

# Automatically match people with jobs

Enrich personal data provided by people to create better matches



**8vance**  
manage your talent

Author	:	Tim Hermens
Place	:	Blerick
Date	:	02-05-2016
Version	:	0

### Student information

Name: Hermens T.  
Student number: 2210369  
Graduation course: Software Engineering (full-time)  
Graduation period: 15-02-2016 to 24-06-2016

### Company information

Name: 8vance Matching Technologies BV  
Department: Research & Development  
Location: Blerick, Kazernestraat 17

### Company mentor information

Name: Keuren P.  
Function: Neural Software Developer

### School mentor information

Name: Schouten G.

### Thesis information

Title: Automatically match people with jobs  
Date of issue: -  
Confidential: Strictly confidential (no distribution allowed)

### Signed and approved by company mentor

Date: -  
Signature:

## Preface

This thesis contains a description of the project I've executed for the company 8vance Matching Technologies BV. This company is active in the data science area (also known as Big Data) of the IT industry. One of their products is called AIMA, which is a digital agent that is able to match profiles of people with jobs. The main objective of my project is to improve the quality of matches of people with jobs by enriching the profile data of people. This project ran from February until June 2016.

This project has been one of the most challenging projects I've done, mainly because of my limited experience with the data science area. With help and insights from the company's data science experts (Sabrina Ziebarth, Lou Cremers, Paul Keuren, Jan Jacobs) I managed to overcome many challenges, for which I would like to express my gratitude. I want to thank Paul Keuren in particular for his thorough feedback and excellent support. And last but by no means least, I would like to thank Gerard Schouten for his close involvement and continuous support throughout the whole project.

## Table of contents

<b>SUMMARY .....</b>	<b>6</b>
<b>SAMENVATTING .....</b>	<b>7</b>
<b>GLOSSARY .....</b>	<b>8</b>
<b>1 INTRODUCTION .....</b>	<b>9</b>
<b>2 THE COMPANY .....</b>	<b>11</b>
2.1 FOUNDATION AND MISSION.....	11
2.2 PRODUCTS .....	12
2.3 BUSINESS PLAN .....	13
2.4 ORGANISATION .....	15
<b>3 THE ASSIGNMENT .....</b>	<b>17</b>
3.1 PROJECT GOAL .....	17
3.2 SCOPE .....	17
3.3 CONSTRAINTS .....	18
3.4 PROJECT CHANGES .....	18
3.5 RESEARCH.....	18
<b>4 THE APPROACH .....</b>	<b>20</b>
4.1 INTRODUCTION .....	20
4.2 METHODS.....	20
4.3 PROJECT PLANNING .....	23
<b>5 ORIENTATION PHASE.....</b>	<b>25</b>
<b>6 RESEARCH AND SOLUTION PHASE .....</b>	<b>26</b>
<b>7 COMPLETION PHASE.....</b>	<b>27</b>
<b>8 CONCLUSION AND RECOMMENDATIONS.....</b>	<b>28</b>
<b>EVALUATION .....</b>	<b>29</b>
<b>BIBLIOGRAPHY .....</b>	<b>30</b>
<b>APPENDICES .....</b>	<b>31</b>
A. PROJECT INITIATION DOCUMENT (PID).....	31

B. OTHER APPENDICES .....	31
---------------------------	----

## Summary

*(An informative summary with a maximum length of ONE page, summarizing the whole thesis.*

*Describe the company and the problem that needed to be tackled.*

*Describe the chosen approach with arguments.*

*Describe the results.*

*Describe the conclusions and recommendations.*

*DON'Ts:*

*Don't refer to other chapters of the thesis.*

*Don't go too in-depth on details.*

*Don't use difficult terms somebody may not understand. Anyone should be able to read the summary.*

*Don't add images or bulleted lists.)*

## Samenvatting

*(The same rules apply from the previous chapter.)*

## Glossary

### **AIMA (Automatic Intelligent Matching Agent)**

This is one of 8vance's products that's able to match people with jobs.

### **ME (Matching Engine)**

This is one of 8vance's products that can find similarities between profiles and create matches between the most similar profiles. For instance, it's able to match jobs and talents for 8vance's AIMA product, and it's able to match real estates and buyers for 8vance's Domizz product.

### **PID (Project Initiation Document)**

This is a document that contains information about a project and the approach and strategy that will be used to reach the project's goal.

### **SAD (Software Architecture Document)**

This is a document that contains information about an application's software architecture.

### **Talents**

People who are looking for a job.

### **URS (User Requirements Specification)**

This is a document that contains the requirements of the stakeholders for an application.

