

Defining tax code logic for purchase invoice process automation

Case Company X

Janina Rintamäki

Examensarbete för Företagsekonomi (YH)-examen

Ekonomiförvaltning

Vasa 2020



EXAMENSARBETE

Författare: Janina Rintamäki

Utbildning och ort: Företagsekonomi, Vasa

Inriktningsalternativ: Ekonomiförvaltning

Handledare: Linda Jönn

Titel: Kartläggning av skattekodslogiken för att stödja inköpsfakturaprocessens automatisering

Datum: 16.11.2020

Sidantal: 50

Abstrakt

Detta examensarbete skrevs för ett fallföretag som behövde en mall för att definiera skattekodslogiken för att stödja inköpsfakturaprocessens automatisering. Syftet var att ta fram ett standardiserat sätt att presentera skattekodslogiken, som kan tillämpas på alla länder som företaget är beläget i. Teorin i examensarbetet förklarar beskattning och handel i EU, Incoterms 2020, ERP och fakturahanteringsprocessen, samt RPA. Teorikapiteln hjälper läsaren att förstå examensarbetets projektdel.

Mellan teorin och projektdelen finns ett kapitel som beskriver hur processen för examensarbetet avancerade. Examensarbetets projektdel innehåller kapitel om hur skattekoder för fakturor fastställdes före projektet, samt definition av den metod som valdes, ett beslutsträdsflödesschema. Resultatet av detta examensarbete är en beslutsträdmall för skattekodslogik som kommer användas för att kartlägga logiken i över 80 länder. Två pilotländer användes under projektet och beslutsträd skapades och implementerades framgångsrikt för båda.

Medan detta examensarbete skrevs har testning, godkännande och validering gjorts för att bekräfta att beslutsträdmallen fungerar som avsett. Det beslutades att mallen ska användas och bli det nya standardiserade sättet att presentera skattekodslogik för både automatisering och bokförare.

Språk: engelska

Nyckelord: skattekod, beslutsträd, RPA, automation, AP

OPINNÄYTETYÖ

Tekijä: Janina Rintamäki

Koulutus ja paikkakunta: Liiketalous, Vaasa

Suuntautumisvaihtoehto: Taloushallinto

Ohjaaja: Linda Jönn

Nimike: Verokoodilogiikan määrittelyminen ostolaskuprosessin automatisoinnin tukemiseksi

Päivämäärä: 16.11.2020

Sivumäärä: 50

Tiivistelmä

Tämä opinnäytetyö kirjoitettiin tapausyritykselle, joka tarvitsi mallin verokoodilogiikan määrittelymiseksi ostolaskuprosessin automatisoinnin tukemiseksi. Tavoitteena oli tarjota standardoitu tapa esittää verokoodilogiikka, jota voidaan soveltaa kaikkiin maihin, joissa kyseinen yritys sijaitsee. Opinnäytetyön teoria selittää verotuksen ja kaupan EU:ssa, Incotermit 2020, ERP:n ja laskujen käsittelyprosessin, sekä RPA:n. Teorian luvut auttavat lukijaa ymmärtämään opinnäytetyön projektiosan.

Teoria- ja projektiosoiden välissä on luku, joka kuvaa opinnäytetyön prosessin etenemistä. Opinnäytetyön projektiosa sisältää luvut siitä, kuinka laskujen verokoodit määritettiin ennen opinnäytetyötä, ja kehitystavasta, päätöspuusta, ja valinnan määritelmästä. Opinnäytetyön tulos on verokoodien määrittelymisen kehitetty päätöspuumalli, jota käytetään kartoittamaan logiikkaa yli 80 maassa. Opinnäytetyön aikana käytettiin kahta piloottimaata, joille päätöspuut luotiin ja toteutettiin onnistuneesti.

Tämän opinnäytetyön kirjoittamisen aikana on tehty testausta, hyväksyntää ja validointia sen varmistamiseksi, että päätöspuumalli toimii tarkoitetulla tavalla. Päätettiin, että malli otetaan käyttöön ja siitä tulee sekä automaatiolle että kirjanpitäjille uusi standardoitu tapa esittää verokoodilogiikka.

Kieli: englanti

Avainsanat: verokoodi, päätöspuu, RPA, automaatio, AP

BACHELOR'S THESIS

Author: Janina Rintamäki

Degree Programme: Business Administration

Specialization: Financial management

Supervisor: Linda Jönn

Title: Defining Tax Code Logic for Purchase Invoice Process Automation

Date: 16.11.2020

Number of pages: 50

Abstract

This thesis was written for a case company that needed a template for defining tax code logic to support the automation of the purchase invoice process. The objective was to provide a standardized way of presenting the tax code logic that can be applied to all the countries that the case company is located in. The theory in the thesis explains taxation and trade in the EU, Incoterms 2020, ERP and the invoice handling process, and RPA. The theory chapters support the reader in understanding the project section of the thesis.

Between the theory and project sections is a chapter describing how the process of the thesis project advanced. The project section of the thesis includes chapters on how the tax codes for invoices were determined prior to the thesis and defining the choice of method for the development, which resulted in a decision tree flowchart. The outcome of this thesis is a decision tree template for tax code logic that will be used to map the logic for over 80 countries. Two pilot countries were used during the thesis and decision trees were created and implemented successfully for both.

During the time of writing this thesis, testing, approval and validation have been made to confirm the decision tree template is working as intended. It was decided that the template will be used and become the new standardized way of presenting tax code logic, for both automations and accountants.

Language: English

Key words: tax code, decision tree, RPA, automation, AP

Table of contents

1	Introduction	2
1.1	Objective	2
1.2	Structure	2
2	Definitions and abbreviations	3
3	Taxation and trade in the EU	4
3.1	Intra-community trade.....	4
3.2	General VAT legislation and reverse charge.....	5
3.3	VAT legislation in the Netherlands.....	6
3.4	VAT legislation in Germany	6
4	Incoterms 2020	7
4.1	E rules for any mode of transport (EXW)	8
4.2	F rules for any mode of transport (FCA)	9
4.3	C rules for any mode of transport (CPT, CIP)	9
4.4	D rules for any mode of transport (DAP, DPU, DDP)	10
4.5	Sea and inland waterway rules (FAS, FOB, CFR, CIF).....	10
5	ERP and invoice handling process.....	11
5.1	Enterprise Resource Planning (ERP).....	11
5.2	Purchase invoice handling process	12
6	RPA and Kofax	18
6.1	Robotic Process Automation (RPA)	18
6.2	Kofax and Optical Character Recognition (OCR).....	19
6.3	Kofax at the case company.....	20
7	Background and introduction to the project.....	23
8	Tax code logic prior to the thesis	26
8.1	How Kofax determines tax code	28
8.2	How accountants determine tax code	29
9	Choice of method – decision tree	30
10	Tax code decision tree for NL companies.....	32
11	Tax code decision tree for DE companies	38
12	Analysis and implementation	43
12.1	Data visualization of decision trees	43
12.2	Choice of pilot countries.....	44
12.3	Expanding country and company code scope.....	44
12.4	Validation	45
13	Conclusion	46
13.1	Reliability and usability	47

13.2	Recommendations for further improvements.....	48
14	References	49

List of figures

Figure 1. Invoice example.....	14
Figure 2. SAP fields.....	15
Figure 3. Purchase order.....	16
Figure 4. Booked invoice.....	17
Figure 5. Kofax process flow chart.....	21
Figure 6. Tax code database table.....	26
Figure 7. Vendor default tax code database table.....	27
Figure 8. Decision tree example.....	31
Figure 9. NLx1 original draft.....	32
Figure 10. NLx1 second version.....	33
Figure 11. NLx1 third version.....	34
Figure 12. NLx1 final version.....	35
Figure 13. NLx1 goods section.....	36
Figure 14. NLx1 services section.....	37
Figure 15. NLx1 decision tree in automation software (removed).....	37
Figure 16. DEx1 final version.....	38
Figure 17. DEx1 decision tree in automation software (removed).....	38
Figure 18. DEx1 goods section.....	39
Figure 19. DEx1 services section.....	40
Figure 20. DEx1 Italy section.....	41
Figure 21. DEx1 Malaysia section.....	41
Figure 22. DEx3 Poland section.....	42
Figure 23. DEx2 Finland section.....	42
Figure 24. Tax code validation.....	46

1 Introduction

The thesis and its result are made for the case company, whose expectation was a decision tree for determining tax code logic, which could then be used as a template to map the tax code logic for as many countries as possible. Information already existed on the tax codes and when to apply them, but that information was presented in many ways, such as in tables, in lists, through text, or even links to government websites. Utilizing the decision tree created in this thesis, the case company has a standardized and unified way of presenting the logic behind tax codes for all the countries that the company is located in.

1.1 Objective

The objective of this thesis is to provide a standardized and unified approach for presenting the logic behind tax code selection of all the countries the company is located in. The aim is that the accountants of the case company can refer to the decision tree and that the automations can directly implement the logic in its presented form. The standardized template supports the creation of future decision trees of other countries that are not covered in this thesis.

1.2 Structure

This is a project-based thesis which is split in a theory section supporting the project and a section with the description of the project and its result. The first theory chapter explains general information about VAT legislation, reverse charge and intra-community trade, and shortly summarizes the Dutch and German VAT legislation. The theory supports the reader in understanding the process of selecting the correct tax code for invoice booking. The next theory at chapter four explains Incoterms and their rules, which turned out to be a vital part of the final result of the project. Chapters five and six give information about different software, programs and robotics that are relevant to the project and has a subchapter explaining step by step the purchase invoice handling process. These chapters support the reader in understanding the decision tree result.

The project section of the thesis begins at chapter seven with background, introduction and the timeline for the project. The following chapter describes how the tax code logic functioned from accountants' and automations' point of view before the thesis project was started. Chapter nine explains and informs the reasoning and meaning of the choice of method used in this project, a decision tree. The chapter helps the reader understand the

result of the thesis. Chapters ten and eleven are each dedicated to explaining the tax code decision trees that were created for the Netherlands and Germany, and these two chapters are the main result of the thesis project. The next chapter analyzes the project and its process and describes how it was implemented in the case company. The final chapter is a conclusion where the reliability and usability of the project is evaluated, and some recommendations for possible future improvements are suggested.

Figures may have red boxes to hide sensitive content. Some figures are fully or partly removed for the same reason.

2 Definitions and abbreviations

In this thesis the reader will encounter several terms and abbreviations, which are shortly explained below to assist and support the reader.

Company code (NLx1, DEx1, DEx2, DEx3) – Company codes are legal entities of the case company located throughout the world. A country may have one or multiple entities that are internally named by a company code instead of its entity name. A company code consists of two leading letters which stand for the country in question, and two numbers, separating different entities within the same country. Company code NLx1 will be the main case company code used in this thesis, with company codes DEx1, DEx2 and DEx3 being used as other examples in chapter 11.

Database table – A database table consists of columns and rows in which the cells are filled in with data to electronically store information.

Decision tree – A flowchart model with a structure that resembles a tree with branches spreading from a single source into decision nodes towards end nodes, see chapter 9.

G/L account – General Ledger account is a 7-digit code to which costs are booked for these to end up at the correct place in the ledger. Two examples are G/L 4500000 for inbound transport and G/L 7580100 for hired-in personnel.

Kofax – Kofax software scans documents to recognize data and fill in data fields for invoice booking. Kofax is described in more detail in chapter 6.3.

Node – Decision trees consist of nodes, which are points in the chart displayed as rectangular boxes. The nodes are connected to each other with arrows directing the flow through the branches. The nodes contain text that informs the criteria for following that specific pathway.

PO – Abbreviation for purchase order. A PO is created by a purchaser of the company code and contains all the details of an agreed purchase. The PO contains data on which items of goods or services are purchased, including quantities and prices. Agreements are included within the PO with for example Incoterms, currency, delivery country and other details.

Robot – Robots are software automations that are developed to process tasks that otherwise are handled manually by a person at a higher cost and lower speed or accuracy.

Tax code – A tax code is a combination of letters, numbers or symbols with a two-character length, which define the procedure for the VAT of the invoice.

3 Taxation and trade in the EU

Chapter three will support the reader in understanding the basics of intra-community trade, general VAT legislation within EU trading and reverse charge, all of which are heavily relied on and used when the case company is doing business globally. In this thesis a project is created for tax code logic for the Netherlands and Germany and to support that, there are short chapters about the VAT legislation and rates in both countries.

3.1 Intra-community trade

Sale and acquisition of goods between VAT registered countries within the EU is called intra-community trading. In general, VAT is not applied in transactions of goods between member countries, if the following criteria are met:

1. The goods are physically transported from Finland to another EU country
2. The buyer is liable for VAT in an EU country other than Finland, and
3. The buyer who is established in another EU country has a valid VAT number.

