see that technical description is important field to define information for item. As the technical description field can be used to indicate specifying information which also should be found from attached datasheet. Besides this purchase text field could be used more to describe information regarding to component purchasing.

Strengths with item use was strongly seen related to reusability and benefits as serving future business. Main points were that the item recognition, tracking and component life cycle management are much easier than without items. Some issues still has to be solved to enable item use for all departments in Outotec e.g. the level of itemization which should be able to vary.

Major question was to think what could be existing bottlenecks in the itemization and challenges where seen in the global way of doing things, item ownership responsibility, item data verification, amount of existing legacy items and change management. Also there are many external lists which are managed outside of Enovia by different departments. Commonly it could be said that there are many roadblocks on the way which are barriers for ongoing change.

The current state of itemization is seen as it could be different at this point; if things would have been done differently in the beginning of OPAL project e.g. system could have been more flexible towards current way of working. As the past cannot be changed future improvement are important to execute. All spare parts should be able to connect to some context to improve traceability, like mentioned earlier in the section. Limitations referring to designed products could be considered as it would decrease component variety by reducing selection of component manufacturers. This would have benefits in the long run considering standardization of designed products and spare part business.

All in all, the interview round revealed that discussion with people by having a common core subject was excellent and efficient way to discover what is the current state of itemization in case company and what strengths and weaknesses relates to it and how would those weakness points be turned to more positive solutions. The following chapters will present found strengths and weaknesses of current state.

3.5 Strengths and Weaknesses of the Current Itemization Process

Following discoveries where made based to interview answers and observation of the current state of itemization process. Table 5 presents current strengths of the itemization.

Table 5. Strengths of the current itemization process.

	Object	**Description**
1.	Item release process	Current process for item release is working well without guidance.
2.	Item quality level definition	Item concept defines quality levels for standard and commercial library items which item information has to meet.
3.	Standard and Commercial library classification	Current new standard and commercial library contains relevant classes to all Outotec commercial and standard policy spare part components. This provides clear classification for component and increases their information level.
4.	New item creation	New items can be created and released to production in short time if all information is available for item specialists.

As seen in Table 5, the current strengths are related to current itemization process but emphasizes to new item creation, as the new items can be created much efficient way than in brownfield item cases where the existing knowledge legacy has to be used for validation of the item. The recently updated standard and commercial library contains organized engineering classes for items and therefore enhances its utilization for design engineers.

Regarding to current itemization process state and interviews of the different departments there are many weaknesses which should be tackled. Table 6 will present these weaknesses.

Table 6. Weaknesses of the current itemization process.

	Object	**Description**
1.	Determination process	Big part of item lifecycle management. When item is either e.g. duplicate, not available, etc. It has to be replaced with new one and the old item will be set to obsolete. This is important for business and their stock values. They have to be noticed what to buy and what to sell first from the stock. Definition of this process is needed.
2.	Lead time	Current process takes too much time and big part is information verification. This affects negatively for item release lead time.
3.	**Item context**	Currently connections are searched for items to define where they are used e.g. project delivery for site. This information however is not stored permanently to item context e.g. connection to specific project.
4.	Upkeep of external lists	Individual lists are upkeep which are not in any certain data base but saved local drives. This information is only managed small group in the department.
5.	**Item information quality**	Especially brownfield items have poor information quality and should be able to have verified information from quotations, order confirmation or purchase orders where exist-

		ing information could be checked to validate item and define it for library.
6.	Item type and policy	Many spare parts are vendor items which can be purchased only from vendor by using their code. These items are not commercial type of items and neither design items. A clear classification and library is missing from these types of items. This would need a new item policy.
7.	**Data management**	There is undefined organization structure which should be defined in concept based on data governance. A more collaboration is needed between Component Engineering and other departments in the case company.
8.	Key User concept	PDM key user concept is not working and should be reorganized and create more efficient internal network for information sharing methods, training and support for employees globally.

As seen in Table 6, there are many subjects that need improvement. The most critical ones (bolded text) considering itemization are context, information quality, master data management and item type and policy selection that are creating top challenges to be enhanced in the future in the company.

Summary of the Key findings

First data management has missing definitions and should be considered to define current model in more exact way for better common understanding in the organization. Data governance and item concept is missing definitions in the case company. This reflects from current organization model and how Component Engineering is seen to support other business units e.g. product lines or spare part service. This key finding concentrated mostly to item data management and how the master data management is currently having lack of usable and incoherent data from other departments for Component Engineering's itemization processes. The data and documentation exists in many different systems and is not accessible for Component Engineering needs. Unlike for new items brown field items need more than datasheet to be verified as the component that has been purchased this also

includes a history check, duplicate check and quotation, order confirmation or purchase order to confirm that correct product has been purchased with defined manufacturer code that exists in the item information. The existing data from legacy systems, which are migrated to current system from acquisitioned companies, is incoherent and does not response to the stage of integrity that it should regarding to item concept. This issue could be reflected to item data quality which would have major benefits for the company if it would be coherent.

Secondly item type and policy related information validation is challenging with current way of working and instructions. New item policies are needed and decision and guidelines should be improved to enable more efficient way to search information, enriching items and classify them to library. Target is to propose a solution for the problem how to validate items with different policies and should they have different validation requirements. Question will be what information should be utilized that already exists and what is really needed to be done and by whom and is there an actual need for everything or should the verification level change regarding different legacy items enrichment process.

Third discover from interviews was the item link to any existing context. There are thousands of project delivered for customers globally but a big challenge is the connect item to existing delivery context. This process should be started from one site and proceed from there to another as the amount of information is huge and take lots of time to execute. Therefore it would be important to carefully prioritize and make road map how it would be processed. Currently context is searched to sold items in other words every time a customer has a need for spare part they will make a request the customer service starts process which will provide knowledge what are the connections for these items regarding to delivered projects. This connection would be really important to keep for future business cases and for that reason an improvement should be made for that process. Focus of these problem areas and results of the CSA are presented in Figure 8.

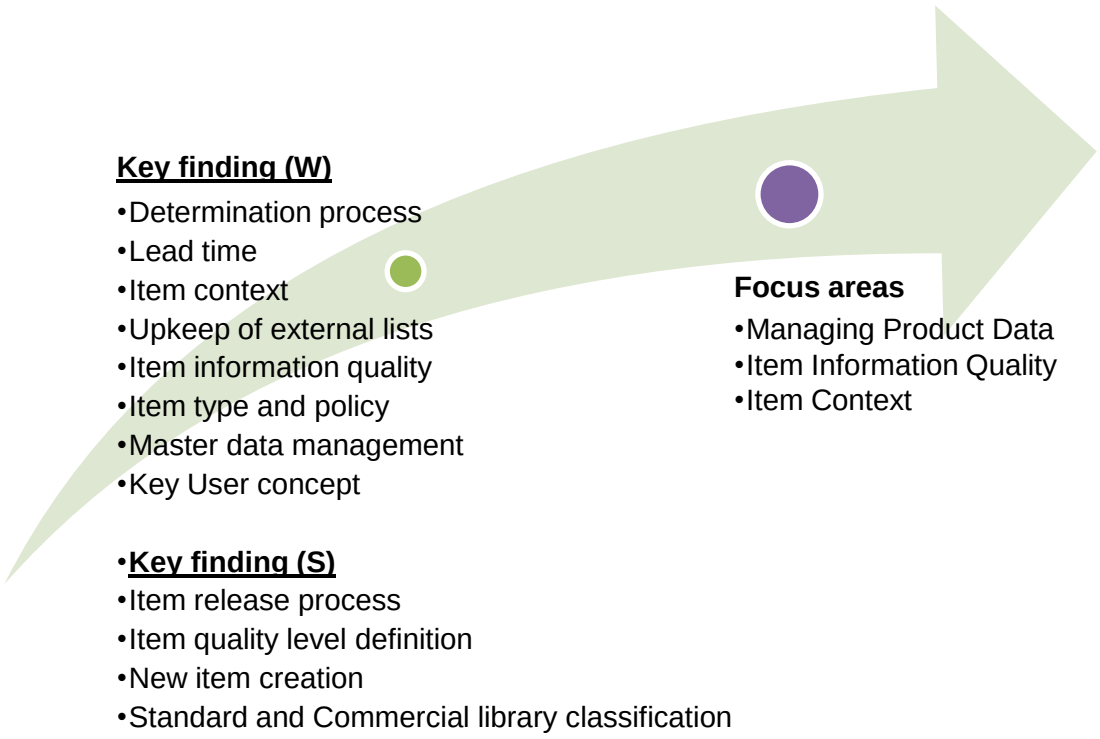


Figure 8. Summary of the key findings from CSA.

As seen in the Figure 8 above, weaknesses that will be focused to improve: (a) item information quality for validation work (b) item connection to existing context, and (c) managing product data processes and information sharing. Section 4 will discover best practices, ideas and support of existing knowledge and literature from topics based to these key findings from current state analysis to support for new proposal and actions which could be executed.

4 Existing Knowledge for Data Management and Improvement

This section discusses the findings from the best practice and literature review for the solutions of improved processes for itemization focusing on the key findings identified in Section 3.

4.1 Information Quality and Data Management in Product Data Management

In Product Lifecycle Management (PLM), information plays the key role as it, for example, is utilized to manage a product and its lifecycle which could include e.g. items, documentation, quality standards, product performance information and so forth (Saaksvuori and Immonen 2008: 2). *Product Data Management* (PDM) could be defined as a system of controlling the evolution of a product and providing other practices and tools with the accurate product information, at the right time and in the right format (Crnkovic, Asklund and Dahlqvist 2003: 19), during the whole *Product Lifecycle* (PLC) as presented in Figure 9.

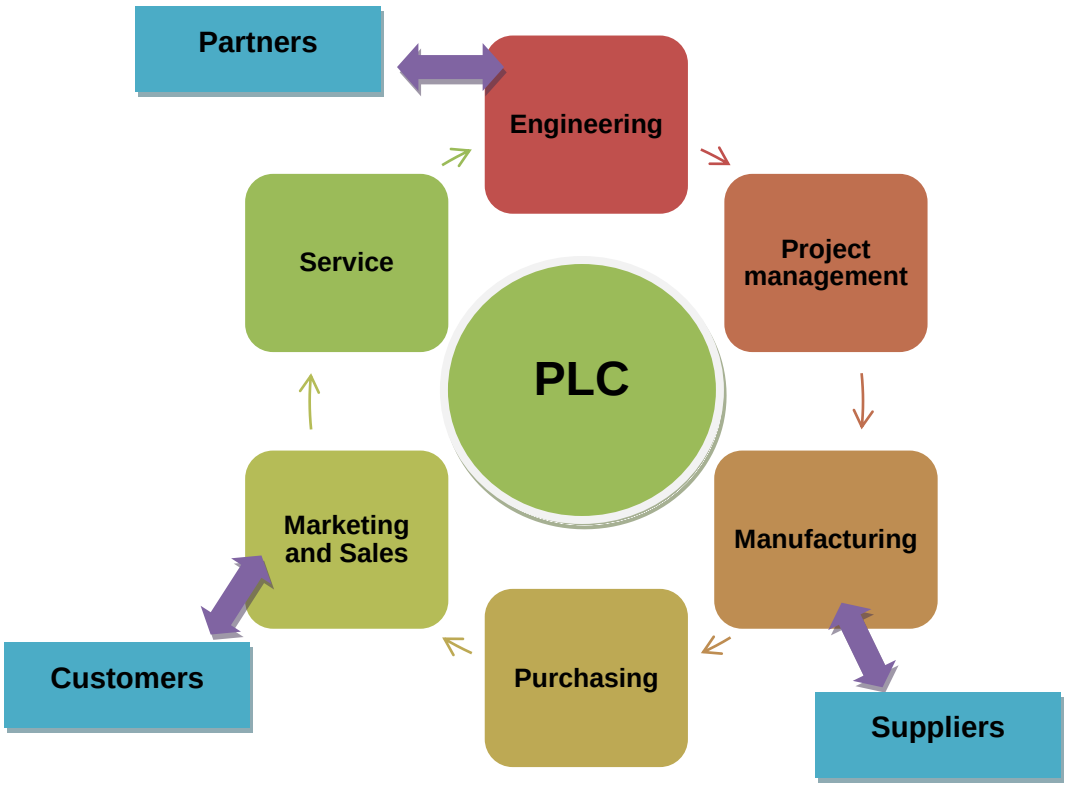


Figure 9. PDM supports the whole PLC (Crnkovic, Asklund and Dahlqvist 2003: 20).

As Figure 9 shows, PLC incorporate supporting processes such as development, manufacturing, marketing and sales, purchasing, and extended enterprises e.g. subcontractors, partners and suppliers. Even though the design phase produces lots of information in the creation stage to PDM system, external parties i.e. partners, suppliers and customers create also information during other phases. Therefore PDM is a base for different systems and an information management system where other systems can be integrated. (Crnkovic, Asklund and Dahlqvist 2003: 20)

Information management could be seen as management for processes and systems that create, organize, store, share and use information. The idea of information management is to provide support for people and organizations to access, process and use existing information efficiently and effectively. The information management process is connected to the information lifecycle with relation of supporting organization to reach its competitive and strategic targets (Detlor 2010: 103). Acquisitions have increased the variety of information resources which organization has to manage. The data includes information stored in databases and unstructured information content found in legacy documents and reports like project documents, purchase orders and order confirmations, this also could be seen as master data that needs to be managed.

Master data management (MDM) could be defined as a process which manages business data entities such as customer, material and supplier data. It can be used for development and maintenance for master data strategy, standards and guidelines, quality and lifecycle activities (Otto and Reichert 2010: 106).

The master data relates indirectly to product data management where the information in the product data is usually connected for immaterial and material aspect which includes creating, developing, and handling and distribution knowledge of the organization. It is important that *the product data* which is included in the objects (items) information is describing physical product's functions and features as they are (Saaksvuori and Immonen 2008: 8). The product data can be used for organizations external and internal functions as it can support internally design and engineering functions and externally co-operations with partners in component, manufacturing and assembly services. There for it is seen that information is critical assist for logistical and financial processes (Abraham 2014: 25).

The information validation depends much of item information quality. Therefore poor data quality produces wrong information for user which makes it difficult to identify and could lead to a failure in the process in information systems (Panahy et al. 2014: 70). Data quality problems often occur when it comes to a collection of information across business actions or organizational boundaries from many sources (Huner et al. 2009: 232). This is especially challenging regarding to legacy items which are enriched to Standard and Commercial library. For this reason data quality should have some kind of standards to be created with same quality. For users it could be difficult to access data that is not readily available, instead of searching and waiting they may to re-invent it by meaning a creation of duplicate item in the system (Stark 2011: 126).

Items are rarely just independent objects meaning that they invariably have relationships, dependencies and dimensions to other items and context in all level of organizations activities (Panahy et al. 2014: 70). To produce accurate information which would be connected to some context e.g. regarding to design structures there are few questions which could be considered: (a) from design point of view what information should data include that it would support design, and (b) what methodological strategies emerges to support the discovered information about information people discovered (Salomon 2002: 249).

The item data information has usually variety of meanings in product data management but despite this, it has one of the highest importances for organizations. Therefore high quality item data or information plays significant role in which should be managed well to enhance networked business. Quality could be described as subjective and therefore despite high or low quality level of information or data always depends on the user's context. Additionally the quality of information is key parameter for approval and success of online service systems (Otto, Lee and Caballero 2011: 80).

4.2 Solving and Improving Quality Problem Definition in Information and Data Management

Although organizations emphasize the significance of information and data management that is structured and standardized, the amount of unstructured information that is either created, acquired, stored, distributed and used within the company usually overtakes the amount of structured data or information which the company manages (Detlor 2010: 103). Most of the engineer's time is used to look information that someone else has produced instead of creative engineering work. Today's manufacturing industries posses four elements; people, information, applications and processes which should be managed electronically therefore PDM and Workflow Management (WM) systems provide support, but only if one of these elements is isolated (Ramanathan 1996: 24). Users are the one who generally uses the data for different purposes and they are the one who will determine the evaluation and improvement of data and processes (Panahy et al. 2014: 71).

Business process improvement needs to improve quality and time and, reduce the cost of the process. There is a need for to manage flows of material, people and information and process improvement which is a systematic approach that will try to increase effectiveness and efficiency of used processes (Panahy et al. 2014: 70). The process contributes organizations to optimize underlined processes to gain more effective results and therefore helps to improve performance in quality, time speed, reliability and customer satisfactions by reducing the costs (Panahy et al. 2014: 71).

Awareness of data and information quality has grown fast during last decades and also its important role in knowledge-based economy. In the past, most of the research focused mainly on multiple data sources and building large data storages, but today it is more focused to data resolution issues that arose when information is integrated from multiple sources with overlapping information (Madnick, Wang, Lee and Zhu 2009: 2: 2). A practical approach for quality improvement would be using a continuous improvement process. The Total Data Quality Management (TDQM) principle relate to the context of data and consists of defining, measuring, analyzing and improving data quality through several improvement cycles (Madnick and Wang 1992: 2). TDQM cycle is presented in Figure 10.

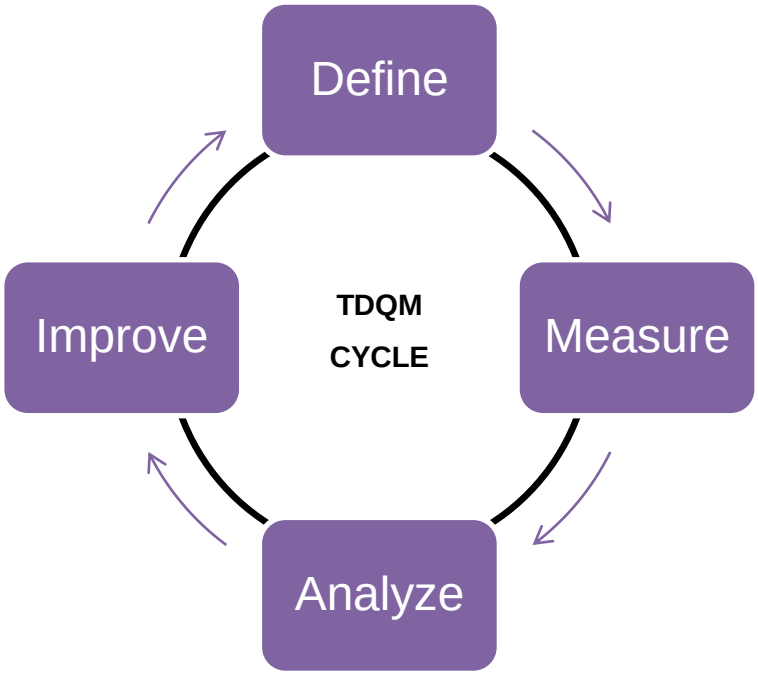


Figure 10. The TDQM Cycle (Madnick, Wang, Lee and Zhu 2009: 2: 3-4).