*(Define unknown words, terms, symbols, abbreviations in alphabetic order. Define abbreviations as follows:*

*<Full term> (<Abbreviation>): <Definition>*

*Explain all the things defined in the glossary in the thesis' text as well when they're first encountered.)*



# 1 Introduction

Imagine you won't ever have to search for a job that exactly matches your skill set and personal desires, but that job will automatically find you instead. It almost sounds too good to be true, but the company 8vance Matching Technologies is convinced this can become a reality. Better yet, they've already developed an early build of this system, which is called AIMA, that's able to automatically match jobs with talents<sup>1</sup> with an impressive accuracy.

8vance Matching Technologies is a relatively young and innovative organisation that's active in the data science (also known as Big Data) area of the IT industry. With AIMA, they aim to render recruitment and career development smart, social, simple, cost-effective and fast. Two other notable products they're working on are Sjerlok and Domizz. Sjerlok is an artificial intelligence system that uses Big Data to track down stolen goods for insurance company Delta Lloyd. Domizz is an online platform for buying and selling real estates, involving an automatic digital real estate agent that automatically finds matches between estates and interested buyers. This project is focussed on AIMA.

All of the company's products are powered by Big Data. However, apart from benefits there're also problems with Big Data. To create matches between jobs and talents, data is collected of the jobs and talents from social networks such as LinkedIn and Xing as well as 8vance's internal database. The data that's acquired of the talents from social networks misses useful information more than 98 percent of the time. The most notable missing information are the talent's owned skills. It's extremely rare that talents have specified all of their owned skills. This is problematic because the matching quality between job and talent is the most dependant on the skills a talent owns. The more skills are missing, the higher the likelihood of a mismatch between job and talent. A mismatch means the failure of AIMA to deliver its promise which leads to unhappy customers. Therefore the focus and the goal of this project lies on complementing the skills of the talent data acquired from social networks to increase the quality of matches.

In order to achieve a qualitative result in this project, a selection of three methods are used. The first one being Agile, which is a work methodology that focuses on incremental development and continuous improvement. Research plays a major role in this project and a majority of the requirements are dependent on the research, which makes Agile a good choice. Furthermore, Fontys' five strategies research method and Tienstappenplan are methods that are used in this project. See chapter 4 to find more information about these methods.

This document contains a description of this project's process, results and key decisions that were made. Information about the company 8vance Matching Technologies can be found in chapter 2. Chapter 3 provides an in-depth description of the project's goal. The approach of this project is addressed in chapter 4. Chapter 5 describes the orientation phase in which the Project Initiation Document (PID) was created and the initial research was executed. The research and solution phase is discussed in chapter 6, including topics such as: the User Requirements Document (URS), the Software Architecture Document (SAD), the creation of the algorithm and the algorithm library. Chapter 7 contains information about the completion phase, including a conclusion and recommendation of the research, the parts of the architecture that still needs to be implemented, and a summary of the user requirements

---

<sup>1</sup> Talents are people who are looking for a job.

that aren't or are partially implemented and thus need attention. And finally, chapter 8 contains the conclusion and recommendations of this project.

*(Start with a nice, catchy phrase that stimulates further reading. Think of interesting questions, a catchy phrase, a comparison/formula, an incredible number, etc.*

*Describe global information of the project. Describe the company, the problem, the assignment and relevance of the assignment. Also briefly mention the approach/strategy.*

*The final paragraph will explain the document's structure and what information can be found in every chapter. You can also explain the structure of individual chapters.)*

## 2 The company

### 2.1 Foundation and mission

8vance Matching Technologies BV is a relatively young and innovative organisation that's active in the data science (also known as Big Data) area of the IT industry. The company was founded at November 2012 with the mission to render recruitment and career development smart, social, simple, cost-effective and fast. There're four main reasons why the company wants to invest to accomplish this mission.

Firstly, unemployment still remains in the top 10 world problems (Hutt, 2016). The partial cause of this problem is the fact that vacancies are spread all over the internet on a variety of websites (e.g. online job boards or a company's website) and the unemployed have trouble finding appropriate vacancies. The company sees an opportunity here to collect as many vacancies and profiles of people as possible from a variety of websites and store them on one central place. Since finding a perfect job (or employee for a company's recruiter) is a challenge in itself, the company wants to provide assistance in the form of an Automatic Intelligent Matching Agent (AIMA) that's able to match people with jobs. This should help to decrease the unemployment rate.