The same rules apply for other EU countries too, Finland was used as an example above for clarity. In cases where the above three criteria are not met, the seller will declare and pay the VAT for the sale as is the normal procedure for sales. (Tax Administration, 2020)

Intra-community trade of goods means that the purchased goods are transported from a member country to the buyer's country, which also must be a member country of the EU. The location of an intra-community trade of goods is the place where the transport of goods to the buyer ends. This place can most easily be defined with the use of Incoterms (see chapter 4). According to the VAT directive, a service is defined as "any transaction which does not constitute a supply of goods". Examples of services are software, patents, cleaning services, online courses, and maintenance. Electricity, heat, gas and the like are regarded as goods and not as services. (The Council of the European Union, 2020)

3.2 General VAT legislation and reverse charge

The VAT treatment of the supply is determined by the VAT legislation of the country where the goods or services are sold. If the supply takes place in Finland, the right to tax belongs to the Finnish state. All countries within the EU follow the same directive, but each state views the directive from that state's legislative point of view. This may in some cases result in both the seller and buyer country paying tax, or neither. It is the Finnish seller's responsibility to find out if the Finnish state has rights to the tax for the supply, in which case tax must be paid even though it would result in double paying of tax. That would be an example of an exceptional situation when the buyer's state also requires tax. (Hakapää, Huhtala, Johansson, Lindholm, & Paavolainen, 2016)

Once it has been decided if the supply of goods or services is taxable or non-taxable, and at which tax rate, the next step is to determine which of the buyer or seller is liable to tax for the supply. According to the EU directive for VAT, the tax is paid by the taxable person carrying out the supply. The premise is therefore that the seller is liable to pay the tax to the seller's country and must register for VAT in the country of sale and charge the buyer a VAT invoice. (Hakapää et al., 2016)

However, reverse charge is an exception to the regular practices. The seller is not required to register for VAT in the country of sale if that country allows the sale to be subject to reverse charge. In that case, the seller will send a VAT-free invoice to the buyer, who then pays no tax for the purchase due to reverse charge. This way neither the seller nor the buyer is liable to pay the tax. If reverse charge applies, the seller can deduct the originally paid input VAT for the goods or materials. The result is that the seller does not receive a disadvantage for losing out on VAT for the sale. For reverse charge to be applicable, the

seller cannot be located, and therefore VAT registered, in the buyer's country. (Hakapää et al., 2016)

In normal circumstances, according to the VAT directive, the taxable person for the supply is the seller and includes on the invoice the VAT of the country of the seller. The invoice is VAT-free when reverse charge applies, and the seller has the right to deduct the original input VAT of the sale. Reverse charge helps simplify VAT on international trade and reduce administrative fees. Reverse charge is very common when services are performed within the EU member countries. For a service to fall under reverse charge, the seller cannot have a business located in the buyer's country. An example for this is when a Finnish business sells a service to a German business, the invoice will have zero percent VAT. (Hakapää et al., 2016)

3.3 VAT legislation in the Netherlands

The Netherlands uses three different VAT rates: 0 %, 9 % and 21 %. The 0 % VAT rate applies for intra-community trade and international transport of passengers by sea and air. The lower 9 % VAT rate is used for several very specific transactions, including food, water supplies, certain medicines and medical equipment, domestic transport, books, hotel visits, restaurants, clothing and ticket costs to certain events. The standard 21 % rate is applied for any other taxable services or goods. (Avalara, 2020)

There are some goods and services that are exempted from VAT, meaning no VAT is paid or deductible. Education and healthcare services, financial services and insurances, childcare and fundraising activities are examples of services that are exempted from VAT. (Belastingdienst, 2020)

3.4 VAT legislation in Germany

Due to the Covid-19 pandemic, the German government decided in 2020 on an approach to support the national economy by temporarily reducing the national VAT rates. Germany has as of current budgeted a total of 286 billion euro for supporting the economy and the VAT rate reduction is 20 billion euro worth of that. In 2018, VAT amounted for over 18 % of Germany's total tax revenue and this reduction of VAT rates is estimated to reduce that to 8 %. (Bunn, 2020)

Normally, Germany has domestic VAT rates of 19 %, 7 % and 0 %. These rates are reduced from 1st July 2020 until 31st December 2020 to 16 % and 5 % respectively. Additionally, the VAT rate for catering food services have been reduced from 19 % to 7 % until 1st July 2021. The 5 % VAT rate is used for medical supplies, books, tickets to cultural events, some foods, domestic transport, flowers, books and short hotel accommodations. The now reduced 16 % VAT rate applies for all other goods and services not taxed at the 5 % or 0 % rates. 0 % VAT rate mainly applies for intra-community trade and international transport. (Bunn, 2020)

4 Incoterms 2020

Chapter four explains in detail Incoterms, their use and their differences. Incoterms play an important part in the tax code logic, especially the rule EXW detailed in chapter 4.1. Currently it is unknown if extra attention should be given to the other Incoterm rules and for that reason, they are summarized in this chapter.

Incoterms are rules that define the obligations, risk and cost responsibilities of the buyer and the seller when trading goods. A key aspect of the Incoterms is that it defines at what moment the goods are considered delivered from the seller to the buyer. Most of the time goods are delivered without any issues and therefore without the need of Incoterms. However, should the cargo be destroyed, damaged or stolen along the way, the responsibilities of both parties are laid out in the agreed Incoterms. There are eleven different rules that use three-letter terms, which are described in this chapter. Incoterms are in themselves not a sales contract or a substitute for such a contract. (International Chamber of Commerce, 2019)

Incoterms are divided in rules based on their first letters. The E rules, which include only Ex Works, mean that the seller places the goods inside their premises for the buyer to use and pick up. In the F rules (FCA, FAS and FOB) the seller delivers the goods to a freight forwarder according to the agreement with the buyer. Using Incoterms in the C rules (CPT, CIP, CFR and CIF), the seller makes a transportation agreement, but is not responsible for the risks after the goods have been delivered to the freight forwarder. Finally, in the D rules (DAP, DPU and DDP) the seller is responsible for all the risks and costs while delivering the goods to the agreed destination. (Railas, 2020)

Seven out of the eleven Incoterms fit any modes of transportation. These seven are Ex Works (EXW), Free Carrier (FCA), Carriage Paid To (CPT), Carriage and Insurance Paid To (CIP), Delivered at Place (DAP), Delivered at Place Unloaded (DPU), and Delivered Duty Paid

(DDP). The other four are additional Incoterms and apply only in transportation by sea or waterways, named Free Alongside Ship (FAS), Free on Board (FOB), Cost and Freight (CFR) and Cost, Insurance and Freight (CIF). (Railas, 2020)

Once an agreement is made between the buyer and the seller, Incoterms are usually stated with the following format: [three-letter Incoterm] [place or port] Incoterms 2020. For example: FCA Rotterdam Incoterms 2020. Stating the correct place is important, since it generally defines where the goods are delivered and thus where the risk of the goods is transferred from the seller to the buyer. (International Chamber of Commerce, 2019)

4.1 E rules for any mode of transport (EXW)

The E rules only contain the Incoterms Ex Works. An agreement with EXW Incoterms gives the seller the least responsibilities and risk, and the opposite for the buyer. The goods are considered delivered when the seller has placed the goods at the buyer's disposal at a named place, such as a warehouse or a factory, which may be located on the seller's premises. Following EXW, the seller is not required to load the goods onto transport or clear them for export. Those responsibilities belong to the buyer. (International Chamber of Commerce, 2019)

The parties should agree who bears the risk of damage when loading the goods onto the selected transportation vehicle. In most cases it is likely that the loading equipment of the seller is used, as the equipment usually is available at the premises, and therefore it is favorable for the buyer to have the seller bear the risk. This is however not stated in the EXW rules, and should the buyer explicitly want the seller to bear the risk while loading goods, the FCA rule should be chosen instead. Once the goods are delivered at the agreed location, all risks are transferred to the buyer. The buyer is responsible for contracting and paying for the transportation of goods from the place of delivery. (International Chamber of Commerce, 2019)

The buyer is responsible for paying all costs relating to the goods once they have been delivered, such as all duties, taxes and other charges that may apply due to export. Furthermore, the buyer will pay additional costs related to the buyer being late for picking up the goods, unless the seller has been noticed in time beforehand. Regarding export clearance, the buyer has the obligation to organize the necessary documents for transferring the goods to a different country. Because of this, EXW is more suitable for domestic trading, where no such documents are required. (International Chamber of Commerce, 2019)

4.2 F rules for any mode of transport (FCA)

Applying the Free Carrier (FCA) rule, the goods are considered delivered to the buyer in two different ways. Firstly, if the agreed place of delivery is on the seller's premises, the goods are delivered once these are loaded on the method of transport provided by the buyer. Secondly, if the agreed place of delivery is located outside of the seller's premises, the goods are delivered when these have been loaded on the seller's method of transport, reached the agreed place, are ready to be unloaded from the seller's transport, and at the disposal of the buyer's transport. The seller must deliver the goods on the agreed date or time, or at the end of the agreed period. (International Chamber of Commerce, 2019)

Following the FCA rule, the seller is required to clear the goods for export, but not for import or transit through other countries except the buyer's country. The buyer is responsible for the costs incurred by import. (International Chamber of Commerce, 2019)

4.3 C rules for any mode of transport (CPT, CIP)

The CPT rule is short for "Carriage Paid To" and has according to the rule the goods considered as delivered when the seller hands over the goods to the carrier, which is contracted by the seller. Even though the transport is made by a carrier chosen and paid for by the seller, the risk is transferred to the buyer at the time of delivery. Once the goods are safely loaded onto the carrier, the seller itself has no responsibility to guarantee the goods will arrive to the buyer's destination in agreed condition or quantity, or even at all. With the CPT rule, the seller pays for the costs related to transport, including documents and security costs. Two places must be agreed to when using CPT; the place where goods are considered delivered, and the destination of the goods where the carrier will transport to. According to CPT, the seller will clear the goods for export, but not import. (International Chamber of Commerce, 2019)

CIP stands for "Carriage and Insurance Paid To" and is almost identical to the previously mentioned CPT rule, except that the seller is required to also have an insurance contract that will cover the risk of loss or damage of the goods, valid from the point of delivery until the carrier has reached the destination. When trading internationally, use of CIP may be challenging if the buyer's country requires locally attained insurance, in which case the CPT rule is more suitable. (International Chamber of Commerce, 2019)

4.4 D rules for any mode of transport (DAP, DPU, DDP)

“Delivered at Place”, the DAP rule, states that goods are considered delivered with the risk transferred to buyer when the goods are placed at the buyer’s disposal on the transport vehicle ready for unloading at the destination agreed on. The seller is responsible for contracting the carrier and pays all costs of transport until the goods have been delivered. The buyer bears the responsibility of unloading the goods from the seller’s transports and is the party holding the related risks. The DAP rule states that the seller is responsible for export clearance, while the buyer’s responsibility is for import clearance. A situation may arise where the goods are held up at the port or at entry to the buyer’s country, due to not being cleared for import, in which case the buyer has not held up its’ end of the contract and is required to pay the related costs. (International Chamber of Commerce, 2019)

DPU is shortened for “Delivered at Place Unloaded” and is very similar to the DAP rule from the previous paragraph. According to the DPU rule, goods are considered delivered once the goods are unloaded from the transport vehicle contracted by the seller and placed at the buyer’s disposal at the agreed destination. The seller bears the risks and pays the costs related to transporting and unloading the goods until delivery is completed. The other conditions are identical to DAP. (International Chamber of Commerce, 2019)

The DDP rule stands for “Delivered Duty Paid” and is also nearly identical to DAP, except the seller clears the goods for import in addition to export. DDP is the rule where the seller bears the most obligation out of all Incoterms and from the seller’s point of view, it should be used with caution. The seller will pay for all costs related to the transport until delivery is completed and additionally any costs for import duty and handle all customs formalities. (International Chamber of Commerce, 2019)

4.5 Sea and inland waterway rules (FAS, FOB, CFR, CIF)

The following four Incoterms only apply to transport that is exclusively made by sea and inland waterways. FAS stands for “Free Alongside Ship” where the rule states that the seller is considered to have delivered the goods when the goods are placed alongside (very close to) the ship chosen by the buyer at the port that the parties have agreed on. The risks are transferred to the buyer once the goods are placed by the ship. The seller is responsible for export clearance and the buyer for anything related to import. (International Chamber of Commerce, 2019)

Using the FOB rule, “Free on Board”, the seller has delivered the goods once they are on board the vessel nominated by the buyer, at the port that is agreed on. Risks transfer from the seller to the buyer once delivery is completed. The seller handles export clearance while the buyer oversees import. The seller is responsible for the costs that may apply for transporting the goods to the port. (International Chamber of Commerce, 2019)

The CFR rule is short for “Cost and Freight” where goods are considered delivered once they are on board the vessel. The risk is transferred when the goods are on board, after which the seller is not responsible for the condition or quantity of the goods. For that reason, it is highly recommended that the buyer arranges some insurance for the goods when the CFR rule is applied. With the CFR agreement, two ports should be named in the contract; the port where the seller loads the goods onto the vessel and the risk is transferred, and the destination port where the seller’s carrier will deliver the goods to the destination. In a situation where multiple different ships or vessels are used for the same transportation, the risk is transferred to the buyer once the goods are on board the first vessel. (International Chamber of Commerce, 2019)

When using the CIF rule, “Cost Insurance and Freight”, the goods are considered delivered once these are on board the vessel, exactly as with the CFR rule. CIF and CFR are very much alike, except that with CIF the seller is responsible for arranging and paying for insurance for the goods and handling all export related actions. (International Chamber of Commerce, 2019)

5 ERP and invoice handling process

This chapter summarizes ERP software and their use, and which systems are the most common globally. A detailed step by step description is included on how the purchase invoice handling process works at the case company. The chapter supports the reader in understanding how the result of the decision tree created in this thesis functions.

5.1 Enterprise Resource Planning (ERP)

ERP stands for Enterprise Resource Planning and is a software application that is very commonly used in business. An ERP system can handle most of the processes a company needs to run smoothly day to day, including planning, sales, marketing, human resources and finance. One of the benefits when large companies use an ERP system is that the

different departments can work together within the same system, allowing the various processes to advance from one department to another whilst keeping the progression easy to follow. (Labarre, 2020)

An example of the benefit of a single system is from the purchasing department. The product data can be stored in the system, which is then used to make a purchase order, to which the accountants book invoices that then go on to the payments department for payment, and finally to the department where the financial reports are kept.

There are many different ERP systems available, with the most common ones being Microsoft Dynamics, SAP, Oracle, and NetSuite. SAP is one of the largest ERP systems with over 50 000 customers in 25 industries. Recently SAP has modernized its ERP system and came out with SAP S/4HANA, while still supporting the older version until the end of 2027. SAP is used for all the financial departments in the case company, as well as many other departments globally. (Labarre, 2020)

5.2 Purchase invoice handling process

This subchapter describes step by step how the purchase invoice handling process in SAP works at the case company. Purchase invoices are handled by the accounts payable (AP) department, that processes over 100 separate legal entities located globally. Each entity has its own company code of which some are chosen as examples and covered throughout the thesis. In this chapter the chosen company code example is NLx1, an entity located in the Netherlands. The sources for this entire subchapter are internal invoice handling manuals and a long-term accountant who handles invoices for the Netherlands.

