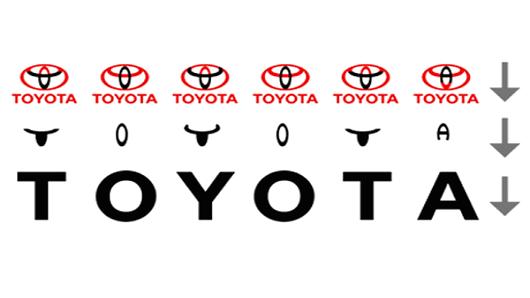
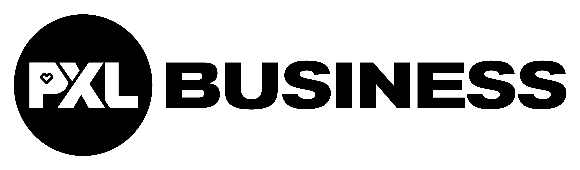


**Optimization of IT systems to conform to Industry 4.0. standards by using Master Data.**

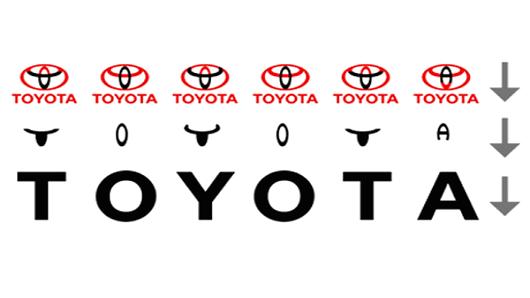


**Sara De Moor 3rd year Logistics Management**

**Promotor: Nadine Mertens Academic year 2019 – 2020  
Mentor: Tom Scheepers, TPCE, Diest**



**Optimization of IT systems to conform to Industry 4.0. standards by using Master Data.**



**Sara De Moor 3rd year Logistics Management**

**Promotor: Nadine Mertens Academic year 2019 – 2020  
Mentor: Tom Scheepers, TPCE, Diest**

# Preface.

In order to help Toyota Parts Center Europe (TPCE) in Diest work towards Industry 4.0. and Master Data management, I made a Critical Data Element analysis and helped to start up the Data catalogue a.k.a. Business Glossary for the Transportation Department. They will use the Oracle Transportation Management (OTM) in the near future, instead of the New Parts Application (NPA) mainframe.

In order to help facilitate this transition and all the next ones, the Atlas project was started.  
This overarching project will help all departments within TPCE and eventually TME (Toyota Motor Europe) transition to a system that is conform to Industry 4.0. standards. The Business Glossary will be used throughout the entire Atlas project, across all departments.

In order to achieve this, I worked as a part of the Atlas Project group, the OTM Project group, the Transportation Department and the Master Data Management (MDM) Project group. The latter which consisted of people from the IT department and me, representing the business side.

**Acknowledgements.**

I would like to thank my mentor team, Tom Scheepers and Kris De Muylder. They always made time to answer my questions and to discuss my progress.   
I would also like to thank my Promotor, Nadine Mertens for her support and encouragement and PXL for presenting me with this wonderful opportunity.

Another big thank you to the following people:   
Rosalie Duivenvoorden, a member of the OTM team, for answering my questions, working with me, basically for letting me function as a full member of the OTM team.  
Sakae Iwasawa, a senior specialist with whom I worked a lot in the MDM team and Pascale De Raes from OTM, for sharing their knowledge and experience with me, to fuel this research.

A blanket thanks to all the office colleagues that helped me by explaining Japanese terminology to me.

Finally, a big thank you to my husband, for everything. Also to my children, Ian and Alexandra, who were super supportive, despite their young age and the rest of my family. Finally, a big thanks to my dear beta readers, Nick Briers, Joyce Streignaerts and Kristine Totten.   
This experience was truly wonderful and I have learned so much.

Sara De Moor  
Wimmertingen, 3rd of June 2020

# Abstract.

The objective of this bachelor thesis was to help TPCE conform to Industry 4.0. standards by ways of efficient Master Data Management (MDM). This will be done in waves across departments.   
First up is Transportation. They are transitioning to the Oracle Transportation Management program (OTM). All data, from the old mainframe (NPA), the Warehouse Management System (WMS), Application System 400 (AS400), various Excel files and Access databases needs to be transferred. In order to achieve this, all parties on the business side and IT-side need to work together to establish which data is important.  
  
The business side determines which data is most important to ensure smooth operations and IT provides the framework in which this can happen. The former is most prevalent in this research.  
The first step here is cleaning up the data, meaning duplicates and irrelevant data -like old addresses and fax numbers- are removed and the validity of data is confirmed.  
The next step is to consolidate all data that made the cut, into a Critical Data Element (CDE) file. This file serves as the base for a Business Glossary, also called a Data Catalogue. All departments will contribute to this catalogue and the eventual end goal is to upload this catalogue into a -to be determined- Data Catalogue Tool, like Collibra. This will provide a single, comprehensive User Interface for use across all the departments.