Secondly, a shocking number of 87 percent of the employees aren't happy at work. A Gallup research has revealed that only 13 percent of the employees are happy and engaged in their work. 63 percent are disconnected and not engaged in their work (130 million employees in Europe). 24 percent are even undermining the company and are actively disengaged in their work (50 million employees in Europe). (Crabtree, 2013) 8vance aims to solve this problem by being one of the first to combine soft and hard skills to find jobs that matches someone's skill set and personal values. This leads to the third reason.

Thirdly, the company distinguishes itself from the competition by offering an even smarter, state-of-the-art matching engine. The company will be the first to combine both hard and soft skills to improve the matching quality between people and jobs.

And lastly, the matching engine that'll be developed to accomplish the mission can be used to accomplish a wide variety of incredible and innovative things (yet to be discovered). For instance, it would be possible to create a career assistance agent that's able to offer suggestions of skills you should achieve and/or educations you should follow to make progress in your career development. In other words, there's a lot of value that can be harvested.

Figure 1 also contains six reasons why 8vance wants to create AIMA.



Figure 1 - Reasons why 8vance want to create AIMA

## 2.2 Products

The company is working on several products. The following products are the three main products they're working on:

- **AIMA.** An automatic intelligent matching agent that's able to match talents with jobs. Since the project is focussed on this product, it'll be discussed more in-depth in the upcoming section.
- **Sjerlok.** An artificial intelligence system that uses Big Data to track down stolen goods for insurance company Delta Lloyd. (Vonk, 2015)
- **Domizz.** An online platform for buying and selling real estates. An automatic digital real estate agent finds matches between estates and interested buyers.

Because of the small number of employees and strict deadlines of the AIMA and Sjerlok products, the company had to cease the development of Domizz. However, an ex-CEO of [www.jaap.nl](http://www.jaap.nl) is now working for 8vance and has taken up the responsibility to continue the development of Domizz.

### 2.2.1 AIMA

As said before, AIMA is an automatic intelligent matching agent that's able to match talents with jobs. This service is available on 8vance's website ([www.8vance.com](http://www.8vance.com)) after registering an account for free.

The target audience of this product are companies and talents around the world. The product will mainly be used by the companies' recruiters and the talents.

The general use of the product consists of the following three steps. (8vance Matching Technologies, 2016)

### Step 1 - Registering

Companies can register with 8vance through a online wizard. The wizard prompts the user to insert the required information about the company it'll need for its matching algorithm. A short leadership test is also part of the registration, which improves the results of the matching algorithm. The company can start posting vacancies of jobs when it's registered. The wizard assists step by step in the creation of a full job DNA. This way AIMA gathers all the required information to find better suited talents for the job.

Talents can also register with 8vance through a online wizard. They have the opportunity to either register with their social network profile such as LinkedIn, or register a new account. When registering a new account, the wizard prompts the user to insert the required information about himself it'll need for its matching algorithm. After the registration, talents can start an automatic search to find the best matching jobs for them.

### Step 2 - Matching and scouting

AIMA uses several methods to find talents. When a company activates one of their vacancies, AIMA searches the internet for public CVs to find talents (this is called scouting). She will establish who has worked where for how long and which competences have played an important role. AIMA may automatically add missing competences which are in line with a talent's working experience and educational background to improve the quality of matching (this part is the goal of this project).

In addition to publicly available information, talents can also create their own 8vance profile. Recruiters have the possibility to upload their database of talents and have AIMA match these data as well. AIMA creates a list of matches containing information about the found job matches or talent matches.

### Step 3 - Advertising and social media

In a few steps, companies can create creative and efficient online campaigns for their vacancies, which includes a link to their own home page. Companies can also create advertisements on job sites and social media. Talents who have seen the advertisements and are interested, can register with 8vance so that they're immediately matched to the vacancy in question.

## 2.3 Business plan

Both the companies and the talents benefit greatly from AIMA's service. However, companies will overall have a greater benefit because of the following reasons (Stoffels, 2016):

- AIMA will reduce the costs of the process to find qualified talents up to 70 percent.
- AIMA accelerates the recruitment process up to 75 percent.
- AIMA reaches up to 50 times more candidates by combining profiles of talents from the companies' databases, 8vance's databases and social networks.
- AIMA minimizes the costs of mismatches because of the high matching quality.
- AIMA offers a better overview and transparency of the available talents.

Because of these added benefits for the companies and their financial interests, 8vance offers the companies a reasonable pricing model, which can be seen in Table 1. (Stoffels, 2016) The pricing is based around the number of jobs a company can submit that will be included in the job-talent matching process.