As seen in Figure 10, the TDQM model consists of four cycles: Define, Measure, Analyze and Improve. First, *Define* means that *d*ata should be defined by consumer's point of view with idea of *fitness for use* and to identify dimensions of data quality. For data consumers fitness for use is defined that data is accurate, believable, objective, relevant, timely, value-added, appropriate in amount, concisely represented, complete and understandable (Lee, Pipino, Strong and Wang 2004: 88). These dimensions have been organized into four data quality categories: accessibility, contextual, representational and intrinsic (Strong, Lee and Wang 1997: 104). Second, *Measure* means that, a data quality assessment instrument was developed for use in research but also in practice too to measure data quality in organizations. The instrument functionalizes each dimension into three to six measurable items to create metrics (Pipino, Lee and Wang 2002: 212). Accuracy can be considered to mean that stored data values do not vary from "real" data values. Iteration between the definition and measurement steps helps to develop metrics that are measurable with fair cost and helpful for improvement actions (Lee et al. 2004: 90).

Third*, Analyze,* interprets the measures and results. The results of the data quality can also be used to reveal differences between data dimension and processing roles. The three major roles, in Table 7, identified by production process: (a) *data collector* (sources

who generate information), (b) *data custodian* (people who manage computing resources) and (c) *data consumer* (people who use data) (Lee and Strong 2003: 17).

Table 7. Data production roles (Lee and Strong 2003: 17-18).

Role	**Part**	**Task**	**Quality Dimension**
Data collector / producer	Data collection processes	Generate information	Accuracy, Completeness
Data custodian	Data storage and maintenance	Manage computing resources for storing and processing data	Completeness, Accessibility, Timeliness
Data consumer	Data-utilization processes	Use of data	Relevancy

As seen in the Table 7, there are three identified roles within data manufacturing systems were the high-quality data could be defined as it is fit for use for data consumers. This refers that usefulness and usability are important aspects of data quality. Roles should be identified during the analyze phase to acquire better knowledge of the origin of the data. Fourth, *Improve* means either to change data values or rather, change the actual processes that generate the low quality data. If actions are not done existing processes continue to generate poor quality data. After this step cycle repeats as improvement activities produce results and as the definition and metrics for data quality are reconfigured. Data integrity rules must be reviewed, redefined, and implemented to produce continuous improvement which is based to iterative process to enhance data quality (Lee et al. 2004: 89). Table 8 presents dimensions that Pipino et al. 2002 defined for high data quality.

Table 8. High data quality dimensions (Pipino et al. 2002: 212).

Dimensions	**Definitions**
Accessibility	The extent to which data is available or easily and quickly retrievable
Appropriate Amount of Data	The extent to which the volume of data is appropriate for the task at hand
Believability	The extent to which data is regarded as true and credible
Completeness	The extent to which data is not missing and is of sufficient breadth and depth for the task at hand

Concise Representation	The extent to which data is compactly represented
Consistent Representation	The extent to which data is presented in the same format
Ease of Manipulation	The extent to which data is easy to manipulate and apply to different tasks
Free-of-Error	The extent to which data is correct and reliable
Interpretability	The extent to which data is in appropriate languages, symbols and unit and the definitions are clear
Objectivity	The extent to which data is unbiased, unprejudiced and impartial
Relevancy	The extent to which data is applicable and helpful for the task at hand
Reputation	The extent to which data is highly regarded in terms of its source or content
Security	The extent to which data is restricted appropriately to maintain its security
Timeliness	The extent to which data is sufficiently up-to-date for the task at hand
Understandability	The extent to which data is easily comprehended
Value-Added	The extent to which data is beneficial and provides advantage from its use

As seen in Table 8, dimensions of assessment instrument can be modified to specific organizational needs. Therefore, it is important that the data integrity rules reflect the dynamic global real-world states and organizations need a process to guide the change in real-world states to updated data integrity rules. First, an organization has to define what data quality means for their data and context by producing data integrity rules. Next, the organization measures the quality of data against defined integrity rules. Third, the violations are analyzed and based to analyze improvement will be made to data to comply data integrity rules which will be redefined, if the actual data is valid (Lee et al. 2004: 90). The effect is support for the changing and global nature of organizational data.

By adopting the definition of the characteristic of high-quality data, in Table 9, they can be formed to consist of four categories: intrinsic, accessibility, contextual and representational

Data Quality (DQ) aspects. These categories give possibility to view data quality problem solving by using certain approach.

Table 9. DQ categories and dimensions (Strong et al. 1997: 104).

DQ Category	**DQ Dimension**
Intrinsic DQ	Accuracy, Objectivity, Believability, Reputation
Accessibility DQ	Accessibility, Access security
Contextual DQ	Relevancy, Value-Added, Timeliness, Completeness, Amount of data
Representational DQ	Interpretability, Ease of understanding, Concise representation, Consistent representation

As seen in Table 9, categories were created by identifying common patterns and sequences of dimensions. The patterns were build up from three problem-solving steps: (a) problem finding (identifying the problem), (b) problem analysis (determination of the cause), and (c) problem resolution (changing the procedures). For example, using these tools organization could during conversion of data recognize poor DQ and initiate an improvement (Strong et al. 1997: 105). Contextual pattern in data quality meaning, relates to data consumers' arguments that current available data does not support their working tasks. These causes could be categorized to: missing (incomplete) data, insufficiently defined data, and data that could not be properly combined. As there are some usable implications for information system (IS) professionals about solving intrinsic, accessibility, representational and contextual DQ problems. All solutions do not provide corresponding result for data consumer's point of view. For data consumers accessibility goes over technical accessibility, in other words they should be able to manipulate existing data to suit their needs.

The contextual DQ meaning is more than just a good data requirement specification. When data produces produce high quality data with "the dimensions of value and usefulness relative to data consumers' task contexts inserts a premium on designing flexible systems with data that can be easily aggregated and manipulated", whereas the other option is regular maintenance of data and systems to meet dynamic data requirements (Strong et al. 1997: 109-110). Reflection context declare knowledge about why an organi-

zation collects certain data, how the data is stored, what limits are determined and how information is used.

The *context,* which differs from contextual IQ pattern, used as unique method, from data quality problem solving and higher perspective can be classified to paradigms, goals, roles, time, and space which construct significant contexts for data quality problem solving. The discussion on context of the data connects otherwise independently managed processes together, and engages it the broader scope of data collection, storage and use over item (Lee 2003: 96).

Importance of knowledge and the equivalent organizational learning has been recognized in ordinary researches, whereas relationships and patterns of different modes of knowledge for specific actions have not been explored (Lee and Strong 2003: 14). The modes of knowledge have two related points. First, all modes of knowledge are combined of all processes, is held by all roles and devote to overall data quality. Secondly each of those modes of knowledge kept by a role devotes to a particular performance measure, a data quality dimension (Lee and Strong 2003: 19). As the purpose of data production processes is to produce data for data consumers, it means that data is fit for use for them and this involves multiple dimensions. To discover relationship between knowledge and data quality, a focus should be in five dimensions of data quality: (a) accessibility, (b) relevancy, (c) timeliness, (d) completeness and (e) accuracy (Lee and Strong 2003: 17). The relationship for knowledge and data quality could be defined by work role. As the data collectors, custodians and consumers have own knowledgeable how to collect and utilize data, in the organization knowing more about certain work processes could lead to performing better to some areas than in others (Lee and Strong 2003: 18).

4.3 Tools for Information Quality Problem Solving in Data Management

First, there is a need to define Information Quality (IQ) although it is an inexact in science terms of assessment and benchmarks. Despite that, there is demand for a methodology that asses how well companies develop their information products and deliver information services to consumers. Whereas information as a product is viewed from engineering point of view and it focuses on the operations needed to set and maintain data in data bases.

This process of changing information values reminds product enhancement. On the other hand information can be seen as a service. A service is defined as an act performed by one party / machine for another, it is an experienced, used or consumed, and it is perishable. It is produced and consumed concurrently. Information as a service focuses to obtain and use information after it is stored as an end-product in the data base, in addition the information service quality addresses the unseen product attributes that become visible during use of data (Kahn, Strong and Wang 2002: 186).

Quality is not easy to define as it is subjective and does not provide practical guidance for improving quality. Two views could be defined for quality: (a) conforming to specification and (b) meeting or exceeding consumer expectations (PSP/IQ) (Kahn et al. 2002: 185). Table 10 columns show the product and service performance model for information quality.

Table 10. Aspect of the PSP/IQ model (Kahn et al. 2002: 185).

	Conforms to Specification	**Meets or Exceeds Consumer Expectation**
Product Quality	Sound information The characteristics of the information supplied meet IQ standards.	Useful Information The information supplied meets information consumer task needs.
Service Quality	Dependable Information The process of converting data into information meets standards.	Usable Information The process of converting data into information exceeds information consumer needs.

As seen in the Table 10, this kind of view of quality tends to be used by product and service designers and sales professionals. Even though this definition may capture the aspect of fitness to use it is still difficult to measure because consumer expectations can change over time. Kahn et al. (2002) determines that product quality has dimensions related to product features which involve tangible measurements and completeness. Whereas service quality includes IQ dimensions to the delivery process and emphasizes more intangible measures, like ease of manipulation, security and added value of the information to consumers. Table 11 shows PSP/IQ model by classifying IQ dimensions.

Table 11. Mapping the information quality into the PSP/IQ model (Kahn et al. 2002: 188).

	Conforms to Specifications	**Meets or Exceeds Consumer Expectations**
Product Quality	Sound information • Free-of-Error • Concise Representation • Completeness • Consistent Representation	Useful information • Appropriate Amount • Relevancy • Understandability • *Interpretability* • *Objectivity*
Service Quality	Dependable information • Timeliness • Security	Usable information • Believability • Accessibility • Ease of Manipulation • Reputation • Value-Added

The four corner stones, as shown in Table 11, of the PSP/IQ information quality model (sound, dependable, useful and usable) create a basis for how organization could develop sound and useful information products and deliver convenient information to information consumers. With this model an assessment could be provided as baseline for determining what kind of improvements could be made. Second, it provides a solution to compare information quality over the organization and to create IQ benchmarks. Coordination between information collectors/producers, custodians and consumers is necessary to deliver quality information as a product or service. (Kahn et al. 2002: 191-192)

Referring to section 4.2, the Table 9 categorized different DQ patterns, which were discovered from different problem-solving projects. The intrinsic DQ pattern helps in problem areas were the same internal data mismatches between sources. Figure 11 show the intrinsic pattern which includes two sub-patterns: (a) Multiple sources of same data and (b) Judgment involved in data production (Strong et al. 1997: 105).

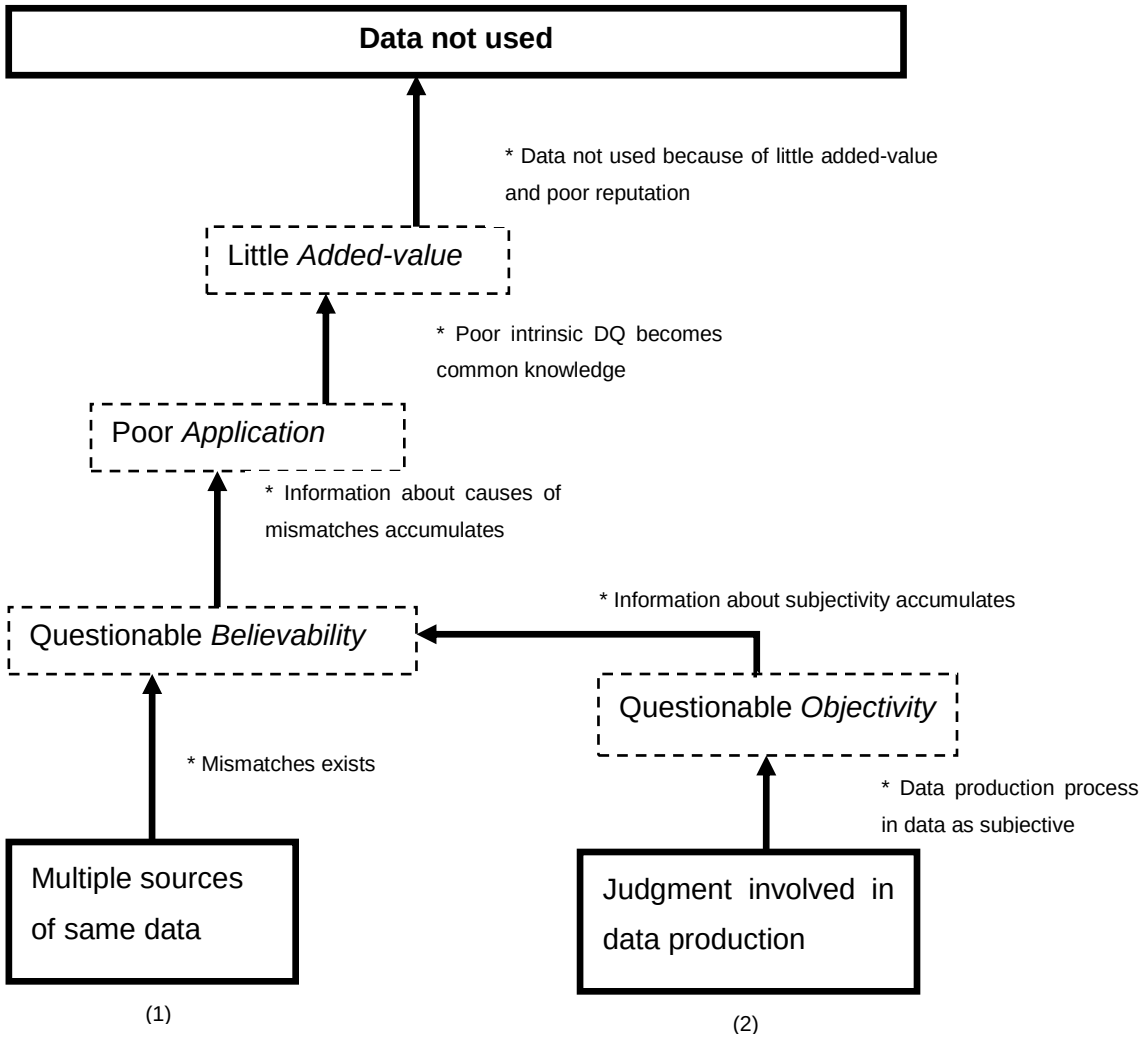


Figure 11. Intrinsic DQ problem pattern (Strong et al. 1997: 105).

As seen in Figure 11, when there are multiple sources of same data, the sub-pattern 1 indicates that they will produce mismatches at some point and its reputation will worsen by degrees until the data is not used for decision making. Sub-pattern 2 prove that data-production forms can impact to objectivity of data by reducing it e.g. if there are many different forms which are used to input data, the variance can influence the believability of this data. (Strong et al. 1997: 106) These solutions express two options to solve the problem either change the system of change the data production process.

When solving contextual data quality problems a need for relevant data is wanted to add value to the tasks of data consumers. Figure 12 presents pattern with three sub-patterns for contextual data quality.

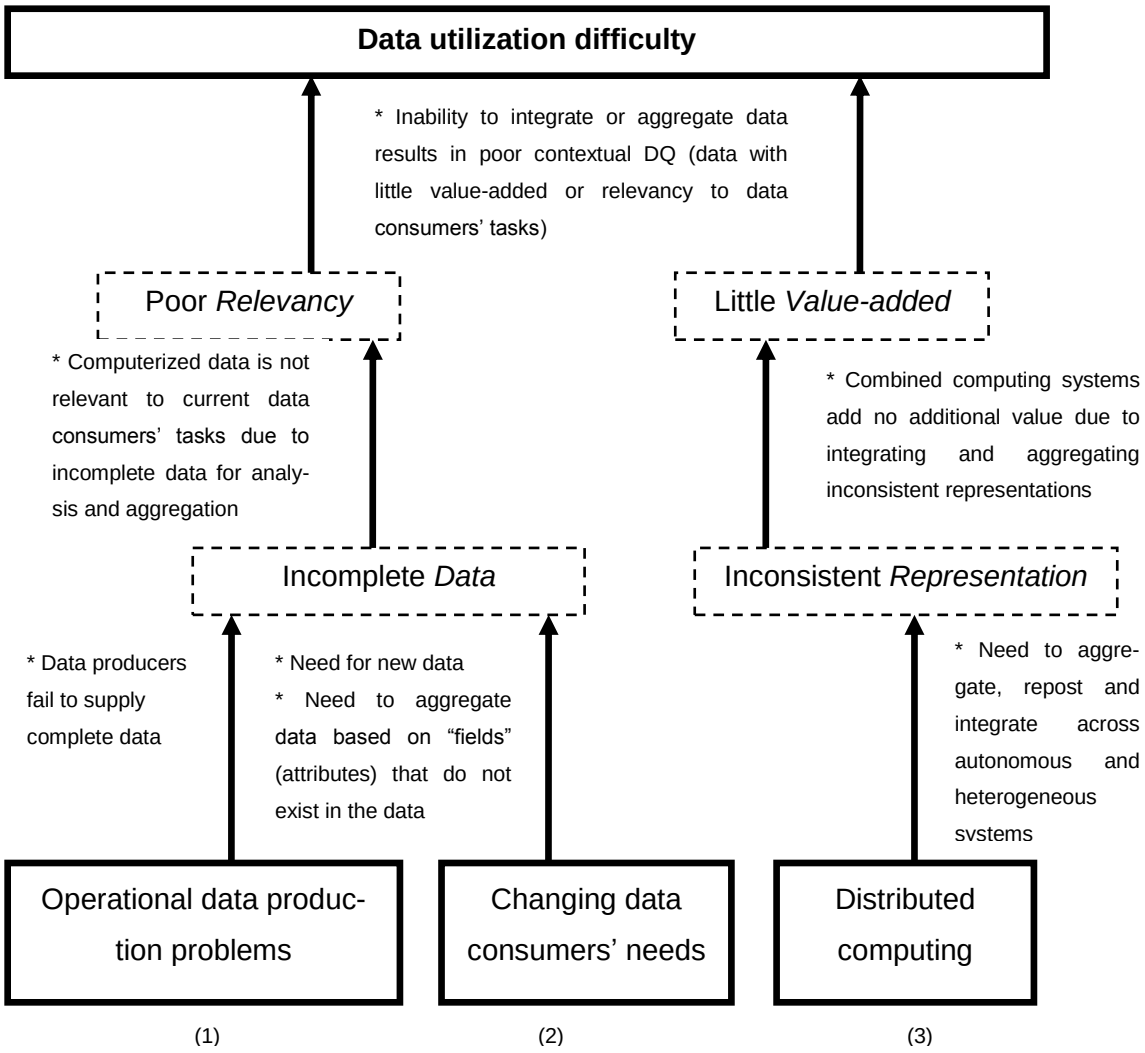


Figure 12. Contextual DQ problem pattern (Strong et al. 1997: 107).

