Bachelor thesis

Improving a data conversion process

Determining a future-proof solution direction by analyzing the data conversion process

Martijn van Kessel

connecting customers

ANVA

UNIVERSITY OF TWENTE.

Bachelor thesis

Title: Improving a data conversion process
Study: Industrial Engineering and Management
Faculty: Behavioral, Management and Social Sciences

Author

Martijn van Kessel
m.vankessel@student.utwente.nl
s2181320

ANVA
Stadsring 201
3817 BA Amersfoort

University of Twente
Drienerlolaan 5
7522 NB Enschede

Supervisor ANVA
Arne van Kalsbeek

Supervisors University of Twente
1st Lucas Meertens

2nd Erwin Folmer

Date
July 28th, 2021

Preface

This report presents my graduation research for the bachelor study Industrial Engineering and Management at the University of Twente. The research was conducted to help ANVA improving their data conversion process. The internship at ANVA was a pleasant experience, despite the restrictions in effect through the pandemic. All colleagues at ANVA were a pleasure to work with, but there are some people I would like to thank particularly.

First, I would like to thank Patrick Keijer for giving me the opportunity to do my internship at ANVA. Next, I would like to express my gratitude towards William Quaadvliet, who constructed the research and put me in touch with the right people for the research. Furthermore, my appreciation goes out to Arne van Kalsbeek for guiding me on a daily basis and thinking along with me.

Lastly, I thank my supervisor from the University of Twente, Lucas Meertens, for giving me constructive feedback during the research. The practical tips and interesting insights were very helpful in shaping the research and report.

I hope you have an interesting read!

Martijn van Kessel,

Lunteren, August 2021

Management summary

ANVA delivers software solutions for clients in the insurance branch. Since the insurance branch is very dynamic, many companies merge or acquire other companies. For most of these events, a data conversion is needed, what implies that two databases are merged. Since increasingly more conversions are needed at the moment and clients demand conversions of high quality, ANVA has great interest in a well-performing data conversion process (DCP). Currently the DCP is facing problems which hold back the performance of the process. This causes unsatisfied clients, which becomes more problematic over time.

The goal of this research is to map the process, find the most important problems and align suitable solution directions. ANVA would like an overview of the data conversion process, what problems it has and what options they have to solve them. To reach this goal, the following research question has been asked: What are the main problems in the DCP and how could these problems be solved?

To answer the research question, the DCP was modelled to understand the DCP well and derive problems directly from the diagrams. Next, interviews and conversions were held with employees and clients who have experience with the DCP. Employees at ANVA were asked to give their opinion on the process, name strong and weak points of the process and suggest improvements. Clients of ANVA were asked to give their opinion on the process, name strong and weak points, point out missing factors or functions, suggest improvements and describe the experience with ANVA.

The analysis and interviews revealed several problems in different fields of the DCP. Most problems were found in the following four fields: problems in the Extract, Transform, Load (ETL) tool, an incoherent system structure, insufficient communication with clients and non-optimal policy. The ETL tool misses functionalities and has weak points in its usability and output. The incoherence in conversion tool and system structure makes the DCP complicated. Clients experience an insufficient provision of information before and during the process. On process-level, unawareness of problems makes the DCP unclear and unsustainable, and reduces efficiency.

Based on the problems found, a list of recommendations is made. Generally, ANVA should reconsider the structure of the DCP, implying to only use the ETL tool, simplify the digital landscape, communicate more with clients, improve the way of working and improve sustainability. ANVA is recommended to only use the ETL tool and not do conversions manually or with Cobol software. This makes the conversion process less complicated and ANVA will only need to invest in just one conversion method. Next, ANVA should look into replacing or improving the current ETL tool to improve continuity and have more functionality. Since the ETL tool becomes the only tool used, it must have all desired functionalities. Therefore, missing functionalities should be added, even if that is expensive. Furthermore, ANVA should provide more information to clients about the possibilities for their conversion and the schedule for the process. By giving more support before and during the conversion, clients will be more satisfied with the process.

At this point, the recommendations made are just solution directions and not highly developed solutions. It is therefore suggested that research is done in further developing the solution directions and finding out what solutions are best for ANVA. Moreover, research can be done on the problematic transfer of data from the ETL tool to ANVA.

Table of contents

List of tables

List of figures

Terms and abbreviations

Term	Explanation
Agent	An agent mediates in the closing of insurance contracts and do things as collecting premiums and handling claims for insurers. An agent does not have a proxy.
ANVA4/5	Old ANVA data environment, to be switched from
ANVA Hub	New ANVA data environment, to be switched to
API	Application Programming Interface, an interface that defines interactions between multiple software applications
DCP	Data Conversion Process
DFD	Data Flow Diagram
ETL	Extract, Transform, Load ; general method to convert data
KPI	Key Performance Indicator
Proxy intermediary	A proxy intermediary has the authority to act directly on behalf of insurers or banks and does everything that an insurer also does, except bearing financial risks
TOC	Theory Of Constraints, method that improves systems by focusing on removing constraints

1 Context

To understand the problem well, it is important to have a complete view of the problem's context. To create that complete view, this chapter introduces the company ANVA in section 1.1, introduces the topic data conversions in section 1.2 and explains the problem statement in section 1.3.

1.1 ANVA

ANVA is a company located in Amersfoort and Bergen op Zoom which operates in the insurance branch. ANVA creates software solutions to efficiently support insurance processes. ANVA has more than 300 clients, which are larger proxy intermediaries, service providers or smaller agents in the insurance branch. For all these clients, ANVA delivers whatever they need in the field of software and digital support. ANVA creates solutions that help with supporting efficient insurance processes, reaching optimal business management and maintaining great customer relationships. With combining collaboration and innovation, ANVA wants to create added value for clients. (ANVA, 2021)

Clients can choose to completely outsource their IT system to ANVA. ANVA has a ready-to-go IT infrastructure to work with, but it is also possible get it adapted to the wishes and needs of the client. Consultants work together with clients to form the optimal digital structure for their company. Consultants help clients with choosing the right package for the client, after which he ensures that the client is optimally profiting from the software. Next to this, consultants maintain the complete infrastructure of the client. (ANVA, 2019a)

Generally spoken, ANVA delivers a great variety of services to companies in the insurance branch, but the key service is making software that helps companies run processes smoothly. Next to delivering the software package, ANVA is very involved in keeping the system up to date during time. This means that any changes in data must also be processed. When this happens, a conversion is needed. Section 2.2 will elaborate further on the concept conversions. (ANVA, 2019a)

1.2 Conversions

The insurance market has been very dynamic over the last couple of years and will be continuing to be so in the coming years. Smaller companies merge, are being taken over by larger intermediaries or join a service provider (Wortell, 2020). If one of the companies involved in the merger is not a customer of ANVA, a conversion is needed. A conversion implies that data from two origins with a different structure must be merged. If one company acquires another company, or they merge, the two databases must be transformed into one new database that includes all data from both origins. The difficulty of such a conversion depends on the size of the companies and thus the size of their systems, and the type of system they are using. To make these differences understandable, we distinguish three types of conversions: (ANVA, 2019b)

1. Two clients of ANVA merge together.

   This type of conversion is the easiest, because both parties already use the software of ANVA. This means that the two systems are essentially the same and are filled with the same tools. However, the systems still differ a fair bit, because of the customizable options. Company A can have a system that differs much from the system of Company B, for example because Company A uses the system for a totally different goal. Despite this, the bases of the systems are equal, which means that this type of conversion is not a big challenge.

2. A smaller company that runs on software of a competitor merges with a client of ANVA.

This type of conversion can cause problems if the software of the competitor is totally different than the software of ANVA. The difficulty of the conversion depends on the size of the companies and the regarding software. If the companies who merge are large and the used software is unknown to the software engineers of ANVA, the conversions is a laborious process that faces a lot of technical challenges. If the companies who merge are small and the used software is somewhat known to the software engineers of ANVA, the conversion is not that hard and should go well. This is mainly because ANVA already has a lot of experience in the field of conversions. Generally, this second type of conversions can be quite problematic, but that is not necessarily the case.

3. A large proxy intermediary that runs on software of a competitor merges with a client of ANVA.

   This type conversions causes the most problems for ANVA. Conversions in this category have the drawbacks of the second type of conversions, but even worse. The companies involved are very big and have giant databases. The conversion faces a lot of technical challenges, which cause a problematic process. Due to the technical difficulty, often a case-specific conversion tool is built. Because of this, the process is very labor intensive, for both ANVA and the client.

It is clear that a conversion is not done easy, when having to deal with the second or third category of conversions. Clients often underestimate the time and labor needed for a conversion. An account manager at ANVA states that this underestimation is unfortunate, because companies can benefit more from a conversion than they think. A conversion is the ideal moment to look closely at the system, in order to improve it. For example, it is very helpful to clean up data before the conversion, so that no contamination occurs after the conversion. With a clean set of data, the employees of clients can easier find the right data in the system. By doing a conversion, you are forced to think about the optimal set-up of the new IT landscape, which allows you to analyze how to do it better. This also fits well with the theory of Antonov et al. and Jain et al.. Although this all sounds great, ANVA has its problems with conversions. These problems are worked out further in section 2.3. (ANVA, 2018)

1.3 Problem statement

There are multiple reasons for the trend of conversions becoming harder for ANVA. As said in section 1.2, conversions of the first category are not that difficult, but conversions of the second and third category cause problems. In table 1, an overview of the types of conversions and the focus of this research is given.

Type of conversion	Level of difficulty	Focus this research
ANVA - ANVA	**I**	
ANVA - other, small	**II**	**X**
ANVA - other, large	**III**	**X**

Table 1 - Overview types of conversions

Now, the most important causes of the difficult process will be discussed.

First, the demand for conversions is increasing rapidly. As a result of that, the number of conversions needed for mergers is so high that dealing with conversions has become part of the daily work activities of ANVA. This might sound as good news, because ANVA has a lot of demand for their services, but it is becoming a problem when combining this raise in demand with other problems.

One of the problems is the size of the conversions. In today's insurance market, companies are becoming bigger (Wortell, 2020). Small companies are being taken over by big companies, making big companies even bigger. With that growth also more data gets involved. This especially the case with category three conversions. The companies involved are becoming bigger, which means they are insuring more consumers, which means there is more data to convert. The IT systems of companies must comply with that data. The amount of data that must be converted is so high, that it makes the process hard to do. A big conversion simply takes more time than a small conversion. So, the process is not only harder, but is also more laborious. This costs ANVA time and money.

The next nuisance factor is the IT application that is used by clients of ANVA. The IT application is called ANVA4/5. ANVA4/5 is already running for a long time, which means the application is not complying with all of today's standards. Next to this, ANVA4/5 does only have a minor standard part, which is the same for every client. A major part of the application can be designed and adjusted to the wishes and needs from the client. This means that every application can have a very different layout, which can be inconvenient for conversions. Luckily, the general set up of the application is mostly the same and the used language is also the same.

However, companies that were not clients of ANVA did not use the software of ANVA. Their systems do not have the same general set up or language as the systems of ANVA. The systems of these 'new clients' can of course be quite known to ANVA, but it can also be completely new. For a known system, the conversion can still be hard when the particular database has a slightly different layout then another one. For a completely new system, the system must be completely figured out by ANVA in order to do the conversion, which is very time-consuming. So, having to deal with the software of competitors causes the conversion process to be laborious and time-consuming, as well as to face technological challenges.

To do the conversion process more efficiently, ANVA bought an ETL tool that is used for conversions. The intention of the ETL tool was that clients could do the conversion via an interface. However, the interface was not good enough for the bigger and more complex conversions. The interface lacked stability and speed and was functionally too limited. The interface was intended as the right answer to the problem of the hard conversion process, but could not live up to its expectations so far. So, the core of the ETL tool was working, but helpful attributes such as performance and usability were lacking or missing. Chen, Ali Babar & Nuseibeh call these attributes Non-functional requirements. Non-functional requirements are requirements that can be used to score a tool or system as a whole. However, they are not necessary for the tool or process to work (Chen, Ali Babar & Nuseibeh, 2013). In this case, the non-functional requirements stability, speed and functionalities are lacking and must be improved.

The last problem is the need of reaching new target applications when converting. Now, all conversions are done to ANVA4/5. As said before, ANVA4/5 is an outdated system, which is not going to be used in the future. ANVA is currently switching to a new system, which should be reached in the future. From now on, conversions must be done to ANVA Hub, which is the main application of the new system. The current ETL tool can only convert to ANVA4/5 and is not able to convert to ANVA Hub. Therefore, the reach of new target application causes technical difficulties in the process.

The problems and their consequences are summarized in the problem cluster, which can be seen in figure 1.