# Index.

[Preface. 4](#_Toc42160237)

[Abstract. 5](#_Toc42160238)

[Index. 6](#_Toc42160239)

[List of figures. 9](#_Toc42160240)

[List of Tables. 10](#_Toc42160241)

[List of abbreviations. 11](#_Toc42160242)

[Introduction. 12](#_Toc42160243)

[1. Toyota Parts Center Europe. 13](#_Toc42160244)

[History. 14](#_Toc42160245)

[Activities and responsibilities. 16](#_Toc42160246)

[Main values and drivers. 17](#_Toc42160247)

[Operational ways. 17](#_Toc42160248)

[Toyota Production System. (TPS) 18](#_Toc42160249)

[2. Materials and Methods. 21](#_Toc42160250)

[2.1. Where did the material come from? 21](#_Toc42160251)

[2.2. Action plan. (PDCA- Deming cycle) 21](#_Toc42160252)

[Chapter 1: What is Master Data and Industry 4.0. 22](#_Toc42160253)

[1. Industry 4.0. 22](#_Toc42160254)

[2. Big Data. 22](#_Toc42160255)

[2.1. Big Data definition. 23](#_Toc42160256)

[2.2. Transactional Data. 23](#_Toc42160257)

[2.3. Log Data. 24](#_Toc42160258)

[2.4. Metadata. 24](#_Toc42160259)

[2.5. Master Data. 24](#_Toc42160260)

[3. Master Data. 24](#_Toc42160261)

[3.1. Master Data definitions. 24](#_Toc42160262)

[3.2. Five categories of Master Data. 25](#_Toc42160263)

[Chapter 2: Master Data Management (MDM) 26](#_Toc42160264)

[1. Processes in Master Data Management. 26](#_Toc42160265)

[1.1. Source identification and collection. 26](#_Toc42160266)

[1.2. Data transformation. 26](#_Toc42160267)

[1.3. Standardization and rule administration. 26](#_Toc42160268)

[1.4. Error detection and correction. 26](#_Toc42160269)

[1.5. Data consolidation. 26](#_Toc42160270)

[1.6. Data storage. 26](#_Toc42160271)

[1.7. Data distribution. 26](#_Toc42160272)

[2. Roles and Responsibilities within Master Data Management. 26](#_Toc42160273)

[2.1. Data Owner. 26](#_Toc42160274)

[2.2. Data Steward. 27](#_Toc42160275)

[2.3. Data Producer. 27](#_Toc42160276)

[2.4. Data Consumer. 27](#_Toc42160277)

[3.1. Master Data Model. 28](#_Toc42160278)

[3.2. Master Data Management System. 28](#_Toc42160279)

[3.3. Governance organization. 28](#_Toc42160280)

[3.4. Business Leader Sponsorship. 28](#_Toc42160281)

[4. Materials and methods. 28](#_Toc42160282)

[Chapter 3: MDM implementation at TPCE: Introduction 29](#_Toc42160283)

[1. Current situation at TPCE. 29](#_Toc42160284)

[2. Ideal situation. 30](#_Toc42160285)

[3. Current situation. 31](#_Toc42160286)

[4. The gap between ideal and current situation. 32](#_Toc42160287)

[5. Materials and methods. 32](#_Toc42160288)

[Chapter 4: Steps in Master data implementation: A Critical Data Element (CDE) file for OTM. 33](#_Toc42160289)

[1. Data obtained from NPA. 33](#_Toc42160290)

[1.2. Data cleaning. 35](#_Toc42160291)

[2. Data obtained from R12. 36](#_Toc42160292)

[2.2. Data consolidation. 38](#_Toc42160293)

[3. Data obtained for OTM. 39](#_Toc42160294)

[3.2. Data consolidation. 41](#_Toc42160295)

[4. Data obtained from additional, smaller sources. 42](#_Toc42160296)

[4.2. Data consolidation. 42](#_Toc42160297)

[5. Making the CDE based on all previous work. 43](#_Toc42160298)

[5.1. Risk analysis. 44](#_Toc42160299)

[5.2. Basis for Rules and Responsibilities determination. 45](#_Toc42160300)

[6. Conclusion. 45](#_Toc42160301)

[7. Materials and Methods. 46](#_Toc42160302)

[Chapter 5: Steps in Master Data implementation: Business Glossary a.k.a. Data Catalogue. 47](#_Toc42160303)

[1. Materials and methods 49](#_Toc42160304)

[Conclusion. 50](#_Toc42160305)

[Reference list. 52](#_Toc42160306)

[Websites. 52](#_Toc42160307)

[Attachments. 53](#_Toc42160308)

# List of figures.

[Figure 1: TPCE on the map. 13](#_Toc42077121)

[Figure 2: Present day TPCE. 13](#_Toc42077122)

[Figure 3: TPCE 1992. 14](#_Toc42077123)

[Figure 4: First expansion in 1999. 14](#_Toc42077124)

[Figure 5: Second expansion in 2006. 15](#_Toc42077125)

[Figure 6: New Supply Concept. 16](#_Toc42077126)

[Figure 7: Push and Pull System. 17](#_Toc42077127)

[Figure 8: Toyota Production System. 17](#_Toc42077128)

[Figure 9: The Toyota Way. 18](#_Toc42077129)

[Figure 10: The Deming Cycle. 20](#_Toc42077130)

[Figure 11: Overview of the industrial revolutions. 21](#_Toc42077131)

[Figure 12: Types of Data. 21](#_Toc42077132)

[Figure 13: Big Data's three V's. 22](#_Toc42077133)

[Figure 14: Complete MDM program. 27](#_Toc42077134)

[Figure 15: Forwarders Master Data- Ideal situation. 29](#_Toc42077135)

[Figure 16: Forwarders Master Data-Current, AS-IS. 30](#_Toc42077136)

[Figure 17: GAP, Forwarder Ideal and Current situations. 31](#_Toc42077137)

[Figure 18: Forwarders-NPA Part one. 33](#_Toc42077138)

[Figure 19: Forwarders-NPA Part two. 33](#_Toc42077139)

[Figure 20: Forwarders-NPA Part three. 33](#_Toc42077140)

[Figure 21:Forwarders-NPA Part four. 34](#_Toc42077141)

[Figure 22: NPA screen 640: Forwarders. 35](#_Toc42077142)

[Figure 23: AS400-SH39 Forwarder screenshot. 36](#_Toc42077143)

[Figure 24: R12 part one. 36](#_Toc42077144)