The first step of an accountant processing an invoice in the AP department is to confirm that the necessary data is included on the invoice, according to the company's country legal regulations, as well as the companywide requirements. The invoice must be addressed to the correct entity of the case company (NLx1). Vendors located in a different EU country than the Netherlands must state the VAT number of NLx1 on the invoice, in accordance with intra-community supply regulations. Apart from certain cases that are not relevant to this thesis, the invoice requires the correct ten-digit purchase order (PO) number created by a purchaser of NLx1. Invoices without a PO number must state the name of the purchaser or contact person at the company code, so that the accountant can forward the invoice to the correct internal person for approval.

Other essential information for booking an invoice is invoice number, date of invoice, total payable amount of the invoice and VAT amount if applicable. It is quite common that some of the required information is missing or incorrect on the invoice, in which case the invoice is sent back by the accountant to the vendor with the request for a revised invoice. It is recommended that invoices include the vendor's VAT number, bank account number and payment terms for the invoice to receive swift processing and therefore a secured payment by the due date, however these details are not obligated to receive payment. The figure example below includes all the required data correctly on the invoice.

Factuur

Pagina 1(1)



Finland

Aflieveradres:

**Bij betaling vermelden:**

Debiteurnummer:

Factuurnummer:

Factuurdatum:

Vervaldatum:

Uw BTWnr:

Aantal	Eenh	ArtikeInr	Omschrijving	Prijs	Per Korting	Totaal €
Geleverd door vestiging:						
Aflieveradres						
Ordernummer:						
Orderdatum: 12-05-2020						
Leverdatum: 12-05-2020						
Uw referentie:						
50,00	meter			14,29	meter NETTO	714,64
			1 inlage			

Nettobedrag	€	714,64
BTW 21% over € 714,64	€	150,07

Te voldoen	€	864,71
------------	---	--------

Figure 1. Invoice example.

The invoice booking process begins once it is confirmed that the invoice meets the requirements outlined above. The first step for an accountant to book an invoice in SAP is to select the correct transaction type from the drop-down list at the top of the window. The list has four choices; invoice, credit, subsequent debit and subsequent credit. Invoice is the most common transaction type and is used for all regular invoices. Credit notes are booked

in a similar way as invoices, except that a previous booking of an invoice must exist in the PO to complete the credit transaction type. Subsequent debit and credit are selected when an invoice is formerly booked and only the price, not quantity, needs adjustment. This applies for when there is a discount or the price in the original invoice was incorrect.

The screenshot displays the SAP 'Park Incoming Invoice' form for 'Company Code NL'. The 'Basic Data' tab is active, showing the following fields:

- Invoice date: 14.05.2020
- Posting Date: 14.05.2020
- Amount: 864,71 EUR
- Tax Amount: 150,07 EUR
- Reference: [Redacted]
- Calculate tax:
- Tax Code: B& B& (NL: 21% Input t...)
- Text: [Empty]
- Paymt terms: 30 Days net
- Baseline Date: 14.05.2020
- Branch number: [Empty]

The 'Vendor' section shows a redacted vendor name and bank details (IBAN, SWIFT, Bank acct, RABOBANK).

The 'PO Reference' section shows a dropdown menu with '1 Purchase Order/Scheduling Agree...' selected. Below this is a table with the following columns: Item, Amount, Quantity, O..., Purchase..., Item, PO Text, Tax Code.

Item	Amount	Quantity	O...	Purchase ...	Item	PO Text	Tax Code
1			M	<input type="checkbox"/>	10	[Redacted]	B& B& (t ^

At the top right, the 'Balance' is shown as 714,64 EUR. The 'PO Reference' dropdown is set to '1 Goods/service items' and the layout is '7_6310 All information'.

Figure 2. SAP fields.

The second step for the accountant is to insert the invoice details to the basic data fields. The SAP fields in the above figure has the required data filled in at the basic data tab. The fields that require data from the invoice are invoice date, reference, amount, currency, tax amount and tax code. In the example of figure 1, the invoice of NLx1 has 21 % Dutch VAT, for which the matching tax code is B& for this specific company code. The basic data details are filled in according to the stated information on the invoice. The next step is inserting the PO number at the PO reference section, which reveals all the item lines the purchaser has created within. An item line displays the goods or services that are ordered by the purchaser and includes the item description, amount, quantity and other details. The PO number 4800180050 in the figure example has only one item line, however it is possible for a single PO to contain hundreds of item lines. Located at the top right of the SAP window is the balance amount, which in the figure indicates that 714,64 euro still needs to be booked for the balance to reach zero, otherwise it is not possible to post the invoice for payment.

Double clicking the PO number in SAP opens the purchase order window, which contains valuable information for booking invoices. The lowest section of the PO overview includes

tabs labelled “Account Assignment” and “Purchase Order History” which are most useful for accountants. In the “Purchase Order History”-tab the accountant can view if the purchaser has marked the goods or materials as received and if there are any other invoices or credit notes previously booked to the PO. The “Account Assignment”-tab displays which general ledger (G/L) account and cost object the cost belongs to.

The screenshot displays the SAP Purchase Order (PO) window. At the top, the document is identified as 'Coupa PO' with a redacted 'Created by' field. The 'Document Overview' section includes options for 'Print Preview', 'Messages', and 'Personal Setting'. The main header shows the PO number, vendor, and document date (12.05.2020). The 'Org. Data' tab is active, showing 'Purch. Org.', 'Purch. Group', and 'Company Code (NL)'. Below this is a table of item lines:

S..	Itm	A	I	Material	Short Text	PO Quantity	O...	C Delv. Date	Net Price	Curr...	Per	O...	Matl Group	Pint
	10	K				50,000	M	D 13.05.2020	14,30	EUR	1	M	COUPA COM.	

The 'Account Assignment' tab is also visible, showing 'Item 1 [10]' and various assignment fields: 'AccAssCat', 'Cost center', 'Distribution', 'Single account assignm...', 'CoCode (NL)', 'Unloading Point', 'Recipient', 'G/L Account (4210000)', 'CO Area', and 'Cost Center'.

Figure 3. Purchase order.

The middle section of the purchase order window displays all the item lines available in the PO and the details for each individual item. A short text description of the item line is written for each item line and frequently a material number when the item concerns goods. The details of these two fields are significant indicators to match the items on the invoice with the correct item in the PO. More importantly, the section shows the quantity of the order and the price per item, which ideally should match with the invoice statements. Accountants match the correct item lines in the PO with the items stated on the invoice and select the corresponding lines. Invoices are immediately posted and automatically approved for payment when the quantities and prices of the invoice items are an exact match with that of the PO. A mismatch requires manual approval and results in the invoice entering the SAP

workflow of the purchaser, where the purchaser either accepts or denies the difference between the PO and the invoice.

The most useful tabs at the top section are “Org. Data” and “Delivery/Invoice”. The “Org. Data”-tab informs for which company code the PO is made and the name of the purchaser or group that created the PO. This information is essential for the accountant to forward any possible questions concerning the invoice. A field at the very top above the tabs displays the vendor account number and name for which the PO was created. The “Delivery/Invoice”-tab includes the agreed information between purchaser and vendor concerning the payment terms and the Incoterms.

The screenshot displays the SAP Park Incoming Invoice interface. At the top, it shows the title "Park Incoming Invoice: Company Code NL" and a balance of 0,00 EUR. The "Basic Data" tab is active, showing fields for Invoice date (14.05.2020), Posting Date (14.05.2020), Amount (864,71), Tax Amount (150,07), and Payment terms (30 Days net). The Vendor information section includes fields for Vendor, IBAN, SWIFT, and Bank account (RABOBANK). Below this, the "PO Reference" and "G/L Account" fields are visible. At the bottom, a table lists items with columns for Item, Amount, Quantity, and PO Text. The first item has an amount of 714,64 and a quantity of 50,000 M.

Item	Amount	Quantity	PO Text
	714,64	50,000 M	

Figure 4. Booked invoice.

The accountant proceeds booking the items of the invoice to the PO when all the necessary data between the invoice and the PO is correct. The appropriate item lines are selected, and the amounts and quantities are inserted to the lines. The balance amount adjusts once an item line is chosen, and the accountant continues this process until the balance reaches zero. Once the invoice booking has no imbalance, as a final confirmation that the booking is made to the correct vendor, the accountant confirms that the VAT number of the vendor invoice matches with the vendor account in SAP. In case no VAT number is stated on the invoice, the address of the vendor is compared for confirmation.

Finally, in the “Payment”-tab of the invoice booking window, the payment terms and correct bank account are selected. The accountant completes the booking by pressing the “Save as Completed” button, which either immediately posts the invoice with a perfect PO match, or the invoice goes to the purchaser’s workflow for review and approval.

6 RPA and Kofax

Chapter six describes the different systems and tools that are relevant for automating a digital business process. The topics that will be explained in this chapter are RPA, Kofax and OCR. The RPA chapter gives the reader insight in what RPA is and its benefits. Next is a chapter about Kofax software in general and OCR. The final subchapter describes how Kofax is used at the case company and a detailed description of how the six-step Kofax process functions in the invoice handling process.

6.1 Robotic Process Automation (RPA)

RPA stands for Robotic Process Automation, which is the technology that makes it possible to create software, or “robots”, that can perform the same actions as a human when interacting digitally in different processes. RPA robots can be programmed to perform repetitive tasks without ever getting tired or making human errors. Examples of actions robots can be programmed to do are; opening emails and attachments, logging in to applications, moving folders and files, copying and pasting data, extracting different data from documents, and filling in forms. (UiPath, 2020)

It is possible for robots to identify and interact with buttons, fields, text and images in most programs and systems in a similar way a human does. However, robots perform the actions faster than a human by for example recognizing and clicking buttons in milliseconds or writing paragraphs of text instantly by using computer power for calculations and decisions.

Benefits of using RPA robots include:

Better accuracy – The robots are programmed to follow the exact rules that were created and implemented. Robots will follow the rules consistently and tirelessly, without making incorrect actions that a human could make, which could result in extra workload for correcting the mistakes. (UiPath, 2020)

Improved compliance – The actions of the robots are executed reliably, which reduces risks. It is possible to monitor and report all actions, and the creator can ensure that all activities are done according to the regulations and standards of the business. (UiPath, 2020)

Cost savings – The use of RPA robots can reduce processing cost by up to 80 %, as the robots can effortlessly take over the repetitive tasks that humans otherwise would spend time and energy on. The initial cost invested in creating functioning robots will often see return in less than a year. (UiPath, 2020)

Increased productivity and speed – Employees benefit from RPA, since it removes activities that are non-value-adding. This makes it possible for the employees to reduce some work pressure and focus on tasks that require human input. (UiPath, 2020)

Forrester forecasted in 2017 that the RPA market will grow to 2,9 billion dollars in 2021, while it was only 250 million dollars in 2016. It is also estimated that by 2021, there will globally be four million robots doing repetitive administrative, sales, and office related tasks. The forecasts are partly based on the reduction in office jobs due to advancing technology and artificial intelligence. (Le Clair, 2017)

6.2 Kofax and Optical Character Recognition (OCR)

Kofax is a software that companies can utilize for processes that are information-intensive, to reduce errors and manual routine work and to lower costs. The software is a combination of RPA, mobility, cognitive capture, analytics and process optimization, which can help the company increase profits and competitiveness, and reduce costs and compliance risks. Kofax has become a large organization with users in over 60 countries globally, and having over 20 000 customers in fields such as healthcare, insurance, supply chain and financial services. (Kofax, 2020a)

Optical character recognition (OCR) is a technology that enables conversion of different types of documents, such as scanned paper documents, PDF files or images into editable and searchable data. OCR recognizes the patterns of individual letters, numbers and characters and transforms these into understandable and usable text. The technology behind OCR has advanced with time, as before programmers had to manually enter the specific patterns that they wanted computers to recognize. Today, the OCR software has built in patterns with different fonts and manual input is rarely required, making it a powerful timesaving tool for companies. (Kofax, 2020b)

Kofax comes with built in OCR software, which is optimized for working with PDF documents. The OCR in Kofax works quickly to search through and convert the text from PDFs and keeps the documents organized and accessible. Using Kofax and its OCR software, a company can organize and store all documents digitally and for example use it to sign legally binding agreements and documents. The user can determine the type of information the OCR software should look for and retrieve that data from the document. Kofax is a helpful tool for increasing the work performance in less time and more accurately. (Kofax, 2020b)

6.3 Kofax at the case company

The source for this chapter is the RPA expert responsible for developing the purchase invoice process automation project in the accounts payable department. Kofax is a process automation software used at the accounts payable department of the case company. Its primary use, in the department, is scanning PDF invoice documents and importing the documents to SAP with invoice data captured from the invoices by Kofax. The recognized invoice data is inserted to the SAP fields of the invoice document that accountants process further by checking whether the inserted data is correct before the invoice document is posted for payment.

Kofax captures the invoice data with OCR, which extracts text from documents by recognizing characters of handwritten or text images and converting the characters into machine-encoded text. Computers can read and search through machine-encoded text, unlike with an image in which no text can be selected. With the invoice text understood, Kofax can recognize invoice data that is necessary to insert to the SAP fields for invoice booking. Kofax assigns text to invoice data such as invoice reference, date, amounts, currency, items and their quantities and more; which then is exported to SAP for the accountant to proceed with the invoice handling process.

The case company works on solutions to reduce the amount of time for the accountant to proceed further with the invoices, by improving the accuracy of the invoice data. In combination with other steps that lead towards RPA robots posting the invoices with occasional assistance from accountants when necessary. Depicted below is a six-step process of Kofax, which provides accountants with invoices to SAP, with pre-filled data to the fields that accountants then confirm or correct the data.