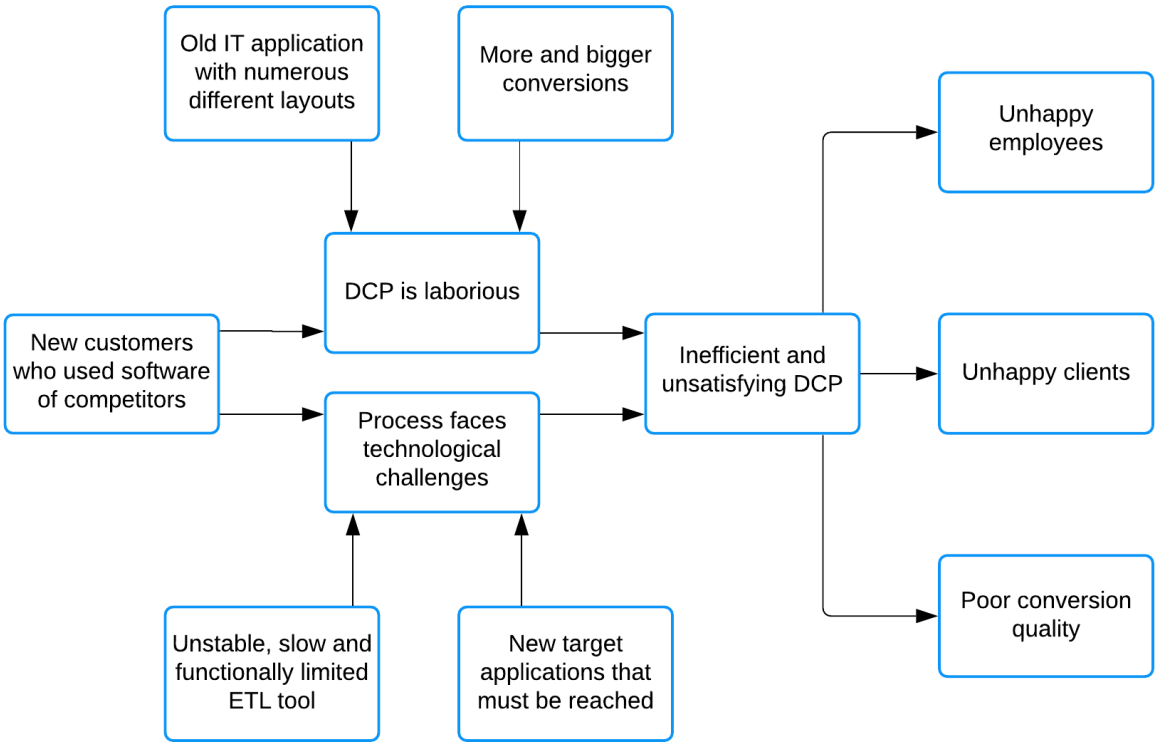

Figure 1 - Problem cluster

According to Heerkens & Van Winden, there are some rules of thumb which can be used to find the core problem or core problems of the problem cluster. First, the core problem must be surely occurring and must have a relation with other problems in the problem cluster. Next to this, the core problem does not have a direct cause in the cluster. It should not directly be the cause of another problem. Furthermore, the core problem should be influenceable. If the problem is not influenceable, there is nothing to do about and it cannot be a core problem. (Heerkens & Van Winden, 2021)

In the problem cluster of the DCP, there are five problems which together cause an inefficient and unsatisfying DCP, which in turn causes unhappy users and poor results. Of these five problems, more and bigger conversions, new customers who used software of competitors and new target applications that must be reached are not influenceable. These problems are given developments which happen to be negative and nothing can be done about them. Therefore, two influenceable problems remain, which are the old IT applications with numerous different layouts and the unstable, slow and functionally limited ETL tool. These problems are both a given situation, but can both be influenced by changing or replacing the application and tool. Since no distinction can be made in importance or impact, these two problems will be the core problems.

The core problems can be made measurable by the variable time it takes to do the conversion. Next to this, total costs of the conversion can be seen as an indicator of how well the process is done. These are the only variables that can make the success or failure of the process measurable. However, the variables are hardly measurable in real life, since it cannot be overseen how much time is actually spent on the process. Therefore, the process will be judged based on experience. Right now, the reality is that employees and perceive experience the process as unsatisfying and inefficient. The norm is that employees and clients perceive the process as satisfying and efficient. The goal of the research is to let employees and clients perceive the process as satisfying and efficient.

2 Research design

This chapter explains how the research is conducted. In section 2.1, the research goal will be described. In section 2.2, the research questions will be listed. When answering all research questions, the research goal should be accomplished. Then, in section 2.3, the research methodology will be explained. It describes the general methodology but also more specifically what methods are used to answer the research questions. Next to that, there is a paragraph about how reliability and validity is incorporated into the research.

2.1 Research goal

The research goal should be in line with the core problem; when accomplishing the research goal, the core problem should be solved. As said in section 1.3, the core problems are the old IT application and the limited ETL tool. It is not clear how these problems came to be. Therefore, the research goal is as follows:

By analyzing and modelling the current data conversion process and mapping wishes and needs of ANVA and clients, find out what the main problems are, why they are there, where are they coming from, and think of a future proof solution that solves the main problems.

In line with this, the deliverable are a list of the main problems and a list of future-proof solution directions. ANVA will receive an overview of the most important problems and how they could be solved. Per problem, one or multiple solution directions will be discussed. They could align these solutions directions with their own interests and find the best solution for them.

2.2 Research questions

Based on the core problems and the research goal, the main research question is formed, which is as follows:

What are main problems in the conversion process of ANVA and how could these problems be solved?

Directly answering this question is not doable. Therefore, a couple of sub questions are needed. The sub questions all answer a part of the main research question. All answers to the sub questions together should form the answer to the main question. The research is split into three topics: the conversion process itself, the experiences and needs of ANVA and clients, and the future proof solutions. For each topic, there will be a sub question. The sub questions are as follows:

1. *How does the current DCP at ANVA look like?*
    a. *How can the current DCP be mapped?*
    b. *What are weak points of the DCP?*

2. *What are the experiences of ANVA employees and clients?*
    a. *How do employees of ANVA experience the conversion process?*
    b. *How do clients experience the conversion process?*

3. *Based on the problems found, how can these problems be solved?*
    a. *Which requirements does the solution have to meet?*
    b. *Based on the found problems and the requirements, how can the problems be solved?*
    c. *What needs to happen to bring the solution into practice?*

2.3 Research methodology

2.3.1 General methodology

To find a solution to the found core problems and answer the research questions, we use the Managerial Problem-Solving Method (MPSM) from Heerkens & Van Winden (2021). This is a method designed to solve action problems. The MPSM consist of the following steps:

1. Problem identification
2. Planning of the problem-solving approach
3. Problem analysis (causes of the problem etc.)
4. Designing and assessing possible solutions
5. Choosing a solution
6. Implementing the solution
7. Evaluating the solution

As the MPSM is a cycle to constantly solve problems, it can be visualized as follows.

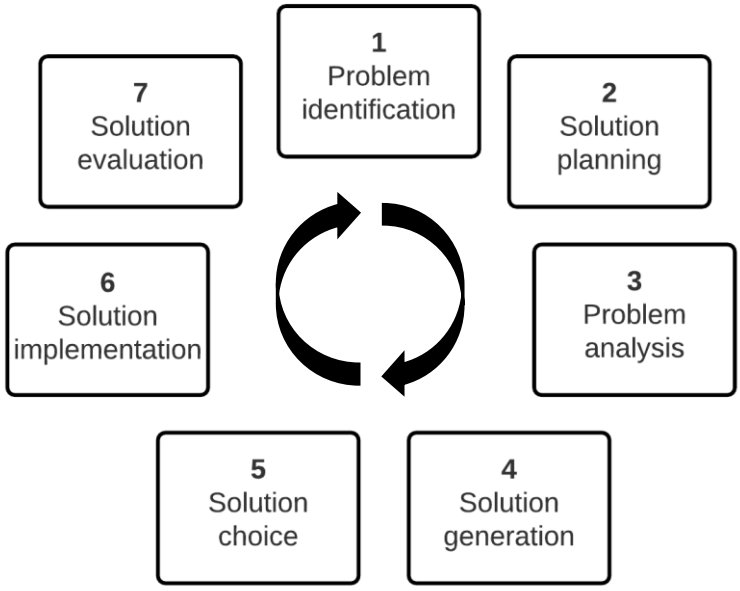

Figure 2 - MPSM cycle

Since we only have to align solution directions and not choose the best solutions, we only execute steps 1 to 4. The steps 5 to 7 are left for ANVA to do themselves. In chapter 1, step 1 was done. The context is researched and the core problems are clear. In this chapter, step 2 is done. The problem-solving approach is formulated. Step 3 is done in chapter 3 and 4, where the DCP is analyzed and problems are identified. In section 2.3.2 is described in detail how step 3 and 4 will be executed. Finally, in chapter 5 step 4 is done. This chapter is concerned with finding solutions to the problems and discussing the trade-off. This is done by using the following seven-step plan developed by Heerkens & Van Winden. The steps are:

1. Defining the decision
2. Defining the decision-making process
3. Establishing criteria
4. Scaling criteria
5. Weighting criteria
6. Coming up with optional solutions or using existing possibilities
7. Evaluating options

As this seven-step plan is a cycle within step 4 and 5 of the MPSM, it can be visualized as follows.

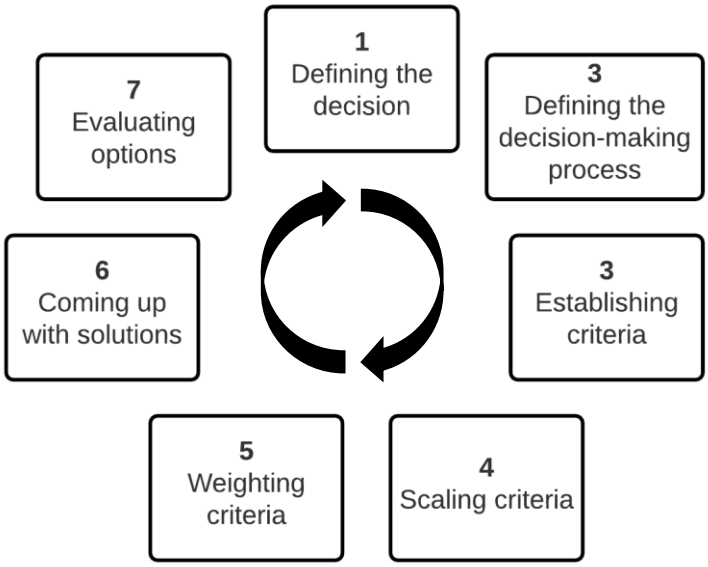

Figure 3 - Solution cycle

This seven-step plan can be used to come up with solutions, given the problems found in earlier sub questions. Again, we do not execute the complete cycle. We only execute step 1, 2 and 6 and for some problems also step 3. In this research often solutions are found without the use of criteria. It is up to ANVA to establish, scale or weight the criteria and evaluate the options based on their interests.

2.3.2 Research methods

In this section will be explained how step 3 and 4 of the MPSM will be executed. For each sub question, the answering method will be explained.

1a. *How can the current DCP be mapped?*

The information gathering requires interviews or observation, since no documented information of the process is available. Therefore, information must come directly from people who know how the DCP works. Because the process takes longer than the available time for this research, no observation method is possible. That is why interviews will be conducted. The people who know how the process works are employees of ANVA who work with data conversions. So, open interviews with employees of ANVA will be done. The interviews are open, because it is not clear what the process looks like beforehand, so it is unclear what to ask beforehand. With the knowledge gained in the interviews, it will be possible to model the DCP clearly and precise.

1b. What are weak points of the DCP?

To find out what are the causes of the core problems, the weak points of the DCP must be identified. To find out the weak points of the DCP, the process needs to be analyzed. This analysis contains several problem identification methods. These methods are the trade-off table, the TOC and identified problems by interviewees. After using these methods, there should be a complete broad overview of what the weak points of the DCP are.

2a. How do employees of ANVA experience the conversion process?

To find out employees of ANVA perceive the conversion process, interviews will be conducted. The interviews will mainly be open, so that the employees can freely speak about his perception of the process. In the interviews, the employees will be asked to form their opinion on the process, to see if they can add something to the analysis. Next, the employees will be asked how they think the process can be improved. After doing a couple of interviews, a clear overview of the experiences and perceptions can be made.

2b. How do clients experience the conversion process?

Just like the experiences of ANVA employees, the experiences of clients will also be found by means of interview. There will be open interviews conducted, where the client can freely speak about his experience with the DCP. The same three main topics will be discussed. The client will also be asked to express thoughts on how he/she thinks the process can be improved. After a couple of interviews, there is a clear view of how clients experience the conversion process.

3a. Which requirements does the solution have to meet?

Some of the requirements are already in the assignment description given by ANVA, which will be the basis of this section. With further personal communication with an expert on the topic at ANVA, the list can be expanded or worked out more extensively. This can be done until the expert approves the list of requirements.

3b. Based on the found problems and the requirements, how can the problems be solved?

Solutions are found by using several methods. By analyzing the DCP and the list of problems, brainstorming in combination with heuristics are used to find and solve problems. According to Michalewicz & Fogel, heuristics like common sense and reversing the problem can be useful to easily solve a problem. Whenever a problem is a wrong action, maybe just do the opposite. Often this can be helpful in solving problems. It only remains to ask how the opposite could be executed (Michalewicz & Fogel, 2004).

Also outcomes of the interviews will be very helpful in finding solutions to the problems. Interviewees will often directly suggest a solution, after they state the problem experienced. These solutions must be considered carefully, since a solution directly coming from a user of the DCP can be very personal or irrelevant. The solution must be checked on neutrality, to prevent the satisfaction of just personal interests. However, all suggestions can be helpful, as they are already invented. The combination of these methods will form the basis of the solution generation.

3c. What needs to happen to bring the solution into practice?

The answer to this question is completely dependent to the chosen solution. Based on the solution, one or multiple ways to bring the solution into practice will be described. It is up to ANVA to choose what method complies best with their interests.

2.3.3 Reliability and validity

According to Heerkens & Van Winden, reliability is about the stability of the results. A similar research at a later stage must yield the same results. For this research, it is hard to ensure that, since most results come from interviews. However, it can assumed that interviewees will give the same answers over time, if the situation remains the same. This is because the interviews are in their own interests, as they are the users who have bad experiences. The results can thus be considered as reliable.

Heerkens & Van Winden state that internal validity is about the extent to which a study has a trustworthy relation between the cause and the effect. It means that the research design is such that alternative explanations to the problems can be ruled out. If employees and clients were given a survey, the results would all over the place, since people all interpret the questions in their way. That is why most sub questions are answered with interviews. With open interviews, the interviewer can steer the answers of the interviewees to what the interviews wants to hear. This can for example be done with follow-up questions. Thus, the interviewer can almost always ensure that he gets the right answers.

Therefore, only one interview with a consultant or client is enough to get an idea of how the process works and what might be weak points of the process. Though, to get a more valid answer, multiple consultants and clients will be asked to express their experiences with the process. As different people with different perspectives on the process can also perceive the problems differently, the use of multiple interviewees helps to form the most neutral look on the process. It reduces outliers and meaningless answers caused by misunderstood questions.

External validity is concerned with the extent to which the results of a study can be applied to a broader context. As this case is quite specific, the results can only be applied to companies with a similar situation as ANVA. This concerns companies with an underperforming DCP. All companies that must transfer data from one database to the other, and in doing so also have to transform the data, can have the same problem. Especially if an ETL tool is used for this purpose, companies can relate. Therefore, the results can be used by all companies that have an underperforming DCP and use an ETL tool with missing functionalities.

Construct validity is about the appropriateness of conclusions based on the observations made. It is about the logical relation between the problem statement and the results obtained. It is important that the problem statement is formulated well and that concepts are made sufficiently specific (Heerkens & Van Winden, 2021). In this case, it is important to verify whether the found ‘problems’ are actual problems for ANVA. It must be taken into account that answers of interviewees can be just personal irritations and not process-wide problems. So, the interviews must be held such that interviewees will only name the process-wide problems.