[Figure 25: R12 Part two. 37](#_Toc42077145)

[Figure 26: OTM part one. 39](#_Toc42077146)

[Figure 27: OTM part two. 39](#_Toc42077147)

[Figure 28: OTM part three. 40](#_Toc42077148)

[Figure 29: EVOLV Template: Forwarders. 40](#_Toc42077149)

[Figure 30: Bank Account Approval form excerpt. 43](#_Toc42077150)

[Figure 31: CDE lay-out example. 44](#_Toc42077151)

[Figure 32: CDE continued. 45](#_Toc42077152)

[Figure 33: Collibra Business Glossary definition. 46](#_Toc42077153)

[Figure 34: Microsoft Azure example. 47](#_Toc42077154)

# List of Tables.

[Table 1: Transposed NPA records. 33](#_Toc42083220)

[Table 2: Updated Table including new records and data sources. 36](#_Toc42083221)

[Table 3: Updated Table, draft for CDE. 40](#_Toc42083222)

[Table 4: Updated with extra platforms. 40](#_Toc42083223)

[Table 5: Blank Business Glossary template. 45](#_Toc42083224)

[Table 6: Filled out Business Glossary Part one. 45](#_Toc42083225)

[Table 7: Filled out Business Glossary Part two. 46](#_Toc42083226)

# List of abbreviations.

|  |  |  |  |
| --- | --- | --- | --- |
| TPCE | Toyota Parts Center Europe. | |  |
| OTM | Oracle Transportation Management. | | |
| TME | Toyota Motor Europe. | |  |
| NPA | New Parts Application. | |  |
| MDM | Master Data Management. | |  |
| CDE | Critical Data Element. | |  |
| AS400 | Application System 400. | |  |
| WMS | Warehouse Management System. | | |
| PSCG | Parts Supply Chain Group. | |  |
| EMC | European Manufacturing Company. | | |
| NMSC | National Market Sales Company. | | |
| TPS | Toyota Production System. | |  |
| JIT | Just-in-Time. |  |  |
| CRM | Customer Relations Management. | | |
| IBM | International Business Machines. | | |
| TDWI | Transforming Data with Intelligence. | | |
| PhD | Doctor of Philosophy. | |  |
| IT | Information Technology. | |  |
| ID | Identification. |  |  |
| UoM | Unit of Measurement. | |  |
| DPMM | Dealer Parts Management Module. | | |
| TRD | Transportation Department. | |  |
| F&A | Finance and Accounting. | |  |
| V.A.T. | Value Added Tax. |  |  |
| IT&D | Information, Technology and Disabilities. | | |
| PUR | Purchasing. |  |  |
| FWD | Forwarder. |  |  |
| EA | Enterprise Architecture. | |  |

# Introduction.

Times change, and every company wants to stay up to date on the latest digital techniques to improve efficiency and reduce costs.   
One of those techniques is a system called Master Data Management, as a part of reaching Industry 4.0. standards.

In order to implement such a system, clear definitions must be found and then tailored to Toyota’s needs. This will be done by a combination of desk research for the global definitions and field research to better understand what the company wants and needs. This data collection and subsequent analysis will lead to a framework that will be the basis for implementation.  
  
At receiving the subject for the Bachelor Thesis, research questions were formulated.

***Research Question one:* What are the specifications a system needs to qualify for Industry 4.0.?**

1. Is there a determined set of requirements to aid in the vetting process?

2. Which changes need to be made in order for the system to conform to these requirements?

Research Question two: How do we create added value and optimize the supply chain by making changes?

1. How cost effective are these changes? If a change costs more than a potential gain, is it worth it?

2. What is the role Master Data can play in this process?

3. After the changes are implemented, is there an evaluation period to see how day-to-day operations are affected?

*Research Question three*: How do we tackle the implementation of those changes?

1. Will the changes be implemented gradually or radically?

2. What is the role of each individual Atlas Team member in this?

These questions need to be answered. In order to do this, it’s good to get a sense of the company and, more specifically, the branch in need of this digital overhaul.

## 1. Toyota Parts Center Europe.

The Toyota Parts Center Europe, TPCE in short, is part of the Parts Supply Chain Group (PSCG).  
They are the central distribution center for Europe, located in Diest.

This site was chosen for a host of different reasons.

* Excellent connection to main ports and direct access to the International Motorway Network
* Large, local pool of young and skilled employees in a good social climate
* Prioritized support from Interleuven to attract international logistics operations
* Site is suitable for future expansion.