As seen in the Figure 12, the sub-pattern 1 refers to incomplete data in consequence of operational problems which were ensued of incomplete transaction data e.g. all needed information is not recorded. Other problem can also come from incomplete design data (sub-pattern 2) which result is that it won't add value towards data consumers. The sub-pattern 3 revealed problem which was cause from integrated data across distribution systems. Inconsistent definitions and data representations are used across divisions' for ex-

ample use of different basic utilization measures or way to present data e.g. decimal points. Most of the problems are cause of autonomous design decisions in each division. (Strong et al. 1997: 108)

For context point of view rules and routines should be common and straightforward in the organization and not create problems. The organizations process will change when to business logic changes. The context could be described as differentiator and relationship builder which specifies the connections between contents and environments. Table 12 clarifies contexts that are defined paradigm, role, goal, time, and (space) place (Lee 2003: 102).

Table 12. Context definitions (Lee 1003: 102).

Context	**Definition**
Paradigm context	Consists of disciplinary principles that form the rules and procedures for data quality problem setting and solving. Refers to "*book*" rule.
Role context	Refers to specific roles that individuals play in the information production system. Identified three key DQ roles: data collector, data custodian, and data consumer. Refers to "*who*" rule.
Goal context	Includes objectives that individuals and organizations aim to achieve through creating, using and processing data. Refers to "*why*" rule.
Time context	Describes the time frame during which, or for which, data is used and processed. Refers to "*when*" rule.
Place context	Refers to site for which data are generated, used or processed. Refers to "*where*" rule.

As seen in Table 12, *context* has many definitions which could be used for data quality problem setting and solving. As the high quality of data is critical for companies the organizations data production process should be divided to three separate areas: data collection, data storage and data utilization. To solve problems effectively the individuals in all three processes must have sufficient knowledge of problems regarding to their process domains. It could be considered that at minimum (a) data collectors must know what, how, and why to collect the data, (b) data custodians must know what, how and why store the data and (c) data consumers must know what, how and why to use the data (Lee and

Strong 2003: 33). These three modes of knowledge with related data roles helps identify and solve problems in the processes and therefore assist to enhance organizational data quality. By understanding the differentiated relationships between work processes, knowledge conditions and organizational performance opens up new paths to improved virtual work environments.

4.4 Conceptual Framework of This Thesis

The conceptual framework of this study consists of three main areas: (a) Information management: Improve data production process and identify individual roles, (b) Data and information quality: A method to create coherent and high quality data and information, and (c) Information context: To define items into context and enhance problem solving. Even though these three key areas are overlapping each other, they are more likely embedded together and have connection to each other and aims to improve the ideology for current itemization and give solution how to enhance or create new processes to increase data quality and its usefulness.

The conceptual framework for this study has to dimensions, topics and methods and it is conducted from simple idea: Any data quality problem is caused by some reason and it can be solved by using certain research methods (Madnick et al 2009: 2: 5). The approach will present methods how data quality problems could be defined and presents roles for individuals how to process data to produce high quality data. The use of this approach will provide overview of research for current itemization situation and data quality problem detection which will focus to individuals and organization. The conceptual framework model is visualized in Figure 13.

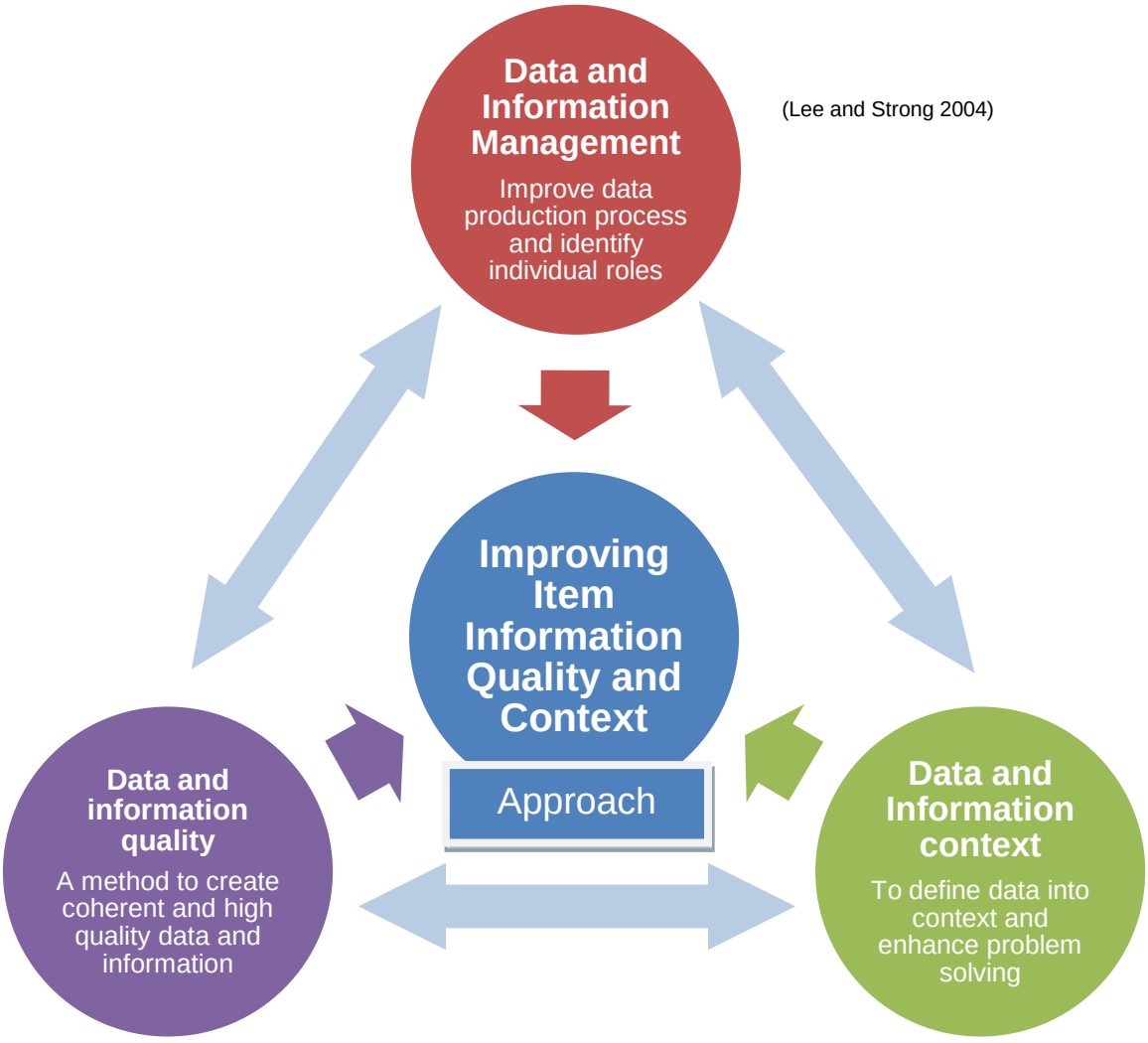


Figure 13. Conceptual framework.

As seen in Figure 13, first area could be seen as basis which has to create a standards and disciplines to enable creation of quality data. Knowledge about work process is a precondition for performing work and work roles and the mode of knowledge do imply (Lee and Strong 2003: 13). Ideally this area will present logic how different groups of individuals could be categorize and define common knowledge of what should be known for actual work process.

Second area presents methods how data quality problems could be detected and how data quality can be categorized for example: *product quality* and *service quality* (Kahn et al. 2002: 185), and what kind of dimensions relate to data quality. Information quality dimension could be used to solve certain type of problems in organizations data and how information quality relates to information producers, custodian and consumers role and processes (Lee and Strong 2003: 13).

The third area will complement the two previous ones. Context will provide another extent to data quality as its importance is growing all the time. The meaning of context has essential role when it is used for carrying out data quality problem solving (Lee 2003: 95). Organizational databases pertain in the larger context of information systems and there can be presented a model how to categorize contextual data quality and what kind of pattern to use for problem solving (Strong et al. 1997: 105).

Combining the three areas, Information and data management, Data and Information Quality and Data and Information Context will produce a generic template and a proposal which detects solutions for item quality problems and to enhance the existing processes which could be used to improve organizations processes. The next section 5 will present a proposal based to sections 4 conceptual framework and discovered findings from section 3 current state analysis.

5 Building Proposal for the Case Company

This section merges the findings of the Current State Analysis with the discovered methodologies on the conceptual framework towards the building of the proposal.

5.1 Overview of Proposal Building Phase

To create this initial improvement proposal, the actual idea of the proposal building phase was received from the results of Data 1 collection, for CSA. In the Current State Analysis, after all the interviews and analysis of bottleneck and classification of strengths and weaknesses, the actual need for improvement came clear. Most of the weaknesses pointed in a way or another to *data information* and *data management* and two important objects: *data and information quality*, and *data and information context*. They were seen as a root cause to all existing problems in data management. As these three key areas were discovered the focus was clearly concentrated to find literature from information quality, information management and information context to create conceptual framework around the data improvement.

Literature findings topics were based to three CSA focus areas: data management, item information quality and context. Based to these subjects the literature gave a variety of options of knowledge, whereof the best practices were shaped into conceptual framework. Thus, it was identified that, to improve current product data management process, the improvement could be approached from the data quality point of view.

Stakeholders were involved to proposal building by the discussion and answers that was received from the interviews in the current state analysis. The proposal was developed from item quality point of view, and based on stakeholders discussions about the itemization in the case company. According to the stakeholders many of the suggestions pointed to item quality problem. As it was seen that problems were related to (a) item enrichment work i.e. incomplete information, (b) creation of the items, whereas all demanded information could not be obtained, and (c) verified information was sometimes difficult to find, which also increased the itemization processing time, and (d) items connection to context e.g. for delivered plant or product, that has been delivered to customer.

5.2 Findings of Data Collection 2

There was a need to have knowledge of the data quality subject, context relations and methods how to actually measure or explore them. As phrased by one of the interviewees:

> *All spare part should be connected to some context (for traceability) e.g. SBOM could be solution (Data 1, Interview 1).*

In the Proposal building discussions, it was mention that, if data quality can be improved, it would a benefit for the whole organization. To create better quality many processes has to be defined in a way that it will serve the correct model to provide quality information on each level. To support this, a notice stood up relating to item context and data roles, as phrased by one of the interviewees:

> *One goal is to maintain current legacy data to clear format that is shows reasonable for the company and will service customers in easy to use format (Data 1, Interview 2).*

In the Proposal building discussions, it was mention that the case company has a lot of existing data, but managing that information is very difficult and certain knowledge is not utilized effectively simply because of lacking knowledge of data processes and roles and responsibilities. Therefore (a) the information management should have improvements made by identifying those processes and roles, (b) secondly data and information quality state must be discovered by recognizing the current state and solution to manage them, and (c) thirdly information context role is not defined and utilized to data and information management. As phrased by one of the interviewees:

> *Service tries to search from context (where the actual item has been used and at the moment discovered context is not connected to delivered items (Data 1, Interview 5).*

In the Proposal building discussions, it was mention that a proposal template could be created for current itemization management to (a) discover to possible information quality problem areas and (b) to provide methodology for different processes which could be utilized more detailed processes in the item management.

As the high data quality would provide a good base for functional processes and bring more value for the organization, it created a need to define and recognize the current state and problems relating to companies data quality. Data quality has been recognized as multi-dimensional concept and organizations have deal with both the subjective understanding of the individuals involved, and the objective measurements based to data set in question (Pipino et al. 2002: 211). There are some unused solutions in data quality problem solving which could enhance current data production processes and determine relations between information. One noticed key element would be the internal transparency between different departments work. As phrased by one of the interviews:

> *Service does not depend who had the ownership for spare part that is sold as long as it had all relevant and verified information for identifying the component or product (Data 1, Interview 2).*

The actual need for this proposal comes from the business challenge as there is “a lack of confirmed information” regarding to the company’s itemization processes. In the practical level the problem occurred as a difficulty to validate the existing information in the items. The most direct solution could be just to improve the current process somehow to make it more efficient, but it would not fix the actual problem, the root cause of the shortage of unconfirmed information. The proposal presents the solution how to detect these root costs and propose a methodology and a tool which could be used to data quality problem solving and therefore provide the confirmed information for operating processes in the organization. Table 13 presents the suggestions from the stakeholder for the proposal.

Table 13. Suggestions from stakeholders.

	Stage of item data processes	***Suggestions***
1	Item traceability is poor and connections between the items are difficult to determine.	The Item Data Manager suggested that traceability could be enhanced by connection the item for certain context e.g. utilizing SBOM (Interview 1).
2	Legacy item data has low information quality.	The Service Quotation Manager suggested that the legacy data should be maintained in a way that it would be in clear format as it would be reasonable for the company and towards customers (Interview 2).
3	All the discovered information is not connected to the item.	Suggestion from Director of Market Units/Spare Parts that the discovered context during service sale process should be utilized and the context should be connected to related item(s) (Interview 5).
4	It depends who creates the information to the item and who has the responsibility of item design.	The Service Quotation Manager suggest that responsible design organizations would provide item information with same quality level, in which case it would be relevant information for use (Interview 2).
5	Item improvement process is difficult to approach.	The external consultant suggested that crating a simple model, would be good approach for itemization improvement and it would be easier to perceive (Interview 6).

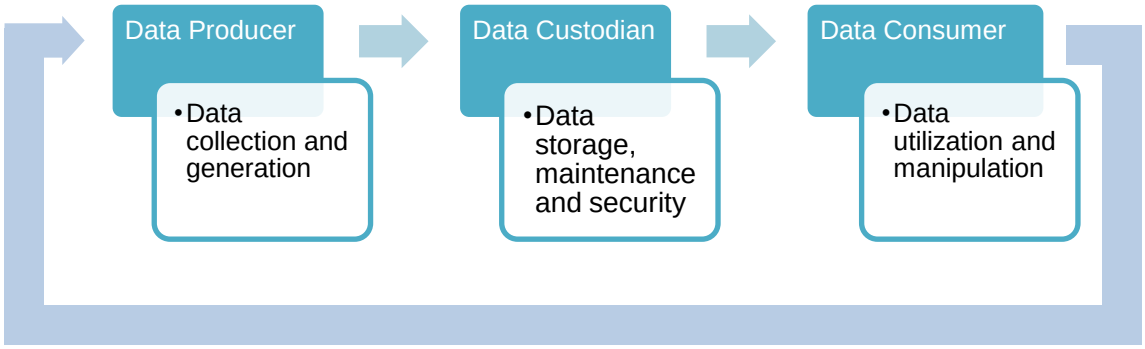
As seen in Table 13, the key point of stakeholder would be to enhance amount of information in the item and manage them more efficiently. Understanding the ideology of information context would provide better understanding to recognize the connection that items have to different contexts. Contexts are commonly used for specifying scopes or as a limit for a study area, but they are also referred to as an external environment. Rules define processes and activities of a practice and they are underlying in the contexts. By discovering these “forgot” rules and with knowledge to help practitioners to reflect-in-action to ena-

ble them to break old rules, revise them or create new rules (Lee 2003: 99). Those rules which could be discovered should be collate and determined to enable creation of new rules which could be used to solve a data quality problem.

Summary of the initial proposal building

Knowledge of how data is generated, stored and used is critical for today's organizations. The huge amount of data is produced continuously and therefore there should be an understanding of what roles are involved to engineering or service information and what kind of knowledge are they holding. Regarding this problem, data production process has definition for certain roles which are collecting, managing and utilizing data in some way and these roles are defined as; data collector, data custodian, and data consumer. These roles are presented in Table 14, which visualizes different roles and their description of the action task.

Table 14. Roles and definitions relating to quality data production.



As seen in the Table 14, there are many connections between data quality problem solving which relates to these roles and their connection to certain context.

Regarding to stakeholders and feedback from key manager there is a clear need to find methods and have better knowledge of how data quality problems should be solved and thereby enhance the existing and produced data quality. This would of course mean changes to be done in the current processes, but it should not be a problem as today's organizations are and should be continuously improving their processes as a part of change management.

The proposed approach for data quality problem solving could help companies to create better overview of their company data quality. Each presented model has self-supporting method to approach for data quality problem solving. This means that proposed approach, Figure 14, consist of four different models which could be used as one to provide comprehensive view of the existing data quality problem and one tool to provide improvements for it. However every model could be also used independently and provide approach for data quality problem solving from certain perspective.

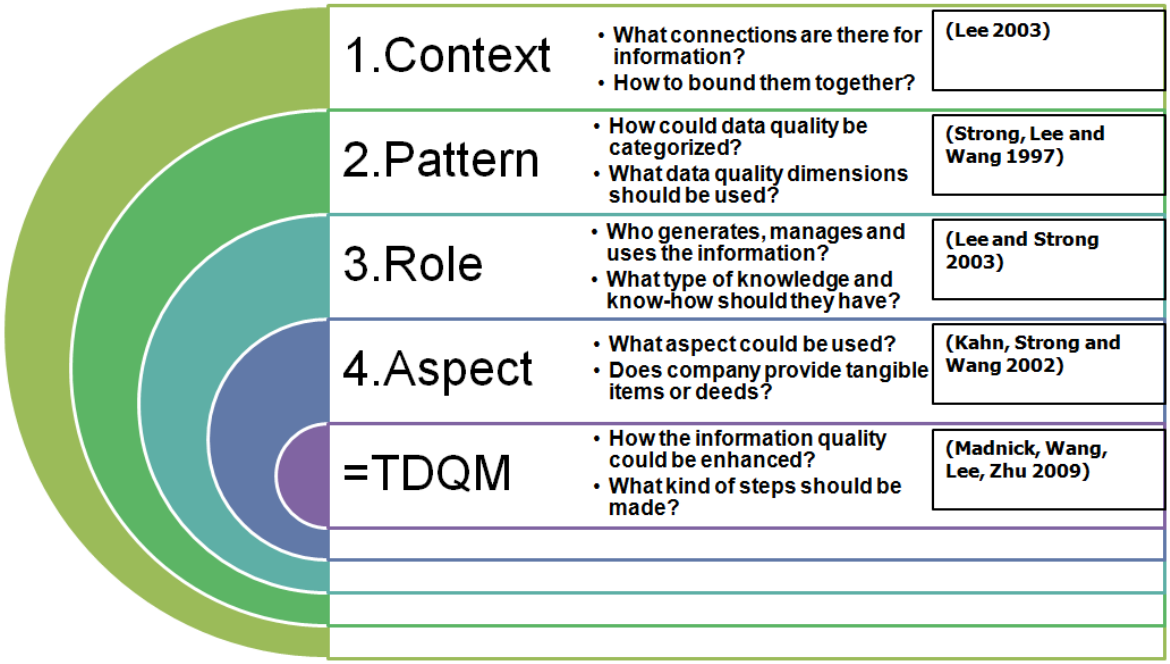


Figure 14. Outlines of Data Quality Problem Solving Cycle.

As seen in the Figure 14, for each model there have been defined questions which provide the meaning for the model in away as it can be used for those purposes. These questions could be used when approaching the data problem solving model or defining the correct model for the data problem solving.

5.3 Initial Proposal for Data Quality Problem Solving

The proposal is constructs of five methods which are divided by the nature of data quality problem solving approach. Five categories are, (a) the *context* method, which presents problem solving regarding to information connections and role in the processes, (b) *pat-*

terns which present models for different categories regarding to data quality dimensions, (c) *roles* that represents different data processing roles for individuals, (d) *aspect* which differs data quality problems between product and service, and (e) *Total Data Quality Management* method that can be used as a tool for continuous data quality improvement processes. These methods could be described as process steps for data quality problem solving. The following sub-sections present these methods and summarize them at the end. The approach is started from the outermost method which is the context-method and from there towards the center to Total Data Quality Management.