3 Theoretical framework

This chapter will form the theoretical basis of the research. The literature supports the research and is helpful in understanding the structure of the research.

3.1 Definition of data conversion

First, it is essential to understand what a data conversion is. It is important to know how a data conversion can be used as a tool to support other processes and how it can further be important for a company. To that extend, a systematic literature review was conducted. In the next paragraphs, just the results are presented. The full systematic literature review can be found in appendix A.

According to Reeve, data conversion is changing data structure to let them comply with changes. If a new application system needs to be implemented or data must be moved from one place to another, the two piles of data have to be brought in harmony. Data conversion is making these sets of data compliable with each other while consolidating data sets or applications. The process of data conversion contains three steps. The first step is extracting data from the source database. Second, the data is transformed to make it have the same format as the data in the desired new databases. Third, the transformed data is loaded into the new database. (Reeve, 2013)

A good data conversion can be essential for a company or a process. Antonov et al. illustrates how the lack of data conversion can cause problems. If there are multiple sources of data that must be merged, there can be multiple problems. Each source will have its own unique set of data and its own layout of their data set. This makes it hard for someone to look up data in those databases. The information stored in the form of data lacks organization and structure, which can lead to negative effects for the company or process. With a proper data conversion done, information is accessible much easier. Jain et al. emphasize the benefits of a data conversion. They state that a data conversion lowers data diversity, makes data more accessible, makes data better readable and supports the reproducibility. All these benefits help a company get more out of their data. (Antonov et al. 2020) (Jain et al., 2021)

3.2 Process modelling method

To understand the DCP, chapter 3 illustrates the DCP with conversion method descriptions, a process walkthrough and a model. Here, the model is a visualized version of the walkthrough. We consult the book Business Process Management of Weske to find out how the DCP should be modelled, given the goal of the model.

According to Weske, a process model represents a set of process instances that have a relation. Process models consist of nodes and arrows which connect the nodes. The node represents an activity, event or gate way. The sequence of nodes, and thus the walkthrough of the process, is denoted by the direction of the arrows. The complexity, or concreteness, of the model depends on the goal of the model. The modelling language also depends on the goal of the model. If the goal is to deeply understand all process instances and their exact relation, the process model can be made very detailed and a formal modelling language should be used. If the goal is to understand how a process works at a general level, the process model can also be made very general and a simple informal modelling language will be sufficient. (Weske, 2012)

Since the goal of our DCP model is to generally understand the DCP and the sequence of activities, the model can be made quite general in terms of relationship arrows. However, there will be many process instances which all represent a small part of the process. So, the model will consist of many process instances but few relationship arrows. The walkthrough of the process is described by one

red thread. Next to this, an informal modelling language is used. Since the goal is only to visualize the sequence of activities, no formal language that is for example able to express a special relations between nodes is needed. To conclude, a simple, informal model is sufficient to reach the goal of the model, so there is no need to make a complex model.

3.3 Problem identification methods

The section will go over the methods used for identifying problems in the DCP. First, a trade-off table is used to illustrate one of the core problems of the DCP. The trade-off table is not a problem identification method itself, but it allows us to understand the structural problem of the DCP. The trade-off table can be seen in section 4.3.

Next, we search for problems in the DCP by using the Theory of Constraints (TOC). According to Aryanezhad et al., the main principle of the TOC is that in every system or process, there is at least one constraint that limits the system or process in achieving higher levels of performance, such as throughput rate or machine utilization. By having the focus on removing this constraint, the system or process will reach higher levels of performance, whereas focusing on a non-constraint aspect of the system or process will not have an effect on the levels of performance. In general, the TOC focuses on improving the qualitative performance of the system, in contrast to many other methodologies which focus on cost reduction. The principles of TOC are therefore also usable for our DCP. (Aryanezhad et al., 2009)

As stated by Inman et al. and Pacheco et al., the TOC contains three general dimensions that form the core of the TOC construction, which are logistics, performance measurement and problem solving. Within the three dimensions, constraints are found with the use of Key Performance Indicators (KPIs). KPIs are measurable variables that indicate how well a process is going in terms of numbers. If a KPI is scoring low, it can be seen as a problem. The results of the TOC can be seen in section 4.4. (Inman et al., 2009) (Pacheco et al., 2021)

Lastly, interviews are used to find problems. Interviewees are asked to give their opinion on the process. To answer that, they have to name problems of the process. With this method, problems that are impossible to identify with just a process model can be identified. The problems found by interviews are in section 5.2 and 5.4.

4 The data conversion process

In this chapter, first the DCP is illustrated extensively, by means of a detailed description per conversion method, a process walkthrough and a model. Then, problems are identified with a trade-off table and the Theory of Constraints. Section 3.1 describes the three different conversion methods. For each method, the most important properties, its pros and cons, will be elaborated on. In section 3.2 will be explained how the DCP at ANVA looks like. The process will be gone through step by step. In the appendix will be a clear overview of the process, in the form of a data conversion process model and a data flow diagram, based on these first sections. In 3.3, a trade-off table will be used to identify a core problem of the DCP. The Theory of Constraints is used in 3.4 to bring more problems to the surface.

4.1 ANVA conversion methods

ANVA has three methods of doing conversions. Conversions can be done manually, with the ETL tool and with custom conversion tools. Generally, a conversion contains three steps: extracting data from the old databases, transforming data and loading the data into the new database. All three conversions methods do these steps in a unique way.

In this section, each of these methods will be explained extensively. For each method will be explained when they are used, how they are used and what the pros and cons of the methods are. Most information of this section comes from personal communication with John IJzerman and Sander van der Burg, both application consultants at ANVA. Unless referenced otherwise, information comes from these conversations.

4.1.1 Manual conversions

Manual conversions are small conversions that are done by hand by of the consultants. A consultant of ANVA manually makes the conversion tables within a couple of days. The conversion tables are made in Microsoft Excel. Conversions tables simply list where all elements (such as name, policy, address) are located in database 1 and where they will be located in database 2. The conversion table is made such that software program can read it, after which it will convert the input data into the ANVA system. With small conversions, the data converted will most of the time only consist of customers' names and policies.

Conversions are done manually when the conversion is relatively small. If the size of the conversion is small, it means that a small amount of data needs to be converted. These types of conversions are done for small clients or small merges. To get an idea of the size of these conversions, an example is that a client asks to put 1200 new clients and 1500 new policies into the ANVA system. This is of course a manageable amount of data that needs to be converted.

The pros and cons of manuals conversions are summarized in table 2.

Pros	Cons
Fast converting	Only small conversions possible
No software programmers needed	Limiting conversion options
Organized conversion process	

Table 2 - Pros and cons manual conversions

The process of a manual conversion can be modelled as can seen in figure 2.

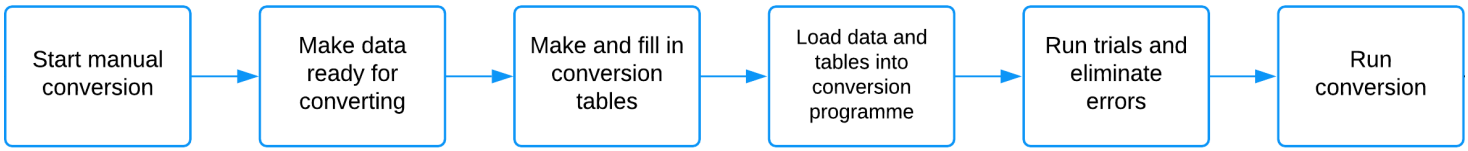

Figure 4 - Manual conversion process

4.1.2 ETL tool conversions

With ETL tool conversions, a consultant uses an ETL tool for converting the data. According to Reeve, an ETL tool essentially is an automated version of a manual conversion. An ETL tool obtains data from a source (extract), changes it to another format that is compatible with the desired storage (transform) and puts the data into the desired destination (load). This how an ETL tool generally works. In figure 5, the general ETL tool process is visualized. (Reeve, 2013)

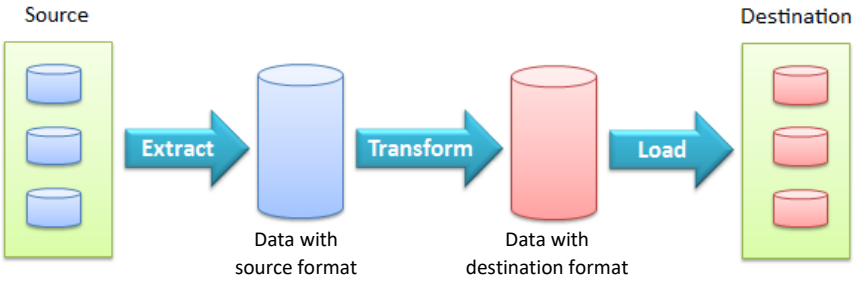

Figure 5 - General ETL tool process

At ANVA, conversions with the ETL tool are mostly used for conversions with provincial offices. The ETL tool process at ANVA begins with extracting data from the client. This data needs to be cleaned up before it can be read in the tool. The cleaning of data implies that the data is generally made into the right format. Incorrect space bars and enters are removed and the format consistency is checked. The data is now ready to be converted. Now, the conversion tables are made. General conversion tables are copied and edited to the specific data of the conversion. These tables form the core of the conversion, they determine what data in the old system will be in the ANVA system. Now, the connection diagrams of the ETL tool can be made. This is essentially the translation of the conversion tables in the tool. Connection diagrams connected pieces of data of the old system to pieces of data in the new system. The connection diagrams let the tool understand how to convert the data.

Now, the complete set up of the conversion is made. At this time, trials can be run to detect errors. Trials are done to test the conversion and detect errors. After each trial, all errors can be solved to improve the conversion. Next to this, data is checked on correctness and accuracy in the ANVA system. After a few trials, the conversion is improved enough, which makes it ready to actually run. Ten, the final conversion is run.

In figure 4, the ETL tool conversion process is modelled.

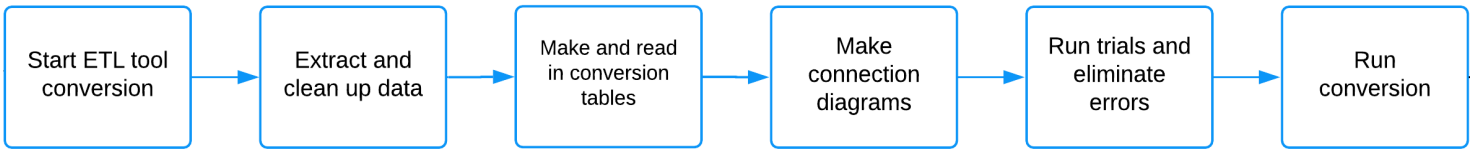

Figure 6 - ETL tool conversion process

Using the ETL tool for a conversion has its pros and cons. The ETL tool can be operated through an interface. On this graphical interface, the user can read in data and conversion tables and can make the connection diagrams. He can also run trials or the actual conversion. Subcutaneously, the ETL tool generates an SQL document to convert the data. The user does not notice any of that, since the whole process is doable from within the interface. This easy-to-use user interface is made to make the ETL tool useable for a lot of people, not only experts of the tool. This is one of the main advantages of the ETL tool. With the interface, the ETL tool can used by just a consultant without the help of a software programmer.

Furthermore, ETL tool works fine in the basis. The tool can be seen as a pleasant tool to work with, provided that the conversion is suitable for the tool. Next to this, the tool has many capabilities in terms of what it can convert and what additional functions it has. The level of usability and tolerance it has for the user, in combination with a solid technological structure forms the appealing basis of the ETL tool.

However, the ETL tool does also have a couple of drawbacks. The drawbacks are mainly in the output of the ETL tool. The tool can generate two types of output: XML and extract.TXT. XML output has a big downside with regard to the processing speed. The ETL tool does not produce XML slowly, but XML cannot be stored in ANVA in a fast way. To send XML in ANVA, a software program called e-ABS is used. e-ABS is also used by clients to send new clients directly into the database of ANVA. e-ABS is therefore mainly built to send individual relations, and not large groups of relations, to ANVA. So, the program takes a lot of time to send thousands of records at the time to ANVA or will even crash in the worst case. This is because the program takes one record, checks all data of that record, then sends it into ANVA if it passes all checks. All these checks per record cause the program to process the records slowly. Therefore, ETL conversions with XML output can only done with small amounts of data.

With extract.TXT output, the downside is that the software X11755WP must be used for sending the data into ANVA4/5. The program X11755WP is very limited in terms of the data it can send. Only customer names and policies can be sent, which is not enough for bigger and more complex conversions. More complex conversions also need for example damages to be converted. Since the X11755WP software cannot do that, using that software is also very limiting. So, both ways of exports have their own barriers. These barriers are the drawbacks of the ETL tool.

The pros and cons of the ETL tool are summarized in table 3.

Pros	Cons
Works well in the basis	XML output processing to ANVA is slow
Easy to use interface	X11755WP output has limited options
Many conversion options	Large conversions not possible
No software programmers needed	

Table 3 - Pros and cons ETL tool

4.1.3 Custom conversions

Custom conversions are done when the conversion is very complicated and large of size. These types of conversions are done with Cobol software. The Cobol software hardly has any restriction with regard to the size or the complexity of the conversion. This is because software programming codes are made manually, completed customized to the specific conversion. The programming is about 70% standard code and 30% completely custom code. It costs a lot of time to write that 30% of code that needs to be completely made up from scratch.

With a custom conversion, first the data is made ready for converting. This mainly entails cleaning the data and making sure a general format is applied to the complete document. Then, the best existing conversion code is chosen from all earlier conversions. An earlier used code is used to already have a great basic code to start with. The choice depends on the client, the software used by the client and the complexity of the conversion. The code with the similarities will be chosen as a start point for the current conversion.

Next, this code will be adjusted and supplemented, to create the right code for the conversion. After many hours of writing, the code is ready, after which trials can be run. Trials serve to identify errors and solving them, in order to improve the code. After sufficient trials, the code has improved such that the conversion will be of good quality. Finally, the actual conversion can be run.

In figure 7, the process of a custom conversion is modelled.