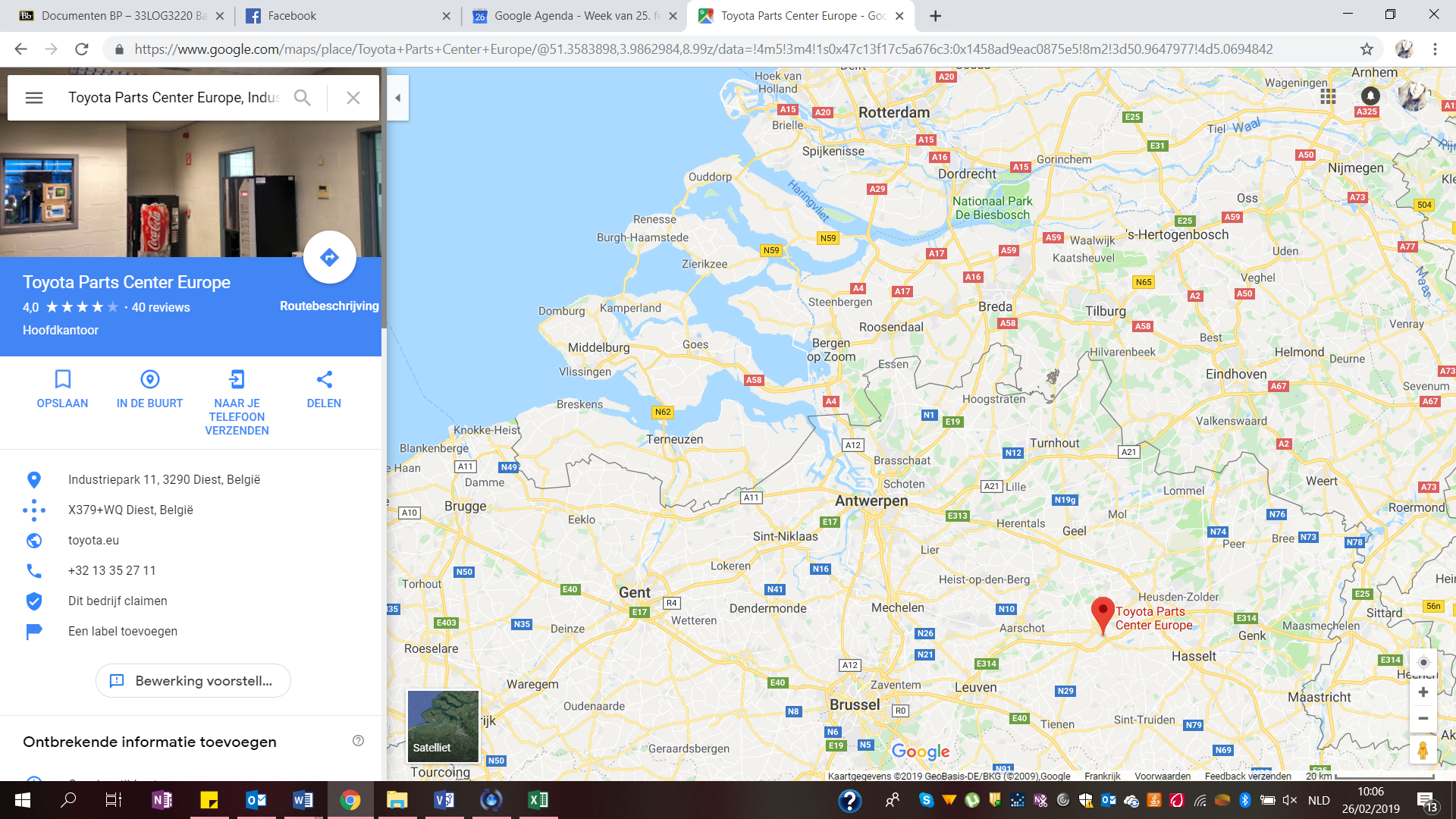


Figure : TPCE on the map.

### History.



Figure : Present day TPCE.

Construction of the Toyota Parts Center Europe started in 1991 and was completed and ready for operations in 1992. It was built on a large stretch of land, 247.300 m². The initial warehouse was 38.000 m², so there was a lot of room left for future expansions.

Some interesting stats:

* 80 employees
* 70.000 part numbers in stock
* Seventeen destinations/Seventeen suppliers



Figure : TPCE 1992.

The first expansion happened in 1999. They built an additional 26.500 m² of surface that would become warehouse 2.  
Interesting stats:

* 176 employees extra
* 100.000 part numbers in stock
* Eight extra destinations
* 44 suppliers



Figure : First expansion in 1999.

The second and last expansion to date added another 30.000 m², happened in 2006 and was called warehouse three

Interesting stats:

* An extra 524 employees, new total 875
* 250.000 part numbers in stock
* 483 destinations
* 682 suppliers.

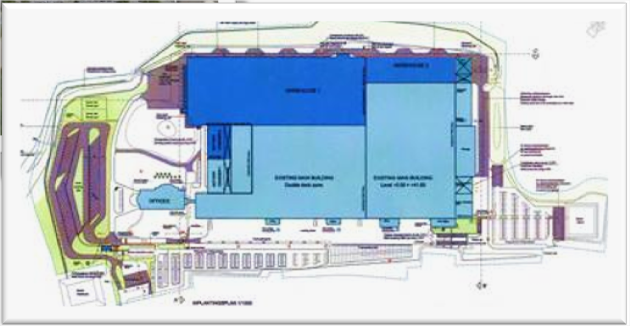


Figure : Second expansion in 2006.

In 2020, that means they have 100 000m² of surface space, 875 employees and an inventory of   
265 000 parts, valued at around 75 million euro.

### Activities and responsibilities.

1. Total Supply Chain Management.  
The entire principle of Toyota’s supply chain is based on the ‘Sell one, Buy one’ principle. Small lots with high frequency. The TPCE gets inbound deliveries from local suppliers, the European Manufacturing Companies (EMCs) and the “POLEs”, the large continental Headquarters.

Outbound they deliver parts to their regional depots, European dealers, POLEs and NMSC (National Marketing and Sales) warehouses. These warehouses stock and distribute TPCE materials but are privately owned. Toyota has no say in how these warehouses are run.

2. Transportation.   
**Central**  
Daily Inbound:

* 103 trucks
* 150 suppliers
* 23 different countries

Daily Outbound:

* 72 trucks
* 42 delivery addresses
* 26 different countries

**Facing** (Direct delivery to Dealers instead of depots.)  
Dealer deliveries:

* 3,348 deliveries
* 23 different countries

#### 3. Regional depots.

They manage the one central warehouse TME (Toyota Motor Europe), thirteen regional depots and they have direct governance over warehouses in Russia, the Ukraine, Kazakhstan and Italy.  
They also manage contact about operations with the NMSC warehouses.

### Main values and drivers.

#### 1. Vision and Mission.

Vision: “Leaders in Part Supply Chain Development”  
Development is a very important word here, as it shows Toyota’s commitment to continual improvement.