Single post plan	Submit less than 50,000 jobs	€ 300.-
	Submit more than 50,000 jobs	€ 900.-
Volume plan	5x Submit less than 50,000 jobs	€ 1,250.-
	5x Submit more than 50,000 jobs	€ 3,500.-
License plan	6 months unlimited job submissions	€ 600.- per month
	12 months unlimited job submissions	€ 500.- per month
Additional tools	Marketing, advertising and social media	Individual pricing, TBD

Table 1 - AIMA's pricing model for companies

The single post plan means a company can submit the given number of jobs all at once at one specific moment. The volume plan means a company can submit the given number of jobs all at once at five different moments. The license plan allows the company to submit new jobs at any given moment during the license period, meaning they can continuously update their available jobs without any further costs.

AIMA is completely free to use for the talents. However, additional future services like a career assistance agent, a generation of an auto-completed profile variant of the talent's profile, or the generation of smart views of the current market with a filtering of the talent's preferences and skills (see Figure 2), can be paid services.

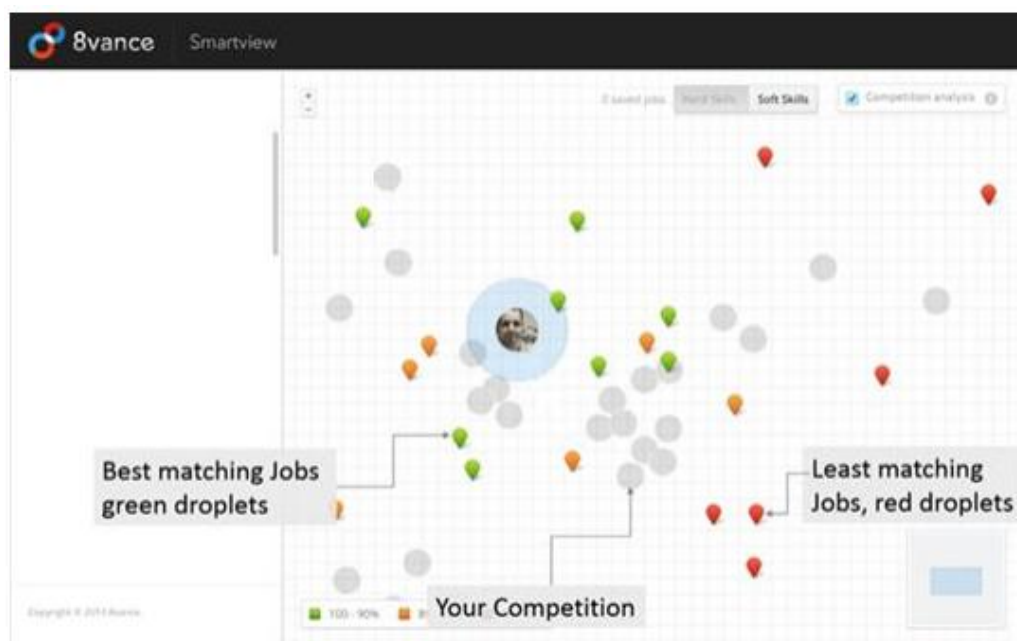


Figure 2 - Illustration of a smart view (job market)



*Figure 3 - 8vance's headquarters in Blerick in The Netherlands*

## 2.4 Organisation

The company currently consists of a total of 15 employees. An organisation chart of this company can be seen in <Figure> below.

<The organisation chart>

The company is situated in the Netherlands and Romania. The software development team is situated in Romania and the data science and marketing teams are situated in the Netherlands. I'm part of the data science team in the Netherlands.

### 2.4.1 Manufactuur

As can be seen in Figure 3, the company's headquarter is situated in the Manufactuur building in Blerick in the Netherlands. The Manufactuur used to be one of the fabric halls of Leolux (a company that specializes in design furniture), but is now an open and transparent work floor for promising start-up companies. The Manufactuur currently houses over five companies, such as Accerion, Yubu and 8vance. (Kickstart Venlo, 2016)

*(A detailed description of the company. When, where and why has it been established? Where are they now? Are there any sister companies? What is their position in the market? How many employees (organisation chart)? What are they building and for whom?*



*Give examples of products they create. Also explain in which section of the company I am working.)*



## 3 The assignment

### 3.1 Project goal

The company is working on a product called AIMA that matches jobs and talents as accurately as possible. Another product called ME (Matching Engine) is responsible for actually creating the matches between profiles. Essentially, AIMA uses the ME to create the matches between jobs and talents.

The ME uses a variety of algorithms to create the matches. The algorithms need to be trained with data initially for them to be able to create matches. The algorithms are trained on company, job and talent data that 1) are scraped from social network websites such as LinkedIn and Xing, and 2) are retrieved from 8vance's internal database. One major problem is that the talent data from the first data source misses useful information more than 98 percent of the time.