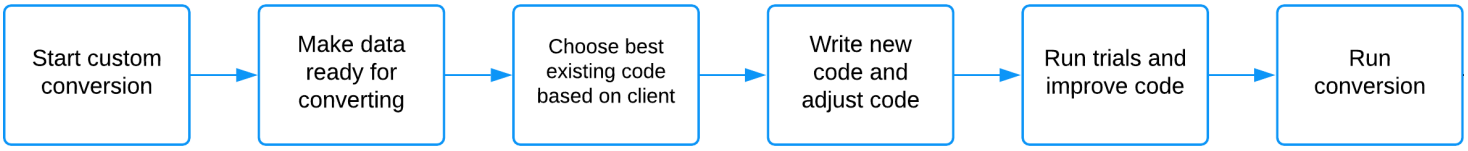

Figure 7 - Custom conversion process

The advantage of custom conversions done with Cobol is that it can do very complicated conversions that are too large for the ETL tool. Cobol is also very fast in running conversions, which means that it can do conversions containing millions of records in a weekend. This is a big advantage, especially compared with the speed of the ETL tool. Overall, the most appealing property of conversions with the Cobol software is the size and complexity it can handle. This results in a carefree conversion process, with regard to those topics.

The disadvantage of Cobol is that complicated conversions cannot be done by the application consultants alone. For more complex conversions, they will need the help of dedicated Cobol programmers. However, ANVA does not have a lot of Cobol programmers employed. This means that there are only a few Cobol programmers available to program conversions for ANVA. On top of this, most programmers do not like to program conversions. They rather program something else than conversions, with which troubleshooting is a big part of the job. This shortage in motivated Cobol programmers is the main drawback of conversions with Cobol. That is the main reason why the ANVA wants to get rid of Cobol conversions in the near feature.

Other than that, writing code for each conversion in Cobol is hard to do. It takes a lot of time to write new code or adjust code to the conversion. The fact that each individual conversion needs its own code for even a small part means that the process is very laborious. Even the use of old code as a basis for the new code cannot help the fact that it is still an inefficient use of time to write new code for each conversion. This inefficiency does not have a positive impact on the process and the employees of ANVA.

In short, custom conversions take a lot of time and effort to do, but also bring a lot in terms of technical possibilities. However, this type of conversions will not be done anymore in a couple years from now. The pros and cons of custom conversions are summarized in table 4.

Pros	Cons
Endless conversion options	Writing code is laborious
Fast converting	Software programmers needed
Capable of doing large conversions	Inefficient use of time

Table 4 - Pros and cons custom conversions

(J. IJzerman, personal communication, 22 April, 2021) (S. Van der Burg, personal communication, 3 May, 2021)

4.2 DCP process

The DCP begins with the choice to do preliminary data research or not. This choice depends on the size of the conversion and the experience with the type of conversions. If the conversion is not too large and the system used is familiar, no preliminary data research is done. When the system is unfamiliar, the data must be research to find out how the data looks like. The data must be made compliable for the conversion software. The research serves to find out how the data differs from already known data layouts.

After the preliminary data research is done, it is time to decide which conversion method will be used. This largely depends on the size of the conversions. Small conversions will be done manually. Middle sized conversions will be done with the ETL tool. Large conversions will be done with Cobol software. The choice can also depend on the software used by the client. If the clients for example uses software X, the conversion will be done with Cobol, because Cobol software has already been used for large software X conversions in the past. Since the software is made for only software X, it can be reused for other conversions with software X. Overall, the choice for the conversion method depends on the size of the conversions and the system from which the data is converted.

The next step in the DCP is to extract the data from the client. This can be done in two different ways. First, the client can permit ANVA access to the system, after which ANVA can extract the data by itself. Second, ANVA can ask the client for all documents needed for the conversion. Then, the client sends all documents to ANVA, after which ANVA can convert the data to ANVA4/5. Either way, the data has arrived at ANVA and is optimized to make it ready for converting. This optimization entails cleaning and restructuring the data such that it becomes general input data of the right format.

Now, the conversion process itself starts. With an analysis of the data, conversion tables are made. ANVA makes the general set up of the table, then the clients fill in the rest of the table. Clients fill in the tables themselves, which allows them to create their own ideal layout in ANVA. They also can decide on conversion dilemmas themselves. For example, in the old system, an employee is labelled as a surgeon. However, the ANVA system only has the category doctor and does not specify more. Here, it must be considered if the category doctor must be specified, so that the surgeon will be labelled as a surgeon instead of a doctor. The client can decide for themselves, by letting them fill in the conversion table. Anything that could not be filled in by the client will be filled in by ANVA. The conversion tables are now ready, which means that the conversion can be carried out.

Before the actual conversion is carried out, some trial conversions are done. The main goal of doing trials is to filter out mistakes in the conversion. The number of trials depends on the size of the conversion. For a relatively small provincial conversion, about two trials are scheduled. For a big proxy, up to six trials will be done before the actual conversion is done. After a couple of trials are successfully executed and enough errors are solved, the final conversion can be run. After the conversion is ran, ANVA will have contact with the client about the conversion to talk about the results. The DCP is now done. The total conversion process will take about three months of time for smaller provincial conversions. For a larger proxy conversion, the total process could take up to two years.

Based on the DCP walkthrough description and the complete process per type of conversion description, a complete model of the DCP is made. This model can be seen in appendix A. With the general DCP model, a data flow diagram (DFD) model can also be made. This model shows how data is transferred throughout the DCP. The DFD can be used to analyze data flows through the process and checking whether this is an efficient process or not. The DFD can be seen in appendix B.

4.3 Trade-off table

Now, we analyze the DCP to identify and structure the most important problems of the DCP. In order to do so, we use multiple problem identification methods. The DCP model and description in combination with output from the interviews forms the input of the analysis.

First, we compare the three conversion methods of section 3.1 by lining up the pros and cons per category. To this extent, we take the pros and cons tables of the three conversion methods and merge that into one table. This results in table 13, which illustrates the differences between the conversion methods.

Category	Manual	ETL	Custom
Conversion speed	High	Medium	High
Labor intensity	Low	Medium	High
Conversion options	Limiting	Many	Many
Data capacity	Small	Medium	Large
Programmers needed	No	No	Yes
Process	Organized	Fairly organized	Cluttered

Table 5 - Comparison conversion methods

From this table, it becomes clear that every conversion method has its own pros and cons and operates best in its own lane. Manual conversions can do small conversions with low demands at a high speed but are very limiting with regard to data quantities and conversion options. ETL tool conversions are suitable for conversions with high demands without needing programmers, but also are limiting in terms of conversion speed and capacity. Custom conversions can handle any size and any demand but do need a lot of time and the help of programmers. It can be concluded that regardless of the conversion method chosen, a compromise must be made. None of the methods is optimal for every conversion. This can be considered as one of the most significant problems of the conversion process. The core of the DCP, the conversion tool, does only comply with the conversion types if multiple methods and tools are used. This creates the need for one non-comprising conversion method.

4.4 Theory of Constraints

We now find problems in the DCP by using the TOC, as explained in section 3.3.

First, we try to find constraints in the logistics of the process. For the DCP, logistics are about managing the transfer of information and data. If we look at the DCP diagram and the data flow diagram, we find that the flow of data is generally efficient. However, with custom conversions, there is an unnecessary back and forth interaction between the consultant and the software programmer in the latest stage. With other type of conversions, this flow is more efficient, since the conversion software for a specific conversion is not made by a software programmer. Another bottleneck of the process is the XML processing module of ANVA. Due to its low speed, partly because of validations, conversions with XML output are slow and therefore considered unusable for larger conversions. This bottleneck is more a technical problem on the surface, but can be considered managerial if no resources are put into the improvement of the XML module.

The second dimension of TOC entails the performance measurement. This part focuses on the performance of a process like throughput time and output quality. Although ANVA do not necessarily want to do as many conversions as possible in the first place, it is always good to decrease the amount of work hours needed for a conversion. In the DCP diagram are two process boxes of analyzing and improving data, regardless of the conversion method. This means that the data

analysis and improvement is not done in one go, which costs extra time. Also, in order to improve quality, it is necessary to have contact with the client before or throughout the process. Now, apart of the first consult where no agreements regarding data quality are made, there is little contact with the client throughout the process. Thus, there are no agreements about the quality or quality checks in the process. This can cause problems with clients and should be improved.

The last TOC dimension is about problem solving and the thinking process behind the DCP. Standing out from the DCP diagram are the three ways of converting, all serving the same purpose. Although there is a reason for existence of all three methods, it still may be an indication of an ill-considered conversion process. So, the conversion process was not thought out well, which resulted in an accumulation of small new conversion software extensions, culminating in three messy conversion methods. Another constraint in this dimension is the difference in opinions and experiences among the employees. For example, while most employees agree on the future of the conversion process, which is using the ETL tool, people are divided on which ETL tool to use and how the ETL tool should be improved. It is hard to solve problems when the employees are not on the same page.

All three dimensions of the TOC have now been discussed. The problems found with the use of TOC are summarized in table 14.

	Dimension	Problem
1	Logistics	Unnecessary back and forth interaction between the consultant and the software programmer
2	Logistics	XML module ANVA is slow
3	Performance measurement	Data analysis and improvement is not done in one go
4	Performance measurement	There are too few agreements about quality or work hours with the client
5	Problem solving	Conversion process as a whole is a mess as a result of ill-considered strategy
6	Problem solving	Employees are not on the same page

Table 6 - Problems TOC

5 User experience

In this chapter will be reported on how users of the DCP experience the process. In order to find out, interviews were held with ANVA employees and clients. Respondents were asked to give their opinion on the process and how they think the process could be improved. Together with the analysis above, the experience of users forms the basis of the package of imperfections of the DCP. That package will be used for finding a solution to improve the process in chapter 7.

5.1 ANVA employees

First, ANVA employees are asked to describe their experience with the process. To this extend, open interviews with the following core topics were held:

- The employee's general experience with the process
- The employee's thoughts on the technical difficulties of the process
- The employee's thoughts on the choice of people doing the conversions
- The employee's thoughts on the general course of the process
- Ideas or suggestions to improve the process in any way

Answers to these questions are processed into a list of experiences, ideas, wishes or needs with regard to the DCP. The employees interviewed have the following role in ANVA:

1. Chapter lead consultancy
2. Application consultant
3. Chapter lead software development
4. Lead architect
5. Product owner ANVA4/5
6. Product owner ANVA Hub

These people play their part in the DCP, all from their own perspective. The differences in perspective are the basis for a broad overview of experiences with the process. This multi-perspective approach ensures that possible problems in the process are judged as fair as possible.

5.1.1 Chapter lead consultancy

The first ANVA employee interviewed is the chapter lead of consultancy. He is responsible for the consultants who do conversions and is therefore an important person with regard to the DCP. Based off the interview, the following experiences, ideas, wishes or needs came up:

- When doing a DCP, often software programmers are needed to do programming chores. Needing the help of software programmers costs time and money. Because the interaction between consultant and programmer is laborious, the process takes much longer than a process ran by only a consultant. Next to this, instead of only one person, now two persons work on the process simultaneously, which is double of costly in terms of salary. Therefore, it is desirable to have a DCP that is made and maintained by programmers, but does not need the help from programmers in day-to-day conversions. This makes the process easier, faster and cheaper, which can also be push forward to client. With a faster DCP, the client's waiting times will be shorter and the client will be charged less for the DCP.

- An important step in achieving a DCP without the need of software programmers is the use of an ETL tool that has a user-friendly interface. The ETL tool may use complicated software subcutaneously but should not be hard to use. This is already the case with the ETL tool, but it is important that it remains the same So, the ETL tool should consist of two components, the tool itself, which can be SQL, and the user interface.

- An addition that would make the DCP better is the possibility to add code at more points in the process. At every stage of the process, there should be gates to add code to that part of the process. That code can for example contain rules that determine how the data is send further or an extra check of the data. The possibility to add code at more stages will reduce a lot of programming issues.

- In line with the last points, it is desirable that certain exceptions or special rules can be added directly from the interface. The interface then facilitates the option to for example let out the agent with agent number 11349. Adding this via the interface can save a lot of time programming the exception in the hard code, by already having a general exception programmed in the code.

- Moreover, it would be great if a solution to these problems also considers the switch to ANVA Hub. Soon, ANVA will be switching from ANVA4/5 to ANVA Hub. It desirable that the direction of the solution is aimed towards compatibility with ANVA Hub, because that would save a lot of time in the future.

- Combining all the points above, a general is wish is to become cheaper and faster when doing a conversion. Now, clients are often not completely satisfied with the conversion and that must be changed. Becoming cheaper and faster, while delivering the same or better quality will rise clients' satisfaction.

- Next to this, ANVA should be making more sustainable and durable conversion tools to use to more conversions than is done right now. A lot of consolidation is happening in the insurance market. Partly because of that, the livelihood of ANVA is that clients become bigger when they are a client of ANVA. Conversions have a direct link with that, as merging companies need conversions. That is also why improvement of the failing DCP is important.

In table 5, all bullet points of the chapter lead of consultancy are summarized in one sentence.

	Experience, idea, wish or need
1	Make the DCP as much 'software programmer-free' as possible
2	Keep on using the ETL tool
3	Make more gates to add code in the DCP
4	Create the possibility to add exceptions or special rules directly from the interface
5	Consider the switch to ANVA Hub
6	Become faster and cheaper
7	Make conversions more sustainable to maintain the livelihood of ANVA

Table 7 - Summary chapter lead consultancy

5.1.2 Application consultant

Next, the experiences of an application consultant are listed. This application consultant is the ETL tool specialist in ANVA, which implies that he is the main ETL tool user in ANVA. Hence, his experiences will mainly be about the ETL tool conversion process.

- A big problem is that ANVA must deal with the state in which data arrives. For automating a conversion, it is essential that data has a general structure that is readable by software. However, data is often delivered with different structures and layouts. This causes problems for the software. That is why consultants of ANVA spend hours to make the data compliable with the conversion software. Although this is not really solvable by changing something in the ETL tool or more general in the conversion process, it is important to take into account when finding improvements for the process. So, the input data is unstructured and therefore unreadable for the software, which cannot be solved by using new software.

- A possible way to improve input data quality is to demand more from clients concerning the delivery of that data. ANVA has the possibility to make clients responsible for delivering data in the right format. However, it could be questioned if this measure will work in practice, since most companies do not have enough knowledge to format the data according to the wishes of ANVA. Still, an advantage of this measure is that consultants will need less time to do conversions and thus conversions will be cheaper.