Mission: “Deliver the Highest Value Logistics for All our Stakeholders. (Qualitative and quantitative)”  
This means everyone involved needs to get the best possible service and performance.

#### 2. Respect for environment and society.

Toyota is already ISO14001 certified, which means they adhere to the international standards for an effective Environmental Management System (EMS).  
In 2050, Toyota strives to be carbon-neutral in all of its operations.

Community outreach is also very important for them. They usually organize carpooling, tree planting, a green month, but now with these difficult Covid-19 times, Toyota employees are encouraged to volunteer for non-profits that help older people get their groceries, or even to volunteer very locally to help your elderly neighbors yourself.

### Operational ways.

Toyota always strives to reduce fluctuations in supply and demand to stabilize the continuous flows (Seiryuka). They do this by separating urgent and non-urgent orders and by picking the right transport, as shown in the graphic below.

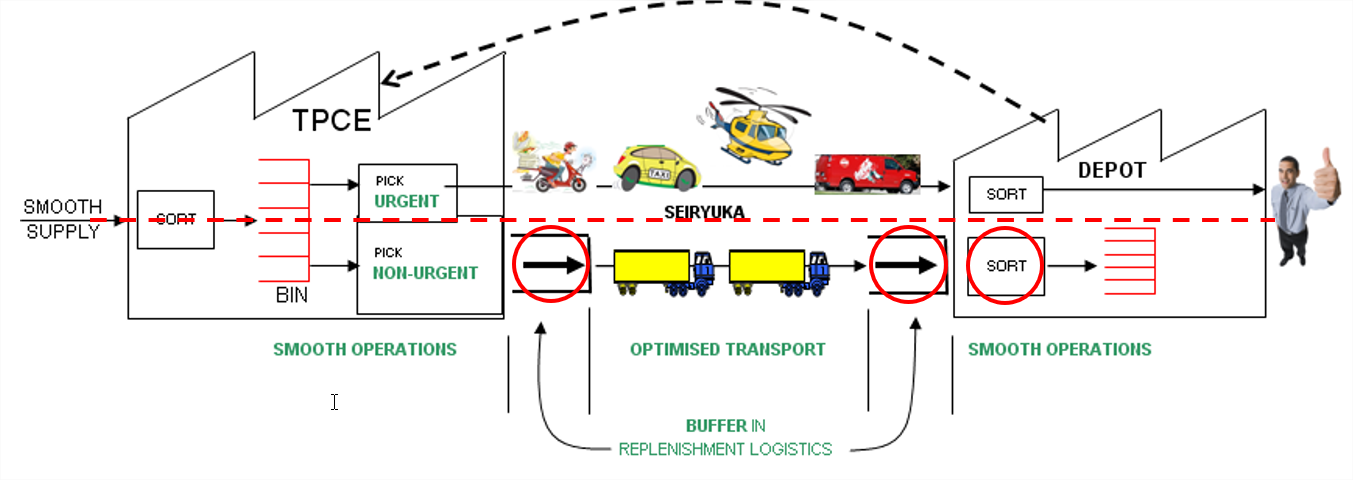


Figure : New Supply Concept.

The “Sell one, Buy one” principle is paramount here, as it explains how Toyota works with the Pull-systems, as opposed to the Push-system in inventory control.  
Push-system: Produce as much as possible, for when it’s needed.  
Pull-system: Only produce when there is demand.

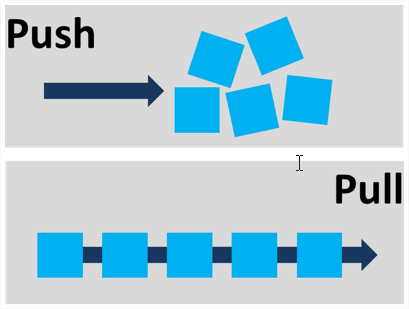


Figure : Push and Pull System.

### Toyota Production System. (TPS)

The Toyota Production System is also known as the Toyota Way, meaning continuous improvement and respect for the people within the organization. This means the right production process will lead to the right products, taking the internal customers and external suppliers into account.

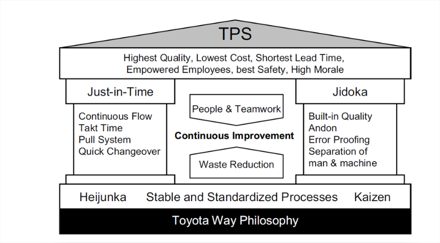


Figure : Toyota Production System.

*Heijunka* is a Japanese word for leveling. It helps to match unpredictable customer demand patterns and eliminate manufacturing waste. TPS also considers all actions that don’t add value for the customer as waste.

*Kaizen* is one of the most important principles in the Toyota Way Philosophy. It focuses on creating a corporate culture where continuous improvement is encouraged and promoted by everyone. Which means putting the employees, customers and suppliers at the heart of the organization.

*Genchi Genbutsu* could be translated to Go and See. Meaning problems on the work floor can’t be solved in an office, but the managers must go and see where the problems are, talk to people and figure out a solution in that way.

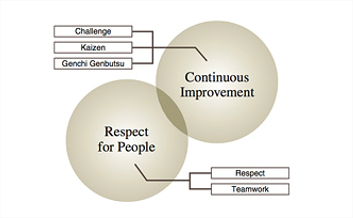


Figure : The Toyota Way.

Lean Manufacturing is an important part of the Toyota Way and is based on creating added value, continuous flows, the pull principle and striving for perfection. Two main principles are identified.  
*Jidoka* or autonomation.