The proposed diagram gives an answer to organization that have need to create definition of what should be observed when solving organizations data quality problem. To define this approach there is a need to understand what is related to data quality problem solving, where it relates to and what questions should be made to create optimal metrics and approach to study organizations data quality. The proposed approach offers an overview how to approach data quality from problem solving perspective. The overview of this study will provide a common knowledge, methodology, and a tool for this approach and share the knowledge and provide a base whereof a possibility to improve a more detailed processes for organizations certain sub-processes which could help to enhance data quality.

5.3.1 Data Quality Problem Solving Methods

The proposed model of data quality problem solving paradigm has been combined from five different approaches, which could be seen as methods, to define data quality problems and analyze them to gather the knowledge for practicable improvements in data production and processes. The idea for these methods comes from logical approach as from top to centre, for the problem solving.

The outermost method of the data quality problem solving diagram contains context which includes different definitions and models that forms major contexts for data quality problem solving. The context method can be used to define what kind on connections they could be solved for information quality. The context has definitions of time, goal, place, role and paradigm and using these definitions the information can be bounded together. The context can be seen as differentiator and relationship creator which specifies the relationships between contents and environments (Lee 2003: 96). In this approach the problem solving

starts from context, Figure 15, which connections should be considered in the phase one of the approach for data quality problem solving.

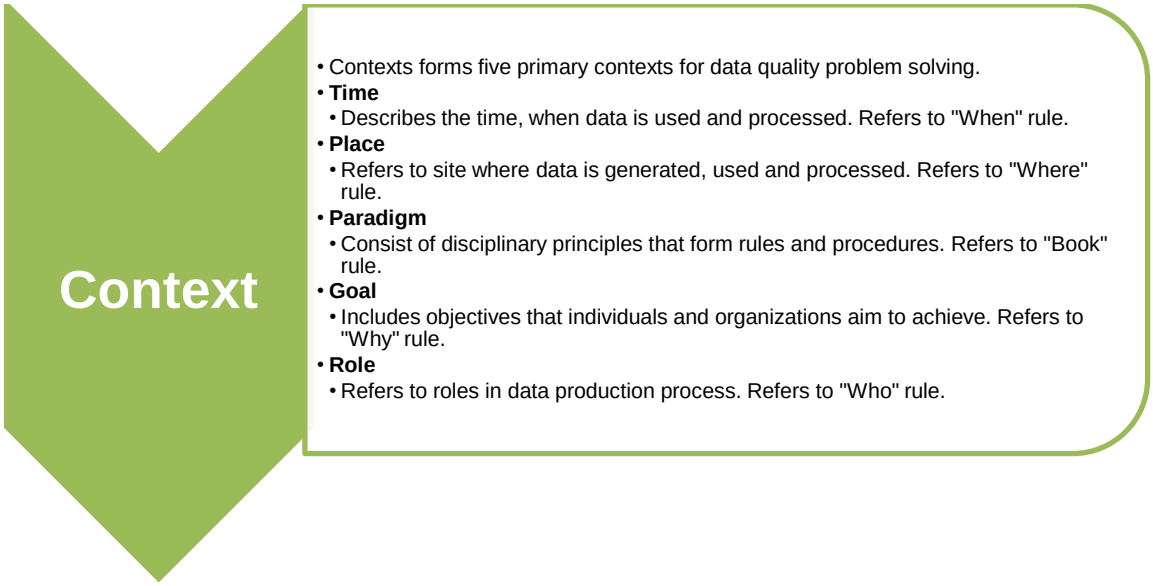


Figure 15. Definition of Context in Data Quality Problem Solving.

As seen in the Figure 15, the context is whole which connects and defines information depending of the nature of the information which is observed. Context-method is seen as the outermost method as it interprets the connections of the information and presents rules for which relates into different variation of contexts. The next method after context provides different patterns for data quality problem solving as showed in Figure 16. They are categorized to four data quality (DQ) problem patterns: (a) Intrinsic DQ, (b) Accessibility DQ, (c) Contextual DQ and (d) Representational DQ, where representational data quality dimensions are seen as underlying causes of accessibility data quality problem pattern. With help of these patterns organizations can provide experimental basis for studying choices and actions for data quality improvement and build data quality theories about the nature of organizations data problems and solutions.

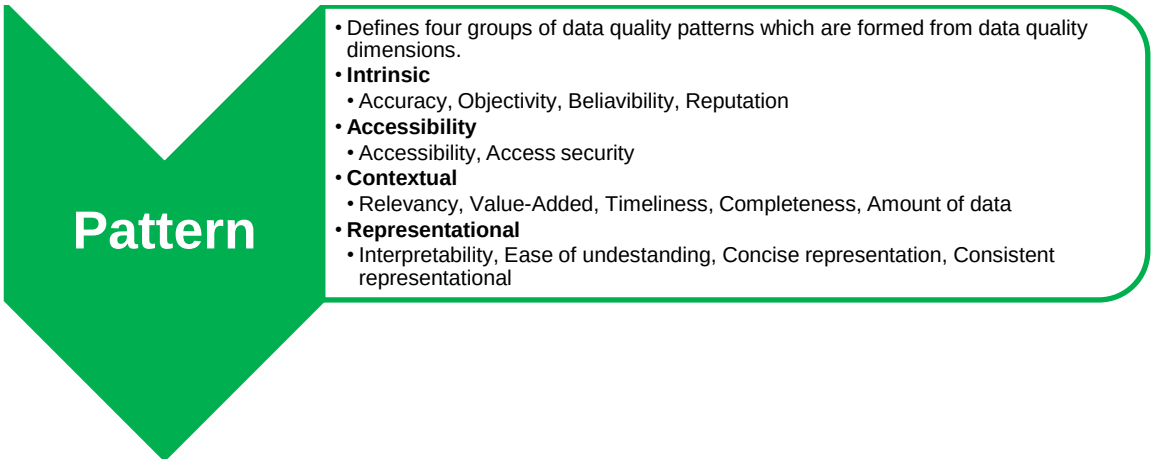


Figure 16. Data quality problem solving patterns.

As seen in the Figure 16, these patterns could be used for organizational decision making and along using the DTQM tool it could establish routines that improve data quality. The next method towards the center distributes data production process roles that could be indentified for individuals. Role-method is presented in Figure 17, there are three roles which are defined to: data producer/collectors, data custodians and data consumers. When company is producing physical products the data production process is divided to clear work processes like collection, storage and use work processes.

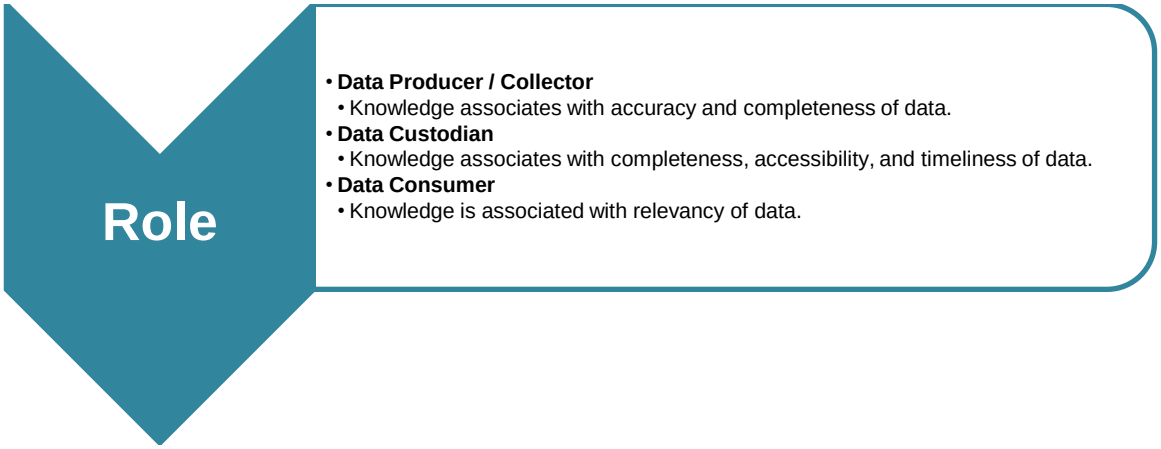


Figure 17. Roles in data production process.

As Figure 17 presents, knowledge about data processes should be identified as different than knowledge about work processes even though both relates within data production process, collection, storage and utilization of the data. When organization are able to iden-

tify these three roles, they are also able to have answer for three modes of knowledge, knowing-what, knowing-how and knowing-why. The knowledge of these and the three data production processed are relating all the way to higher data quality and can be connected to five data quality dimensions: accuracy, completeness, accessibility, timeliness and relevancy. The next method supports coordination for these three roles could be enhanced with information quality (IQ) benchmarks which can be received by using product service performance model (PSP/IQ), in Figure 18, which defines four quadrants models of information: sound, dependable, useful and usable. These can provide basis for how good sound and useful information products organization is developing and delivering dependable and usable information services to information consumers.

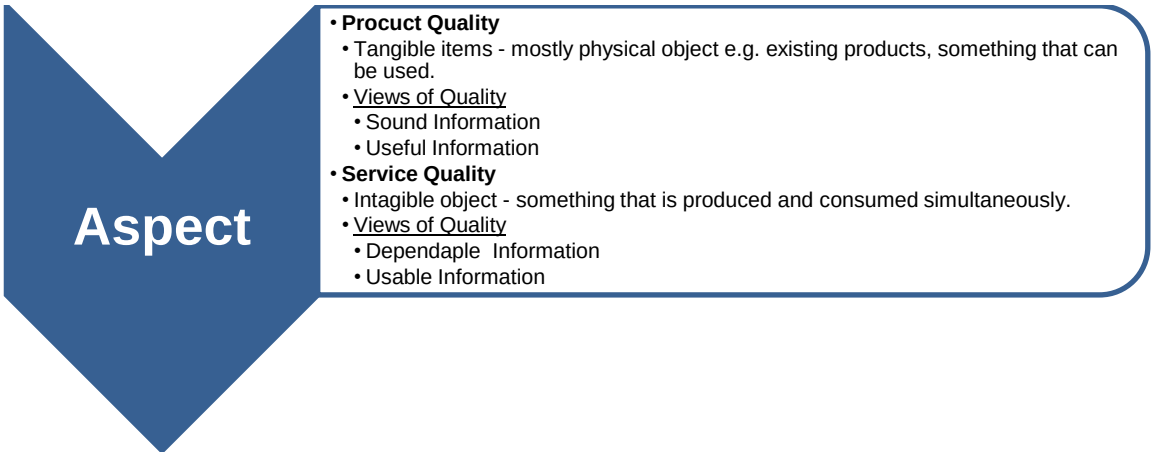


Figure 18. Data Production Process Roles.

As Figure 18 defines there are four quadrants; two quality views for product quality and service quality. The ultimate goal for organizations should be to consistently deliver all four quadrants but it could be very challenging. Therefore organizations should be mainly focus on the soundness quadrant (information as a product that conforms to specification). Regarding to PSP/IQ model and three defined data product roles there would a solution to coordinate and manage information quality activities by using concept of Information Product Manager (IPM).

Total data quality management tool (TDQM) represents continuous data quality improvement, see Figure 19, with four step cycles of *Define, Measure, Analyze*, and *Improve* which incorporates also other methods which were described previously. It can be used to

address organizations data quality concerns. The key insight has been that, data is actually a product which is manufactured by most organizations but it is not treaded or considered as such.

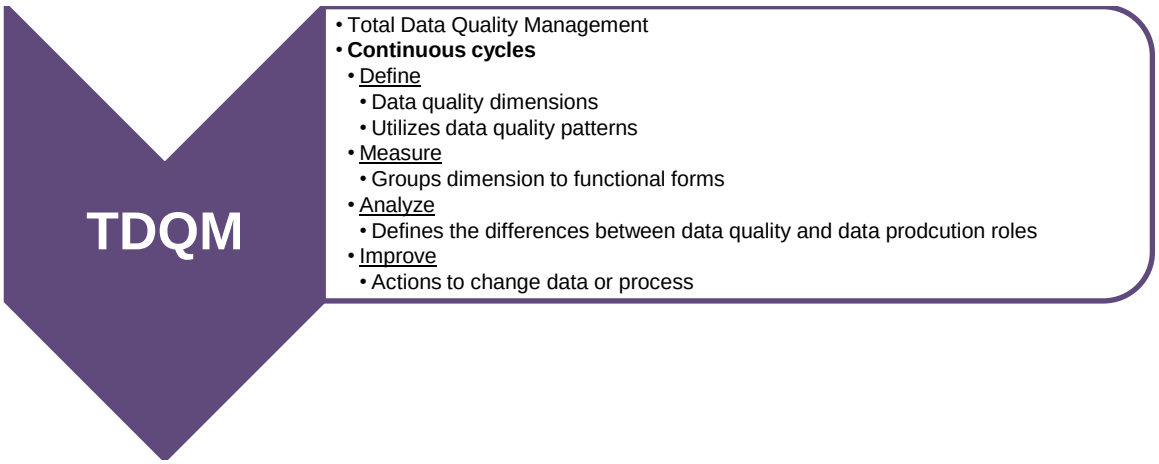


Figure 19. Total Data Quality Management Tool.

As seen in Figure 19 the TDQM tool could be used for continuous data quality improvement. *Define* cycle includes all data quality dimensions which can be defined to four categories: accessibility, contextual, representational and intrinsic as described in Pattern method. The *measure* instrument divides each data quality dimension to into four or five measurable items and it can be adapted to certain organizational needs. *Analyze*, step will decode to measured results and present differences between data quality dimensions and data production roles which refers to Role-method in the proposed diagram. It also identifies the most need of improvements and possible root causes of organizations data quality problems. The *improve* step contains actions which are take to change data values or more preferred to change actual processes which are producing the data and data quality problems.

All these methods could be utilized as steps of one compound process as seen in Figure 20. For organization it would be a good start to first understand the nature of data quality and how the data is produced in the company and who actually participates to it. Then select the best way to analyze it, as it depend a lot of the organizations type and the data type which will be defined and measured

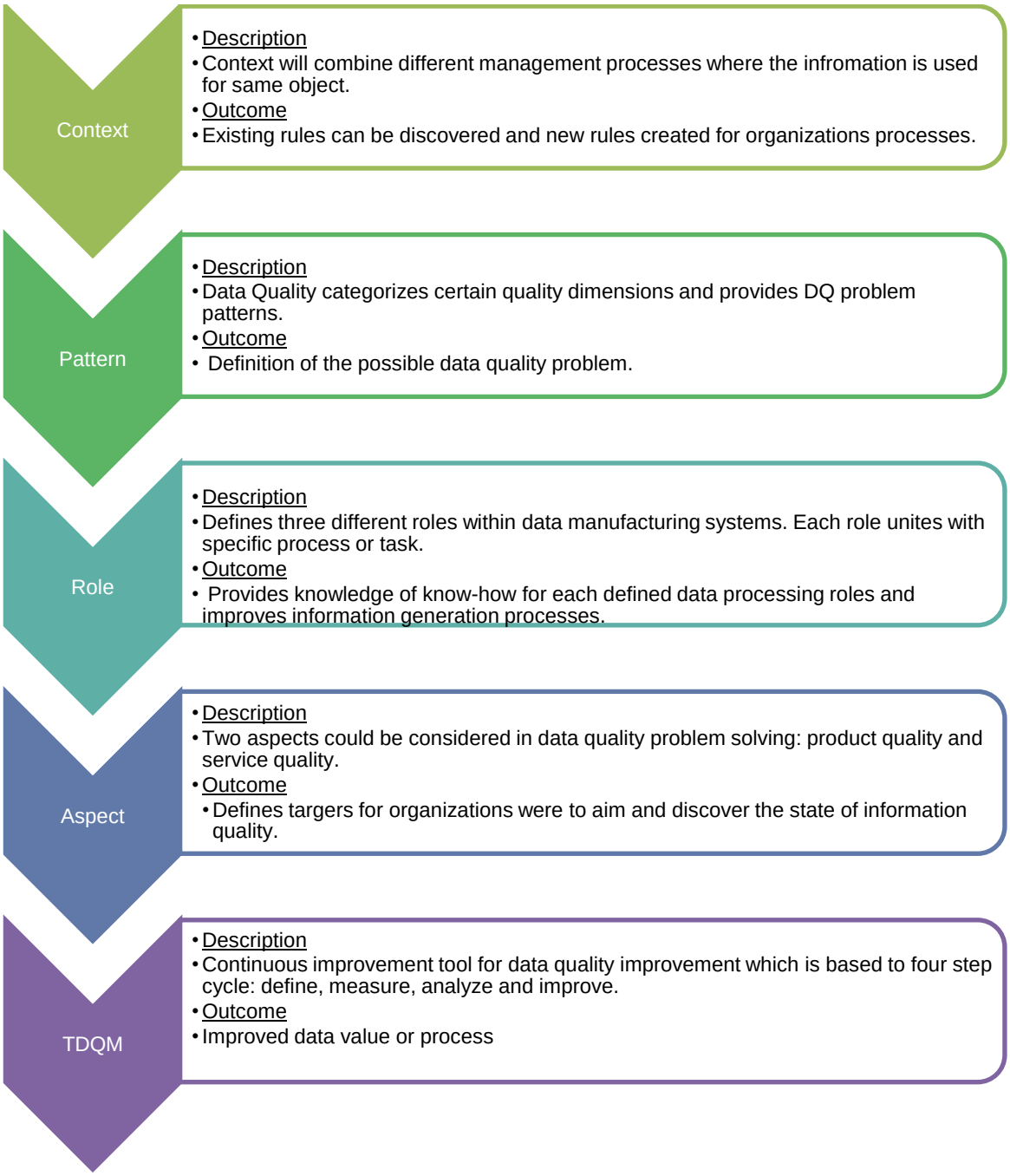


Figure 20. Steps for data quality problem solving.

As the Figure 20, shows the approach could be represented as a five step process. The process starts from definition of context and from there definitions will be made to category, role, and aspect and finally utilize TDQM tool for executing the data quality problems which were discovered from the other methods. All primarily methods complements to us-

age of Total Data Quality Management as it uses to same dimensions, patterns, roles and ideology in the same way, but actually implement it to processes.