- The ETL tool is not fast because of the slow data entry program of ANVA. The program called e-ABS cannot handle the XML output of the ETL tool well. It tears apart every piece of information and gets it through numerous kinds of checks, which takes a lot of time. If one would put a new relation in ANVA, he would need to go through several steps before he is done. The e-ABS system does the same, but robotized. In combination with the number of checks, it takes much time to load in thousands of relations.

- In order to create a faster ETL tool conversion process, the XML module of ANVA4/5 must be upgraded. A suggestion to improve the data entry speed is to send data parallelly to the XML module. Now, the relations are sent sequentially to the module, so it processes only one relation at the time. A better way to do this is to gather about five relations and send these simultaneously to the module. The module then processes five relations at the time, which makes the process five times as fast if the processing speed does not decrease.

- Since the ETL tool conversion process has its difficulties with both XML and x11755WP, I would suggest looking for something else in the future. Both options are hard to improve and will take a lot of time and resources. It would be better to invest that effort into coming up with a new solution that takes away weak points of the current options.

In table 6, all bullet points of the application consultant are summarized.

	Experience, idea, wish or need
1	Realize the problem of input data formats and the software-insolvability of it
2	Consider demanding input data of higher quality from clients
3	The slow data entry of the ETL tool is due to the incapable XML module of ANVA
4	The data entry speed to ANVA could be improved by sending data parallelly to the module
5	Put effort in finding something new and not in improving current options

Table 8 - Summary application consultant

5.1.3 Chapter lead software development

Next, a conversion with the chapter lead of software development was held. As the chapter lead of software development, he is responsible for the development, modification and maintenance of software in ANVA. So, he can play a big role if new software must be developed for the future DCP. His experiences, ideas, wishes or needs are:

- ANVA could have done better with regard to the focus of conversions. ANVA focused on doing individual conversions as quickly as possible. Conversions are made as ugly and simple as possible, since they will be used only once anyway. So, conversions are done such that it is just sufficient for the client to not complain. This attitude of ANVA should change drastically. ANVA should focus on a more sustainable DCP, with a tool that can be used for almost all conversions.

- In line with the last point, the ETL tool should have been set up much more extensive, such that it could have been used for many conversions, which is much more sustainable. This implies that the tool can do almost all conversions with very little customized code per conversion. This tool must be able to convert from and to every application needed. A tool like that ensures that ANVA can the deliver the total package with regard to conversions. ANVA missed the chance to make the ETL tool completely viable first time. The ETL tool is often customized for individual conversions, each time adding a different function. It would have been more sustainable to make an extensive tool right from the start. However, this is still possible, since ANVA would still profit from a completely renovated ETL tool. The chance to do that lies ahead of ANVA.

- ANVA has no limitations with regard to technical possibilities. Everything is possible to make, as long as enough money is spent on the right people. Putting ideas into practice is the easy part, coming up with good ideas is the hard part. ANVA finds it hard to come up with the elaboration of ideas, the so-called paradise on paper. When trying find new ideas, it should be considered that there are no technical limitations.

- ANVA has problems with creating the big picture at the start of conversions. If you know how a house looks like, you do not start building the attic. ANVA first built the attic, after which they realized they need a foundation, so they build a foundation on top of the attic. Next, they realize a shed is needed. Then they put a man between the shed and the house holding them together. This way of building is not thoughtful and can be improved drastically. To do that, it is important to realize it is all about the paradise on paper. The concept should be thoughtfully elaborated and worked on paper, before it is built. This concept should be as big possible. It is not important to fully build the whole concept, but is about the cohesion between different parts, connected with each other in the big picture. The big picture view allows ANVA to carefully research weak points of the process and strategically invest in these weak points only.

- Dependency with software developers should be addressed. It is not favorable for ANVA to be dependent on software developers for updating programs or codes. Software developers have their own tension fields which makes it hard to make time for ANVA specifically and should therefore be avoided. ANVA should consider doing something about this dependency.

In table 7, all bullet points of the chapter lead of software development are summarized in one sentence.

	Experience, idea, wish or need
1	Focus on a sustainable DCP instead of individual conversions
2	Renovate the ETL tool such that it becomes a sustainable conversion tool
3	With finding solutions, consider that there are no technical limitations
4	Sketch the paradise on paper before implementing an ill-considered concept
5	Try to reduce software supplier and maintainer dependency

Table 9 - Summary chapter lead software development

5.1.4 Lead architect

To get more ideas from someone who has a technical background, a conversion with the lead architect of ANVA was held. As lead architect, he is mainly concerned with the technical structure and environment of ANVA, especially focused on the perspective of the client. Because he has a substantial share in building ANVA Hub, he has knowledge about how ANVA Hub is designed to cope with conversions well and how conversions are made easy for clients.

- To improve the DCP in general, ANVA should work towards using only one application for all conversions. The current diversion in conversion tools is fatal for the DCP. For every tool or method used, a complete process is set up to do conversions with that tool. Different tools are used slightly different, but they all serve the same purpose. It would be much more organized, faster and cheaper to do all conversions with one tool.

- The diversion in conversion tools caused diversion in data in ANVA4/5. Because every conversion tool converts data in another way than the other, the output data that ends up in ANVA4/5 differs in format and quality. For example, with a conversion to ANVA4/5 policies with the date February 31st ended up in ANVA4/5. Now, when data is exported to either another system or ANVA Hub, the conversion tool will give errors because the data is invalid. So, differences and errors in a validation process, but also in any other component, cause problems. The bad quality of data is a good reason for ANVA to look into this.

- Realize that within a couple of years all data in ANVA4/5 must be converted to ANVA Hub. This huge project will consist of more than 400 conversions. These conversions should not be stowed as unimportant. At some point components of ANVA4/5 will be shut down and before that all data must be converted. Software engineers should be engaged with this project, they should think about how the data will be converted to ANVA Hub. This is about issues as which parts will not be converted, which parts will be created from scratch, which parts remain the same.

- ANVA currently uses Matillion for data warehouse purposes, mainly because of the underlying databases warehouse Snowlake. Essentially, Matillion is a ETL tool which can be used to convert data. It may be an idea to also use Matillion for conversions in ANVA, in combination with Snowflake. Then standard generalized conversion process can be made in Matillion. Matillion can be used to convert raw data into raw data with the right format. The new data can be imported in either ANVA4/5 or Hub. Maybe it could be researched if this way of converting could work for ANVA. Next to this, what does ANVA currently not have and what is needed to make it work? What is potential for conversions with Matillion? With an answer to these questions, ANVA would know if this a good option to consider. The use of

Matillion would reduce differences in the technical landscape, since it is already used for data warehouse purposes in ANVA hub.

- Using another ETL tool than the one currently used might be a good idea, since the current tool is made and maintained by one person. If some happens with that person, ANVA are left empty-handed. Because of this, continuity of this tool cannot be guaranteed. If the ETL tool needs an update or data needs to be changed, ANVA hardly has any guarantees.

In table 8, all bullet points of the lead architect are summarized in one sentence.

	Experience, idea, wish or need
1	Use one universal tool for all conversions
2	Improve data quality by using that tool
3	Realize that shortly data from AVNA4/5 must be converted to Hub
4	Look into Matillion as a replacement of the current ETL tool
5	Realize the possible lack of continuity of the current ETL tool

Table 10 - Summary lead architect

5.1.5 Product owner ANVA4/5

To get a grip on the perspective of the destination of converted data, the product owner of ANVA4/5 is interviewed. As product owner, he is concerned with all business around ANVA4/5 and thus also with conversions to ANVA4/5. As someone with another perspective relative to other interviewed employees, he experiences the conversion process from the other side, which can be interesting for his opinion on the process. His experiences, ideas, wishes or needs are as follows.

- Usually when ANVA develops software, there is one big team that consists of mainly software developers and consultant which together create the software. With conversions, the data analysis is done by consultants, then a software programmer makes the software and then the consultants test the software. After an error is discovered, the error is communicated with the software engineer and he solves the error. After that the consultant tests the software again and brings it back to the software engineer if is not good enough. This results in a constant back and forth motion between the software developer and the consultant. Conversions can and should be done more efficiently by working smarter. The complete conversion should be completely thought out before any software building is done. The conversion should be thought out and built as a team and not as collaborating individuals from different fields. If the task division of analyzing, building and testing remains the same, this way of working would be more efficient.

- The best way of converting is with the ETL tool and the XML output. Data validation is an important factor in converting and should be made one of the highest priorities. The data validation of conversions with XML is regulated well and therefore XML should be used.

- ANVA should set higher standards on provided data of clients. Now, provided data is sometimes full of errors which cannot be interpreted by the conversion software. Consultants lose a lot of time with analyzing and improving data to make it ready for converting. ANVA should demand higher quality data from the clients, to make the conversion process cheaper, easier and shorter.

- In line with this, ANVA should make more agreements with clients in advance of the conversions. These agreements should cover topics as data quality and hours of workload. For example, ANVA will devote 300 hours of work to the conversion and nothing more, regardless of the outcome after 300 hours. This reduces the risk of needing significant more working hours than declared or getting backlash for delivering bad quality of data.

- A possible way to speed up conversions with the ETL tool and XML output is to distribute data over time when converting. The client will probably not need all data just after the conversion. To speed the conversion process, ANVA can only convert data from the past year and do less relevant data at a later stage. The slow processing speed of XML will than be less of a bottleneck, since only a part of the data needs to be processed.

- Instead of putting much work into individual conversions, put more time into revising the general ANVA landscape, which also benefits conversions. For example, the validations registered in the APIs might be outdated. It can be profitable to revising the validations, which also helps with the XML output problem, since the transferring of data is faster due to less validations. Next to this, the APIs are cluttered and should be rebuild. Now, there is an API for small labels in a part of ANVA4/5. These APIs should be generalized more, so that for example there is only one API for relation number in the complete ANVA landscape. These types of revisions would make the ANVA landscape much clearer and that will also benefit conversions.

A summary of the experiences, ideas, wishes or needs of the product owner ANVA4/5 can be seen in table 9.

	Experience, idea, wish or need
1	Bring software developers and consultants together as a team when building conversions
2	Use the ETL tool with XML output for conversions
3	Demand higher quality data from clients
4	Make more agreements regarding data quality and work hours with clients
5	Do not convert data all at once, but distribute less relevant data later
6	Revise several parts of the ANVA landscape

Table 11 - Summary product owner ANVA4/5

5.1.6 Product owner ANVA Hub

Lastly, the product owner of ANVA Hub is interviewed. Just as the product owner of ANVA 4/5, he is positioned on the other side of the process, but this time on ANVA Hub. At his position, he may have interesting ideas about how the conversion process should be shaped in the future, so that it matches well with ANVA Hub. The experiences, ideas, wishes or needs of the product owner of ANVA Hub are:

- Especially in the first years, the client base of ANVA Hub will differ quite a bit from the client base of ANVA4/5. The clients using ANVA4/5 are mostly larger provincials or proxy intermediaries whereas the clients of ANVA Hub will be small agents. This requires a different conversion approach, since agents do not have the same demands as proxy intermediaries. For agents, just information about relations is enough, since they just secure an insurance for a provincial or proxy intermediaries. Unlike that, proxy intermediaries do need information about for example past policies and extensive policy information in general. Therefore, conversions to Hub are free to have limited functionalities or to leave out data from previous years. This holds for at least the coming few years.

- The difference in client base does also have an effect on what is generally demanded from the conversion. Large proxy intermediaries have a large budget for a complex conversion, small agents have a small budget for a simple conversion. Since agents are the target clients for Hub in the coming, it is important to have a tool which can do simple conversions in a small amount of time. This will allow ANVA to keep prices low, which is the key in acquiring clients. Acquiring a lot of agents is important for ANVA, since that will keep proxy intermediaries also running on ANVA.

- The current or another ETL tool is good for a cheap conversion process needed for these agents. For ANVA, an ETL tool can be a cheap way to do simple conversions to Hub. It would therefore be a good idea to invest in the ETL tool, making it ready for conversions to Hub.

- ANVA should be careful with demanding higher data quality from clients. The quality of data is important for the conversion process and for the client in the future. However, especially small agents should not be charged with an extensive data cleaning process, since that would deter them from converting to ANVA. Next to this, the principle of quality in = quality out holds in general. The quality of data is the responsibility of the client. If the client delivers bad quality data, he can also expect bad quality data back after converting. However, the sore point is here that converting is harder with bad quality data, since the conversion program will give a lot of errors. A compromised solution may be the use of an error list. With this, the clients delivers its data to ANVA, after which the data is scanned on errors by a program. The program makes a list of errors, after which the client can improves its data in a small amount of time.

The points of the product owner ANVA Hub are listed in table 10.

	Experience, idea, wish or need
1	Conversions to Hub are free to have limited functionalities or to leave out data from previous years
2	A simple conversion process with low prices is essential for Hub conversions
3	The use of an ETL tool is good way of converting for the future
4	A compromised solution for the data quality issue is the use of an error list

Table 12 - Summary product owner ANVA Hub

5.2 Employee problems

We now identify problems based on the experiences of employees of ANVA. Summarizing table 5 up to 10, we get an overview of al experiences, ideas, wishes or needs from ANVA employees. Overlapping statements are merged to one more general statement. The overview can be seen in table 15.

	Type	Statement
1	Experience	The quality of input data is a problem and cannot be solved with software
2	Experience	The current ETL tool supplier has a lack of continuity and is therefore risky
3	Need	Before building something, first overthink the complete concept
4	Need	Make more agreements regarding data quality and work hours with clients
5	Need	Decrease the use of software engineers needed for conversions
6	Need	Creates more gates to add code in the DCP
7	Wish	Make the DCP more sustainable and reusable by using only one tool
8	Wish	Consider the switch to Hub in the near future
9	Wish	Keep using the ETL tool, renovate if needed
10	Idea	Bring software developers and consultants together as a team
11	Idea	Do not convert all data at once, but distribute less relevant data later
12	Idea	Revise parts of the ANVA landscape such as APIs and validations
13	Idea	Demand higher quality input data from clients
14	Idea	Consider Matillion as a replacement of the current ETL tool

Table 13 - Overview experiences ANVA employees

Each experience, need, wish or idea propagates from a problem in the process. Thus, from each statement a problem can be deducted. In section 4.1, these problems are all worked out, but sprawled over a lap of text. To make the problems clearer, an overview is made in table 16, by deducting the problem from the statement in table 15.