* Discover an abnormality
* Stop the process
* Fix the immediate problem
* Search for and solve the root cause of the problem

*Just-in-Time (JIT)* production is one of the most important principles. This combines most of the other principles, like having a steady flow and be able to supply the customer just-in-time so there isn’t too much stock and those costs can go down. Only downside is everything has to run perfectly, as the slightest delay can cause big problems. So here the Jidoka principle is important too, as problems need to be fixed immediately, as to not disrupt flows.

## 2. Materials and Methods.

### 2.1. Where did the material come from?

This information was mainly supplied by the PowerPoint presentation TPCE uses as an introduction for new employees and a yearly reminder for current employees. To make sure everyone stays aware of Toyota’s drives and values.  
The employees supplied a lot of the information too, especially pertaining company specific Japanese terminology. The Internet provided information to fill eventual gaps.

### 2.2. Action plan. (PDCA- Deming cycle)

Plan.  
Finding out the root of the problem. What is Master Data, how does it relate to Industry 4.0. and how can it be implemented in a way that is most beneficial for TPCE?  
Proffered solutions will be tested and at the end, there will be results. What is a good, workable definition of Master Data and how will the new systems operate from then on?

Do.  
Action is the main word in this stage. How do different departments and people of different ranks react to upcoming changes? Is there clarity?  
This is a good stage to create a platform so everyone can feel and be involved in this process, since they will become the main users of this new system.

Check.  
Have the milestones, defined in the Planning stage, been met? Were there difficulties? Was everything achieved, or did some parts not make it?  
Achieved results must be objectively evaluated so changes and tweaks could be made if the outcome was unforeseen.

Act.  
Analysis and results from the previous stages produce conclusions. Results are best considered in short and long term. Short for immediate changes, long for problems that may surface after prolonged use of new systems.

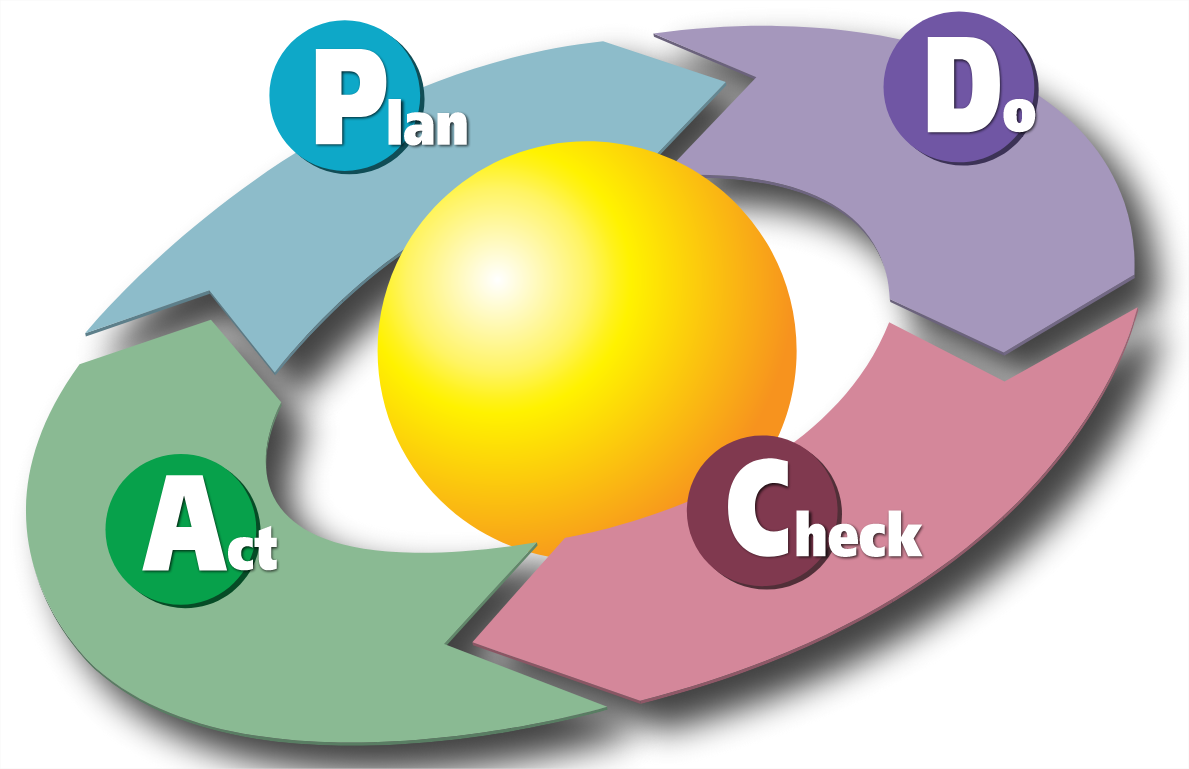


Figure : The Deming Cycle.

# Chapter 1: What is Master Data and Industry 4.0.

## 1. Industry 4.0.

Industry 4.0. is also referred to as the fourth industrial revolution. Building on the previous waves it seeks to add Cyber Physical systems to existing automation and computer systems.