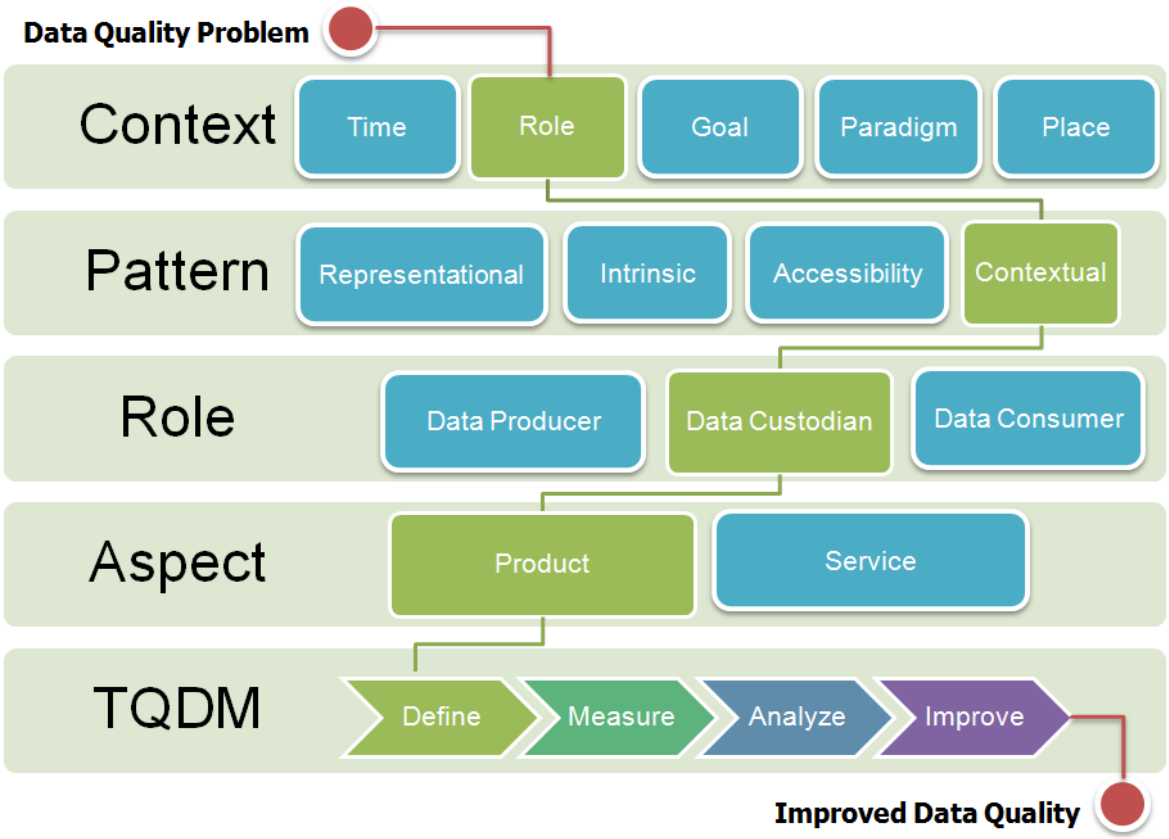
5.3.2 Proposed Approach for Data Quality Problem Solving

This study only presents an approach for data quality problem solving which the case company can utilize. As there are many methods available which enable many possible variation of the data quality problem solving. For better understanding of the proposed approach, is to demonstrate it with simple road map which organizations could use when defining data quality problem.

When the data quality problem has been recognised or wanted to be researched there could be use following approach: Defined path could be simply be chosen from the existing methods inside the main methods which would be selected with the idea, what is need to be involved. Using this approach would involve several different high quality data dimensions which could be analyze one-by-one or as a whole, but it could be challenging. Despite that the approach gives opportunity to select many different methods and dimensions at the same time, but for clearer approach only one selection of each step is chosen.

The option for approach is presented in Table 15. This demonstrates the path from outermost (context) method to last total data quality management (TDQM) tool. In this case the first definition would be done in *context* method where the role-option would be chosen. The next step would be to select fitting *pattern* as in this case the contextual pattern. From role method a data custodian *role* would be selected as it has connections to role-context. The *aspect* is chosen to be product as physical products are involved in the business. After these definitions of four different methods the actual definition is primarily done for the last TDQM step. With help of the continuous cycle of TDQM, the defined data would be measured and then analysed. Depending of the results and conclusions the needed improvement would be made straight to the data or more preferably to existing processes that actually produces the low quality data.

Table 15. Initial proposal for approach implementation.



As seen in the Table 15, the selected methods would provide certain knowledge about certain quality dimensions. One of those quality dimensions would be *completeness* which is common factor for three of presented methods above: pattern, role, and aspect. The role-context connection to completeness comes via the role of data custodian which is defined in the role context, among data producer and data consumer. Though this approach would require some knowledge of the data quality dimensions in the first place and therefore could be challenging to utilize it in the beginning. The following Figure 21 presents the common connections for completeness dimension i.e. the data quality dimension for different methods.

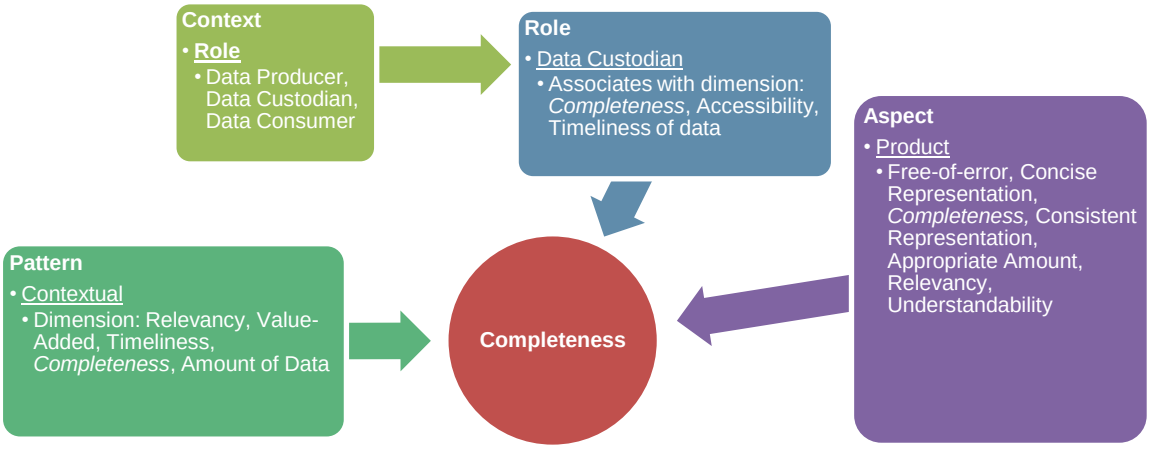


Figure 21. Common quality dimension: Completeness

As seen from the Figure 21, the completeness is one of the many data quality dimensions which are included to different methods. This quality dimension could be seen as important, as it describes the information which is usable for task at hand i.e. this dimension is referring to information which is useful as it exists in the current format. Another common dimension was also *timeliness*, which was founded from pattern and role methods, but not from aspect method. As this approach ended to point out only one data quality dimension regarding to this conclusion, the solution could also be used another way around. Regarding to the previous Table 15 example, the completeness could also be selected first from the dimension group (see section 4.2, Table 8) and then discover in which different methods it is involved.

5.3.3 Summary of Data Quality Problem Solving

The actual initial proposal combines few different data quality problem solving methods which also could be used individually to solve specific problem are in data quality. Either way the company must have rules and routines how to manage data production. This also means that pattern-based solution must document, understood and used in the data quality problem solving. The ideal situation would be that data quality knowledge, problem context and solution tools would be working together in problem solving. Regarding this it would be necessary to incorporate both technical and organizational observations into solving data quality problems. Figure 22 presents the approach for Data Quality Problem Solving diagram as whole and defines the different data problem solving models in meth-

ods which have some kind of connection to each other that gives opportunity to utilize the received results of data quality between multiple methods.

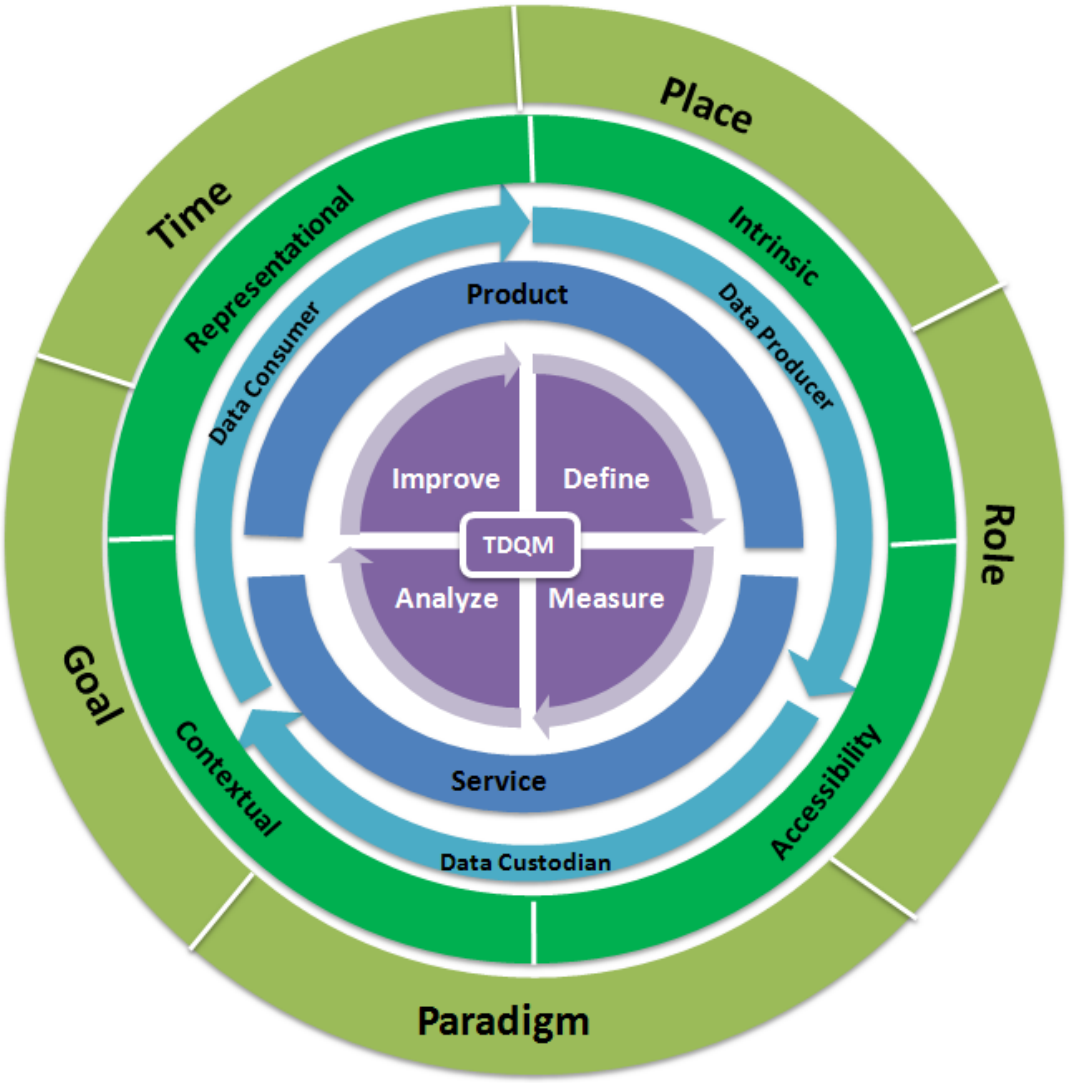


Figure 22. Approach for Data Quality Problem Solving -diagram.

As seen in the Figure 22, the approach for Data Quality Problem Solving constructs of four methods and one tool, which all can be used independently. From the perspective of outermost method (Context) to Aspect method, all four models can be used for data quality problem solving. TDQM cycle can be used as an actual tool for continuous data information quality improvements. The meaning and focus of this Data Quality Problem Solving approach is to create and bring new ideas for organizations, to question the current data production processes, and also to evaluate their information quality and based to received results and metrics establish improvements to generate high quality information.

6 Validation of the Proposal

This section discusses the validation of the hypothesis of the proposed approach process steps for the case company. The outcome of this section is the final action plan approach which defines the next steps for the case company to approach to data quality problem solving.

6.1 Overview of the Validation Stage

The third and final data collection step involves presenting the proposal for Item Data Manager and few other independent consultants of the case company who validate the proposal and give feedback of the initial proposal. Best feedback was received from these people because they have the best knowledge of the itemization processes that is ongoing in the case company. The interviewed two independent consultants also have been working for the case company for several years and have strong background relating to data and information systems and ideology of the product data management. All the interviews were tape recorded but very specific field notes was not written of received feedback for this study.

The final proposal draft was build regarding to received feedback and question that appeared during the first proposal round. The methods themselves were not seen as an issue, by contrast the initial proposals approach brought up question that how the proposed approach would actually be utilized. There were question regarding how to actually use this proposed approach and what methods especially could be measured from data quality point of views, and would it be relevant to measure the state of service product information. This was a relevant question from the case company point of view as their interest is more in the practical level. A big question also related to difference between Aspect levels product and service model and what they actually meant in this case. Although there were many questions mostly regarding to approach, the green light was given for the proposal and it got approved as a valid approach after improvements.

6.2 Findings of Data Collection 3

For the final proposal some changes had to be done and clarify some steps of the proposed approach and especially the definition for the approach to actually utilize the different methods. Regarding to key Manager the first implementation could be concentrate for service product information perspective to analyze information quality with idea to improve it someway.

Feedback from first proposed approach model was encouraging, but still something was missing from the actual proposal. There was missing the actual clear proof how the actual approach would be done. The initial proposal presented it by just picking certain fields form different levels, but this would not have been very logical without a good knowledge of the data quality dimensions and their connections to different fields in the levels. Utilizing this feedback some changes to the approach was done and the layout of the presentation was improved to be more logical and to be clearer to understand.

Based to feedback from key manager, there should be a clear understanding of the definition of product and service aspect and what kind of knowledge can be gained to investigate these areas from the information quality point of views. As phrased from the feedback:

> *The next step regarding this proposal would be to actually do a field test in practice and analyze service information quality (Data 2, Feedback 1).*

The information and data quality should be improved and especially all the data from company acquisitions that has been migrated from old data base to new one. Big question will be how the actual installed base data could be enriched to use or should a decision made that it won't be taken to use at all. Answer for these things could be the utilization of the quality data problem solving methods and tools.

Despite the relevant need from the case company point of view would be to have actual practical model for use, this study only concentrates to presents the proposal as an approach solution and point of view which could be utilized and amend as needed for the

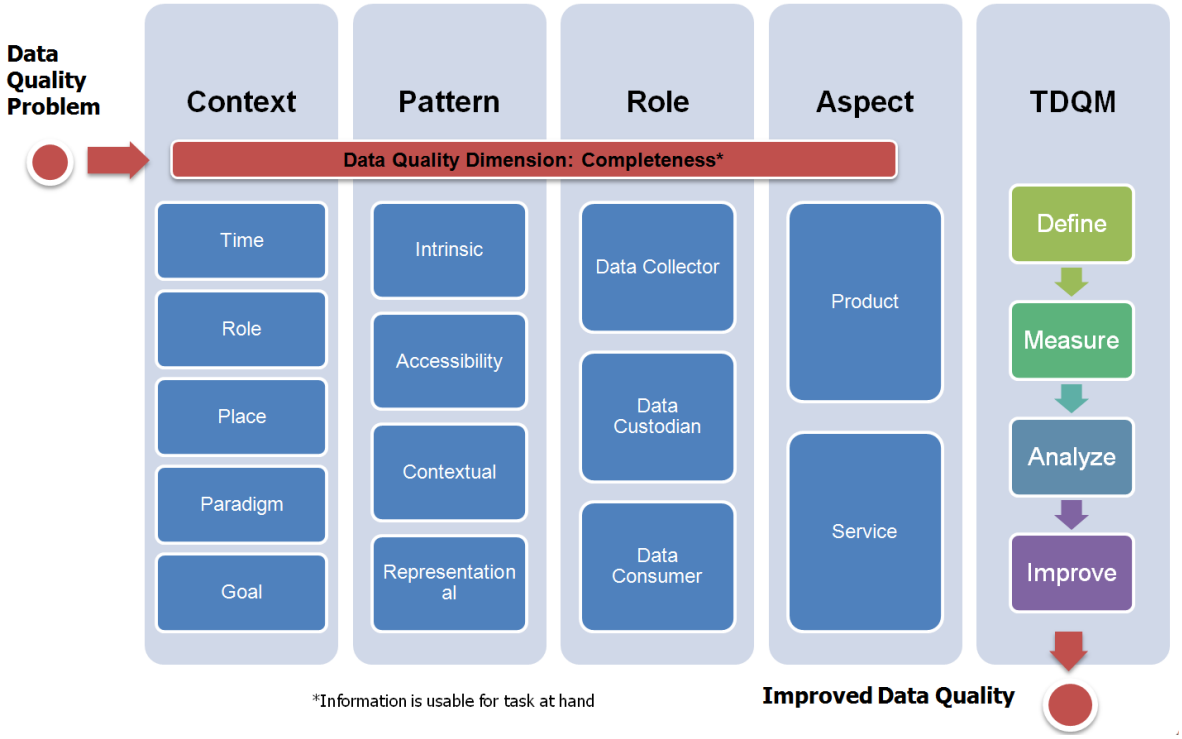
case company. Even thought to practical proposal was not included to this study, the case company stakeholders saw that this proposed approach could be still exploit.

6.3 Final Proposal Approach

The key things of final proposal is actually utilize the ideas of the data quality problem solving and implement the presented idea in a way that it recognizes the information quality state regarding to current itemization processes. The proposed methods of the approach will stay the same, but implementation approach of the final proposal will present improved version of how the ideology of presented proposal works and can be used to “real life” situations when approaching data quality problem solving. During this study the actual implementation was not done and relevant field test will be done later on. As the management approval has been received the actual use of the proposed model should be executed to see does it work as planned. This execution would be own project as it would need major preparations and research before it could be tested.

Regarding to the initial proposal (Section 5.3.2, Table 15) example, the *completeness* should be selected first from the quality dimension group (see section 4.2, Table 8) and by researching existing methods to discover in which different methods the completeness dimension is involved. Better option would be to contemplate the approach from the high data quality dimension perspective. This would be very clear approach as in the first phase researchers choose the desired high quality dimensions what to measure. By utilizing defined high data quality dimension and discover the connections of select dimension to different problem solving methods would provide transparent overview how and where the information is connected and where does it have an effect. Using this kind of approach would be also a good base for iterative processes to systematic and continuous data quality improvement. The actual approach would be begin by simply selecting one of the high data quality dimensions and define it as an essential problem or unclear dimension in existing information quality problem which is wanted to be solved. For example, to approach the problem one high data quality dimension: completeness is selected as it describes information that is usable for task at hand, as seen in Table 16. In the first phase there is no need to or cannot be know exactly which methods would be related to this dimension.