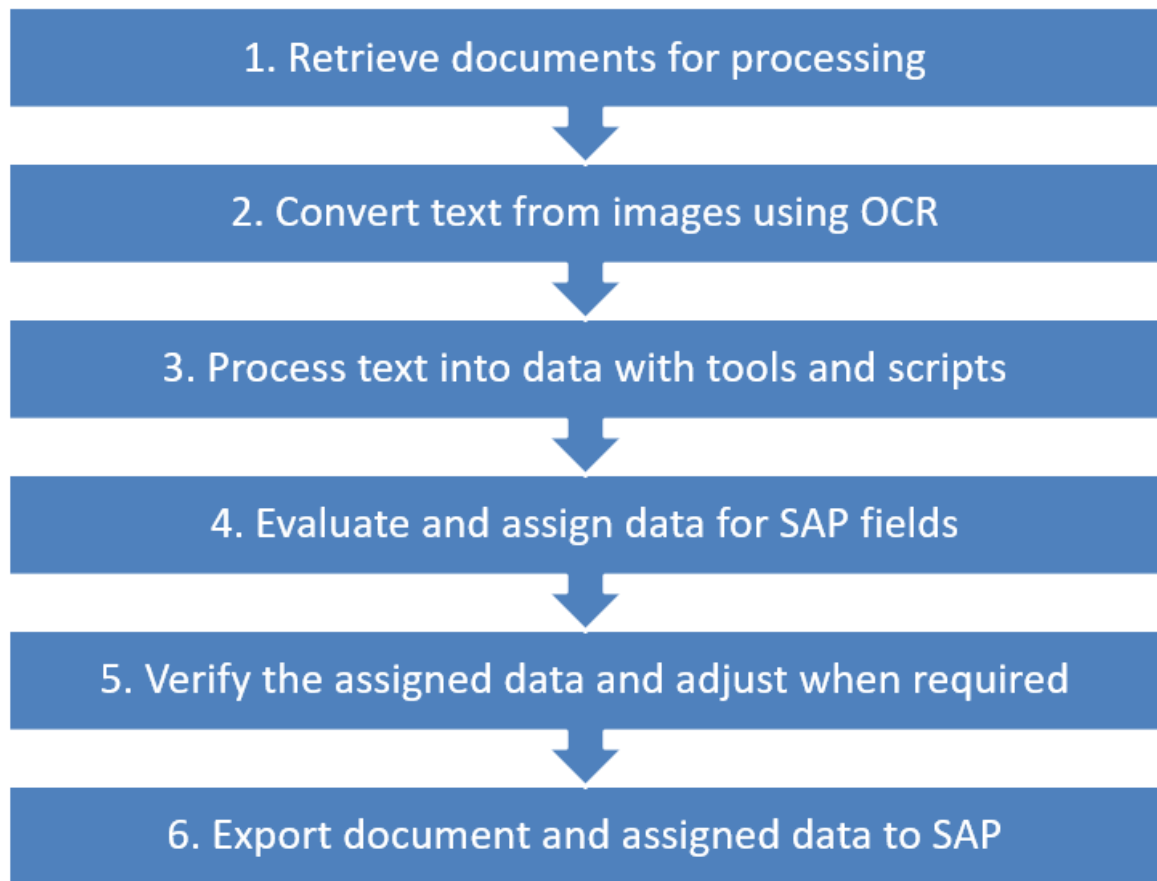


Figure 5. Kofax process flow chart.

1. Kofax retrieves invoice documents from multiple different sources, such as emails, folders and scanners located inside the office. The PDF documents are removed from the source and are transported through the next process steps of Kofax.
2. Computer programs cannot yet process the text on the documents for a majority of the invoices as a significant amount of the received PDF documents were scanned by the sender as an image, creating an image-only PDF in which the text cannot be selected, searched or edited. Kofax converts the images into machine-readable text data with OCR, through which Kofax acquires knowledge on the text that the documents contain. (Brandeis University, 2020)
3. Different methods are used for locating data and inserting the correct details to the appropriate fields in SAP. The vendor details on the invoice are compared to existing vendor data stored in SAP, in which the highest match determines, in step four, the vendor that Kofax assigns to the invoice. Invoice reference numbers are located mainly based on keywords, coordinates and trained structures of the invoice. Amounts are numeric with the total, net, and tax amounts often grouped together at the end of the invoice.

The PO numbers begin with a small quantity of numeric combinations and have a fixed length of ten digits. Kofax extracts the PO numbers and retrieves from databases the PO items that are created under the PO to reduce the possible matches of items with the item costs on the invoices. The PO items already have a tax code assigned through RPA, in which the decision tree, created in this thesis, is the most dominant and determines the tax code that Kofax assigns to the invoice.

4. The processed text from step three results in a substantial amount of diverse options that still require evaluation to reach a final result to fill the SAP field with. Vendor data for example compares different details such as vendor name, city, street, bank account et cetera. The data of vendors in SAP does not always match exactly with the details on the invoice. The name of the vendor may be written slightly different or the name has a close match with that of another vendor. A considerable number of vendors are located in the same cities, or perhaps Kofax could not find a match with any of the examined address details. The top results of vendors are evaluated, and comparisons are made to reach a percentage confidence in which one vendor outperforms the others. The vendor with the highest confidence is assigned to the vendor field and is an important variable that is used to conclude the results for other SAP fields.

Another example are the end results of tax codes that belong to domestic invoices. PO items that are purchased domestically require the VAT rate to conclude the exact tax code out of a small range of possible codes. The VAT rate is unknown until the invoice is received and the tax amount and/or percentage is read from the document. Sometimes the VAT percentage is not stated on the invoice, in which case Kofax calculates the percentage based on the net, VAT and total amounts. Kofax evaluates the different tax percentages and links the matching percentage to the corresponding tax code for that company code that is then assigned to the tax code field.

5. The assigned and provided data from Kofax is in a final stage verified if the data is accepted for export to SAP. For example, the fields with amounts are verified whether the filled in data is numeric and contains a maximum of two decimals. The decimal character is verified whether it is a comma and not for example a period. The maximum character lengths for the fields are verified to be within the accepted length that SAP allows. Invoices containing data that fail to comply with set standards are denied from transport and could have the failed data either overwritten

by default values, or an employee investigates the case before the invoice is exported further to SAP.

6. The final stage of Kofax is transferring the documents and data to SAP. SAP receives the invoice PDF document that accountants can view within SAP, as well as the data assigned to the fields. RPA processes the invoice documents and either posts the invoice for payment, or it may insist that accountants process the invoice if not all data was possible to confirm completely by automation.

7 Background and introduction to the project

The case company has process automation developments in many of its departments, with one RPA expert dedicated for the AP department. One continuous large RPA project is for robots to process and post invoices for payment. A new feature was developed for Kofax to learn how and from where on the invoices specific data should be taken. The feature enabled accountants to teach invoices to Kofax and correct the system if the data was identified incorrectly. I was one of three accountants who were chosen from pilot company codes to test the new feature before it was rolled out to over 50 other accountants. Since then, I wanted to continue develop and be involved with automations where possible.

A challenge for RPA to process more invoices of other company codes outside the current scope are the tax codes. An invoice must have the correct tax code at the posting and every country has at least between seven to twenty different codes, and that for over 80 company codes each. It became clear, from discussions with the RPA expert, that a solution was necessary for a simplified transition between accountants' tax code instructions and robots' tax code logic. No solution existed yet as reviewing and mapping the tax codes was known to take time and effort that was not available with the current resources.

The whole thesis project started in May and finished in November. During this time, there have been numerous meetings, interviews, discussions, brainstorming sessions and cooperation with managers, RPA expert, accountants of different company codes, and others. These persons were in contact with other specialists, experts and higher management from their side, and returned to me with answers, feedback or ideas. The cooperation during the thesis project was extremely successful and I am very pleased with the result that came out of it.

The idea for this thesis was pitched to my manager at accounts payable and the RPA expert for the same department in May 2020. For that meeting I compiled the original draft of tax code logic in the Netherlands, figure 9, and a very rough draft for the Finnish tax codes. Initially it was intended that the two pilot countries for the project would be the Netherlands and Finland, but due to Finland's complex tax code system, it was decided Germany would be the other pilot country instead. The Netherlands was chosen by myself as I had been handling invoice booking for the Dutch company codes for a year and I knew the tax codes well.

My manager and the RPA expert both showed support for this project and a meeting was arranged to include the other AP managers, a tax expert, a G/L specialist and other interested parties. Prior to this meeting I was requested to summarize the project in an email. In the email I wrote a short description of the first draft and the reasons why the Netherlands was chosen as the case company code and my thoughts on which different checkpoints could be in the decision tree. I received positive feedback from the meeting along with ideas and suggestions for enhancement.

During the process of making the second version (figure 10) I wrote down the following comments and concerns:

- G/L account: Based on what the purchaser has written in the PO. If there are both materials and shipping on the invoice and PO, the material tax code is used for both. If the invoice only has freight to be booked to the PO, service tax code is used. Robot could check if invoice matches one material G/L in the PO and decide the whole invoice is to be booked with material tax code. A possible and realistic issue is that some purchasers use incorrect G/L accounts frequently. How to deal with this?
- Vendor country: Based on information from supplier master data. The data is maintained in SAP by another department and robots retrieve the data to a database used by RPA.
- VAT%: Mainly domestic vendors, some exceptions exist with certain other vendors depending on the delivery address. Exceptions go case by case, not recommended to program the robot to handle these situations. VAT% is calculated with net amount, VAT amount and total amount that are all three trained to Kofax.
- Delivery country: Currently not checked by accountants when booking NL and not mentioned in company code instructions. Various other countries do check the

delivery country and have their own instructions accordingly. Could possibly be programmed for the countries that require the delivery country as a checkpoint.

- Incoterms: Will this be needed? Currently not relevant when booking NL and not mentioned in company code instructions. At least some other countries check this. See above comment about delivery address.
- Multiple tax codes: Certain domestic vendors apply up to three different tax codes on the invoice. Book manually or try to program the robot for it?
- Framework orders (FO): Framework orders are quite common for NL. The robot posting invoices is currently programmed to not handle these since a validation requirement cannot be made. What to do about these invoices?

The concerns in the comments were addressed in a meeting with managers and the RPA expert, and I proceeded to work on the next version of the decision tree. Decision trees are explained in more detail in chapter 9. It was decided that Incoterms and delivery country should be included in the tree template and these were added as seen in figure 11.

In October 2020 the final version of the Dutch tax code decision tree was approved by management and RPA. The final version has all the current requirements for determining tax code, however the version might still change with additions, modifications or enhancements as new company codes enter the scope. The Dutch decision tree is quite simple, in comparison to some other company codes, with only seven tax codes in use. For this reason, it was decided that another more complex template should be created as well to test whether the developed decision tree template could be applied to other company codes and easily expand the scope.

The German company codes were selected for a more complex template, of which the German tax code decision tree is explained in chapter 11. An accountant handling the German company codes was chosen to create, in cooperation with me, a tree that involves a bit more complex logic. German company codes, like many other company codes, are VAT registered in other countries and are for that reason a good candidate to include in the pilot scope. In November 2020 I have been working together with management, RPA and accountants of Germany, Norway, Sweden and Denmark to create tax code decision trees for said countries.

In the following chapters the tax code logic is first explained as how it was before this thesis project, with the succeeding chapters introducing the Dutch and German decision trees.

8 Tax code logic prior to the thesis

The information source for Kofax are databases that store data from SAP. The data of SAP is taken from multiple SAP tables and other transactions within SAP and imported by RPA to SQL where Kofax can retrieve this information. Prior to this thesis project, not all necessary data from SAP was stored to enable Kofax to determine the correct tax codes for invoices. Some key examples of data that were not stored are the delivery address of the item costs and the G/L accounts of the PO items. Kofax concluded the tax code based on very few data criteria that were stored in a database table through a script that instructed how the table was read in order to receive a tax code result. The concluded tax code was not guaranteed to be correct due to the limited data that Kofax had available and could use as variables for criteria and checkpoints. The case company has many company codes that each have different criteria and rules that need to be taken into account concerning taxes. Company codes, that for certain tax codes need to consider variables that were not available to Kofax at the time, would notice a lower success rate for invoices with this specific tax code field.

	CompanyCode	DocType	VATRate	Vendor	VendorCountry	TaxCode	Comment
1	NL [REDACTED]	*	9.00	*	NL	ZC	NL Rate 9.00
2	NL [REDACTED]	*	*	*	NL	B&	NL Default
3	NL [REDACTED]	*	*	*	*	RF	Non-EU
4	NL [REDACTED]	*	*	*	EU	P&	EU Default
5	NL [REDACTED]	*	21.00	*	NL	B&	NL Rate 21.00

Figure 6. Tax code database table.

The database table that Kofax used contained columns of the company code, document type (MM or FI), VAT rate, vendor account number, vendor country, and tax code. Kofax consults the table to match row data in the columns to insert the tax code. Kofax receives input from other systems or users to which company code an invoice belongs and of which document type the invoice is and deliberately does not overwrite the input although it might come to a different conclusion concerning these two variables. The VAT rate and vendor account number are read and evaluated by Kofax from the invoice document, with the vendor country linked to the vendor account that is stored in SAP.

A Kofax script matches in the tax code database table certain conditions between the variables of company code, document type, VAT rate, vendor account number and vendor country. The table has a final column that contains the tax code that is selected as the result when a match occurs. An invoice that has all variables matching with the conditions in the database table will receive that specific tax code of that table row. The company code column contains in each row a company code that must always match with the receiver of the invoice, however not all columns require a specific value for a match to occur. Most rows in certain columns contain an asterisk character that is used to indicate that any value is accepted as a match. The vendor account number column, for example, would otherwise require every single vendor for the company code and every possible combination to process the invoice with a tax code.

CompanyCode	DocType	VATRate	Vendor	VendorCountry	TaxCode	Comment
NL [REDACTED]	*	*	[REDACTED]	*	ZC	Vendor Default
NL [REDACTED]	*	*	[REDACTED]	*	ZC	Vendor Default
NL [REDACTED]	*	*	[REDACTED]	*	O&	Vendor Default
NL [REDACTED]	*	*	[REDACTED]	*	RG	Vendor Default
NL [REDACTED]	*	*	[REDACTED]	*	RG	Vendor Default

Figure 7. Vendor default tax code database table.

For invoices where not all conditions are met with explicit data, the asterisk character is placed strategically within the columns to still have a tax code result for nearly all invoices. The columns have different priorities for situations where Kofax finds more than one row to match due to the asterisk characters. For example, the vendor account number has the highest priority over any other column, as it is specifically informing the tax code used for that individual vendor. The vendor country column has a lower priority, since the same country is used for many vendors. There is also a row that only contains asterisks if no specific data can be gathered and this results in a default tax code for that company code, as seen in line three in figure 6 above.