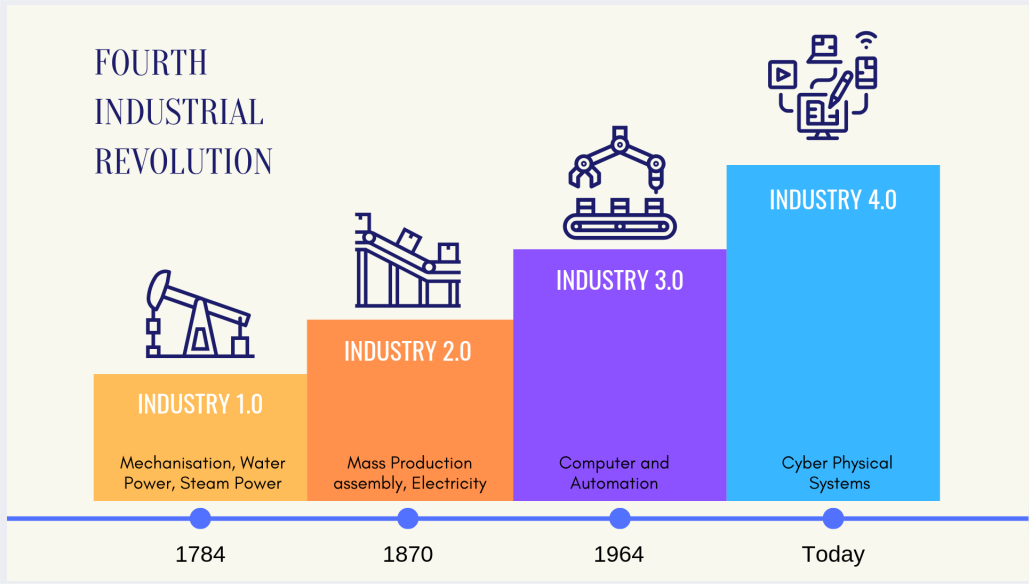


Figure : Overview of the industrial revolutions.

Some of these Cyber Physical systems are mobile technologies, 3D printing and Cloud computing. This also means using Big Data for analytics, predictive maintenance, Customer Relations Management (CRM) and Master Data Management (MDM). In order to make use of these technologies, it’s important to understand them better, so it’s important to formulate clear definitions.

## 2. Big Data.

Big Data is the collective name for data with facts and statistics that are used for analysis and reference. The following graphic will clarify the types of data and how they relate to each other.

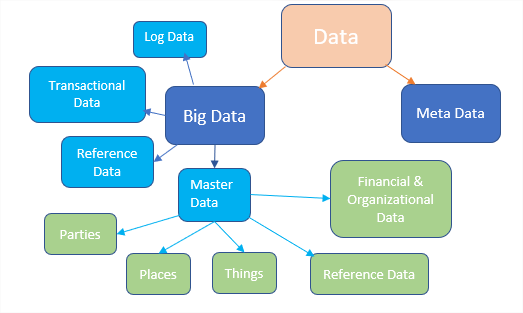


Figure : Types of Data.

### 2.1. Big Data definition.

There are many definitions, but the one offered by Gartner's Doug Laney can be summarized by the three V's: Volume, Variety and Velocity. Gartner is a well respected website concerning data management. More about them later.

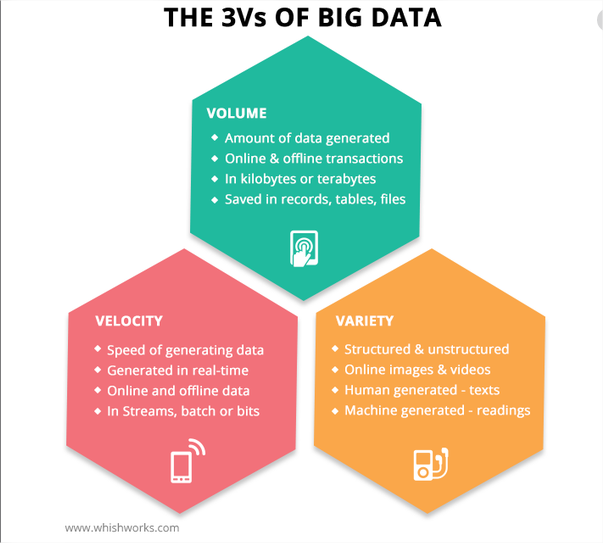


Figure : Big Data's three V's.

Of course, as usual, there are many different definitions of Big Data. Some go as far as to identify six V's, rather than Laney's three. It can be cause of some confusion sometimes.

However, it's always good to consider different perspectives, and the IBM (International Business Machines) Data and Analytics Hub offers a Four V's definition that adds an important V.

Aside from, Volume, Variety and Velocity, they also consider Veracity, or the uncertainty of Data.

Where is the information coming from? Is it correct? Who verifies it? In a survey by IBM, it shows one third of business owners don't trust the information they use to make decisions.

Imagine having to decide to close a branch, as a company, based on faulty sales figures. They could cut into the company’s profits as a whole and the social ramifications of people losing their jobs could be dramatic.

Big Data is a combination of different types of data.

### 2.2. Transactional Data.

This is a huge quantity of data. It documents an exchange, an agreement or a transfer between individuals or companies.

For instance: A receipt in the grocery store. Very important for CRM, a loyalty card that registers all the products a customer has bought, to offer tailored reductions

### 2.3. Log Data.

Log data is a summation of events and snapshots of processes.

For instance:   
Maintenance logs. If a machine breaks down every so often, these logs can determine what the pattern in breakdowns is and a solution can be found.

An even simpler example is an electricity meter. It indicates the amount of electricity consumed at any time, so the electrical bill won’t be a complete surprise.

### 2.4. Metadata.

Metadata is data, about data. For instance, a Word document lists who the author is, where and when it was written. Even though Metadata is not part of Master Data, it has its uses within a Master Data Management System.

For instance:   
When working together on a document, it's good to see which author made changes, on which date. That way it’s possible to determine responsibilities, in case something went wrong or when more explanation is needed.

### 2.5. Master Data.

Master Data is a part of Big Data, but Metadata, Log Data and Transactional Data are not a part of Master Data. Log and Transactional Data are very fast-moving types of data, Master Data is a slow moving type, not prone to too many changes.

There are many definitions, which will be explored further.

## 3. Master Data.

There are many definitions of Master Data and reaching a consensus is not always easy. In that case it might be advisable to consider a few different definitions by authorities in the field and try to find the similarities.