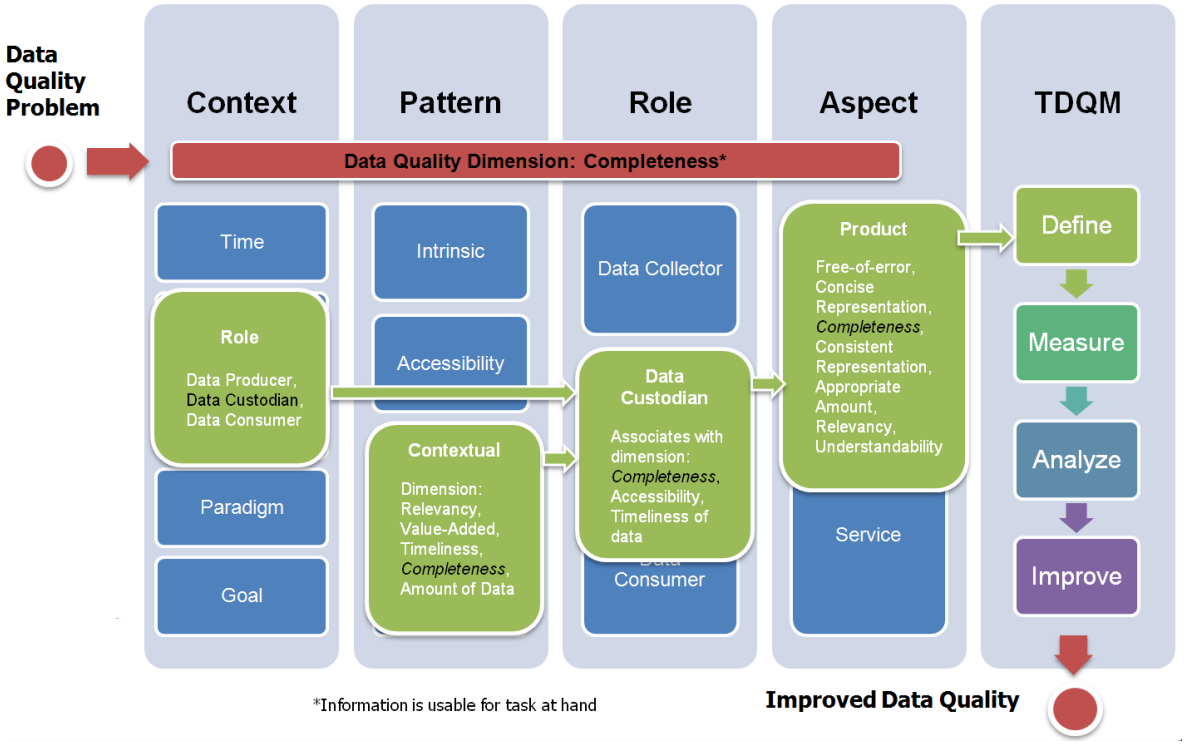
Table 16. Final approach, using Quality Dimension: Completeness – Phase 1.



As seen from the Table 16, all definitions of different methods are available before the specified high quality data dimension is chosen. In the first phase the survey concentrates to four different methods as the actual TDQM tool will be used in the second phase after first results are received from phase one.

Using this approach the phase 2 will give a good overview and discovery how the high quality dimension is connected to all methods and how they are connected together and which way. The Table 17 presents how *completeness* high quality dimension is related to context, and pattern, role, and aspect methods.

Table 17. *Completeness* dimensions connections to different methods – Phase 2.



As seen in the Table 17, the completeness high quality data dimension can be discovered inside different methods which are combined in proposed approach. The completeness can be found from three different methods: pattern, role, and aspect. The dimension leaves the fourth context-method out from these three because it does not directly include quality dimension definitions; by contrast the role-context is connected to completeness dimension via role-method which incorporates data custodian role. As there are sixteen defined high quality data dimension they all have specified meaning. This definition helps to understand these connections between different methods as the actual purpose of the dimension stays the same despite in which method it can be discovered. In this example the common factor for these methods is that information is not missing and it is sufficient breadth and depth for the task at hand (completeness).

Using one of high quality data dimensions for approach the definition for define-state of total data quality management –tool can be stated. This means that certain amount of items could be viewed from certain point of view and measure how the included information response to defined task at hand i.e. is the existing information useful or not for defined task. Next step with TDQM would analyze the received results e.g. from hundred

items only twenty percent are usable. According to analyzed result the final step would be taken either as an improvement of the data or the entire process which provides the actual low quality data for itemization process which involves all the information producers.

As result of this one dimension connection between methods the organization could either concentrate to utilize only one method which would also involve other high quality data dimensions or just one dimension e.g. completeness, to start improve their processes for better information creation. The questions for these would be: how the completeness is seen from contextual-pattern point of view, data custodian-role point of view and product-aspect point of view, and how the role-context is connected to completeness dimension? These are the matters that organization should be recognized and posses' knowledge and understanding to be able to improve their itemization processes to create high quality data.

The final proposal discovers the actual root causes of itemization processes, which could be improved using the information quality problem solving methods and approach the problems as described previously. These improvements should provide standardized work actions, training and common knowhow of the intentionality of information quality to employees and this also would create more collaboration for different departments and make the information connection (within items) more transparent and easier to utilize for different tasks in different units of the case company.

7 Discussion and Conclusions

This section discusses about the conclusions of the Thesis by summarizing the findings, validating the results of the study and by suggesting directions for further studies on the subject.

7.1 Summary of the Thesis

This Thesis focuses on improving data management for better itemization processes in the case company. The case company is currently improving its product lifecycle management to increase profitability. The strategy for it is to increase component knowledge accuracy, reduce delivery time of the products, and increase the cost competitiveness. As the product lifecycle management involves itemization processes, it is important that the information used for it has certain quality to meet the requirements of the itemization processes. Especially the lack of confirmed information creates challenges for the itemization team.

Therefore, the focus of this study was to identify and solve the root causes behind the problem why confirmed information in the product data management is not systematically received. This problem affects the processes where items are created and maintained and where there is a critical need for complete and high quality information.

The objective of this Thesis is thus to create an approach to improve the itemization process which would increase the quality of data in terms of component knowledge accuracy, reduce its delivery time, and would, in the end, increase competitiveness towards competitors. The study is conducted by, first, analyzing the current state of itemization process, combining the best existing knowledge, and creating an approach for information quality problem solving. This approach is then applied on an example case showing how to solving the data quality problem for the benefit of the case company.

The outcome of this Thesis is an approach for data quality problem solving. The approach consists of different methods based on: context, role, pattern, and aspect of information quality problem solving, which could be utilized in data management, and that can be used

together or individually to improve the information quality, and a TDQM (total data quality management) tool which can be used for continuous information quality improvement. The approach presents how to utilize the approach by showing the connections of high data quality dimensions between the methods.

The approach proposes to utilize one by one the high data quality dimensions for improving data quality. First, the selected dimension is determined, and then it is used to guide the data production from the point of view of all methods merged in the proposed approach. Second, after the connections of the selected dimension are discovered among the different methods of the approach, the next step is to utilize the final TDMQ element of the problem solving approach as the definition part is ready. Third, after the definition, a certain bunch of items are measured regarding the quality of information that is included to them. For example, hundred items are measured by the definition of completeness i.e. the fit of the information for use and if it is complete enough. Fourth, after the measurement, they are analyzed, and a decision has to be made to either update the information or more preferably change the actual process that creates the low quality information for items. The last phase is to implement the decided actions into production or existing processes.

Using this kind of approach, the biggest advantage for the case company would be first, to increase their knowledge about their information quality and discover how the information is connected to different processes, and how it is related to the current data production roles and how these matters are connected to different responsible design organizations. Secondly, the proposed approach could improve the case company's processes related to itemization by increasing the information quality of produced item data, which could have essential impact for company's productivity and competitiveness.

7.2 Managerial Implications from the Thesis

The information is currently scattered all around, but still abundant, in the case company. The company produces daily so much information in different forms that it is impossible to storage all of it, and it is not even necessary. The top level management's knowledge and comprehension of the information that is produced in different departments of the organi-

zation would help to create more clear and logical picture of the organizations model. As the information is key element for every action and managing products in the future, like it will be the case with internet of things (IoT), there should be clear definitions for how the certain data is used and stored. This should be first priority and defined to common operational handbook where everyone could discover their place and work actions what to do, how and where to do it and why to do it. Meaning of this is that employees should be aware what information they are producing and why it is important that it is done in a certain way and how does their actions effect for the actions where that specific information is needed.

The proposal approach for the data quality problem solving should provoke question for managers to actually reconsider what information actually is valid for certain operations and how much effort should be placed to enrich information to a stage where it can be seen as high quality information in the items. The enrichment question is a challenging situation where there is a need to make a decision, that is the information usable or would a better solution just to create new information from scratch by defined quality standards. This decision is quite complex and would need evaluation from many different angles e.g. is the history important, what would be the benefits to migrate legacy data to new systems and enrich it, and with what costs.

Better information quality would provide better service, analyzes and increase value of the company. High quality information would provide better transparency between departments and increase the liability of the information as it would be managed everywhere as confirmed information. Ideology for the management should be: invest now and save in the future. Discovering data quality and problem areas relating to it could bring major cost savings for company in long term. The result of this research can be seen that information quality has major influence to in different fields and actions where the information is used. With increased information quality there will be more efficient processes for item process management and also to cost management.

Role context has major part in data quality problem solving. Obviously individuals play important parts in data process management and therefore only people can improve or fix the existing problem which they have caused at the first place. An important thing would

be also to know what is the data quality level and what organization could do to improve it from the creation to obsolete state e.g. when considering to improve product lifecycle management. For future information standardization rules there should be considered to use concept of information production manager which would provide managed information determinations. Nonetheless, producing high quality information would provide more efficient and valuable itemization processes which would lead to better products and services and therefore increase productivity and also value towards customer.

7.3 Evaluation of the Thesis

The benefit for the company is that this approach can point to methods how to improve the quality of data and the processes related to managing it. In the context of the case company, it could provide a more efficient itemization processes that would benefit the case organization and its stakeholders.

7.3.1 Outcome vs. Objective

The Thesis objective was to create an approach to improve itemization process. This requirement was met and the objective was addressed through two outcomes: an approach for data quality problem solving methods, and an approach how to utilize it in the organization. First, the approach was build up from the discovering of the current state analysis which leads to actual root cause of the business challenge by pointing out information quality. Approach was created to give more understandable overview to understand item information quality and arouse questions what classifications and definitions are related to it. Secondly, an approach was present how to actually utilize the approach and how an organization could use it to measure their information quality level. The approach explains how to define certain quality dimension and how the connection is seen from the approach.

The challenge in the beginning of this study was that there was not a clear visibility what was the actual cause for lack of confirmed information. The proposed approach, although it is a more ideological point of view, supports the actual need to solve the existing problem by discovering different methods. Even though the implementation and practical way

to improve itemization process was not done. Regarding to the outcome, the final proposed data quality problem solving approach, which was set as a target for the study, was produced along with an approach. It was seen as a good result as they create questions for organization to how actually use their information and what aspects should be taken to notice, and therefore the objective of this study can be perceived to be achieved.

7.3.2 Reliability and Validity

In this Study the section 2.4 defined the reliability and validity plan which are important aspects that the study is truthful and can be declared to have valid context.

To increase the *validity*, this study used multiple instruments for data collections which included discover of internal documents, interviews and feedback / discussion sessions with the case company informants. This increases especially the internal validity as each of the five interviewees were working in different departments of the case company and closely with itemization processes and all the interviews constructed from same questions regarding to itemization and were carefully transcribed, analyzed, summarized and added to the appendixes. Likewise, the best knowledge was searched out to from several reliable sources which mostly were articles from different researches. For example Kahn, Strong and Wang (2002) presented survey of PSP/IQ instrument (section 4.3) which has been developed, tested and used for another research study and has received validity and reliability. Other described methods included in the final proposal's data quality problem solving approach are also experimented and tested in another research study and are therefore seen as increasing validity for this study.

Quinton and Smallbone (2006) argued that key test of validity could be presented as following: Was what was found a response to the questions originally asked? As this study was a qualitative research and during its progress the approach emphasized to be more inductive the internal validity can be seen as particular strength of this study. Because there was so much data collected, that it can be seen in itself be sufficient to tell the subject of this study. Therefore it can be seen that the answers what were found answered to the question how to improve data management, although the answers were received by

approaching the challenge from upper scale, and from information quality problem point of view.

Reliability of this study is increased as the proposed approach is transferrable to any other organization which would need to exanimate their information quality and trying to improve it. This thesis also includes appendixes from all five interviews which are transcribed as accurately from the tape recordings of the interview and are therefore visible to see what was actually answered to each question. From the *authenticity* point of view it can be said that the thesis educative authenticity as it presents different methods to learn how to approach and examine data quality problem and how to solve it in a level which could be used to variety types of information. Catalytic authenticity comes from the new aspect to actually analyze and measure information quality utilizing proposed approach and approach. This creates curiosity to change methods to provide higher quality information.

Trustworthiness could be addressed by use of different data sources: internal documents, process maps, interviewees, and feedback from different fields of the case company. Secondly, involving people with many years experience of data management for the feedback and who could validate the proposed solution. Significant is also to apply established theories and utilize them as a one approach for problem solving which could be used to same problem solving without depending will it be done next week of after several years.

7.4 Closing words

Information is more and more increasing around us and continuously there are developed new ways to utilize it. During this thesis a whole new world relating to item data and information opened for the researcher and increased his understanding of the things how information can be defined, measured, categorized and improved. By now it can be said that better knowledge of the information as a subject could provide significant discoveries in the future. Although this was just a small study in a big pool of information it has inspired the researcher to gain more knowledge in the future of how information could be used more for different things in business and everyday life.

References

Abraham, J. (2014). Product Information Management: Theory and Practice. Management for Professionals. 1st ed. Switzerland: Springer.

Baxter, P. and Jack, S. (2008). Qualitative Case Study Methodology: Study Design and Implementation for Novice Researchers. *The Qualitative Report* Vol. 13 pp. 554-559.

Coughlan, P. & Coghlan, D. (2002), Action research for operations management. *International Journal of Operations & Production Management,* Vol. 22. pp. 220-240, URL: http://dx.doi.org/10.1108/01443570210417515

Crnkovic I., Ulf, A., Dahlqvist A. P. (2003), *Implementing and Integrating Product Data Management and Software Configuration Management*. 1st ed. Norwood: Artech House.

Detlor, B (2010) Information Management. *International Journal of Information Management.* Vol. 30(2), pp. 103–108.

Huner, K., M., Ofner, M., & Otto, B. (2009). Towards a maturity model for corporate data quality management. *SAC '09 Proceeding of the 209 ACM symposium on Applied Computing*. pp. 231-238

Kahn, B., Strong, D., M., & Wang, R. (2002). Information quality benchmarks: Product and service performance. *Communications of the ACM*, Vol. 45(4), pp. 184-192.

Lee, Y. (2003). Crafting Rules: Context-Reflective Data Quality Problem Solving. *Journal of Management Information Systems*, Vol. 20(3), pp. 93-116.

Lee, Y., & Strong, D. (2003). Knowing-Why About Data Processes and Data Quality. *Journal of Management Information Systems*, Vol. 20(3), pp. 13-39.

Lee, Yang W;Pipino, Leo;Strong, Diane M;Wang, Richard Y (2004). Process-Embedded Data Integrity. *Journal of Database Management,* Vol. 15(1); pp. 87-103.

Madnick, S. & Wang, R. Y. (1992) Introduction to Total Data Quality Management (TDQM) Reasearch Program (TDQM-92-01): *Total Data Quality Management Program, MIT Sloan School of Management*. p. 1-8.

Madnick, S. E., Wang, R. Y., Lee, Y., & Zhu, H. (2009). Overview and Approach for Data and Information Quality Research. *Journal of Data and Information Quality (JDIQ)*, Vol. 1(2), pp.1-22.

Näslund Dag, Kale Rahul;Paulraj, Antony (2010) Action Research In Supply Chain Management-A Approach For Relevant And Rigorous Research. *Journal of Business Logistics*, Vol. 31(2) p. 331-355.

Otto, B., & Reichert, A. (2010). Organizing master data management: Findings from an expert survey. *SAC '10 Proceeding of the 2010 ACM Symposium on Applied Computing*, pp. 106-110.

Otto, B., Lee, Y., & Caballero, I. (2011). Information and data quality in networked business*. Electronic Markets*, Vol. 21(2), pp. 79-81.

Outotec Internal Documents (2016). *Item Concept*. [Accessed 16th February 2016].

Panahy, P. H. S., Sidi, F., Affendey, L. S., & Jabar, M. A. (2014). The impact of data quality dimensions on business process improvement. *Information and Communication Technologies (WICT)*, pp. 70-73.

Pipino, L., Lee, Y., & Wang, R. (2002). Data quality assessment. Association for Computing Machinery. *Communications of the ACM*, Vol. 45(4), pp. 211-218.

Quinton, S. and Smallbone, T (2006). *Postgraduate Research in Business: A Critical Guide*. 1st ed. London: SAGE Publications Inc.

Saaksvuori, A., & Immonen, A. (2008). *Product Lifecycle Management*. 3rd ed. Berlin: Springer.

Shields M. D. and Young S. M. (1991). Managing product life cycle costs: An organizational model. *Journal of cost management for the manufacturing industry*, pp. 39-51.

Solomon, Paul (2002). *Discovering Information in Context*. Annual Review of Information Science and Technology Vol. 36 pp. 229-264.

Stark, J. (2011). Product Lifecycle Management. 2nd ed. London: Springer.

Strong, D., M., Lee, Y., & Wang, R. (1997). Data quality in context. *Communications of the ACM*, Vol. 40(5), pp. 103-110.

Turner, Daniel W. (2010) Qualitative Interview Design: A Practical Guide for Novice Investigators, *The Qualitative Report* Vol. 15(3), pp. 754-760, URL: http://www.nova.edu/ssss/QR/QR15-3/qid.pdf

Yin, Robert K. (1994) *Case Study Research Design and Methods*. 2nd ed. Washington : SAGE Puplications Inc.

	Current State Analysis Interview Questions
1	How itemization is connected to your work?
2	How are items used in your department's daily processes?
3	What software's are used for daily processes? (E.g. JIRA, ENOVIA, EDM, Vault, EBX, SAP, Maximo)? - How they are used?
4	How Component Engineering supports your processes?
5	What information you especially use from item information? (e.g. MEP code) - How information can be verified? (How to know that correct product has been purchased with manuf. Code which defined to item.)
6	What are the strengths of using items? - Why we itemize?
7	How should the item information be described to be clear? (e.g. should technical description contain specific information) - Where information is stored (e.g. special lists not in Enovia? Etc.)?
8	How would you describe the strengths/weaknesses of Key User concept?
9	How concept guidelines / used tools are informed/presented to employees?
10	What could be biggest "bottle necks" of itemization concerning your department?
11	How itemization supports the way of working in your department? (Difference between past and current) - What could be done differently? - What could be improved in the future?