This method of determining tax codes is the current process for Kofax to fill in the tax code field in SAP.

8.1 How Kofax determines tax code

Kofax attempts to recognize the invoice final amounts and VAT percentage from every other number on the invoice. This type of data is trainable by accountants for Kofax to improve and recognize the amounts and percentage with high accuracy. An invoice without the VAT percentage present in the text of the document has the percentage calculated based on the total amount, net amount and VAT amount. The VAT percentage is by default set to zero for invoices that do not contain a VAT amount or have the same net amount and total amount.

Kofax is continuously trained on many invoices, particularly the most common vendors for all company codes within the case company. The trainings improve the accuracy of various recognitions, including the final amounts and VAT percentage. Vendors that are not yet trained may not recognize the correct or any amount, in which case Kofax concludes the VAT rate at zero percent. However, domestic invoices with a zero percentage are rare and do not occur for almost any vendor. Due to the rarity of zero VAT domestic invoices, the default tax code is set to the most dominant code for that company code, to increase the chance of it being correct. Vendors that are confirmed to have zero VAT rates on their invoices are in the table individually added to the vendor account number column data with the respective tax code. Invoices with zero VAT received by the vendors listed in the vendor account number column have as a result the correct tax code applied, while new untrained invoices and vendors with domestic non-zero VAT percentages have a high success rate of tax code.

By using the Dutch company codes as an example, this means that if Kofax cannot determine the VAT rate, it will automatically default to 0 %. However, most domestic Dutch invoices have 21 % with a small number of vendors that do have zero VAT rate for domestic purchases. The tax code for 0 % on domestic invoices is RB, which would be incorrect to apply if the invoice has an unrecognized 21 % VAT rate and in fact requires B& as the correct tax code. If Kofax can read VAT rate 9 % or 21 % from the invoice and establishes that the vendor country is the Netherlands, it will select the correct tax codes based on the table displayed in figure 6.

Kofax attempts to determine the correct VAT rate by recognizing the amounts displayed on the invoice which results in accurate tax codes when the recognition is successful. However, it is not possible to guarantee that an invoice is with zero VAT rate as it is more likely that Kofax failed to recognize the correct rate or amounts, since the tax code RB is very uncommon. As shown in the table of figure 6, Kofax will put B& instead as the default tax

code if the vendor country is the Netherlands and it cannot find the VAT rate on the invoice. Invoices are regularly received from new vendors and these may not have the recognition accurate until trained to Kofax, so by making the most common tax code the default code at possible failed recognitions, the tax code is more likely to be correct.

Kofax does not read the vendor country from the invoice. Instead the information is retrieved from a database table where the recognized text from the invoice is matched against the table for Kofax to choose the most likely vendor. Depending on whether the vendor is located within or outside the EU, the tax code differs based on if the costs are for goods or services. Prior to the thesis there was no data to support Kofax in deciding if the costs on the invoice were for goods or services, and for that reason a database table containing specific vendor account numbers and their default tax codes were necessary as a temporary solution. Invoices for goods are more common than those for services in the Dutch company codes, therefore the tax code for non-domestic invoices are set as default to the cost type goods. When an automation or accountant discovers that for a specific vendor the service tax code is most used based on the posting history of the vendor, that vendor account number is added to the database table and will as default receive the services tax code on invoices.

8.2 How accountants determine tax code

This subchapter shortly describes how accountants determine which tax code should be used when booking an invoice. The sources for this section are two accountants who are handling purchase invoices for the Dutch company codes. The first step for the accountant is to check that the legal requirements for the invoice are met on the invoice document. The second step is to confirm that the data, such as reference, date and amounts inserted by Kofax are correct or if the accountant needs to adjust the fields in the booking. These steps were discussed in more detail in chapter 5.2.

Once the invoice and basic data are confirmed to be valid, the accountant checks from which country the vendor is based on VAT number and vendor address. It is a domestic purchase for the Dutch company codes if the vendor is from the Netherlands and one out of three tax codes is selected based on the VAT rate. The three VAT rates are 21, 9 and 0 percent, and their respective tax codes are B&, ZC and RB. If multiple VAT rates are present on the invoice, then the tax code field in the basic data of SAP is left empty and the different tax codes are filled in per item line.

If the vendor is from an EU country other than the Netherlands, the accountant checks if both the seller's and the buyer's VAT numbers are correctly stated on the invoice, otherwise the vendor needs to revise the invoice. For the EU invoices, the accountant must determine if the invoice costs are for goods or services as the two possible tax codes are both of zero VAT rate. Tax code P& is applied when the invoice is for goods and when the whole invoice is service related the tax code O& is selected. Often invoices for goods have freight costs as well, and even though freight is a service, the whole invoice will be booked with the goods related tax code, since these costs are classed as ancillary costs. When the vendor is from a country outside of the EU, the same logic applies where the cost type determines the choice of tax code, except the tax codes used are RF for goods related invoices and RG for invoices that are for services.

When accountants decide on the tax code for the Dutch company codes, neither delivery country or Incoterms are examined, which may lead to incorrect tax procedures that the accountants, purchasers or vendors may not be aware of.

9 Choice of method – decision tree

The choice of method for the decision-making process for the various types of tax codes needs to meet some key objectives. The first objective is to replace the existing inconsistent instructions between the company codes with a single template that is functional and logical. A single template with a clear structure and rules enables a smoother transition to new company codes for which accountants can create the decision steps based on already existing examples of other company codes. The second objective is that the template needs to be applicable by both accountants and automations, in which it is straightforward and simple to understand the activities involved in the decision-making to reach the final outcome of tax code. Finally, the third objective is to clearly display the types of data required to conclude a tax code, since the currently used data is insufficient to accurately decide on the tax code.

The objectives were decisive in the choice of method as multiple methods were examined for this project. At the beginning a table method stood out as a first choice as some tax code instructions of company codes were written in tables that looked promising but proved to be insufficient for automation purposes. RPA executes a series of if-then-else statements for all the conditions that apply to the logic of tax codes. Observing the automations and possibilities in RPA brought forward the decision tree method as a clear winner that meets all the objectives.

A decision tree is a flowchart model with a structure that resembles a tree with branches spreading from a single source into decision nodes towards end nodes. The method of choice handles effectively non-linear data that is seen in the logic of tax codes. The decision tree contains in each branch split the conditions that need to be met to proceed further down the branch. Decision trees are simple to understand by both accountants and RPA developers without requiring knowledge and experience of the company codes' tax codes. Another advantage is that the decision trees are possible to make in common tools such as Microsoft Excel, which all accountants in the case company are familiar with. Excel is easy to prepare as it is very flexible with its customizations unlike for example SmartArt (see figure 8) that limits the user to its settings. This enables experienced accountants to create a tax code decision tree for the company codes they handle. (Corporate Finance Institute, 2020)

Decision trees consist of nodes that are units containing data and linking to other nodes with arrows. A decision tree begins at a starting node that triggers the process of the flowchart into either a downwards or horizontal direction. The starting node connects to decision nodes where a decision is made based on specific conditions at every node. The conditions from the nodes determine the next branch of direction and leads to another decision node or to an end node. End nodes are the final stage of the decision tree and provide the outcome of result. (Segal, 2019)

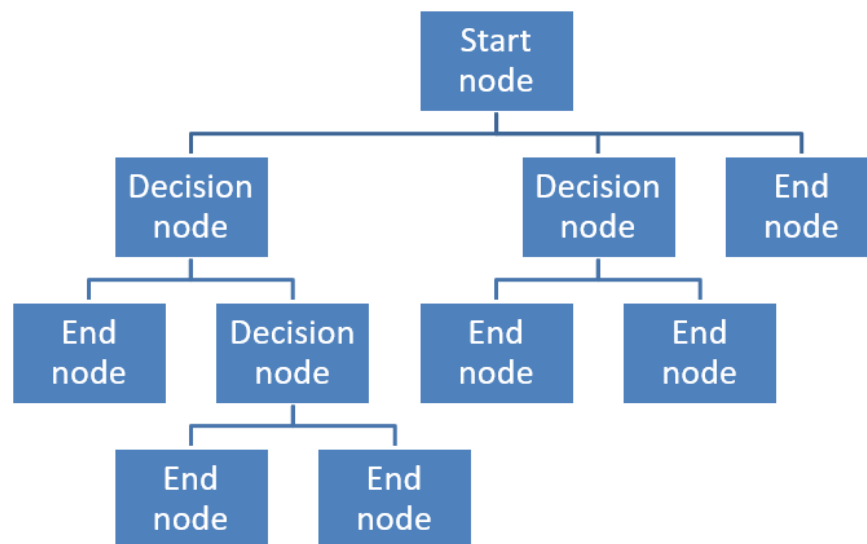


Figure 8. Decision tree example.

The presentation of a decision tree is possible in various forms. The evolution of the decision tree throughout this project is shown in the following chapters, where the structure is modified over time with its reasons described in more detail in chapter 12.1.

The major benefit of a decision tree structure in this project is that the model uses if-then-else statements at the decision nodes, which is possible to recreate within the RPA program using the exact template. This means that when a decision tree is created by an accountant, the RPA developer can copy the structure and its decision nodes into the robotics programs. This can be seen when comparing the Excel decision tree in figure 16 with the decision tree in the automation software, as seen in figure 17. A clear visual legend would support the developer in communicating the types of data required to flow through the nodes towards the end nodes.

10 Tax code decision tree for NL companies

The Dutch decision tree was the first in development and has transformed gradually with time into different versions, where each version is more advanced and evolved from the previous. The progression came as meetings were held, as well as when the project also moved to other company codes for which the template of the tree needs to be user friendly to understand and create by other persons.

The very first decision tree was made before the thesis project. I made the original draft to demonstrate my idea to the AP managers and propose the decision tree method for a visual and logical approach to defining tax codes for invoices. I pitched it as a possible thesis project in support of the invoice automations for if the managers would be interested.

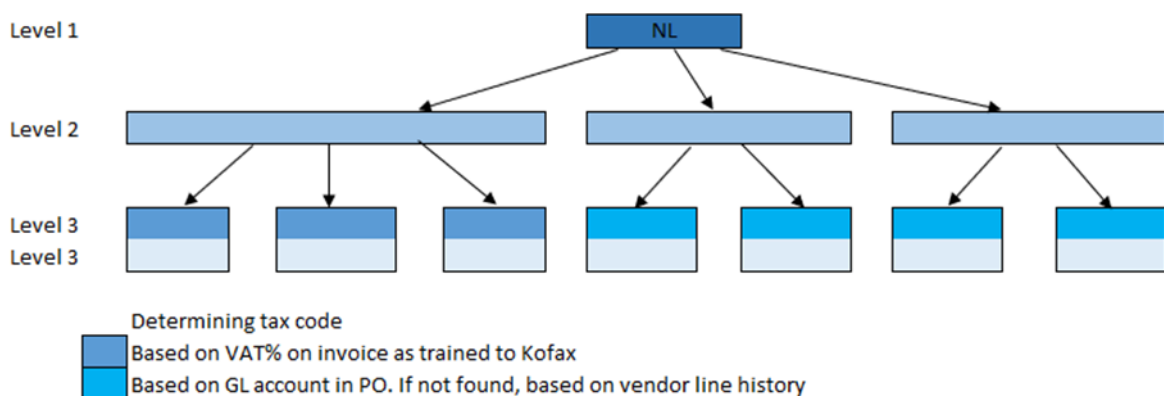


Figure 9. NLx1 original draft.

The above original draft of the tree is based on the exact steps currently in use by the accountants handling the Dutch company codes to determine tax codes, as described in chapter 8.2. The original draft was still incomplete as it displayed the general logic and the main checkpoints that accountants use to quickly decide on a tax code, which for most

invoices is enough to result with the correct code. The original draft was presented 12.5.2020.

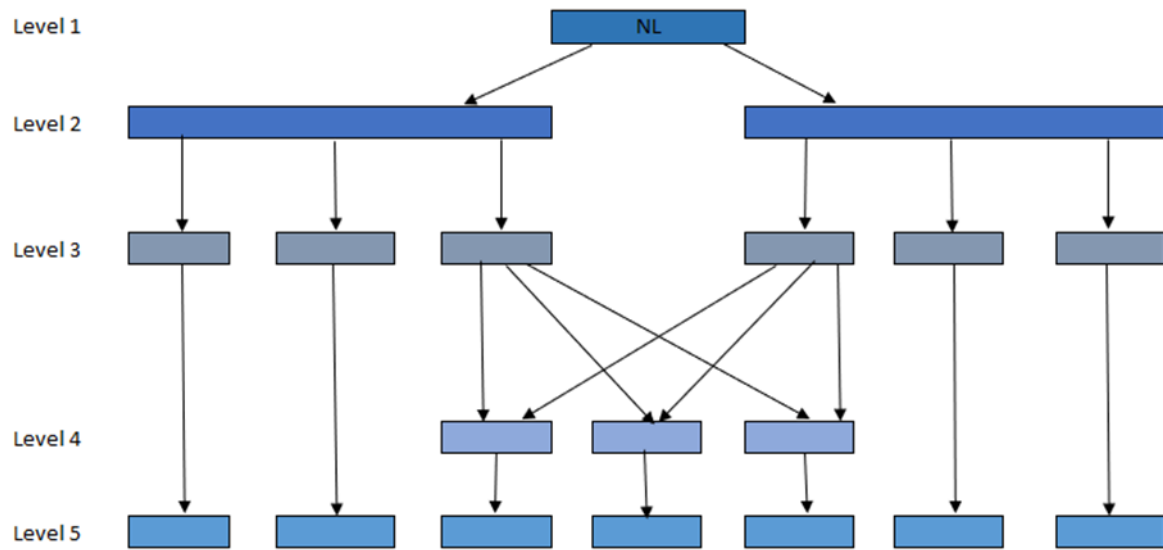


Figure 10. NLx1 second version.