The information that's missing varies from talent to talent. The missing information can consist of followed educations, work experience, owned skills, the work industry, dates, and more. The most notable missing data in this list are the owned skills, because the matching quality is the most dependant on the skills a talent owns. The more skills are missing, the higher the likelihood of a mismatch between job and talent. This doesn't necessarily result in the failure of the ME, for it has found the best matches based on the incomplete information. But it does result in the failure of AIMA to find the best matches between talents and jobs, because the lack of skills leads to an incomplete picture of the talent's skill set which severely limits the possibilities of matches. Therefore the focus and the goal of this project lies on complementing the skills of the scraped talent data to increase the quality of matches.

The preferred solution to this problem comes in the form of an algorithm that's able to complement the missing skills for the provided talents. This solution can be used in a stand-alone tool to provide skill suggestions on request, and can be integrated in the scraping process to automatically complement the skills for newly scraped talent data.

### 3.2 Scope

The scope of this project is specified in Table 2, which is directly taken from the appendix Project Initiation Document (PID).

Belongs to the project	Doesn't belong to the project
A project plan.	Creating predictions of skills someone could own in the near future.
A research document where found solutions are discussed.	Initially, the information provided by the talents isn't checked on validity. Only when time allows it, this check may be done.
A solution that complements the missing skills for the provided talents.	The integration of the solution in the scraping tool.

A stand-alone tool in which the solution is integrated that provide skill suggestions and can be used to test the solution.	
A user requirements document.	
A software architecture document.	
Only the missing skills will be complemented for the talents.	

Table 2 - Scope of the project

### 3.3 Constraints

- Because the scraping of the talent data is done in a Python application and the solution must be able to be integrated in this application, means the solution must be compatible with Python.
- The stand-alone tool that's developed must be platform-independent.
- Complementing missing skills for one profile cannot take longer than one second.

### 3.4 Project changes

There's been one change during the project. A software test plan first was also part of the deliverables, but was removed because of two reasons. The first one being there wasn't enough time to create a full-fledged test plan next to all the other, more important deliverables. The second reason is that primarily the quality of the found solutions need to be tested. Testing these solutions already is a part of the research document, meaning the most important tests are already covered.

### 3.5 Research

During this project's research, the following main question needs answering: *Which solutions can be used to complement the missing skills for talents as accurately as possible?* To answer this question, the following sub-questions are answered:

- What data is available of the talents?
- Which selection of the talents' data can be used to help to determine the missing skills?
- What are possible solutions to complement the missing skills with help of this selection of talents' data?
- Which solutions perform best and is the company 8vance satisfied with?

The research framework that's taught in the research course at Fontys is used in this project. This framework entails five different strategies that can be used to answer the research questions. Read the next chapter to find more information about this framework and why it's chosen.

*(Provide a detailed description of the assignment. Explain:*

- *The initial situation. What's the problem? Why is that the problem? What are the consequences?*
- *What's the project's goal? What are you and the company trying to achieve? What's the preferred solution?*
- *Define the precise assignment description.*

*Describe the scope and constraints of the project (mandatory languages, documents, methods like scrum/prince?).*

*Describe any changes during the project.*

*Describe the research questions and research framework.*

*Be concrete in the assignment/problem specification. Don't say "This caused delays.", but say "This caused forty percent of the customers to receive their orders four months later.")*

## 4 The approach

### 4.1 Introduction

The approach of this project is described in this chapter. The work and research methods that were used in this project as well as the project planning and contact moment will be discussed. The original approach is written in the PID which can be found in appendix A.

### 4.2 Methods

The methods that were used in this project are: Tienstappenplan (TSP), Fontys' five strategies research method, phasing, and Agile. Each method will be described below.

#### 4.2.1 Tienstappenplan (TSP) and phasing

Tienstappenplan is used for setting up the project so that the project's executor has more control over the project's process. This method helps in securing the quality of the project. This method is commonly used for graduation projects in the IT industry because of its effectiveness. (Kempen & Bennink, 2016)

TSP consists of 10 steps, which are separated over three different phases. This can be seen in Figure 4. The phases are:

- Oriëntatiefase (Orientation phase);
- Onderzoeks- en oplossingsfase (Research and solution phase);
- Invoeringsfase (Completion phase).