Summary of the Current State Analysis Interviews					
Interview question	**Item Data Manager**	**Quotation Support Manager**	**Senior Product Engineer**	**Head of Automation &Instrumentation (Delivery)**	**Director of Service Spare Part Marketing/Sales**
1.	-Connected to service and engineering departments who are closest co-operation departments	-Quotation support. Recognize legacy items and upload them to SAP. - Towards Component Engineering itemization model (valid quality items for library)	- Global equipment delivery center – all core components have been itemized	-Instrumentation does not use itemization at the moment. -Automation has some itemization done for some plant deliveries -Equipments are managed by Outotec's own units (they use more components)	-items used daily
2.	-Component Engineering creates and upkeeps standard and commercial library. Library contains only validated items.	-Items needed for quotations. - Way to communicate -After quotation item should be contain enough information that it should be only used for moving components.	-Core components are designed thoroughly using items -Core components: all products that have technology in side them -> excluding steel structures etc. -Core components are included to eCatalogues Upper level design components (assemblies) are only defined to Design-library. Only this can be found in SAP -Flotation (product line) will define themselves what electric motor will be used	-Itemization is seen as too complicated and providing extra work -Itemization is not used (plant deliveries) - Purchases are done with category items. One "row" is for the whole purchase (all needed parts and components are listed under it). Attached specification documents will indicate the technical details of the products.	-Quotation and selling items -All transactions are done with items (quote, deliver, invoice items) -Item usage is mandatory
3.	-ENOVIA	-SAP (main software)	-ENOVIA	-Plant documentation with	-SAP

	-SAP -JIRA -EBX	-JIRA (for quotation orders) -EBX -Maximo (workflow) -Legacy systems -SAP material request tool	-EDM -In SAP a WPS structure is defined	DMS Meridian -SAP used for purchases -AVEVA,SPI,ePLAN for design	-Maximo for workflow -JIRA for RFQs -Aton (filters) -eCatalogue (upload from excels not all from Enovia)
4.	-Component Engineering only offers the support for internal customers.	-Determination request are handled via CE -Enrichment of item data -CE work seen as positive	-A request for new item creation is made for Component Engineering team	-Not at the moment	-Not visibility how CE currently supports spare part -CE does not add visible value for Spare parts unit -Library full of stuff that Outotec won't sell
5.	-Sales description, MEP code, weight (Enovia) Lead time, price, Country of origin (SAP)	-Sales + technical description, purchase order text -Basically item could be purchased without datasheet -For service it does not matter who has the item owner ship -> information needs to be valid	-Technical description -E.g. when exploring different valves the MEP code does not tell what type the valve actually is. -MEP code supports items recognition and will be compared to product lines internal lists (documents) which are defining the correct project code – like in FESTO valve project code. -E.g. in technical description; frame size, frequency, voltage and Sales description; power, rotational speed (decision can be made from these attributes) all other optional adds have to be checked from datasheet	-I&A has codes in their product but the code does not contain all needed information to define the product. -Besides the code there are many details to be defined e.g.: 1.Calibration / measurement area (for instruments) 2.Oxygen washed part (for oxygen use) 3.ATEX certification (for electrical equipment in explosive atmosphere) 4.A tag (which is used for part recognition in the plant and contains project specific information)	-Identifier is DWG number or MEP code

6.	- Will serve future business. - Possibility to recognize all the spare parts sold and enables use of eCatalogue for online sales for customers use. - Trace ability and for storage management.	-The whole process from quotation to delivery can be tracked -Item lifecycle easier to manage - controlled in warehouse management system, prediction of usage and estimation of re-order points - Possibility to specify things relating to vendor -> Different vendors have different kind of systems to specify things e.g. some has their own item numbers, others might use only specified descriptions	-Product management is much easier -Costs can be monitored much more precise -Possibility to reuse items (transactions with same code)	-At the moment too complicated to itemizes instruments	-to reuse items -recognize items -link to project and product context
7.	-Sales description should be clear.	-Possibility for errors increases if technical description is not used efficiently	-Technical description is important -All needed information is found from sales description and technical description -Sometimes info text is used for comments -Purchase text is used when an actual purchase will be made with items that is owned by product line	-Purchase process: 1.An excel will be made which defines the need for equipment 2.Supplier makes offer of the required components (there can be more than one suppliers) 3.Offer is compared and validate 4.Offer can be used later for purchasing (contains valid information)	- In the past description was more free form – context was added to description (e.g. for this project/part/component) -> things were found… -technical description -> in SAP desc.1 and 2. -
8.	-Good to have specified people for the position. -Current concept not	-A good thing when it will work. Knowledge can be shared more efficiently.	-Enovia key users unknown	-KeyUser concept seen as good thing -SPI and ePLAN keyusers	-How that works in practice is not visible -Challenging situation

	working properly - chosen people don't have time to do the work.			exists	
9.	-Item management training- -Clear guidelines. -Change management as top development. -Support should be always available through key user concept.	-People are not trained enough -There has been lack of instructions -More practice is needed before it will have better operative effect	-Place for instructions "not clear" -Long time from trainings – should be repeated -Item Wiki – unknown -Item request tool – use started beginning of 2016	-Own processes and tools could be improved more -Change management should be improved (AVEVA could be solution) -During OPAL – many things were deployed in unfinished format E.g. UAT	-A three year journey where guidelines are made are instructions for every step -Instruction how to find item without Enovia etc. -Problem situations are flagged to process and concept owners -Item guideline for service – by Jeremy Spencer (modified to role specific needs for service)
10.	- Organizational model - Replacement process (who is responsible of the decision) - Change management - Amount of legacy items and their enrichment work	-Demand of items can be seen as "bottle neck" for Quotation Support -Itemizations "bottle neck" is to demand certain information to items -Instructions -Global harmonization – the same way to do things globally -Product lines as Responsible design organizations - at the moment only few product lines are taking care of their design items and have support to create and maintain them. -Training, comprehension and operative action model has to be	-When item is created – specifications for the item is/has been a problem ->Previously this took too much time (sometimes two weeks)	-Too slow way of working. E.g. one instrument item takes about 30 minutes.	-Find verified data for items – -Much easier to create a new item -Lead time is quite long for quoting - Legacy way of enhancement seems to be inefficient way to do item

		establish			
11.	- Product line involvement for itemization process - All spare part should be connected to same context (for traceability) e.g. SBOM could be solution.	-Old legacy systems are lacking of information of used items -> in the past company rely more to vendor knowledge (who know Outotec products) and did not see importance to have more quality data in their own systems. -There has been difference how should legacy data has been managed and used logical features e.g. how to read and find data in the legacy system (hard drive / network drive). -One goal is to maintain current legacy data to clear format that is shows reasonable for the company and will service customers in easy to use format.	-External lists are maintained by product line which contains items that are not in Enovia. This was started couple years ago (same time with OPAL started) -Clear workflow should be done to Enovia (in EDM at three stages) -Product lines have their own internal lists that are used for easier recognition of the components.	-Should be considered is it necessary to give “all choices” for customer or could there be some limitation or product where to choose. -Productization could be improved more in the future (as one goal in OPAL process) Some efficiency has been noticed because of productization	-Service tries to search from context (where the actual item has been used) 1.At the moment discovered context is not connected to delivered item 2.Lack of time – biggest challenge to make structures -Not possible to change suppliers because of lack of information -After spare part engineers have identified items and its context 1.Feels waste of time that they fill a form to create an item which they could do them selves 2.Should be more specific what the library is for and what should be added to library 3.”Vendor items” – a new type / policy definition for supplier items

Research Interview 1
Master Thesis – Data 1

TOPIC: Current State Analysis of the Case Company

Information about the informant (Interview 1) *Table 1*

Details	
Interviewers	Antti Korpiniemi
Name (code) of the informant	Informant 1
Position in the case company	Item Data Manager
Date of the interview	Feb 15th 2016
Duration of the interview	1h 6min
Document	Field notes + Tape recording

Field notes (Interview 1) *Table 2*

	Track time	QUESTIONS FIELD NOTES
1	**2.20-4.40**	How itemization is connected to your work?
		- Responsible of Outotec's the whole global itemization process - Defining how items support and serve Outotec's business - Connected to service and engineering who are closest co-operation departments - Regarding to company strategy: o Installed base program – which includes all Outotec's products that has been delivered to customers globally and collect the information to one place. - Installed base item are enriched to use to increase service business
2	**6.00-8.45**	How are items used in your department's daily processes?

		- Product management needs defined third party items in their structures
		  o Third party items needed with specification and verified information
		- Spare part service
		  o All the sold components must be recognized and
		- Items provide information of the structures were the products are actual build from and identification of components to the sensor-level-stage
		- With digitalization, information's gathered from plants and it will be used to link correct items to correct deliveries in correct plants.

3	**8.50-12.50**	What software's are used for daily processes? (E.g. JIRA, ENOVIA, EDM, Vault, EBX, SAP, Maximo)? - How they are used?

		- All mentioned soft wares are used in Outotec's itemization processes - Plant information is mastered in SAP - For instrumentation -> ENOVIA - For service -> SAP - Maximo is new software that is taken to use and it must be taken to notice when **defining requirements for itemization**
4	**12.55-14.30**	How Component Engineering supports your processes?
		- Valid question for other departments
5	**14.35-17.30**	What information you especially use from item information? (e.g. MEP code)
		- Same item information fields do not contain usable data for all departments - Every department have their specific focus area in item information - We need common information for items as a base o New information will supplement to item afterward (when used) - **Sales description, lead time, price, weight, country of origin** - **Lead time and price** – Priority! 1. MEP code for commercial items
6	**23.55-27.55**	What are the strengths of using items? - Why we itemize?
		- It will serve future service business - Possibility to recognize all the spare parts that is sold. Create spare customized spare part packages in advance and prizing them - Enables use of eCatalogue and online store concept where customer can buy product independently - As item numbers are Outotec's own it will "guide" customers to buy products from Outotec. - From product management perspective; items reuse ability i.e. possibility to use same item number when building products. - Better traceability and visibility for **storage management – quantities** - Best utilization of items comes from service (service business starts from product management – were it float from upstream to downstream).
7	**29.25-34.00**	How should the item information be described to be clear? (e.g. should technical description contain specific information)
		- Identification is number one priority and ability to find correct items - E.g. sales description should be clear and not reveal specifying information because it will be presented to customer. Customer should not be able to identify part from sales description. 1. Description should be recognizable for library users
8	**34.10-37.15**	How would you describe the strengths/weaknesses of Key User concept?
		+ Good to have specified people for the position - **Not enough time to do actual Key user work - biggest challenge** - Current system does not work in real life - Cannot say what would be a better model

9	**37.20-43.02**	How concept guidelines / used tools are informed/presented to employees?
		- Item management training - Clear guidelines - Change management will be top development - If problem appears – contact should be key user -> working key user support would be helpful
10	**43.10-51.05**	What could be biggest “bottle necks” of itemization concerning your department?
		- **Organizational model (old model where separated departments, Component Engineering and Engineering – not supporting each other) -> created difficulties** - Replacement process deploying -> difficulties with decision making (who’s responsible) - Change management -> Requests from JIRA (e.g. what changes could be done and who’s responsibility is to make decision for wanted action) - Item search from systems - One way to describe items in Library – a common format - **The amount of legacy items (Brownfield, installed base) and enrichment work to library** - Product lines are lack of managing skills for their own items and they don’t have enough support how to handle situation at the moment - Total of 23 product lines -> about half of them are using items - **How involve product lines to itemization process – major issue**
11	**60.05-66.02**	How itemization supports the way of working in your department? (Difference between past and current) - What could be done differently? - What could be improved in the future?
		- Process should be changed - Willingness to change appears - For better itemization product lines should be involved to co-operate more with Component Engineering - All spare part items should be connected to some context - (Context: meaning that item could be linked to certain product spare part which is part of delivered equipment to specific plant in certain site.) -> SBOM will be solution for this

Research Interview 2
Master Thesis – Data 1

TOPIC: Current State Analysis of the Case Company

Information about the informant (Interview 2) *Table 1*

Details	
Interviewers	Antti Korpiniemi
Name (code) of the informant	Informant 2
Position in the case company	Quotation Support Manager
Date of the interview	Feb 17th 2016
Duration of the interview	1h 25min
Document	Field notes

Field notes (Interview 2) *Table 2*

	Track time	QUESTIONS FIELD NOTES
1	**5.35-9.20**	How itemization is connected to your work?
		- Job description: Spare parts (Alpha) – to be spare parts and ware parts of Finland and Sweden - About 16 to 20 product lines - Quotation support: Goes through Spare parts legacy data and uploads legacy items to SAP. Service has its own itemization in spare part department. - Own itemization is for efficient and fast item delivery to spare part use (brown field) -> **idea is to go towards a more unified process as Component Engineering and Product lines have**. - Point is that items should be in same ”quality” level despite where it is made. - Itemization is all in all in their work.
2	**9.35-10.25**	How are items used in your department’s daily processes?
		- **Item is needed for quotations and created for that purpose.** - **After quotation item should be contain enough information that it only can be used to move components.** - **Items are way to communicate.**
3	**10.35-21.30**	What software’s are used for daily processes? (E.g. JIRA, ENOVIA, EDM, Vault, EBX, SAP, Maximo)? - How they are used?
		- SAP main software for delivers (for spare parts priority number one tool) - Enovia – document and information; to support SAP usage

		- EBX – for search tool - Maximo controls work flow -> tasks to quotation support are created from this system - **JIRA:** filters itemization and **quotation orders for suppliers management** (Request For Quotation i.e. RFQ process) – additional information requests handled via JIRA 1. **Items should be created for quote attachment** 2. **Document of the purchase will remain to procurement (email or own local hard drive / network drive – no specified process for purchase order storage)** - Other legacy systems where legacy data is stored (IFS, Pyramid) – not in operative use - SAP material request -> item request tool (should this be taken to use)
4	**21.35-32.35**	How Component Engineering supports your processes?
		- By enriching the used data (items) -> Big support - Quotation support items should be also be enriched (process still pending – no solution for this) – Component Engineering could help with this in the future. - **Determination requests are managed via Component Engineering (most visible support)** - **Quotation support sees Component Engineering work in positive sight**
5	**32.40-45.23**	What information you especially use from item information? (e.g. MEP code)
		- MEP code is very important, if valid code the whole item can be purchase with that information - The most important thing: all information together in the item will describe it perfectly / all the way - **E.g. some pumps do not have existing data sheet but attributes could be defined for identification (which exists in some legacy system) -> this information does not fit the sales + tech description.** 1. **Purchase order text information field** 2. **Basically item could be purchased without datasheet** 3. **Datasheet contains the best information to identify product** - **If configured commercial component** 1. **Should have clear description that contains programming, configuration, custom work, etc.** - **Third party components exists that contains specified information of e.g. installed software to item (in case where manufacturer has few options for this) -> then it is not configuration** - It does not matter who is has the responsibility of the items (Product line or Component Engineering – only item information matters 1. This only indicates who has the responsibility to maintain and validate the item data
6	**45.30-51.45**	What are the strengths of using items?
		- Example: 1. Good item, few minutes of work 2. Bad item, days of work and still not clear what the item is - At the stage when everything is itemized.

		1. **The whole process from quotation to delivery can be tracker from history (1. When item is created? 2. Who has created the item? From there item information can be delimited by the time frame, project, legacy system, market and where could it be related to? Etc.)** 2. **Item lifecycle can be followed, controlled in warehouse management system, prediction of usage and estimation of re-order points** - **Vendors' information** can be assigned to SAP. Quotations and purchased documents can be attached to these vendors. - Way to think: If I buy a pencil, what information should I have to purchase exactly the same pencil second time? (Color, shape, price, model, type etc.) - **Possibility to specify things relating to vendor** 1. **Different vendors have different kind of systems to specify things e.g. some has their own item numbers, others might use only specified descriptions (e.g. pump model X with fluid flow 100litres/minute).**
7	**53.35-64.30**	How should the item information be described to be clear? (e.g. should technical description contain specific information)
		- For sales description (only what customer sees) it is enough to know general information of the product. In some cases information could be dimension e.g. length which will separate items from each other. (Warehouses are outsourced so drawings are not given to their use). - Technical description should tell specific attributes clearly relating to item. If these fields are not enough long. Necessary text could be added to info text field also. - For commercial items the most important field is technical description to have attributes. For example: 1. If 100 pumps have same sales description and no technical description added, it impossible to select correct pump based to seen descriptions 2. **Possibility for errors increases if technical description is not used efficiently**
8	**71.35-72.35**	How would you describe the strengths/weaknesses of Key User concept?
		- A good thing when it will work. Knowledge can be shared more efficiently.
9	**72.36-75.30**	How concept guidelines / used tools are informed/presented to employees?
		- People are not trained enough - There has been **lack of instructions** - **More practice is needed before it will have better operative effect**
10	**75.32-78.35**	What could be biggest "bottle necks" of itemization concerning your department?
		- Demand of items can be seen as "bottle neck" for Quotation Support - Itemizations "bottle neck" is to demand certain information to items - Instructions - Global harmonization – the same way to do things globally - Product lines as Responsible design organizations - at the moment only few product lines are taking care of their design items and have support to create and maintain them.

		- Training, comprehension and operative action model has to be establish
11	**78.40-81.45**	How itemization supports the way of working in your department? (Difference between past and current) - What could be done differently? - What could be improved in the future?
		- Current place has been ready for itemization when the process began in the case company Outotec Oyj. - **Old legacy systems are lacking of information of used items -> in the past company rely more to vendor knowledge (who know Outotec products) and did not see importance to have more quality data in their own systems.** - There has been difference how should legacy data has been managed and used logical features e.g. how to read and find data in the legacy system (hard drive / network drive). - One goal is to maintain current legacy data to clear format that is shows reasonable for the company and will service customers in easy to use format. - Guidelines and instruction should be clear and complete for future use.