The second version of the tax code chart is pictured above. A request was made to have the decision tree start by checking which G/L account is used in the PO to determine whether the item is a goods or a service. Once the item type is established, the next step is to determine the country of the vendor. This information will then lead to the correct tax code, with the domestic tax codes overlapping each other for goods and services at the VAT percentage checkpoint. The template of the tree also evolved with this request, as now the nodes of the same data type are positioned on the same level of height and will continue to do so for all future versions.

Other persons got involved with the project for feedback and task assignments. The G/L account specialist created a new list in which all the G/L accounts used in AP are defined as relating as goods, services, or both if applicable. After the second version was presented to the management team, it was decided to include some additions to the tree. The feedback was addressed where the key points were to add the delivery country and Incoterms, as well as some nodes where the automation will stop and not proceed further to conclude a tax code for the invoice.

I mentioned the concern about possible incorrect G/L accounts inserted in the PO by purchasers, which would affect the flow and correctness of decisions at the beginning stages of the tree. New trainings for purchasers were scheduled regarding the importance of correct G/L accounts. The meeting where the second version was presented was held on 18.8.2020.

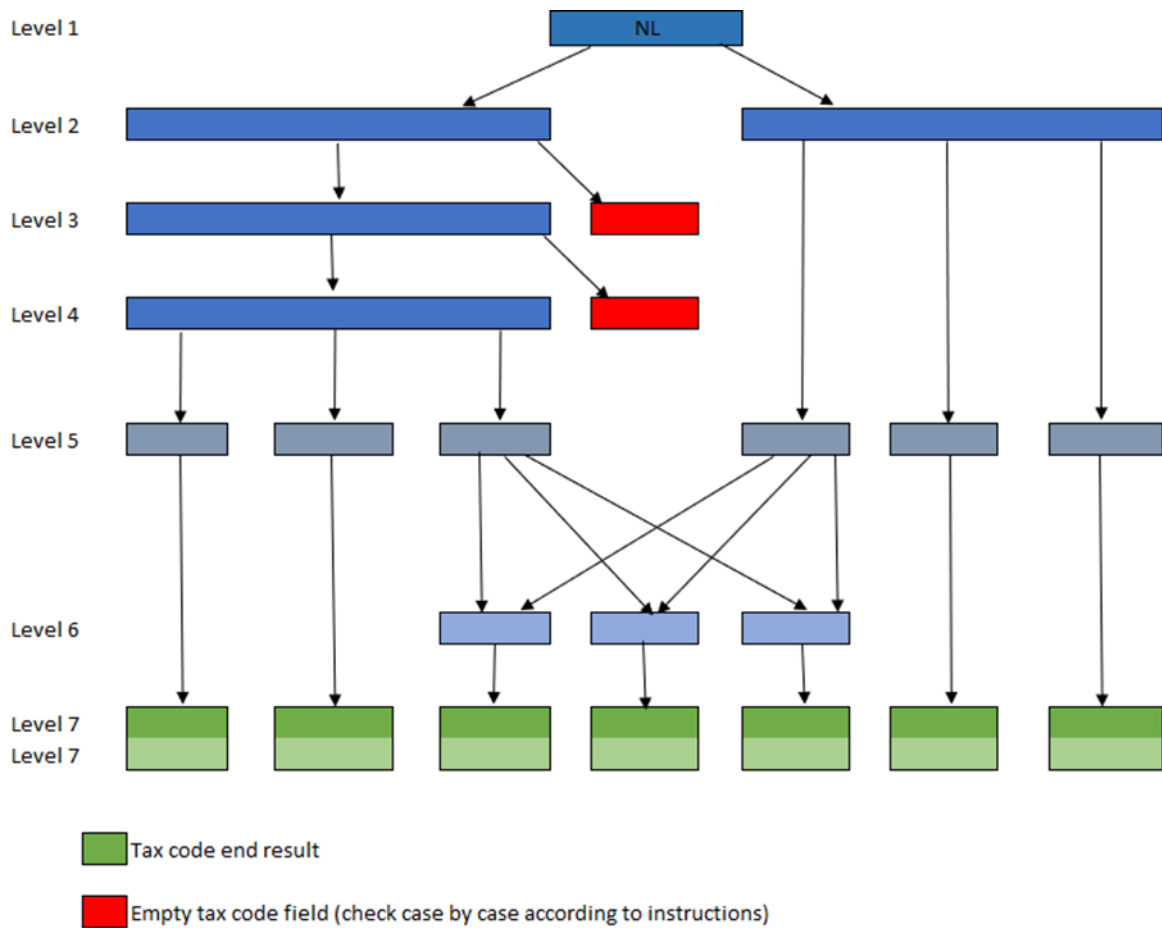


Figure 11. NLx1 third version.

The figure above is the third version of the Dutch tax code tree, which was approved by my manager and the RPA expert. In this new version, the services branch remained unchanged and is the same as in the second version. The goods branch has the addition of Incoterms under the level of goods. Ex Works (see chapter 4.1) is a special kind of Incoterm that is not recommended for use in the case company, however it is still quite commonly applied. If the Incoterms in the PO are anything else than EXW, then the robot can continue further down the branches towards a tax code. Otherwise the robot should stop, and an accountant will insert the tax code for that invoice, which may require assistance from the purchaser.

The other addition is the country of delivery, which currently is not checked by accountants when booking invoices for the Dutch company codes. Most of the invoices addressed to the Dutch company codes have the Netherlands as the delivery country and therefore it does not affect the choice of tax code for most of the invoices. However, the PO data of the past half year does contain delivery countries that are outside the Netherlands, for which the incorrect tax code is a certain result if the delivery country is not considered. With the delivery country added to the decision tree, the automation will check if the goods are delivered to the Netherlands, in which case the robot will continue down the branch. If the delivery country

is anything else than the Netherlands, the automation will stop and let an accountant insert the tax code to the invoice.

A final addition was made at the very bottom of the chart, which is the reporting country that is linked to the actual tax code. The reporting country is useful for company codes that are VAT registered in multiple countries and are selected in SAP by accountants to select the tax code that belongs to that country. Every country has different tax codes and by selecting which reporting country should be used, the options for tax codes are narrowed down and it is easier to find the correct one.

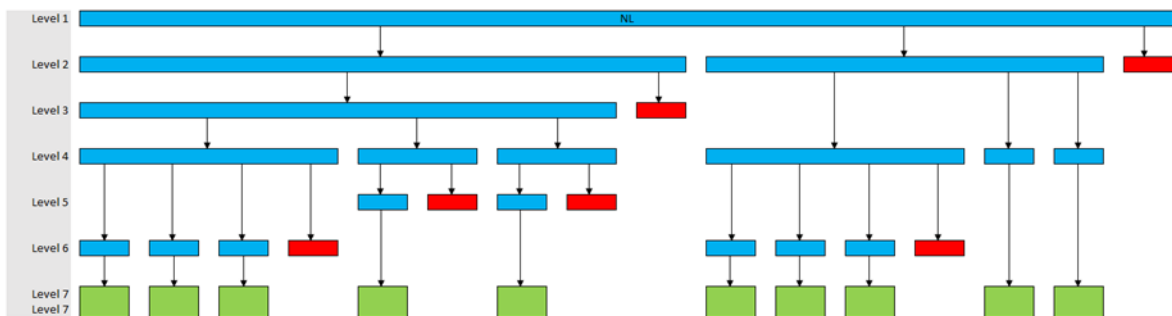


Figure 12. NLx1 final version.

The Incoterms had to be studied to gain knowledge on how they are applied properly to the tree. After realizing that the EXW rule does not affect domestic purchases, the Incoterms checkpoint was moved further down below the vendor country level. The decision tree figure above displays the complete chart and a split between goods and services in the following two figures for clarity. The width of the nodes was modified to clarify data sections, with straight arrows leading towards the next node, instead of diagonal arrows which could become visually confusing for other company codes. Another addition suggested by the RPA expert are red nodes that indicate a dead end for the robot to prevent the automation from running stuck in a loop in cases where it could not find certain data correctly.

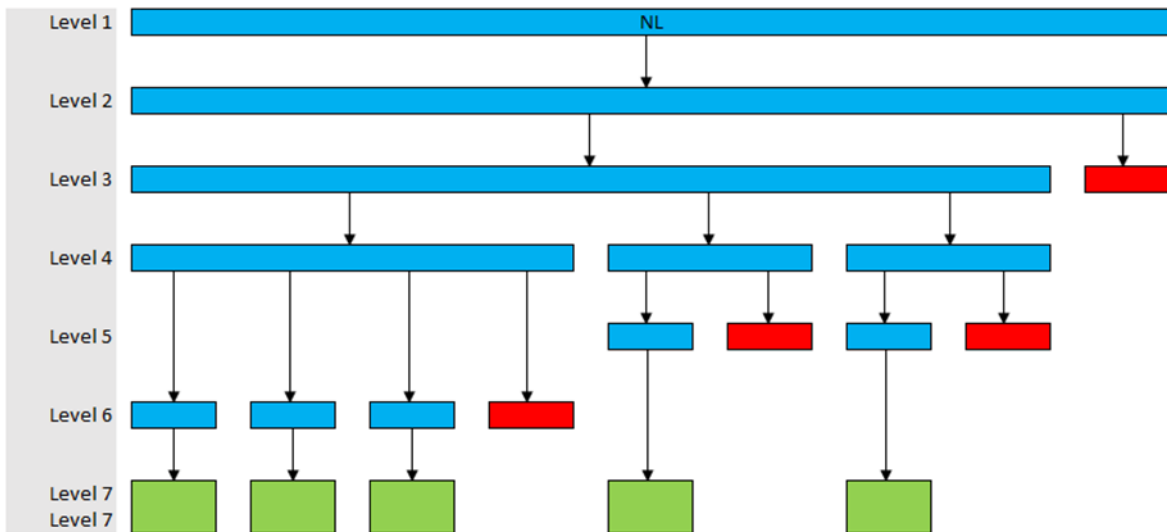


Figure 13. NLx1 goods section.

If the G/L account in the PO is one for goods, the next step is to check the delivery country. If the country of delivery is the Netherlands, the robot proceeds to the vendor country, and if it is a non-domestic country, the automation stops without a tax code result. Continuing at the vendor country, if the vendor is in either the EU or outside of the EU and the Incoterms are EXW, the automation will also stop without a tax code. If the Incoterms are anything else than EXW, the automation can confidently determine the correct tax code as displayed in the green node at the bottom. The Incoterms have no impact when the vendor country is domestic and can therefore skip this level. The automation checks which VAT percentage is used on the invoice and concludes the correct tax code based on that. In case any amounts are written incorrectly by the vendor on the domestic invoice and the VAT percentage does not match with any of the country's VAT percentages, the path leads to the red "Other" end node, where the automation will stop.

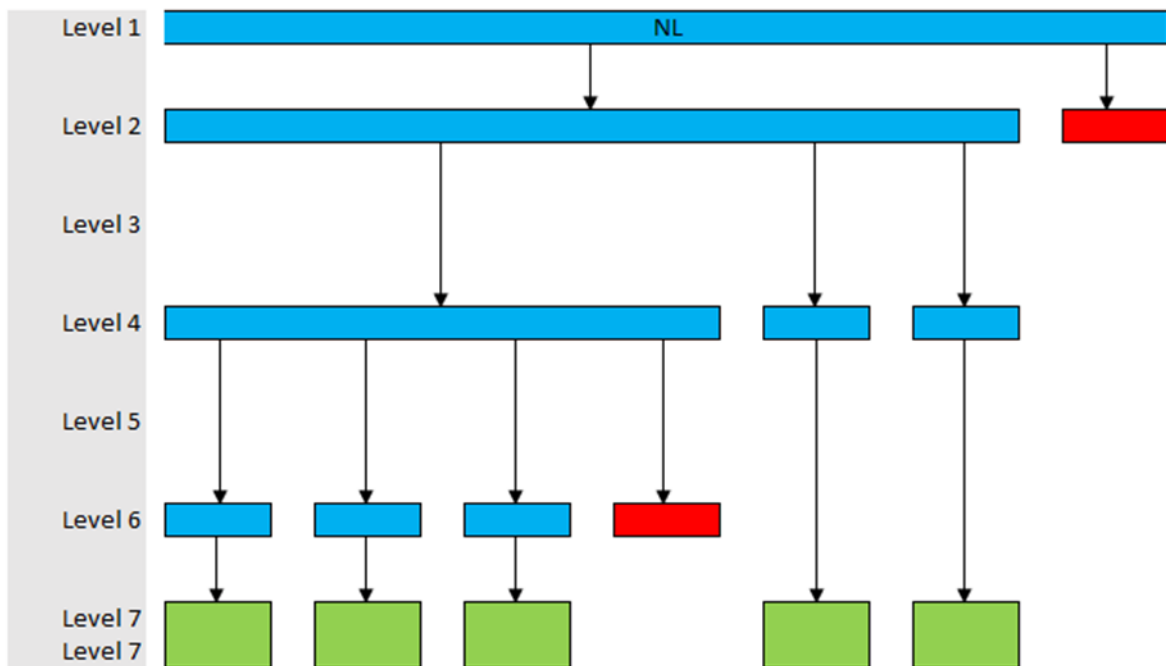


Figure 14. NLx1 services section.

When the automation decides that the G/L account inserted to the PO belongs to services, the process is almost identical to the goods section, with the exception that no Incoterms are involved when services are performed. It is possible that a G/L account is not categorized as belonging to either goods or services, or even that the account is used for both categories. In that case there is a red end node which stops the automation at an early stage from concluding a tax code for the invoice. Currently an investigation is ongoing where these G/L accounts are examined with the aim to identify if new G/L accounts should be created to improve the clarification of the accounts and eliminate the possibility for the automation to end early on the decision tree.

Figure 15. NLx1 decision tree in automation software (removed).

The figure above is a small sized screenshot from the robot in automation software, where the layout of the decision tree is identical to the decision tree made in Excel. The layout of the created decision tree was confirmed by the RPA expert to be successful and simple to transfer to the automation software. The nodes in the automation have matching text with those in Excel and the legend categories are named at the lines between the nodes in the

automation. The connections to the nodes have the data criteria that the robot follows downwards to the tax codes. With the matching layouts in both Excel and the automation, any necessary adjustments are easy to copy over to either templates.