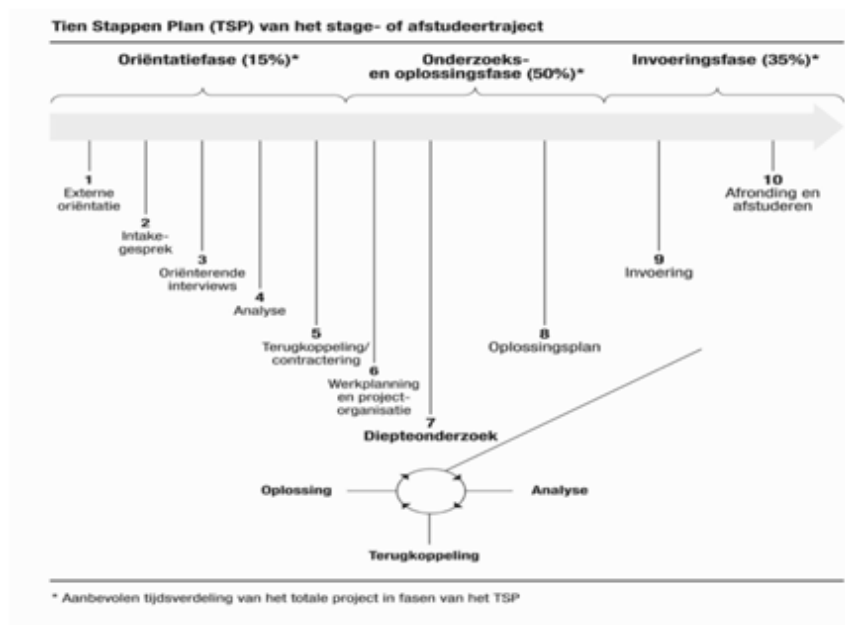


Figure 4 - Tienstappenplan

## Orientation phase

The orientation phase has the PID as the final product. This phase exists of five steps:

- **External orientation.** The goal of this step is to acquire insights in the desires and goals of the company for the project. This step is performed in preparation for the intake interview.
- **Intake interview.** The company is interviewed in this step in order to determine the exact desires and goals of the company for the project.
- **Orientation activities.** The situation of the company at the moment before the project started is described. All the information that was obtained from the previous steps are gathered and an own opinion is formed on the desires and goals of the company.
- **Analysis.** An own interpretation of the project's goal and approach is formulated in the PID.
- **Feedback.** The company and school mentor evaluate the PID and determine whether or not the project may be started.

When all steps are taken, all of the project stakeholders will know what to expect of the project.

The orientation phase also includes an initial research which is focussed on obtaining knowledge of the data that will be worked with.

## Research and solution phase

The majority of the products are finished in this phase. This phase exists of three steps:

- **Planning and project organisation.** In this step, the planning and organisation as defined in the PID are realized.
- **Research and solution.** In this step, the project as described in the PID is executed and the Agile work method is followed.
- **Solution plan.** The goal of this step is to motivate why a particular found solution is better than another. This way, the company can better assess the value of the solutions.

## Completion phase

This final phase exists of two steps:

- **Conclusion.** A conclusion and recommendation of the research is formulated which the company can use to continue the research and possibly improve AIMA.
- **Completion.** The thesis and products are finalized and handed over to the company. The concluding presentations for the company and school are prepared and executed.

### 4.2.2 Agile

During the *Research and solution* step in the research and solution phase, the Agile work methodology is used. This methodology is used because of several advantages.

## Incremental development and continuous improvement

Agile realises the possibility for incremental development. The research is a long-running process where good solutions may or may not be found. Therefore it's impossible to say at the start of the project which things can and cannot be realised and/or achieved. Hence it's a good idea to start small and realize the most important requirements first. As the research progresses and more knowledge is required, requirements can be added or modified accordingly. (Moran, 2010)

## Well-known

Agile is a well-known work methodology in the IT industry. Up to 37% of the companies use this methodology. Both companies and schools are familiar with Agile and acknowledge its strengths and usefulness. (Langley, 2016)

## Transparency

Transparency plays a crucial role in Agile and is an important aspect why it's so effective. Agile forces a close communication between every stakeholder. Through communication, problems and weak points are discovered, discussed and resolved faster. This brings higher transparency and helps to prevent escalating problems. (Moran, 2010)

## Quality

Testing and validating is an integral part of Agile. A requirement is only considered to be successfully implemented when it satisfies all the demanded quality requirements. (Waters, 2007)

## Risk management

Agile development seeks to avoid the issues of "the customer got what they asked for, but it isn't what they wanted" because of a misunderstanding in the requirements. With Agile, working solutions are frequently delivered and inspected in order to avoid these issues. If there're misunderstandings, immediate corrections can be made. (Moran, 2010)

### 4.2.3 Fontys' five strategies research method

The research framework that's taught in the research courses at Fontys is used in this project. This framework entails five different strategies that can be used to answer the research questions. Look at Figure 5 to see an overview of these strategies.