	Problem
1	The quality of input (received from client) data is bad
2	The current ETL tool supplier has a lack of continuity
3	ANVA has built its landscape without overthinking the bigger picture
4	Too few agreements about data quality and work hours are made with client
5	Too much help of software engineers is needed for conversion
6	There are too few gates to add code in the DCP
7	The DCP is not sustainable
8	Soon also conversions to Hub must be realized
9	There are different conversion options while there should only be one
10	Software developers and consultants are not working as a team during the DCP
11	XML conversions are slow because the XML module cannot handle big conversions
12	The ANVA landscape is non-optimal environment that needs an update
13	The current data quality responsibility of ANVA costs time and money
14	The current ETL tool has its limitations in numerous views

Table 14 - Employee problems

5.3 Clients

It is also important to let clients who have done conversions with ANVA describe their experience with the DCP. This was done in an open interview with the following core topics:

- The client's general experience with the DCP
- Strong and weak points of the DCP
- Ideas or suggestions to improve the DCP in any way

The extensive questionnaire used for the interviews can be found in appendix D.

5.3.1 Company X

The first client is company X. Company X is a big client of ANVA and have experienced many conversions with ANVA. They also use the ETL tool for internal conversions. They have experienced external conversions (of category 2 and 3) and internal conversions (of category 1), with all types of conversion methods. On top of that, their daily use of the ETL tool can generate some helpful insights. The experiences, ideas, wishes or needs of company X are:

- In the ETL tool, it is inconvenient that every time a source file is uploaded, and you try to edit the coverages, you first need you make a file with all unique policy numbers. If you do not do this, the policies will not be on one line, but on multiple lines. This messes up the conversion and thus making a unique policy number file is necessary. It would be great if the tool automatically puts one policy with all coverages on one line.

- The X11755WP program is a great program for sending data from external locations into ANVA. The application is easy to use and can send data at a high speed. It is easy to see what the program is actually doing and what is does to the data. The program also generates a file during the process, which makes it possible to do a manual check on the data. This all makes it a handy connection between the output of the ETL tool and the database of ANVA.

- The XML output of the ETL tool is not used. This is because the system is unusable when data is sent towards ANVA. Because the XML module is very busy with sending relations one by one to ANVA plus a lot of validations, the system becomes so slow it is not usable anymore. Since a lot of conversions are run, conversions cannot only be ran in weekends. So, the system would become unusable during working hours, which should be avoided. Next to this, the program does not generate a file during the process. So, the data cannot be checked manually, which is a need of company X. Therefore, the XML output of the ETL is not used.

- A missing function in ANVA, and a profound desire, is the function to directly compare two screens in ANVA. The function for example makes an Excel sheet where the old and new screens are compared. When selected the information needed, the sheet shows all information of the two systems next to each other, so that similarities and differences can be recognized immediately. This would save a lot of time in setting up the conversion.

- Another desired function is the possibility to filter on policy level in ANVA. Now, policy conditions are checked randomly because it is unknown which policy condition numbers are filled. For example, there are thirty numbers which could include a condition, but in a certain conversion the numbers 1, 2, 3, 4, 5, 6, 7, 8, and 28 contain a condition. Now only numbers 1 till 8 are checked on a condition, since most of the time there are no conditions after 8. However, this leaves number 28 unconsidered. It would be convenient if it was possible to

only display the numbers which actually hold a condition. This would save time in setting up the condition and would increase the quality of data.

- The ETL tool can only convert a certain number of data types. For example, it cannot do damages and invoices. However, our clients, mostly large proxy intermediaries, do want to have damages and invoices converted. As a result of that, we have to tell our clients those things are not possible to convert, which can be a setback to them. An increase in conversion possibilities would make the tool better.

- We are missing a clear step-by-step plan for conversions from ANVA to ANVA. Since a lot of tools are needed for a conversion, it can be hard to find out what has to be done in what order. It would be great if ANVA makes a step-by-step plan which describes the exact sequence of actions for every type of conversion. Now consultants must be contacted every time an action is not clear. This would save time and make the conversion process easier.

- A very time-consuming property of ANVA is the existence of unlabeled policies. This is caused by the difference in ANVA systems of different companies. When two ANVA databases need to be merged, the converter finds a great difference in how the two databases are structured. Companies can all makes their own labels, which creates a totally different set of labels. It is very time-consuming to make all labels match with each other. ANVA could improve this by tolerating less customization in ANVA systems.

The experiences of company X are summarized in table 11.

	Experience, idea, wish or need
1	The ETL tool should automatically put one policy with all coverages on one line
2	The X11755WP program is a great program for sending data from external locations into ANVA
3	XML is not used because it makes the system slow during working hours
4	A missing function in ANVA is to directly compare two screens
5	A missing function in ANVA is to filter policy files on filled or not
6	The ETL tool is too limited in what it can convert
7	A clear step-by-step plan for conversions from ANVA to ANVA is missing
8	ANVA should tolerate less customization to make label matching easier

Table 15 - Summary company X

5.3.2 Company Y

The second client is company Y. Company Y is a big client of ANVA and need many conversions per year as a result of takeovers. Whereas earlier they made use of multiple services of ANVA, including conversions done by consultants, they now only use the online infrastructure of ANVA. This is due to bad experiences with conversion done by ANVA. As a client who no longer makes use of the conversion services, it is very interesting to hear what made them make that decision. The experiences, ideas, wishes or needs of Company Y are as follows.

- The tools used for conversions are too limited in flexibility. There are too few functionalities to select, move or merge certain parts of data. This causes extra work, since little extensions must be built by hand to make things work. During the conversion such inconveniences add up, which causes a more laborious and less satisfying conversion process.

- The information given by ANVA at the beginning of a conversion is insufficient. The planning given is a standardized planning used for multiple conversions and thus not custom made for one conversion. This is not satisfactory, since it is often unclear how the conversion will precisely look like day by day or week by week. This results in an inefficient back and forth interaction between us and a consultant.

- Next to this, the expected possibilities of the conversion are not communicated well. For us, it is not clear what is possible for what conversion. For example, with a certain conversion no damages and invoices can be converted. This becomes clear during the conversion and is not being communicated with us. We then have to disappoint or client and say their damages and invoices will be lost after converting. This lack of communication causes dissatisfaction among our clients and inconveniences for us.

- The level of support during the conversion process is not adequate. Especially at the start of the cooperation with ANVA, we had no experience with conversions and did not know how it worked. We could have used more support and guidance during the process. A to-do list would have helped with going through the process and improving the quality of the conversion. Also thinking along a bit more would prevent errors and better the process.

- We would like to do conversions ourselves in the future. For that, it would be great if ANVA made an ETL tool available which we could use. That tool must be able to convert damages and invoices, as that is not available right now. An easy-to-use tool with these functionalities would make us independent and satisfied.

- The conversion tool needs more validations. This would save a lot of time correcting errors. If the program would warn the user for trying to convert an empty field, the user will win a lot of time correcting that error. It would be great if it was possible to select a certain group of required fields which all must be non-empty. Before running the conversion, the program then gives the list of fields empty. This is an easy way of detecting and solving errors in the data.

The experiences of company Y are summarized in table 12.

	Experience, idea, wish or need
1	The tools used for conversions are too limited in flexibility.
2	The information given by ANVA at the beginning of a conversion is insufficient
3	The expected possibilities of the conversion are not communicated well
4	The level of support during the conversion process is not adequate
5	It would be great if ANVA made an ETL tool available for us
6	ANVA should build in more validations

Table 16 - Summary company Y

5.4 Client problems

As the last part of the problem identification, we identify problems based on the experiences of clients of ANVA. Summarizing table 11 and 12, we get an overview of al experiences, ideas, wishes or needs from clients. Overlapping statements are merged to one more general statement. The overview can be seen in table 17.

	Type	Statement
1	Experience	The X1175WP program is a great program for sending data from external locations into ANVA
2	Experience	XML is not used because it makes the system slow during working hours
3	Experience	The conversion tool is too limited in functionalities and flexibility
4	Experience	ANVA does not communicate well during and before conversions
5	Need	We need a clear step-by-step plan for every type of conversion
6	Need	More validations are needed to improve the conversion
7	Wish	It would be great if ANVA let us use the ETL tool for doing conversions ourselves
8	Wish	ANVA misses certain functions which would make it easier for us to convert

Table 17 - Overview experiences clients

Each experience, need, wish or idea propagates from a problem in the process. Thus, from each statement a problem can be deducted. In paragraph 4.2, these problems are all worked out, but sprawled over a lap of text. To make the problems clearer, an overview is made in table 18, by deducting the problem from the statement in table 17.

	Problem
1	XML is unusable because it makes the system slow during working hours
2	The conversion tool is too limited in functionalities and flexibility
3	ANVA does not communicate well during and before conversion
4	The exact procedure for each type of conversion is unclear for us
5	The conversion tool lacks validations
6	We do not have a conversion tool to do conversions ourselves
7	AVNA misses certain functions which make the conversion process more efficient

Table 18 - Client problems

We now have gathered problems using the analysis and the results of the interviews. In tables 14 and 18, all identified DCP problems are listed. In the next chapter, solutions to these problems will be discussed.

6 Solutions

In this chapter, solutions to the found problems will be discussed. In section 6.1, requirements to which the solution must comply with are listed. In section 6.2, several methods are used to find optimal solutions based on the analyses and the solution requirements. It is also discussed how these solutions can be brought into practice. 6.3 contains the summary of 6.2, in the form of a problem-solution couples table.

6.1 Requirements

At this point, the DCP is analyzed extensively to identify weak points and bottlenecks. Now, solutions can be found to solve the problems. This solution does not only have to solve the problems, but it must also comply with other requirements. These requirements will be discussed in this section.

According to the official bachelor assignment given by the chapter lead of consultancy, the solution must comply with the following requirements:

- The solution must be in line with the wishes and needs from ANVA. The solution should be as much as possible in the interest of ANVA. Coming up with a solution, these requirements, experiences of employees and improvements from employees should be taken into account.
- The solution must be in line with the wishes and needs from clients. This mainly implies that the experience and feedback of clients has been taken into account when finding the solution.
- ANVA wants to convert with the ETL tool in the future, unless that turn out to be the worst choice. Therefore, the solution should be based around the ETL tool. This tool should be easy to use, such that it can be used by functional consultants and clients, reducing the amount of software engineers needed in the process.
- In line with this, in the end the tool should ensure a fast and easy conversion process. To reach that, components in the tool need to be reusable.
- ANVA is currently moving over to new applications. Currently conversions are only done towards ANVA4/5. Shortly, conversions should also be able to reach ANVA Hub. The solution should consider that these new target applications must be reached in the near future. Concretely, the first conversions to ANVA Hub are planned on July 1st this year, which means that small conversions must be possible by then. Therefore, any modifications or additions in the DCP must consider the need of converting to ANVA Hub within a couple of months.
- Conversions must have audit trail as standard.

6.2 Solutions

Now that the solution requirements are clear, solutions to the problems can be found. Solutions are found by using a variation of methods, as described in section 2.3.2.

First solutions to the problems are discussed, then a list of all problem-solution couples is made in table 19. Next to the solution discussion, also trade-offs are elaborated, if multiple solution directions are possible.

6.2.1 DCP structure

The first issue is the unclear structure of the DCP. This problem is mainly caused by having multiple conversion methods and tools. There are three conversion methods, all having their own conversion path. This is makes the process inefficient and unclear. According to Weske, the processes which have do not have value in the business process model, in this case the DCP model, should be eliminated (Weske, 2012). As there are three conversion methods running parallelly, always two of them are valueless.

Therefore, the number of conversion methods should be reduced to one. As has been come clear in the interviews, reducing the number of conversion methods to one will make the DCP more clear. So, it can be concluded that only one conversion method should be used from now on. Since a requirement of ANVA is to do conversions without software engineers, ANVA should only convert with the ETL tool and get rid of manual and custom Cobol conversions. Using an ETL tool, everyone can do conversions after a small introduction or training.

6.2.2 ETL tool

Before this can be implemented, the ETL tool will need some innovations. First, the lack of continuity of the ETL tool can be a problem for ANVA, since a long or permanent absence of just one man results in an ETL tool with no administrator. On top of this, the ETL tool is missing functionalities and needs several innovations to make it appropriate for being the only conversion method. With the outcome of the problem analysis which is based on the process analysis and interviews with ANVA employees and clients, a requirement list can be made for the future ETL tool of ANVA. The ETL tool must:

- Be able to convert any wished type of data, even things such as invoices and damages
- Be able to convert large amounts of data at a fast rate
- Have an easy to use interface, so that no software engineers are needed for conversions
- Be safe regarding the continuity of the owner
- Have the possibility to add functions or smart tricks as wished

For the sake of complying with these requirements, ANVA could consider switching to another ETL tool. When switching to another ETL tool, the problem of continuity is solved and more functions become available. The downside is that ANVA has less saying in what happens with the tool and what additional functions can be added. There is also a chance that employees and clients do not like the new tool, which is a risk to take. Switching to another ETL tool has both upsides and downsides and should be investigated thoroughly before executing. The list of requirements can help with comparing available ETL tools on the market. Here the decision depends on the criteria weights of ANVA and the supply of the market.

The decision on the ETL tool can be impactful. Multiple factors play a role in the decision. There is a tension field between the time available, the resources and the quality. ANVA has certain deadlines which must be respected. At the same time the tool ages which brings down quality and increases the need for new software. ANVA can choose to hold old software or switch to new software, which can be risky. In this case the old ETL tool could stay or a new ETL tool could be used. ANVA must weigh the advantages against the risks, which depends on the interests they have.

If ANVA does not want to make the risky move, they can remain using the current ETL tool. However, the ETL tool will needs improvements to make it compatible for larger conversions and to improve the user experience. According to experiences of users, there are missing functionalities which cause an incomplete conversion. For example, invoices or damages cannot be converted, which leaves them in the old database. If a client wants to have invoices or damages converted, it is not possible and they have to do it without.