11 Tax code decision tree for DE companies

It was decided that there should be two different templates for tax code logic mapping to test if the established main framework of the chart would be applicable to other company codes. Germany was chosen as the country for the next company codes, which involve more complex decision points and has at the time of writing this thesis temporary tax codes in use that would require solutions in the decision tree for visual indicators. The entities of the case company located in Germany are VAT registered in a few different countries, meaning that in specific cases different tax codes of other reporting countries will have to be used, resulting in a much larger decision tree which is visually more complex to follow. The German tax code decision tree was made in cooperation with an accountant responsible for booking invoices for the German company codes. The tree is split into a few separate parts as it would otherwise be too large and confusing if all the parts were displayed as one.



Figure 16. DEx1 final version.

Above figure displays how the main tax code decision tree looks like. Comparing the German tree to the Dutch one, there are many similarities with a few important differences.

Figure 17. DEx1 decision tree in automation software (removed).

The figure above is another small sized screenshot from the automation software, depicting how the German tax code decision tree was implemented on the automation side and is

identical to the layout in Excel. As the picture is so wide the text is unreadable, but the structure is clearly visible to correspond to the decision tree displayed in figure 16.

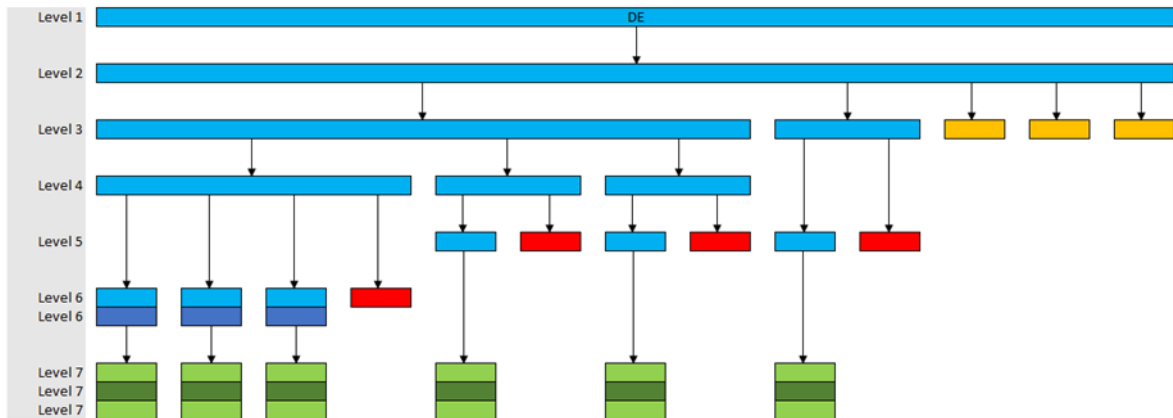


Figure 18. DEx1 goods section.

As with the Dutch chart, also the German has as the first checkpoint to determine if the invoice is for goods or services. If it is for goods, the next checkpoint is delivery country. While in the Dutch chart the automation stopped if the delivery country is anything besides domestic, for Germany the automation can proceed further. When the delivery country is Germany, the same exact logic applies as in the Dutch decision tree. The German tax code chart is fascinating as well as it has temporary VAT rates that are valid until the end of year 2020, due to the Covid-19 pandemic. It is not yet sure if the VAT rates will change back to what they were before the pandemic after the year is over or not. When the delivery country is outside of Germany and the Incoterms are not EXW, the result is one specific tax code for the invoice.

In the delivery country checkpoint, there are additional nodes of specific countries for specific German company codes, which trigger for the automation to continue from another separate decision tree that is not displayed in the main chart. The criteria for this occurrence are when the company is VAT registered in a different country where the delivery is being made, thus making the purchase domestic for that country. The nodes that trigger the robot to proceed further with another part of the tree are colored in orange and explained at figure 20 in this chapter.

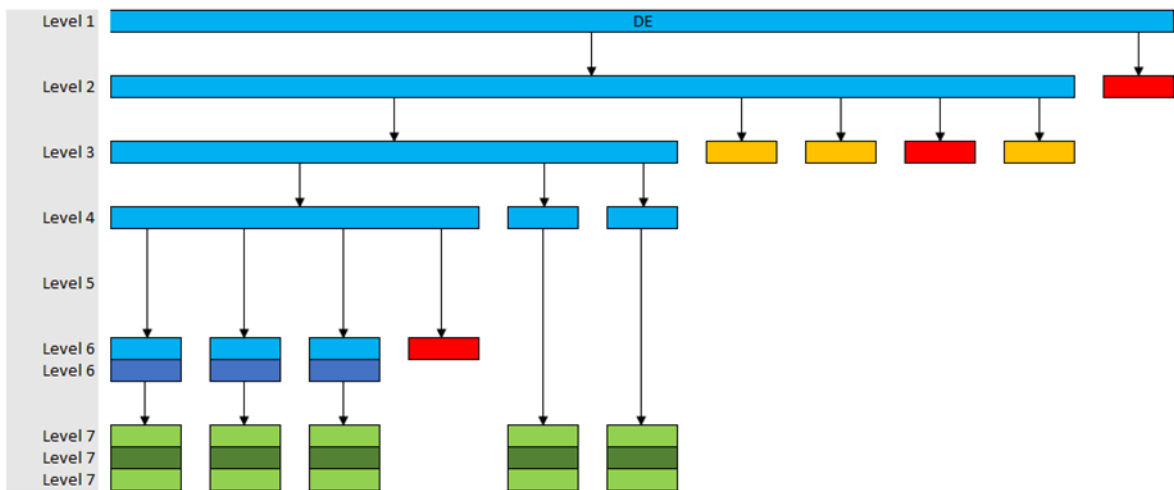


Figure 19. DEx1 services section.

The differences in the service branch between the Dutch and German trees are few with both charts being very similar to each other. The delivery country nodes contain another specific country, Malaysia, because the invoices related to that country are always of services and are therefore not found at the section under goods. The node that triggers for company code DEx2, with Finland as delivery country for the services costs, is indicated as an end node with the reason that the tax code instructions for this occurrence are still under discussion with the local persons employed by the company code.

The other difference is that for services by domestic vendors a reverse charge can be applied, which is indicated on the invoice by a specific text, “Steuerschuldhaft des Leistungsempfängers “Gem. §13b Abs 5”. The reverse charge applies for costs related to scrap metal and purchases of electronics with a value of over 5 000 euro in a single invoice. When an invoice is found to fulfil these criteria, the automation will stop without a tax code. Implementation of this difference is still under consideration with a possible solution for teaching Kofax to recognize the text from the invoice PDFs or a vendor specific approach.

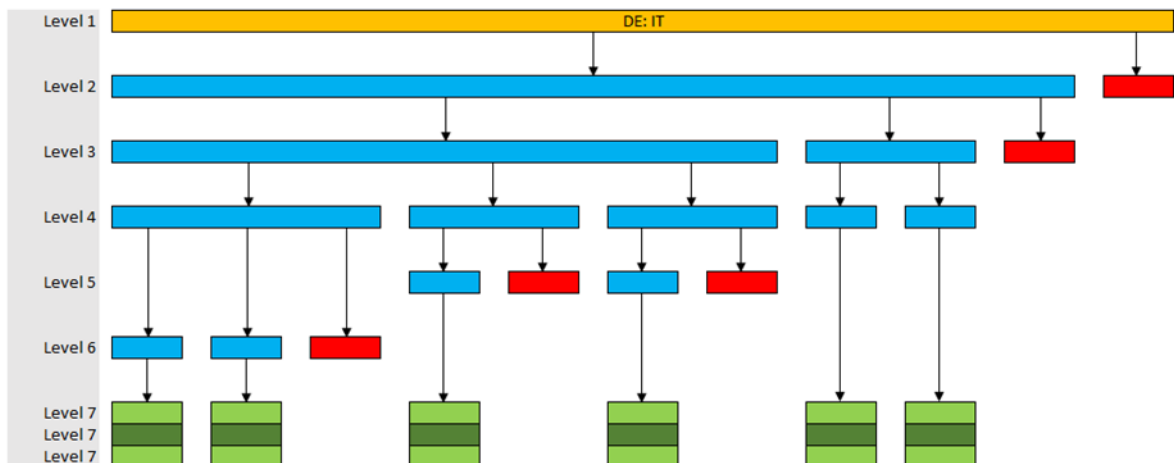


Figure 20. DEx1 Italy section.

The figure above is the decision tree part that the automation follows when the data of the invoice is of company code DEx1 with the delivery country Italy. The orange node found in the main chart links to this displayed decision tree that is added separately to keep the sections compact and simple to follow the data criteria. The company code DEx1 is VAT registered in Italy, so the first checkpoint is to confirm that the company's Italian VAT number is correctly mentioned on the invoice as legally required. The automation will stop if the requirement is not recognized on the invoice and an accountant will instead investigate if the invoice fulfills the requirements or if the invoice is denied for payment.

If the purchase is for a service, there is one tax code for if the vendor country is domestic (Italian) or an EU member country, and a different code for vendors located outside the EU with a temporary tax code. When the invoice is for a purchase of goods and the vendor is not domestic, the same rules apply as before. Interesting to note is that if the vendor is from the EU, the tax code to apply is of Italian reporting country and not German, even though the company code is German. The purchase is considered entirely domestic when goods are purchased from a vendor located in Italy and delivered to Italy. In this case the Italian VAT rates apply, and subsequently Italian tax codes of reporting country Italy are to be inserted for the invoice.

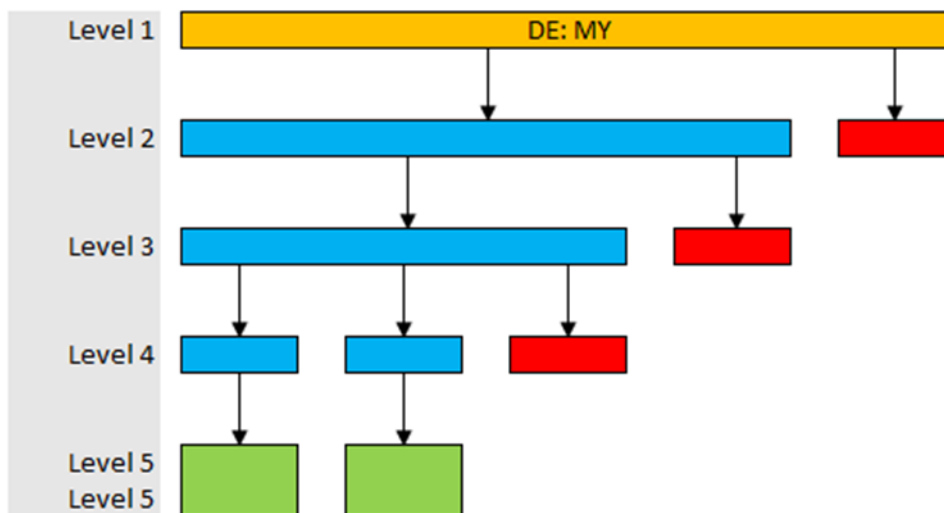


Figure 21. DEx1 Malaysia section.

The other separate section for DEx1 criteria is for when the automation finds that the purchase is for services and is addressed to the Malaysian branch of the DEx1 company code. If the invoice is addressed to Malaysian branch and the vendor is situated in Malaysia, the automation will choose the correct Malaysian tax code depending on whether the VAT

rate on the invoice is of six or zero percent. If the automation runs into other data, such as where the vendor is located anywhere but in Malaysia, the robot will not insert a tax code as the invoice most likely is invalid for processing.

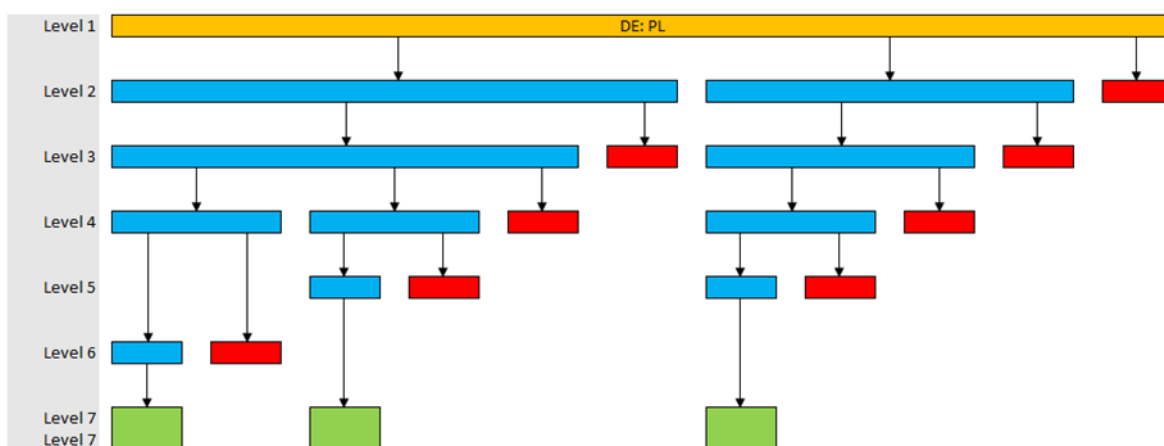


Figure 22. DEx3 Poland section.

Another example for the German tree extensions is when the company code DEx3 has the delivery country Poland. The figures present well to the imagination that if all these extensions were inserted to the main tree, the decision tree would have been very extensive and discouraging for accountants to study the tax code logic.