Figure 5 - The five research strategies

Let's have a brief look what each strategy means. The strategies are Dutch terms. The translation of each term are placed between brackets. (Kempen & Bennink, 2016)

- **Veld (Field)**. You dive into the field to gather information around the application domain. The gathered information isn't always reliable, so it needs to be validated accordingly.
- **Lab (Lab)**. In this strategy you test an aspect of your solution. This always involves measuring so you know your solution is the right one.
- **Werkplaats (Workshop)**. You get your hands dirty to explore new ways or to validate a solution. The way you go about this is methodical and structured.
- **Bieb (Library)**. You search for available work (literature or software) or knowledge (domain experts) that could serve as the foundation of your solution. In contrast to the Field strategy, this obtained information is reliable to a large extent.
- **Showroom (Showroom)**. The purpose of this strategy is to compare your solution to other solutions in order to measure the quality of your solution.

These strategies can be used to answer the research questions from different viewpoints. Each strategy also involves its own appropriate and unique sources (sources of Lab can be tables with test results, and sources of Library can be interviews with experts or research papers, etc.). Providing answers from these different viewpoints and using different sources make the answers more credible. (Kempen & Bennink, 2016) These are the main reasons why this research framework is chosen.

The Lab, Workshop and Library strategies are the followed strategies to answer the research questions. The Library strategy is mainly used to acquire information from 8vance's data science experts and their research as well as published research papers on the internet. The Workshop strategy is used to try out own or found solutions, and to validate them. The Lab strategy is used to compare the found solutions with each other and with the company's current skill suggestion solution.

Read the PID in appendix A for more information about the discussed methods.

### 4.3 Project planning

The planning is created based on the pre-known data and situation. The phasing in the TSP method made for a great foundation to be used for the planning. Table 3 contains a simplified version of the project planning which is based on the project planning of the PID, which can be found in appendix A.

Task	Days	Start	End
<b>Orientation phase</b>	<b>10</b>	<b>15-02-2016</b>	<b>26-02-2016</b>
PID	10	15-02-2016	26-02-2016
Initial research	5	22-02-2016	26-02-2016
<b>Research and solution phase</b>	<b>75</b>	<b>22-02-2016</b>	<b>03-06-2016</b>
Research document	75	22-02-2016	03-06-2016
URS	5	07-03-2016	11-03-2016
SAD	25	14-03-2016	15-04-2016
Create analysis tool	50	28-03-2016	03-06-2016
Implement algorithm	50	28-03-2016	03-06-2016
Implement algorithm library	50	28-03-2016	03-06-2016
<b>Completion phase</b>	<b>15</b>	<b>06-06-2016</b>	<b>24-06-2016</b>
Finalization of the products	15	06-06-2016	24-06-2016
Conclusion and recommendations	5	06-06-2016	10-06-2016
Final presentation preparation	13	08-06-2016	24-06-2016
Company presentation	1	10-06-2016	10-06-2016
Deadline thesis	1	14-06-2016	14-06-2016

*Table 3 - Project planning*

*(Start with a introduction, explaining the details of this chapter.*

*Describe the chosen methods I've used during the project (phasing, agile, tsp) and why. Also mention the research methods (inquiries, interviews, proof of concepts, architecture) and research strategies (field, library, lab, etc.) with explanation.*

*Describe relevant topics from the PID like the project planning and contact moments.)*



## 5 Orientation phase

*(Start with a introduction, explaining the details of this chapter.*

*Topics of importance, in this order:*

- *PID. Why is it useful?*
- *Initial research. Mention its results and why this was done. This includes profile data analysis and scraping method analysis.*

*End with a conclusion, briefly mentioning the results of the topics and the next steps to take.)*