Research Interview 3
Master Thesis – Data 1

TOPIC: Current State Analysis of the Case Company

Information about the informant (Interview 3) *Table 1*

Details	
Interviewer	Antti Korpiniemi
Name (code) of the informant	Informant 3
Position in the case company	Senior Product Engineer
Date of the interview	Feb 23rd 2016
Duration of the interview	1h 45min
Document	Field notes + Tape recording

Field notes (Interview 3) *Table 2*

	Track time	QUESTIONS FIELD NOTES
1	**1.45-3.25**	How itemization is connected to your work?
		- Global equipment delivery center – all core components have been itemized - Daily usage of items - Product line will make own drawings and components (pro-priority items i.e. design items) and uses 3rd party commercial components
2	**3.40-18.05**	How are items used in your department's daily processes?
		- **Core components are designed thoroughly using items** - E.g. metal frames are built only from drawings by third party manufacturer – these structures are only with "top level" item i.e. assembly which contains the whole design structure and is sold with one item number - E.g. tanks are always different 1. Meaning no need to itemize - Flotation cell tank sizes vary from 5 to 650 cubic meters - **Core components: all products that have technology in side them -> excluding steel structures etc.** - **Core components are included to eCatalogues** 1. Customer are able to make spare parts orders independently via eCatalogue - Service will make eCatalogues - **Upper level design components (assemblies) are only defined to Design-library.**

		- **Only upper level items are connected to specific plants in SAP e.g. all service plants etc. for sales in SAP.** 1. **Only this can be found in SAP** - **Flotation (product line) will define themselves what electric motor will be used** - Exception is in ball bearing units – Turula could order whatever they want if product line won't define exact ball bearing unit. - Same structure in EDM can be used for every similar project and needed changes are done in PDM. All items are not in EDM structure. 1. At the moment it is not considered to define structure all the way in EDM 2. Only the drawing is brought from EDM to items which is already created in PDM - New structures are copied from "Master structure" in PDM – to new projects 1. EDM is able to upload correct drawing to item (item number is defined to EDM before structure upload)
3	**18.12-30.00**	What software's are used for daily processes? (E.g. JIRA, ENOVIA, EDM, Vault, EBX, SAP, Maximo)? - How they are used?
		- EDM, ENOVIA, SAP (not in daily use, more like weekly) - EDM and ENOVIA for designing - **In SAP a WPS structure is defined** 1. **Items are connected to their own slots** - Project controllers opens project to SAP for the product lines 1. Defines order recognition templates for customer and vendor side 2. Product line can start their project before this 3. This is for cost tracking - No knowledge of Maximo
4	**30.15-36.35**	How Component Engineering supports your processes?
		- If need comes to use specific component e.g. Siemens motor and it is known that it will be used for several times (knowledge) 1. A request for new item creation is made for Component Engineering team -> library team - If component includes customization and configuration it is created as design component - There is a thing line when item is chose to be design project item or commercial item (mostly it is seen that motor will be used only once) – no need to be library item
5	**36.40-47.45**	What information you especially use from item information? (e.g. MEP code)
		- **Regarding to library search the <u>technical description is important</u>** - **E.g. when exploring different valves the MEP code does not tell what type the valve actually is.** - **MEP code supports items recognition and will be compared to product lines internal lists (documents) which are defining the correct project code – like in FESTO valve project code.**

		- **E.g. in technical description; frame size, frequency, voltage and Sales description; power, rotational speed (decision can be made from these attributes) all other optional adds have to be checked from datasheet**
6	**47.47-54.20**	What are the strengths of using items? - Why we itemize?
		- **At the moment all "commercial" components don't have existing datasheets connected to them** - **Product management** is much easier - **Costs** can be monitored much more precise - Possibility to **reuse items** (transactions with same code) - Most of the electrical motors are created as project items 1. Regarding projects a case can be that customer buys motors by themselves with defined specifications. For these "motors" an item is created in the structure for future recognitions. This type of item is not used for purchasing
7	**54.25-58.20**	How should the item information be described to be clear? (e.g. should technical description contain specific information)
		- Technical description is important - All needed information is found from sales description and technical description - Sometimes info text is used for comments - Purchase text is used when an actual purchase will be made with items that is owned by product line
8	**58.22-62.30**	How would you describe the strengths/weaknesses of Key User concept?
		- Flotation has EDM key users - **Enovia Key Users unknown** – should be? This concept does not have visibility. - In the problem situations product line contacts to Enovia administrators if problem regards to Enovia.
9	**62.34-72.02**	How concept guidelines / used tools are informed/presented to employees?
		- Place for instructions "not clear" - Long time from trainings – **should be repeated** - **Item Wiki – unknown** - Item request tool – use started beginning of 2016
10	**72.05-76.15**	What could be biggest "bottle necks" of itemization concerning your department?
		- When item is created – specifications for the item is/has been a problem 1. Previously this took too much time (sometimes two weeks)
11	**76.17-102.10**	How itemization supports the way of working in your department? (Difference between past and current) - What could be done differently? - What could be improved in the future?
		- **External lists are maintained by product line which contains items that are not in Enovia. This was started couple years ago (same time with OPAL started)** - Product lines have their own internal lists that are used for easier recognition of the components.

		- Few years ago – Component Engineering did not demand datasheets for all components, it was "optional". - SBOMs are not used at the moment in product lines - Current process supports idea that datasheets are required at the same time when a **quotation request is send to supplier.** 1. This has been changes from previous years - Demand for specification is seen as a good thing – a stage of quality - Use of ECR – should be established process 1. This would need guidelines for correct work around - Current "idea" for the process ECR -> ECO is too complicated and slow for the need which it should be used. 1. Change management should be done with "one ticket" like ECR with approval workflow 2. In EDM when component is revised – a pop up window appear which need a description for revision (this will be saved to change history) - ECR/ECO object should not be an extra object – the more there are different objects the possibility increases not to use them. Too heavy structure. - Clear workflow should be done (in EDM at three stages)

Research Interview 4
Master Thesis – Data 1

TOPIC: Current State Analysis of the Case Company

Information about the informant (Interview 4) *Table 1*

Details	
Interviewers	Antti Korpiniemi
Name (code) of the informant	Informant 4
Position in the case company	Head of Automation and Instrumentation
Date of the interview	8th March 2016
Duration of the interview	1h 15min
Document	Field notes + Tape recording

Field notes (Interview 3) *Table 2*

	Track time	QUESTIONS FIELD NOTES
1		How itemization is connected to your work?
		- Our department does plant automation and instrumentation delivery and part of Engineering - Delivery hubs (in Finland and Germany). Is consisted of three departments: 1. Engineering 2. Project implementation 3. Purchasing - When Outotec delivers plant 1. Supports sales with their expertise 2. Offers solutions to Outotec customers - Creates quotations of instrumentation and automation 1. Scope defining 2. Prizing 3. Cost accounting - Own team about 15 employees. - Three phases for customer delivery: 1. Study; creation for customers that includes a. Disciplines 2. Planning 3. Delivery - **Instrumentation does not use itemization at the moment.**

		- **Automation has some itemization done for some plant deliveries** - **<u>Equipments</u> are managed by Outotec's own units (they use more components)**
2		How are items used in your department's daily processes?
		- Itemization is not used **(plant deliveries)** - **Itemization is seen as too complicated and providing extra work** - With instrumentation parts 1. It is seen that there are too many possibilities and products vary a lot and change in a way that they cannot be used in another project. I.e. plant deliveries are strongly project orientated 2. **Purchases are done with category items. One "row" is for the whole purchase (all needed parts and components are listed under it). Attached specification documents will indicate the technical details of the products.**
3		What software's are used for daily processes? (E.g. JIRA, ENOVIA, EDM, Vault, EBX, SAP, Maximo)? - How they are used?
		- Plant documentation management is done via DMS Meridian - Enovia not in daily use (because they are not itemizing) - SAP is used for purchase proposals – actual purchase order is done by purchase team for specific project - Plant itself does not have item number - Designers can export needed specifications to excel sheet when e.g. for a specific purchase - AVEVA for basic designs - SPI – for instrumentation detail - ePLAN – for cabinet designs
4		How Component Engineering supports your processes?
		- **Not at the moment** - **There is a decision that instrumentation won't be executed at this point** - In the future an upper level item is created with datasheet
5		What information you especially use from item information? (e.g. MEP code)
		- I&A has codes in their product but the **<u>code does not contain all needed information to define the product</u>**. - Besides the code there are many details to be defined e.g.: 1. Calibration / measurement area (for instruments) 2. Oxygen washed part (for oxygen use) 3. ATEX certification (for electrical equipment in explosive atmosphere) 4. A tag (which is used for part recognition in the plant and contains project specific information)

6		What are the strengths of using items? - Why we itemize?
		- Needed for Service business - **At the moment too complicated to itemizes instruments**
7		How should the item information be described to be clear? (e.g. should technical description contain specific information) – is the currently used information clear?
		- **Purchase process:** **1. An excel will be made which defines the need for equipment** **2. Supplier makes offer of the required components (there can be more than one suppliers)** **3. Offer is compared and validate** 4. Offer can be used later for purchasing (contains valid information)
8		How would you describe the strengths/weaknesses of Key User concept?
		- AVEVA for basic designs 1. In pilot state – needs massive improvements - SPI – for instrumentation detail (one keyuser) 1. Works well in its current state - ePLAN – for cabinet designs - **KeyUser concept is seen as good thing**
9		How concept guidelines / used tools are informed/presented to employees?
		- Own processes and tools could be improved more - Change management should be improved (AVEVA could be solution) - During OPAL – many things were deployed in unfinished format o E.g. UAT
10		What could be biggest “bottle necks” of itemization concerning your department?
		- Too slow way of working. E.g. one instrument item takes about 30 minutes.
11		How itemization supports the way of working in your department? (Difference between past and current) - What could be done differently? - What could be improved in the future?
	58.50-75.20	- Productization could be improved more in the future (as one goal in OPAL process) 1. Some efficiency has been noticed because of productization - Products should be tried to sell more (different solutions for customers) - **Should be considered is it necessary to give “all choices” for customer or could there be some**

		limitation or product where to choose. - Siemens – is what mostly is used and a contract partner (limits a little bit of the range of selection) - **Price could be lower** - Regarding OPAL – a feeling in the PDM system selection is that technology and product where forgotten and system (Enovia) was chosen based to system itself not for what actually is produced. 1. People first heard of the PDM in the UAT – not the best way to start presentation of the new system - Most of the money flow comes from projects and service - PDM system seems to missing the point of actual practical usage i.e. it is not serving product lines well enough in the current way of working.

Research Interview 5
Master Thesis – Data 1

TOPIC: Current State Analysis of the Case Company

Information about the informant (Interview 5) *Table 1*

Details	
Interviewers	Antti Korpiniemi
Name (code) of the informant	Interviewee 5
Position in the case company	Director, Supply & Distribution Network, Markets Unit/Spare & Wear parts
Date of the interview	22nd Mar 2016
Duration of the interview	1h 20min
Document	Field notes + Tape recording

Field notes (Interview) *Table 2*

	Track time	QUESTIONS FIELD NOTES
1		How itemization is connected to your work?
		- Team 4+4 people - Supply chain specialist, distribution network manager, transportation manager and people from sourcing - Items are used daily - Value stream mapping -> quotation process is long. Finding item takes time. 1. Where the time is used 2. Can we reduce waste (Lean thinking)
2		How are items used in your department's daily processes?
		- Quotation and selling items - All transactions are done with items (quote, deliver, invoice items) - Item usage is mandatory
3	**-14.45**	What software's are used for daily processes? (E.g. JIRA, ENOVIA, EDM, Vault, EBX, SAP, Maximo)? - How they are used?
		- SAP - Maximo for workflow → when customer send email to local service center Maximo automatically creates a workflow (includes several tasks) - Basic workflow case:

		1. Customer needs something and sends a picture 2. Local service center look at the picture (they know or not know what it is) 3. If don't, they will past ticket to global unit to quotation specialists (technology specialist (Outotec has 42 technologies)) 4. Quotation specialist will try to recognize the product from drawings (have access to many systems, network drives, legacy systems) - JIRA is used for procurement RFQs - Aton (filters) - eCatalogue (important – only way to see structure) 1. **most of them are not in Enovia – uploaded using excel** 2. all installed base structures can be found from eCatalogue (but not so "structured" format) - access to legacy system (Pronto and IFS are still used) - Sweden is still using Pyramid at some stage (in Sweden they haven't reuse items back in the days)
4	**14.50-**	How Component Engineering supports your processes?
		- **Not visibility how CE currently supports spare part** - CE does not add visible value for Spare parts unit - Library full of stuff that Outotec won't sell - Biggest advantage would be if CE would not just add parts to library but actually **do more standardization and even "limit" amount of components which are used by designers.** 1. Only few manufacturers should be preferred. (Bosch Rexroth for hydraulics and FAG in roller bearings. 2. Filtering the number of choices -> too many choices at the moment 3. Would help to standardize products (productization) 4. CE could be solution maker limiting scale of used components ▪ Design for cost ▪ Design for service - Outotec has three **global distribution centers** (Finland, Nederland, USA) 1. These stock are full component from product line who has done some kind of productizing (or standardization) – Filter, automation 2. Parts which have lots of inventory turns (can be stocked – and some kind of forecast can be made for items) - Local stocks - Customers stock (only parts stored for that customer – ideally critical spares) - E.g. grinding mills are all different and parts are stocked all over the world. - *When Outotec delivers a mill to customer – after that Metso is offering their solutions to standardize the mill and taking care of the service business.* - **CE could be supporting spare part service by guiding/affecting product selection/range which product lines are using**, if possible. (e.g. certain electric motors etc.) 1. Filter, Flotation, automation -> are ongoing with standardization 2. Flotation design has been done in three difference places (because of the acquisitions) -> **a problem with installed base**

5	**-38.20**	What information you especially use from item information? (e.g. MEP code) - How information can be verified? (How to know that correct product has been purchased with manuf. Code which defined to item.)
		- Identifier either **drawing number or manufacturer number** 1. With help of the drawing it can be connected to specific project and recognized from the documentation relating to it 2. eCatalogue 3. All documents have been digitalized and can be founded from network drive - Service tries to **search from context** (where the actual item has been used) 1. **At the moment discovered context is <u>not connected</u> to delivered item** 2. **Lack of time – biggest challenge to make structures** - If searching from Enovia library – the correct component cannot be recognized - Item information can **be verified by looking history** what and when has been sold last time the same customer. (e.g. some customers order same motor every year) - For item verification in very detailed level should be considered very carefully (where to draw the line) 1. Basically -> **if correct datasheet is found and it matches with existing description a decision can be made that it is valid item for use.** 2. Order confirmations are not always available 3. In case of unknown item (lack of information) – it should not be used. -> new item
6	**38.25-49.25**	What are the strengths of using items? - Why we itemize?
		- **Not possible to change suppliers because of lack of information** from Swedish department (Swedish supplier has the information of the products that Outotec uses – meaning that supplier has to exact knowledge of the product that Outotec is needs for its projects.) 1. Outotec does not have access to those drawings anymore 2. Suppliers know more these days - A real case: Customer has demanded Outotec item codes for his system for all spare parts and if this will happen – all the spares will be bought from Outotec 1. If customer has **easy access to correct spare part code** – they will definitely buy it from Outotec 2. If spare parts are not listed – customer starts to search "correct" model and could discover original manufacturer and purchase the part straight from them. - Capex projects are mostly delivered without item codes (seen as too complicated) - Enovia to eCatalogue does not work – an excel has to be used to upload information to eCatalogue 1. Customers will have DVD 2. **eCatalogue should be online and updated to all the time – even for customers (if item code has been changed)** - At the moment big problem is that maintain people don't know where to assembly the spare parts… because information and documentation is missing. - It needs to be considered what should be itemized. What is the level that we need (e.g. assemblies with many sub-assemblies) 1. All technologies should be noticed separately – what would be the best level
7	**49.28-**	How should the item information be described to be clear? (e.g. should technical description contain specific information) - Where information is stored (e.g. special lists not in Enovia? Etc.)?

		- In the past description was more free form – context was added to description (e.g. for this project/part/component) -> things were found… - SAP desc. 1 is searchable but desc.2 is not (?) - Enovia search does not work well - Special lists still lists – customer has their own lists (customer agreement)/price lists/offline lists 1. Kalle upkeeps price list with he's team 2. Customer specific price list is not in SAP ▪ Difficult to handle SAP at the moment ▪ Offline lists - PO (purchase orders start - **Order confirmation** 1. Comes from B2B solutions 2. "Electronic receipt" 3. **Access needed to RFQ's? Happens in <u>JIRA – RFQs</u> are managed there.** 4. **<u>Quote document will have technical details of the part</u>** ▪ Item number has created before this ▪ Special occasion RFQ is done first and the itemized when information has been received from supplier ▪ **After RFQ case item can be created quite correctly** - Service items are most of the cases – **make to order "Vendor items"** 1. These should **not be mixed with commercial parts** (e.g. hydraulic cylinders or FESTO actuators which could be assembled in 30 different ways) 2. **Solutions specific (not really commercial) – a design part but not Outotec design part** 3. Designers give parameters to supplier and supplier will design and deliver needed cylinder but all keeps all the details (Outotec has only layout picture). Supplier owns the design for that cylinder (**vendor item**) 4. <u>Make to order</u> and designed especially **for** Outotec by supplier. ▪ <u>Basically it cannot be sold to anybody else</u>
8		How would you describe the strengths/weaknesses of Key User concept?
		- How that works in practice is not visible - Challenging situation
9		How concept guidelines / used tools are informed/presented to employees?
		- A three year journey where guidelines are made are instructions for every step - Instruction how to find item without Enovia etc. - Problem situations are flagged to process and concept owners - "operational handbook" quite useless – impossible to find anything o Mess of lotus notes - Item guideline for service – by Jeremy Spencer (modified to role specific needs for service)
10		What could be biggest "bottle necks" of itemization concerning your department?
		- **<u>Biggest issue is the find the verified data for items</u>** 1. **It is much easier to create new one** 2. **Itemizing on the quote takes time → lead time**

		- When prizing items we use 1. proprietary items which are Outotec's own design 2. Commercial + standards which are of the shelf – OEM is clearly somebody else's and can be identified 3. **Solution specific components (vendor items)** –e.g. hydraulic cylinders. These should be separate somehow, because they ▪ **Are mixed in different places** ▪ **Different prizing logic** ▪ **Different supply chain and strategy** ▪ **Different suppliers and setups are needed** ▪ **Different distribution** - **Legacy way of enhancement seems to be inefficient way to do item** 1. Are correct items enriched to library 2. This would need customer point of view – not item or sales order ▪ For what customer should be focused on ▪ Which legacy data first – prioritize ▪ Installed base mapping project ongoing – few customers are already picked up where to begin
11		How itemization supports the way of working in your department? (Difference between past and current) - What could be done differently? - What could be improved in the future?
		- In the past catalog was not created for customer along the deliver 1. Only way to know what has been delivered is the original project documentation - After spare part engineers have identified items and its context 1. Feels waste of time that they fill a form to create an item which they could do them selves 2. Should be more specific what the library is for and what should be added to library

Research Discussion 1
Master Thesis – Data 2

TOPIC: Discussion and Suggestions for Initial Proposal

Information about the informant (Discussion 1) *Table 1*

Details	
Interviewers	Antti Korpiniemi
Name (code) of the informant	Interviewee 1, 6 & 7
Position in the case company	Item Data Manager, Data Architect, external consultant, Data Analysis, external consultant
Date of the interview	22nd & 26th Apr 2016
Duration of the interview	60min + 60min
Document	Field notes + Tape recording

Field notes *Table 2*

	Track time	QUESTIONS FIELD NOTES
1		Feedback / Questions
		<u>Item data manager</u> - Not sure what the definition on service and product means in this case? Tangible and intangible? - Approach methods themselves can be understood in common level as presented and seen as valid proposal. <u>External consultants</u> - The approach for the approach is quite difficult to understand. What does it exactly do and how it works? - The presented approach could be a very simple model, which would be easier to perceive.

Research Discussion 2
Master Thesis – Data 3

TOPIC: Case "Improving Data Management for Better Itemization Process"

Information about the informant (Discussion 2) *Table 1*

Details	
Interviewers	Antti Korpiniemi
Name (code) of the informant	Interviewee 1, 6 & 7
Position in the case company	Item Data Manager, Data Architect, external consultant, Data Analysis, external consultant
Date of the interview	26th & 27th Apr 2016
Duration of the interview	60min + 60min + 30min
Document	Field notes + Tape recording

Field notes *Table 2*

	Track time	QUESTIONS FIELD NOTES
1		Feedback / Questions
		<u>Item data manager</u> - New approach is clearer, and can be seen as valid approach, but still would be need sharpening to practical example. <u>External consultants</u> - The new approach give clearer picture how to utilize the proposed approach.