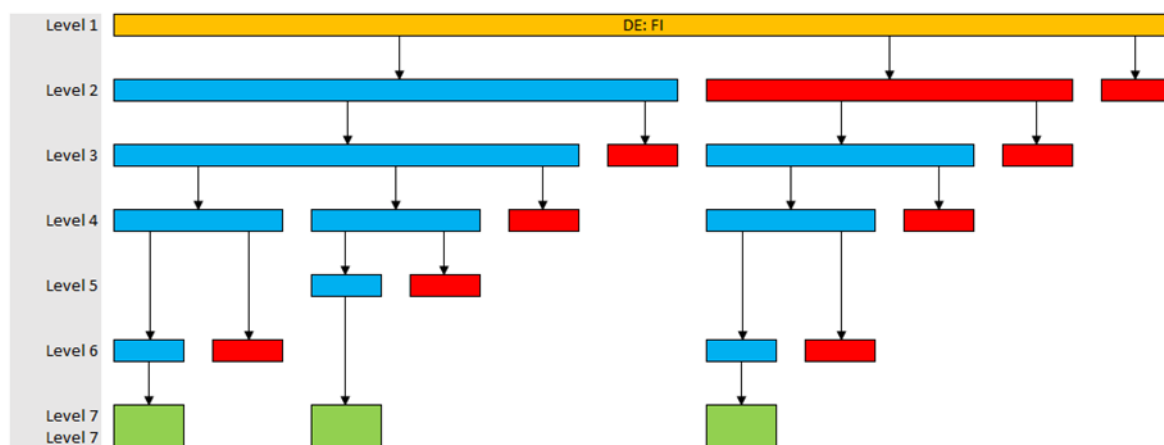


Figure 23. DEx2 Finland section.

A final extension of the German decision tree is with company code DEx2 for Finland as delivery country. The nodes are at this point self-explanatory; however, the node of services is displayed in red with below the node other sections that still flow into a possible tax code Z&. The sections under the end node “Services” are still under discussion and may be changed. Any current invoices for DEx2 that meet the criteria up to this point are pending for processing until accounts payable receives the instructions from the local entity’s employees.

12 Analysis and implementation

This chapter details the logic behind the data visualization of the decision trees, such as its layout and choice of color codes. One subchapter analyses the choice of the pilot countries for this project, their similarities and differences, and which countries could be easy or challenging to map a decision tree for in the near future. Additionally, there is a subchapter describing how the expansion for company codes and countries advanced and continues forward. The final subchapter explains the validation method created for the decision trees to ensure the automation is operating correctly with the provided information and structure.

The decision tree templates for the Netherlands and Germany have been approved by the RPA expert, accountants handling invoices for both countries, three AP managers as well as their superior. The VAT expert of the German company codes has approved the German tax code decision tree and as a result of the tree, accountants recognized mistakes made by either the supplier or purchaser regarding missing VAT when using the Incoterm EXW. The project is considered successful by the persons involved as it has improved the process and progress in the accounts payable department.

12.1 Data visualization of decision trees

The final template of the decision tree uses data visualizations to quickly distinguish between the different types of nodes for both accountants and RPA developers. The width of the nodes is of a specific length, as well as the space between the nodes. The nodes on the higher levels have their width increased to that of the total width of the nodes below. The flow of the nodes is downwards, which is chosen to keep the template compact where it is possible to insert text to the nodes without affecting the size of the chart as it would otherwise if the flow of the nodes were horizontal from left to right. A legend of the node levels is written at the left side of the decision tree in a gray colored panel. The reason behind a legend, as opposed to having text written at the arrow lines, is to remove clutter that would make the decision tree more difficult to comprehend quickly.

The nodes come in a total of four different colors. The nodes that flow to another node on a lower level are colored in blue and require a logical decision based on data variables to decide to which node it will continue next. The red color is used for end nodes that do not result in a tax code and instead result in a blank conclusion, for which the accountant is required to request the local tax expert for advice. The green color represents end nodes that do have a tax code as a result and are located at the lowest end level of the decision tree.

Finally, the orange colored nodes connect to another decision tree that is separated from the main chart to keep the decision tree at a reasonable size and visually representative with the most common factors.

The data visualization supports the aim in creating a standardized agreed template for all new future decision trees by maintaining the structure and modifying the nodes according to the tax code logic of the company codes.

12.2 Choice of pilot countries

The Netherlands and Germany were selected as pilot countries for the tax code decision trees. The two countries are similar in some ways, such as both being member countries of the EU and therefore following the same rules for intra-community trade and the VAT directive. Both countries also have similar domestic VAT rates, however within the case company there are some differences in these two countries. Germany has temporary tax codes in use due to the Covid-19 reduced domestic VAT rates. This made the country more interesting from the automation point of view as well, since it may have posed a challenge that as a result could have required large changes to the decision trees or automations. Nonetheless, the German tree was implemented successfully with the minor adjustment being an extra level to the legend. The German company codes are also VAT registered in different countries within the EU which added for a challenge and more complex decision tree as a result.

With this project there are now templates for the simpler and more complex company codes, depending on the number of tax codes in use and VAT registrations in different countries. The company codes with fewer tax codes will be easier to make decision trees for, while for example Finland will be a great challenge. The Finnish company codes are VAT registered in many different counties and the Finnish VAT legislation is quite complex. The Finnish company codes have 16 tax codes in regular use, which is a big difference in comparison to the seven in use by the Dutch company codes. Countries outside the EU might prove challenging to map with the decision tree, as the VAT legislation may differ greatly from the counties within the EU.

12.3 Expanding country and company code scope

The case company has legal entities at over a hundred locations in 80 countries. Taxation is for most countries the largest source of government revenue and heavily regulated

differently in each country. It is therefore an important process to increase the scope of company codes and have decision trees made for all companies handled in the accounts payable department. The Dutch and German decision trees are both for countries that are within the EU and have therefore a similar structure.

New decision trees for Norway, Sweden, and Denmark are, as of writing this thesis, in development where Norway is associated with but not a member country of the EU. The decision trees are in the making by the accountants who process invoices of the three countries with the completed decision tree template as guidance and my assistance. The addition of new countries increases the company scope to a total of five countries that cover a total of 14 company codes that receive a high overall percentage of invoices in the accounts payable department.

The expanding scope of countries strengthens the knowledge and understanding as accountants must truly understand the logic of the tax codes rather than applying a tax code according to pre-existing instructions. The positive aftereffect of this project is an accurate conclusion of inserted tax code by Kofax and RPA, as well as a decrease in human errors where the accountant is required for tax code input to the SAP field. The accountants that are involved with the decision tree process of their own company codes will also be able to assist other colleagues that join the project once new countries are included after the pilot.

The next projection of plans is to include a country that is not within Europe to identify possible new challenges of data variables to decide the tax codes, while keeping the existing structure of the decision tree.

12.4 Validation

An automated method of validation was necessary to confirm whether the decision trees are concluding the correct tax codes, since the number of invoices and PO item lines are well in the hundreds of thousands and would be impossible to manually validate due to time resources. A solution was developed where an automation imports data of paid invoices that were booked by accountants during the past few months. The data excludes invoices that were processed by robots to eliminate the comparison of data where both inputs came from automations.

The automation examines per vendor which tax code is most used out of the imported data and has the code displayed in the third column of the database table, as shown in figure 24.

A fourth column displays the number of times the most common tax code is used, and a fifth column presents the percentage of use out of all the data. With the table below as example, if invoices of vendor account number 1200 used tax code B& more often than other tax codes, then this tax code is presented in the column “Common TaxCode”. If the imported data contains a total of 50 invoices of which 45 invoices used tax code B&, then the percentage is calculated at 90 percent.

Vendor	#Invoices	Common TaxCode	#Common TaxCode	%Common TaxCode	Expected TaxCode
1200	50	B&	45	90.00	B&
1500	80	B&	60	75.00	?

Figure 24. Tax code validation.

Vendors with a percentage above 80 percent receive in the sixth column the expected tax code which is used to validate the tax code given by the decision tree automation. Vendors with lower percentages are deemed inconclusive and no expected tax code is possible to compare for validation with the decision tree without a closer examination of the data variables or the invoice PDFs. The described developed validation method offers an immediate alert of any mismatches between the tax codes that were inserted by accountants for the invoices and the tax codes that are concluded by the decision tree automation.

Not all invoices of a vendor are posted with a single tax code, thus the method cannot confirm and validate all cases. Vendors with an inconclusive tax code by the validation are displayed with a question mark, as displayed in figure 24, and require manual validation. Any mismatches between the decision tree tax code and validation method tax code require manual validation by the accountant who created the decision tree for that company code. Accountants investigate the reasons for mismatches and report these to the RPA expert, who looks for possible solutions to better the accuracy.

13 Conclusion

The thesis project started with the aim to create a template for tax codes that visually displays the logic in a simple to follow form for both accountants and RPA. The best template for the objective was identified as a decision tree, which indirectly created the decision process of

the automation, as the exact structure is possible to recreate within the automation software by the RPA developer. The created decision tree templates were taken into use for the pilot company codes in the Netherlands and Germany and are being rolled out to company codes in Norway, Sweden and Denmark. Accountants that are involved with the rollouts will be able to assist other accountants in applying the templates to their company codes until all countries are covered.

Once other persons beside myself are able to support and introduce the new templates, the development will speed up a lot more and every company code will have the tax code logic ready for automation. With that in mind, accounts payable has a team of accountants that consist of over 50 persons with a single responsible RPA expert for the department. It is therefore important that the template can be directly copied over into the automations without the need to study the instructions of so many company codes. The chapters covering the decision trees for the Netherlands and Germany demonstrate and verify well that these objectives were met. Below is an evaluation of the reliability and usability of the created tax code logic decision trees and suggestions for possible further improvements in the future.

13.1 Reliability and usability

The decision tree template is considered reliable, as the checkpoints are well thought out and precise, so the robot will not conclude a tax code for an invoice if it does not fulfil the specific criteria outlined. With the developed validation process, it is possible to follow up on the tax code logic's accuracy and correctness. The tree has been approved by management and the RPA expert, and received positive feedback from the accountants that were chosen from the pilot company codes. The accountants had few challenges to implement the new company codes to the template, which speaks for a well-developed structure of the tree that is possible to use by the department.

Statistics on automation data allows the RPA expert and accountants to review situations where for example specific vendors are not given a tax code. It is possible to analyze and identify required improvements to the process or modifications to the decision trees created and sent in by the accountants, thus increasing the accuracy and successfully enabling the robots to post more invoices.

The resulting decision tree template of this thesis is highly usable and is already taken in use. It has saved a lot of time and resources that were not available at the time in the department. The thesis project offered a solution to a roadblock that was halting the progress of invoice

automation that still required correct tax codes to the invoices. The decision trees will be used by the automations as well as accountants, supporting them in choosing the correct tax code by following the data criteria decision rather than studying the mixed and sometimes incomplete instructions present for most company codes without this template. The template structure will standardize the tax code logic for all countries, which makes it easier for new employees or substitutes to book invoices with correct tax codes, without a steep learning curve.

13.2 Recommendations for further improvements

The project is still ongoing at the time of finishing this thesis. Improvements can be made to the template in the form of finetuning or adding new checkpoint levels when seen necessary once other company codes are taken into the scope. The booking of invoices is very heavily dependent on the data in SAP, especially the PO data. Accountants and RPA cannot improve the quality of the existing data as the PO data is maintained by purchasers and vendor data by another department. The correctness of the data is the main challenge for automation as it is already for accountants booking the invoices manually. Automations will now be able to collect all relevant data and identify mistakes made outside the accounts payable department. A main focus would be to improve the accuracy of data related to the tax code logic project, provided by purchasers, which are Incoterms, delivery country, G/L account, and quantity and price.

As the project continues and more collected data is analyzed, then based on those statistics it may be decided that certain G/L accounts may need to affect the automation differently. Currently some G/L accounts are categorized as both goods and services, or neither. When the automation runs into one of these G/L accounts, the robot will not be able to assign a tax code for the invoice. Should this become a common occurrence, the solution could be to define the accounts differently or even split them into separate categories.

14 References

- Avalara. (2020). *2020 European Union VAT rates*. Retrieved from <https://www.avalara.com/vatlive/en/vat-rates/european-vat-rates.html>
- Belastingdienst. (2020). *Exemptions*. Retrieved from https://www.belastingdienst.nl/wps/wcm/connect/bldcontenten/belastingdienst/business/vat/vat_in_the_netherlands/calculating_vat/exemptions
- Brandeis University. (2020). *Fixing Image-only PDFs: Text/Character Recognition in Acrobat DC*. Retrieved from <https://www.brandeis.edu/web-accessibility/pdfs/image-to-text-pdfs.pdf>
- Bunn, D. (2020, June 16). *Germany Adopts a Temporary VAT Cut*. Retrieved from Tax Foundation: <https://taxfoundation.org/germany-temporary-vat-rates-cut/>
- Corporate Finance Institute. (2020). *Decision tree*. Retrieved from <https://corporatefinanceinstitute.com/resources/knowledge/other/decision-tree/>
- Hakapää, S., Huhtala, H., Johansson, A., Lindholm, M., & Paavolainen, J. (2016). *Kansainvälinen arvonlisäverotus*. Helsinki: Hansaprint Oy.
- International Chamber of Commerce. (2019). *Incoterms 2020*. Bielsko-Biała: Dimograf.
- Kofax. (2020a). *About Kofax*. Retrieved from <https://www.kofax.com/About/our-company/about-us>
- Kofax. (2020b). *What is OCR and How Does It Work in Kofax Software?* Retrieved from <https://www.kofax.com/Products/power-pdf/Learning/what-is-ocr-and-how-does-it-work-in-kofax-software>
- Labarre, O. (2020, July 7). *Enterprise Resource Planning (ERP)*. Retrieved from <https://www.investopedia.com/terms/e/erp.asp>
- Le Clair, C. (2017). *The RPA Market Will Reach \$2.9 Billion By 2021*. Retrieved from <https://www.forrester.com/report/The+RPA+Market+Will+Reach+29+Billion+By+2021/-/E-RES137229#>
- Railas, L. (2020). *Incoterms 2020 käyttäjän käsikirja*. Helsinki: Hansaprint Oy.
- Segal, T. (2019, September 1). *Decision Tree*. Retrieved from <https://www.investopedia.com/terms/d/decision-tree.asp>
- Tax Administration. (2020). *International trade within the EU*. Retrieved from <https://www.vero.fi/en/businesses-and-corporations/about-corporate-taxes/vat/international-commerce/intra-community-trade/>
- The Council of the European Union. (2020). *Council Directive 2006/112/EC of 28 November 2006 on the common system of value added tax*. Retrieved from EUR-

Lex: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A02006L0112-20200101>

UiPath. (2020). *Robotic Process Automation (RPA)*. Retrieved from <https://www.uipath.com/rpa/robotic-process-automation>