## 6 Research and solution phase

*(Start with a introduction, explaining the details of this chapter.*

*Topics of importance, in this order:*

- *Requirements specification. Explain why and mention a range of important and unimportant requirements (for competence reasons).*
- *Software architecture document. Explain why (kruchten, ISO 25010) and refer to a simplified model of the chosen architecture (must be simple enough to understand for anyone). Also mention the architectures that were also considered. Also mention the things of the architecture I will be responsible for to implement and the things 8vance will be responsible for.*
- *Building the algorithm library. Started with it after the architecture was partially created. This library contains four major features: pre-processing the data, using an algorithm to predict skills, and post-processing the data.*
  - *Pre-processing of the data. Explain why it's crucial. This includes the degree model (interview), using existing taxonomies (interview), and normalizing the data. Also mention the problem of creating an own model and the lack of a 'major' model.*
  - *Algorithms to predict skills. Mention the algorithms that were considered and tested. Explain why one algorithm was better than the other and which one was the best. Explain how the algorithms were tested and briefly explain how the best algorithm works.*
  - *Post-processing of the data. Explain why it's crucial. Maybe I'll leave this out because there isn't much interesting to say about this because it's very similar to the pre-processing section.*
- *Building the algorithm analysis "tool". Mention the problem with the first architecture of this tool (a complete GUI application with a 3-layer architecture) and why this architecture wasn't necessary. Explain the use of this tool and why it's important.*

*End with a conclusion, briefly mentioning the results of the topics and the next steps to take. Important questions to answer are:*

- *Does the found solution solve the problem?*
- *Is the found solution satisfactory to the company?)*

## 7 Completion phase

*(Completion phase might not be the best translation of "invoeringsfase", but it will do for now.*

*Start with a introduction, explaining the details of this chapter.*

*Topics of importance, in this order:*

- *Conclusions and recommendations of the research. Briefly mention the best found solution and its flaws. Mention recommendations for the company to get possibly better results.*
- *Finalization of the architecture document. (especially the deployment view which isn't created yet) What part of the architecture is and isn't implemented?*
- *List of implemented, partially implemented and not implemented requirements.*

*End with a conclusion, briefly mentioning the results of the topics. A conclusion may be unnecessary for this chapter though.)*

## 8 Conclusion and recommendations

*(The reader must be capable to understand this chapter without reading any of the intermediate chapters. This chapter cannot suddenly introduce new information out of nowhere.*

*Determine whether or not the project has been successful by reviewing the found solutions. Mention recommendations to improve the solutions.)*

## Evaluation

*(Use of "I" is mandatory in this chapter.*

*Reflect on your own work process and experiences. Describe what I've learned, what I enjoyed, and what were the most important learning moments. Emphasize especially how I solved my mistakes.*

*Also mention and reflect on my personal learning goals.)*

## Bibliography

8vance Matching Technologies. (2016). *How it works*. Retrieved May 17, 2016, from 8vance: <https://www.8vance.com/howitworks/>

Crabtree, S. (2013, October 8). *Worldwide, 13% of Employees Are Engaged at Work*. Retrieved May 18, 2016, from Gallup: <http://www.gallup.com/poll/165269/worldwide-employees-engaged-work.aspx>

Hutt, R. (2016, January 21). *What are the 10 biggest global challenges?* Retrieved May 8, 2016, from World Economic Forum: <https://www.weforum.org/agenda/2016/01/what-are-the-10-biggest-global-challenges/>

Kempen, P., & Bennink, H. (2016). *Competent afstuderen met het Tien Stappen Plan*. Retrieved March 7, 2016, from Noordhoff: [http://hoadd.noordhoff.nl/sites/7745/\\_assets/7046d91.pdf](http://hoadd.noordhoff.nl/sites/7745/_assets/7046d91.pdf)

Kickstart Venlo. (2016). *Huisvesting*. Retrieved May 17, 2016, from kickstartvenlo: <https://www.kickstartvenlo.nl/huisvesting/>

Langley, M. A. (2016). *The High Cost of Low Performance*. Retrieved May 23, 2016, from pmi: <http://www.pmi.org/~media/PDF/learning/pulse-of-the-profession-2016.ashx>

Moran, D. (2010, April 27). *Top 10 Reasons to Use Agile Development*. Retrieved May 23, 2016, from devx: <http://www.devx.com/enterprise/Article/44619>

Stoffels, H. (2016). (T. Hermens, Interviewer)

Vonk, M. (2015, October 13). *Sjerlok spoort gestolen goederen op voor Delta Lloyd*. Retrieved May 17, 2016, from amweb: <http://amweb.nl/branche-724322/sjerlok-spoort-gestolen-goederen-op-voor-delta-lloyd>

Waters, K. (2007, June 11). *10 Good Reasons To Do Agile Development*. Retrieved May 23, 2016, from allaboutagile: <http://www.allaboutagile.com/10-good-reasons-to-do-agile-development/>

(APA notation literature reference list) bibliography

## Appendices

### A. Project Initiation Document (PID)

(Contains the whole PID document.)

### B. Other appendices

(This section is just a placeholder for other appendices. Only add documents or preferentially parts of documents if they offer relevant information for the reader.)