Next to this, the ETL tool misses non-crucial, but handy functions which improve the user experience. These functions are:

- Automatically putting one policy with all coverages on one line
- Possibility to directly compare two screens
- Possibility to filter policies on filled or not

Employees and especially clients miss these functions and would like to see them added to the tool. The functions ensure a faster and more satisfying conversion process, improving the general experience with the DCP. ANVA should consider adding these functions, even if it costs a fair amount of money. The costs, possibilities and feasibility depend on the tool owner, but there will surely be possibilities with any tool chosen. An alternative is that software programmers of ANVA can be asked to build the desired functions in the tool, if that is possible. For that, a good conversation with the software programmers will give sufficient insights in the possibilities.

Furthermore, the two output options of the ETL tool have constraints. The largest downside of XML output is the slow speed with which it can be loaded in ANVA. This is because the XML module of ANVA, e-ABS, is made for sending one relation at the time to ANVA. If the module must send thousands of relations, it cannot handle the numbers and will take a long time to send the data to ANVA. A solution for this problem can maybe be found in the way of sending day. Relations are now sent sequential to the module, resulting in a sequential processing process. If relations were sent parallelly to the module, for example in batches of ten, the module would process ten relations at the time, making the processing ten times as fast. Hereby, it is assumed that the module can handle multiple relations at the time and that the module does not become ten times as slow when processing ten relations at the time. Both assumptions are fairly reasonable, making this an interesting possible solution to the slow XML speed.

Another possible solution which does not solve, but bypasses the XML speed problem, is to convert in parts. By converting only relevant and necessary data first and less relevant data at a later time, the processing speed of XML will be less of a bottleneck in the process, since less processing power is need. According to interviewees, this is possible since clients often do not need all data directly after the conversion. For example, invoices of five years ago are hardly needed and if so there is often no rush. This is especially the case for smaller agents, who only need information about the relations and its policies. Applying a conversion in parts can therefore make a difference in process speed, without harming interests of clients.

6.2.3 Data responsibility

To solve the problem of bad data quality, there are some solutions which could be considered. Instead of taking full responsibility for the data quality, ANVA could put more responsibility on clients. Then, the principle of quality in = quality out holds. However, if a client does not care about data quality and delivers bad data, ANVA still has to deal with that bad data quality. The bad quality will cause errors in the ETL tool, which delays the process. It costs extra time to do a conversion with bad input data quality, regardless of the output data quality. So, just communicating the quality in = quality out principle may not work.

Therefore, ANVA could consider demanding it from clients, by contractually define a minimum data quality level. This would force clients to deliver a certain level of data quality. A higher level of data quality would make the process more efficient and less time consuming. However, demanding a certain level of data quality can be unjustified in the face of small clients with no knowledge of data managing. Considering ANVA Hub will be targeted on small agents in the first years, the demands of small agents should be respected. Small agents will probably find it too much to ask, resulting in them finding another company for doing their conversion. That is why ANVA should carefully consider this solution, especially for smaller clients.

A compromised solution to this issue is the use of an error list. Clients deliver data to ANVA, then ANVA runs a check on errors in the documents. The program automatically makes a list of errors, after which the clients can solve all errors in a small amount of time. In this situation, ANVA hardly loses any time spent on data quality, while also the client does not have to spend a lot of time. The time and responsibility are divided over the client and ANVA. A disadvantage is that the error checking program must be made. Overall, the solution has potential to be good compromise.

If ANVA decides to (partly) move responsibility to the client, the DFD will change. The process “improve data format and structure” is now in the lane of the client, instead of the ANVA lane. The data flow diagram with input data responsibility change can be seen in appendix C.

6.2.4 Communication

As for the communicative side of the conversion process, there are some flaws. Clients name several bad points regarding the communication before and during the process. However, once realizing the existence of the problem, it is easy to solve.

First, too few agreements about data quality and work hours are made with the client. The lack of agreements on data quality impairs both ANVA and the client. The client does not know what level of quality to deliver at ANVA and what level of quality to expect after the conversion. If the client is not satisfied with the quality, ANVA must improve the quality with free hours. To improve this, ANVA should make more precise agreements on the quality of data and the work hours needed to reach that quality. That will prevent surprises for the client and free work hours for ANVA.

Second, the conversion planning given to clients is not sufficient. The planning is very general and is used for multiple conversions. The client cannot rely on the planning, since it is not made specifically for the conversion. To improve this, ANVA should invest more time in making a detailed per conversion. This may cost extra time at the start of the process, but it will likely pay back later in the process.

Lastly, ANVA provides too little information about the properties of a certain conversion. For every type of conversion, there is a unique set of possibilities, which is not communicated with the client. In order to improve the satisfaction of the client, ANVA should provide a document with all possibilities for every conversion. Next to that, not enough information about the procedure per type of conversion is provided. Especially when new, clients do not know how conversion work and what they have to do. A to-do list per type of conversion would have clients going through the process. Providing more support and guidance, along with thinking along a bit more would prevent errors and better the process

6.2.5 Policy

Some problems are caused by a non-optimal policy. Certain decisions have led to difficulties in the DCP structure, the ANVA landscape and cooperation during the process. These problems can (partly) be solved by improving the general policy or executing some simple steps.

The DCP will become clearer and more sustainable by only using an ETL tool in the future. The digital landscape will also become less confusing over time by switching to ANVA Hub. For these problems, big decisions are needed. The decision to switch to Hub is already made and will help to improve the landscape. To improve the DCP as well, the decision to only use an ETL tool will be essential, according to the analysis.

As experienced by employees, there is unnecessary back and forth motion between software engineers and consultants. By working smarter, it is possible to reduce the inefficient communication. The complete conversion, including details, should be thought out before any software building is done. Next to that, a detailed time plan of the process should be made. This will reduce back and forth motion to the minimum, as both the software engineer and consultant will not face surprises during the process. A strong preparation improves the efficiency of the process.

To further increase the level of motivation and cooperation among employees, it is helpful to have all employees facing the same way. The current difference in opinions and experiences can cause several problems. It is therefore important to make sure all employees agree on the roadmap for the future. This could be done by a collective session, where employees discuss the future together. The outcome should be a compromise which is accepted by everyone.

6.3 Problem-solution couples

Now all problems and their possible solutions have been discussed. The solutions have been determined based on the scope of the research and the requirements given. Other considerable solutions have been left out, since they did not fit into the scope of the research or comply with the requirements given. The problem-solution couples are summarized in table 19.

Problem	Solution(s)
Unclear DCP structure	Only use ETL tool
Lack of continuity ETL tool	Switch to new ETL tool
Software engineers are needed for conversion	Only use ETL tool
ETL tool misses functions	Switch to new ETL tool Build new functions
XML module is slow	Send data parallelly instead of sequentially Convert in parts
Input data quality is bad	Hold and communicate quality in = quality out principle Contractually define minimum data quality Use error list program
Too few agreements are made	Make more agreements
Planning of process given to clients is not sufficient	Make a unique planning which suits the conversion
Too little information about conversion types is provided	Provide a document with possibilities and to-do list per type of conversion
ANVA landscape is messy	Eventually switch to Hub
The DCP is not sustainable	Only use ETL tool
Staff not working together optimally	Together extensively think out process before starting
Employees are not on the same page	Organize collective discussion session

Table 19 - Problem-solution couples

7 Conclusions and recommendations

This research was aimed at formulating an answer to the question: 'What are main problems in the conversion process of ANVA and how could these problems be solved?' To do that, both a technical and user experience analysis was conducted. The technical analysis was used mainly for identifying problems on a process model level, while the interviews were used mainly for identifying problems not directly visible from the surface.

The results showed that there are several problems at different stages of the process. Problems were found in the structure of the process, the ETL tool, the data responsibility policy, the communication with the client and the overall policy. For every problem, one or multiple solution directions were found.

The unclear structure of the process is caused by the use of multiple conversion methods. The best way to make the process clear is by reducing the number of conversion methods to one. The ETL tool finds its downsides in the continuity of the tool and the limited possibilities it has in some stages of the conversion. To improve that, ANVA could either find a new tool which has a better continuity and more functions or improve the current tool by adding new functions. The choice depends on the exact wishes of ANVA and the amount of risk taken by ANVA.

ANVA's current data responsibility policy is indecisive, causing inconveniences with clients. More agreements on a contractual level are needed to improve the situation for both clients and ANVA. ANVA could decide to take full responsibility for the quality of data, give clients full responsibility or compromise by using an error list.

Clients are not satisfied with the provided information before and throughout the process. ANVA should improve their conversion planning and provide more information about the possibilities and procedure per type of conversion. Furthermore, a non-optimal policy causes the DCP to remain unsustainable and unclear. The efficiency of the process and employees could be improved by acknowledging the problems and basing decisions on that.

All problems are technically or practically possible to solve. They are therefore a cause of certain policy and choices. By understanding and acknowledging the problems, small changes in the policy could make a huge difference in the DCP. ANVA should therefore reconsider their process as a whole while taking into account the areas of concern mentioned in this report. Based on their interests regarding the found problems, they should think about restructuring parts of the process or taking a different approach in certain areas. It is recommended to reduce the amount of options and thus simplify the process, provide more information to users of the DCP and rethink the working methods used.

Furthermore, ANVA is recommended to find out which of the identified problems are most important to them and invest time into solving those problems. The solution directions and trade-offs described could help them to right direction. Since this thesis only treated problems superficially, it is recommended to further dive into the most important problems. This is a suggested topic for future research.

Additionally, difficulties were found in transferring data from the conversion tool into ANVA. The different systems used for that purpose, for example e-ABS, cannot do the job satisfactory. Since it did not fit into the scope of this research, it was not included in this report. However, for future research it is an interesting topic to dive into, since solving that problem will be very helpful for ANVA.

References

Antonov, E., Lopatina, E., Ionkina, K., & Tretyakov, E. (2020). Agent data merging. *Procedia Computer Science*, *169*, 473–478. https://doi.org/10.1016/j.procs.2020.02.222

ANVA. (2021). *Over ons*. https://anva.nl/over-ons/

ANVA. (2019a). *Services*. https://anva.nl/services/

ANVA. (2019b). *Conversies*. https://anva.nl/services/conversies/

ANVA. (2018). *Een overname biedt zoveel meer mogelijkheden*. https://anva.nl/nieuws/een-overname-biedt-zoveel-meer-mogelijkheden/

Aryanezhad, M., Badri, S., & Rashidi Komijan, A. (2009). Threshold-based method for elevating the system's constraint under theory of constraints. *International Journal of Production Research*, *48*(17), 5075–5087. https://doi.org/10.1080/00207540903059505

Code of Ethics | National Society of Professional Engineers. (2012). NSPE. https://www.nspe.org/resources/ethics/code-ethics

Chen, L., Ali Babar, M., & Nuseibeh, B. (2013). Characterizing Architecturally Significant Requirements. *IEEE Software*, *30*(2), 38–45. https://doi.org/10.1109/ms.2012.174

Heerkens, H., & Van Winden, A. (2021). *Solving Managerial Problems Systematically*. Taylor & Francis.

Hızal, A. (2021). Frequency domain data merging in operational modal analysis based on least squares approach. *Measurement*, *170*, 108742. https://doi.org/10.1016/j.measurement.2020.108742

Inman, R. A., Lair Sale, M., & Green, K. W. (2009). Analysis of the relationships among TOC use, TOC outcomes, and organizational performance. *International Journal of Operations & Production Management*, *29*(4), 341–356. https://doi.org/10.1108/01443570910945819

Jain, S., Jain, V., & Balas, V. E. (2021). *Web Semantics: Cutting Edge and Future Directions in Healthcare* (1st ed.). Academic Press.

KNAW. (2018). *Netherlands Code of Conduct for Research Integrity*. https://www.knaw.nl/shared/resources/actueel/bestanden/netherlands-code-of-conduct-for-research-integrity-2018-uk

Ladyman, J. (2001). Understanding Philosophy of Science. Routledge.

Michalewicz, Z., & Fogel, D. B. (2004). *How to Solve It* (2de ed.). Springer Medizin Verlag.

Pacheco, D. A. D. J., Antunes Junior, J. A. V., & de Matos, C. A. (2021). The constraints of theory: What is the impact of the Theory of Constraints on Operations Strategy? *International Journal of Production Economics*, *235*, 107955. https://doi.org/10.1016/j.ijpe.2020.107955

Principles of research ethics | Lærd Dissertation. (2012). Laerd Dissertation. https://dissertation.laerd.com/principles-of-research-ethics.php#fifth

Reeve, A. (2013). *Managing Data in Motion: Data Integration Best Practice Techniques and Technologies (The Morgan Kaufmann Series on Business Intelligence)* (1st ed.). Morgan Kaufmann.

Van der Poel, I., Royakkers, L., & Pieters, A. (2004). *Ethiek & techniek : morele overwegingen in de ingenieurspraktijk* (1st ed.). ThiemeMeulenhoff.

Weske, M. (2012). Business Process Management: Concepts, Languages, Architectures (2nd ed. 2012 ed.). Springer.

Wortell. (2020). *Trendrapport 2020 voor verzekeraars*. https://www.wortell.nl/media/Trendrapport-2020-voor-Verzekeraars.pdf

Appendix A

In order to find the right definition of a data conversion, we perform a systematic literature review. To optimally search for definitions of data conversion, we define inclusion and exclusion criteria. Inclusion criteria are criteria that a study must have to be included in the review. Exclusion criteria are criteria that make a study ineligible to be used in the review.

We first define the inclusion criteria:

- Time frame: a study should not be more than ten years old, in order to be relevant for the review. Studies of more than ten years old will not give an up-to-date definition of data conversion. The technology described in these studies will be outdated in comparison with current technology. That is why the data conversion of then may be totally different than the data conversions of today.
- Study design: the study should be an explanatory research, so it will have the goal to explain something, in this case the concept data conversion.
- Language: the study should be in English or Dutch, in order for it to be understandable for the reviewer.

Now we define the exclusion criteria:

- Research population: if a study interviews people, these people must not be random people. If interviews are held with people other than experts of the field, the study is not eligible for the review. This only holds when the topic of the interview is about data conversions. This criterium ensures that a professional scientific definition is given in the article.
- Accessibility: if the source is limitedly accessible and is thus only accessible after paying, the source is not suitable for this review. For the review, there is no budget available.

For the review, two databases will be used, which are Web of Science and ScienceDirect. These databases are mainly chosen because of their focus on science and thus more towards technology. I think I will find the most valuable studies in these databases. Next to this, I use Web of Science because of the sheer size, there are many studies to be found. For ScienceDirect, another reason to use it is the familiarity I have with it. I used ScienceDirect a lot in the past, which makes it for me a relied database to work with.

My first search was on Web of Science. I think it is handy to search for matching words in the title, because that will give the best results. So, I used the search terms ‘data conversion’, with the option to only search in titles enabled. Next to this, I selected the time frame 2010-2021, to only display articles that are not outdated. This gave a small 500 hits, which is not bad. Some articles are seem quite good to use based of the title, but it could be better. Most hits namely do have ‘data’ and ‘conversion’ in the title, but the words are separated from each other. This means that the article is about something with data and something with conversions, but not with data conversions.

Therefore, I now search with the same strategy, but adding the NEAR feature. The new search looks like: data NEAR/1 conversion, with the same bounds as last search. This gave me 136 results, which is a good number. Next to this, I now only have articles with ‘data conversion’ or ‘conversion of data’ in the title, which can all be relevant for this review, in contrast with earlier results. Then, I replaced ‘conversion’ with ‘merg*’, to get new results but with the same topic. Using that synonym generated 152, which is almost the same amount.

The last search refinement on Web of Science included a combination of previous searches. As the search term, I used both ‘data conversion’ and ‘data merg*’, to include the most articles possible. I

used the same time frame, 2010-2021. To reduce the amount of hits significantly, the hits were filtered on category. I only let in several topics within the category Computer Science. This resulted in 25 relevant hits.

The second search was done on ScienceDirect. With the same reason as before, I am only searching in titles. Also, the time frame 2010-2021 was selected again, to only get recent and relevant articles. To get different results, I use the search term 'data merge' in stead of 'data conversion'. While it means the same, I will now get different results. The search generated 97 results, which is already a good amount. To lower this amount of hits even more, I also filter on subject areas. The subject areas that seem the most relevant to the topic are Engineering and Computer Science. After applying this filter, 22 hits were found. The different searches are listed below in table 1, the search log.

Database	Search terms and parameters	Amount of hits
Web of Science	TITLE: (data conversion) Timespan: 2010-2021	492
Web of Science	TITLE: (data NEAR/1 conversion) Timespan: 2010-2021.	136
Web of Science	TITLE: (data NEAR/1 merg*) Timespan: 2010-2021.	152
Web of Science	TITLE: (data NEAR/1 (merg* OR conversion) Timespan: 2010-2021. Categories: Computer Science: theory methods, hardware architecture, software engineering, information systems and interdisciplinary applications	25
ScienceDirect	Title: "data merge" Years: 2010-2021	97
ScienceDirect	Title: "data merge" Years: 2010-2021 Subject areas: Computer Science and Engineering	22

Table 20 - Search log

Now that only a few dozen relevant articles are left, the searches are complete. The most relevant articles will be tested on their compliance with the inclusion and exclusion criteria. If that test is passed, the article can be added to the conceptual matrix. The conceptual matrix lists four articles and shows which concepts are discussed in which article. The conceptual matrix can be seen below in table 2.

Articles	Concepts			
	Definition data conversion	Data conversion methods	Essence of data conversion	Data conversion analysis/reflection
Antonov et al., 2020			**x**	**x**
Hızal, 2021		**x**		
Jain et al., 2021		**x**	**x**	
Reeve, 2013	**x**			**x**

Table 21 - Conceptual matrix

As can be seen above, the conceptual matrix categorizes four sources based on what concept is processed in the articles. The concepts are four important parts of the topic data conversion and together form a broad view of what data conversion is. It is interesting that no source covers all four of the concepts. However, with a combination of the sources, all concepts can be covered.

The systematic literature review is now completed. Four helpful sources have been found to support the theoretical framework. The sources are used together in section 3.1, to help defining the concept data conversion.

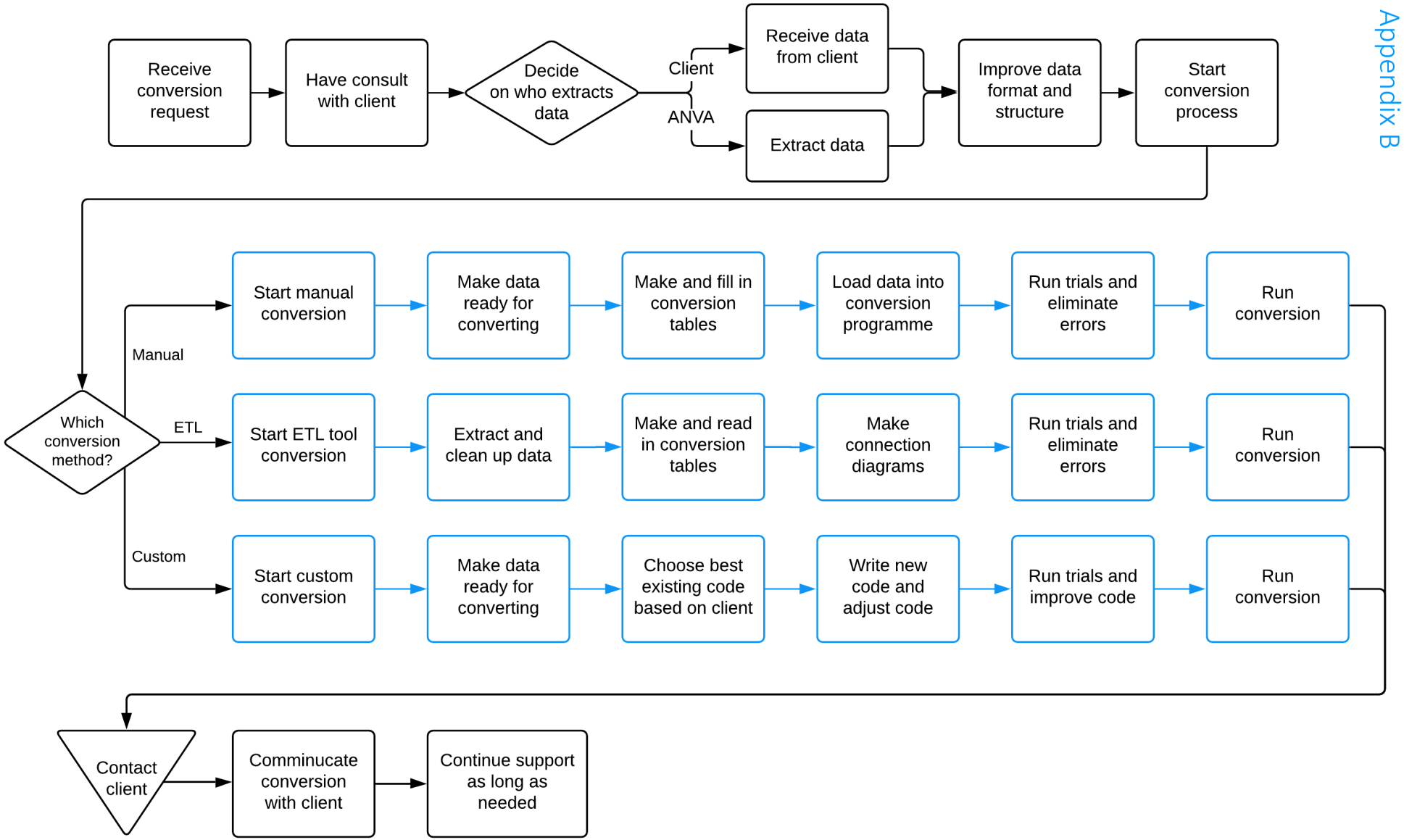

Figure 8 - Data conversion process

Appendix C

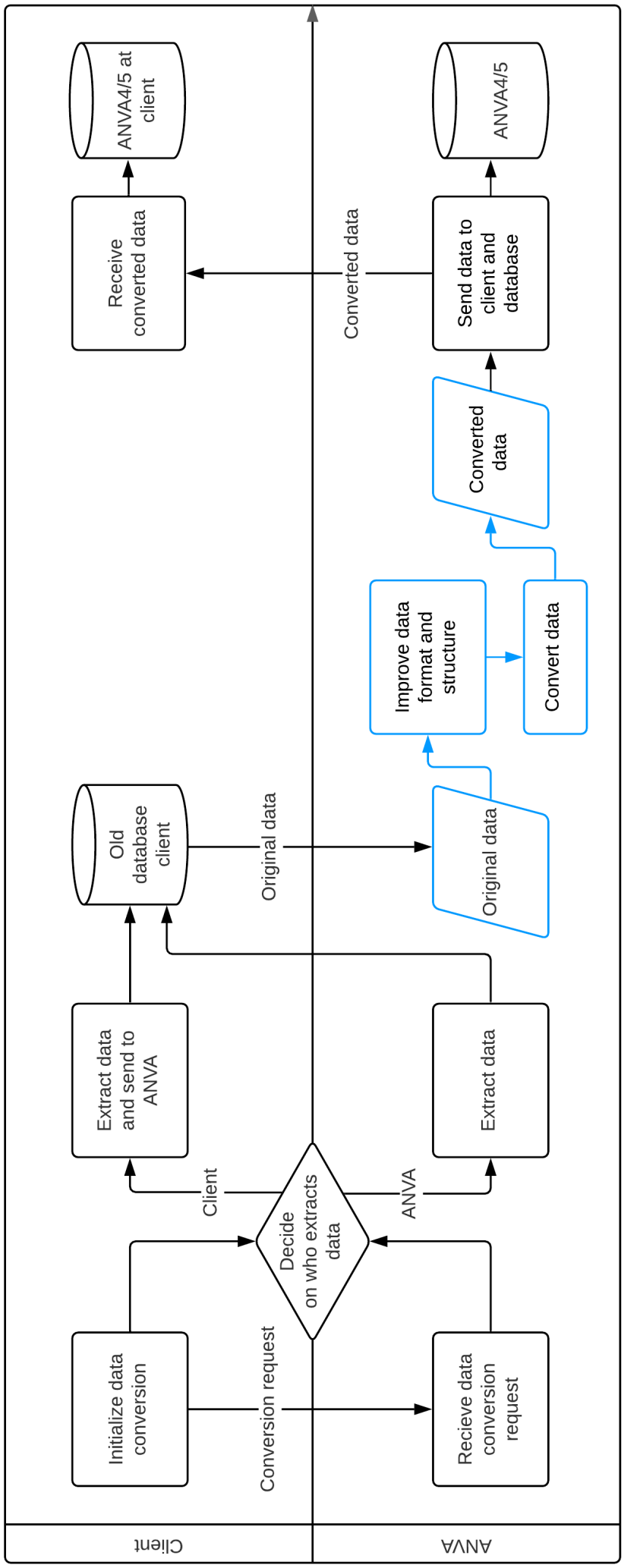

Figure 9 - Data flow diagram

Appendix D

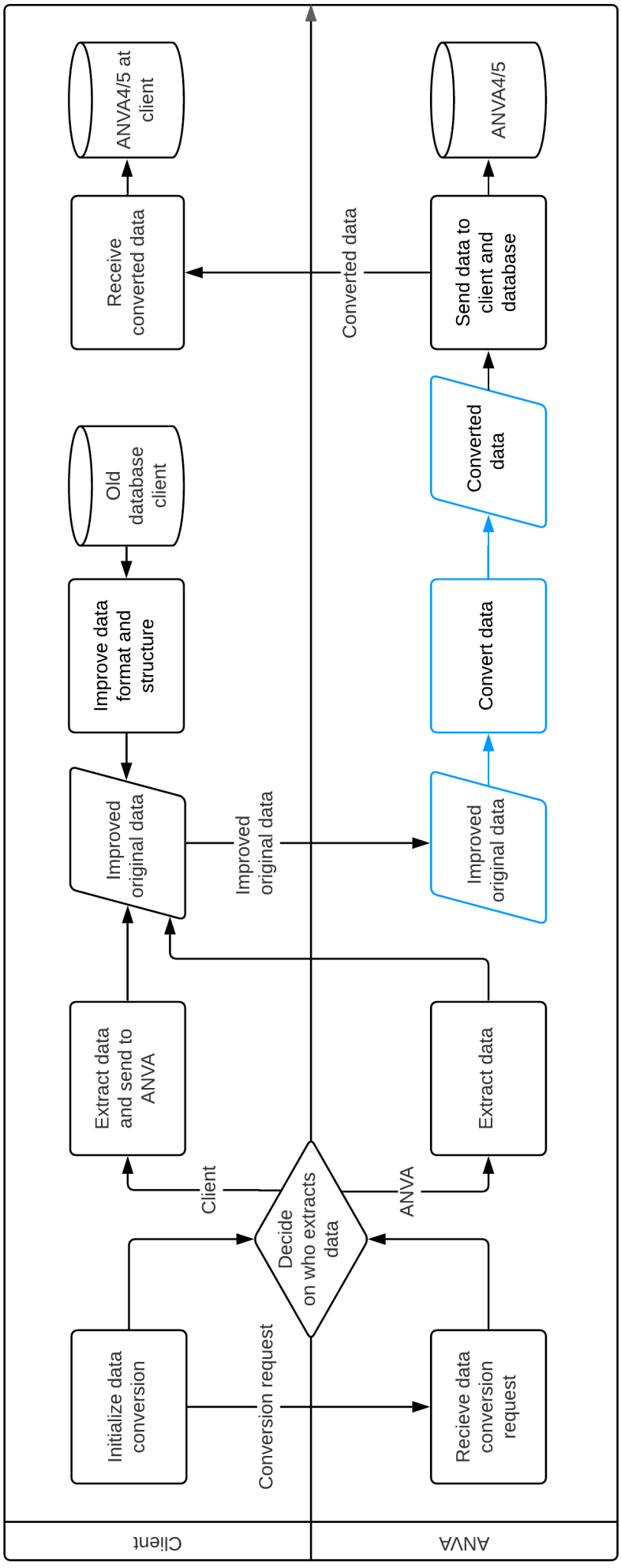

Figure 10 - Data flow diagram with input data responsibility change

Appendix E

Questionnaire used for interviews with client X and Y

1. What is your general experience with the DCP?
2. What are strong points of the DCP?
3. What are weak points of the DCP?
4. How could the DCP be improved?
5. Is the current task division optimal?
6. If not, what should be changed?
7. What functionalities are you missing in the tools you use?
8. What new functionalities would you like to see?
9. What are general points of improvement technically speaking?
10. Are you satisfied with the communication with ANVA?
11. Does ANVA make (clear) enough agreements with you at the start of the DCP?
12. What are general points of improvement communicationally speaking?