### 3.1. Master Data definitions.

#### Gartner definition.

“Master Data is the consistent and uniform set of identifiers and extended attributes that describes the core entities of the enterprise, including customers, prospects, citizens, suppliers, sites, hierarchies and chart of accounts.”

Gartner is a firm that specializes in advice and research pertaining to business and all that it entails. It’s seen as a trusted source that organize workshops and webinars in a lot of different subjects. One of those being Master Data Management.

#### Philip Russom- TDWI definition.

“Master Data is having consistent definitions of business entities (e.g. customer or product) and data about them across multiple IT systems and possibly beyond the enterprise to partnering businesses.”

TDWI means Transforming Data with Intelligence. Philip Russom is a Doctor of Philosophy (PhD) specialized in Data Warehousing and Integration. The definition is largely same as above but adds that Information Technology (IT) systems could be expanded beyond the enterprise, to its partners.

#### Technopedia definition.

“A clear description of Master Data includes customer data items, like ID’s (Identification). This data type is considered Master Data, versus Quantitative Data related to a single transaction, customer ID or other data (such as addresses and phone numbers), which are continuously used by a business to analyze customer behavior, establish contacts or drive high-level research.”

Again, a clear description, but it focuses even more on smaller details and makes a distinction between Master data <-> Quantitative data. Log Data is a good example of Quantitative Data.  
Techopedia is a trusted website, full of resources for professionals in IT.

Keywords distilled from the three definitions: Clarity, uniformity, well-defined and consistent.

Conclusion: Master Data is slow-moving. Things like name, phone number and address changes don’t happen too often. That’s why it’s important to have a single, correct reference to them.

### 3.2. Five categories of Master Data.

#### A) Parties.

Everyone who conducts business with the enterprise, including, customers, prospects, suppliers and partners.  
For TPCE this means, suppliers, dealers, regional depots and carriers.

#### B) Places.

Actual places and how they are segmented, such as locations, subsidiaries, branches, sites, etc.

#### C) Things.

Everything that the enterprise sells or manages, such as goods, services and assets.   
In TPCE's case the parts themselves, as a prime example.

#### D) Financial and organizational data.

It's about reporting and accounting categories, including organization structures, sales territories, chart of accounts, cost and profit centers and price lists.   
TPCE specific examples: Part price lists, hierarchy within a department, etc.

#### E) Reference Data.

Smaller lists of values that give context to bigger ones, for instance, a list of countries where the customers come from. Also, country codes, corporate codes and Units of Measurements (UoM).

In conclusion: Master Data is a single, clear, uniform, slow-moving and well-defined piece of data that will become the company reference. Also referred to as the ‘Golden Record.’  
Master Data consists of five categories, Parties, Places, Things, Financial/Organizational Data and Reference Data.

# Chapter 2: Master Data Management (MDM)

The first step is to determine what the company’s Master Data is.  
Once Master Data has been identified, it’s paramount to manage it correctly.  
For instance, which department and who within that department deals with address changes?

So, the second step must be to determine specific roles and responsibilities within Master Data Management.

For reference, a definition by Gartner.  
“Business and IT work together to ensure consistency and accuracy in all assets of an enterprise, like parties, places and things and have a good grip on financial and organizational structures and reference data.”

## 1. Processes in Master Data Management.

### 1.1. Source identification and collection.

Where is the data coming from? From which departments?

### 1.2. Data transformation.

Once data is available, it’s time to start cleaning it up. This means checking for, relevancy, clarity and duplicates.

### 1.3. Standardization and rule administration.

The remaining, cleaned data will undergo standardization. For instance, every client number will now be written in the same way, to avoid confusion and mistakes.  
This will become a rule.

### 1.4. Error detection and correction.

People within departments and companywide, will be responsible for detecting mistakes, suggesting corrections and approving those corrections.

### 1.5. Data consolidation.

Corrections will replace the previous, faulty data.

### 1.6. Data storage.

The data will be stored centrally, in a dedicated database.

### 1.7. Data distribution.

This data will be accessed by many different parties inside and outside the immediate company (like Toyota Dealers, carriers, etc.) So, this needs to happen via a clear Interface, like TPCE’s current Carrier Portal.

## 2. Roles and Responsibilities within Master Data Management.

The responsibilities here fall to two, very distinct departments.  
The business side produces the data and the IT-side builds and maintains the systems the business side uses. The main focus here is the business side. Producing in this case also means processing the data received from either customers, suppliers and other relevant third parties.

### 2.1. Data Owner.

A Data producer is a senior person in their department. For instance, a senior manager of the finance department can become the Data Owner of all the Master Data pertaining to the finance department. They know the relevance and accuracy of the data. This person also determines who gets access to the data. They are often assisted by Data Stewards and review correction suggestions by these Data Stewards.

### 2.2. Data Steward.

A Data Steward is a knowledgeable person within a department and performs the following tasks.

* Check Data fitness: is the used data fit for its purpose? A purely manufacturing plant does not need to know the names of the end consumer, the people that will actually drive the car.
* Prunes unused data, for instance, no use in keeping the old address of a dealer in the system.
* Guards the consistency and correctness of the data.

Difference between a Data Owner and Steward? Data Stewards usually deal with the day-to-day operations. They will draw up data quality rules that could be approved by the Data Owner.

### 2.3. Data Producer.

Usually all the people within the department that are responsible for logging the data from new customers, carriers, suppliers and dealers.

### 2.4. Data Consumer.

These same departments also make up the bulk of the data users/consumers. Master Data like phone numbers are used to call dealers, but it also works for the dealer themselves.  
TPCE is working on DPMM (Dealer Parts Management Module), which will allow the dealers to easily order the parts they need and check their availability.   
  
3. Master Data implementation.   
The definition and roles are clear, which paves the way for the hardest part, the implementation. This is what TPCE considers a completed MDM program.

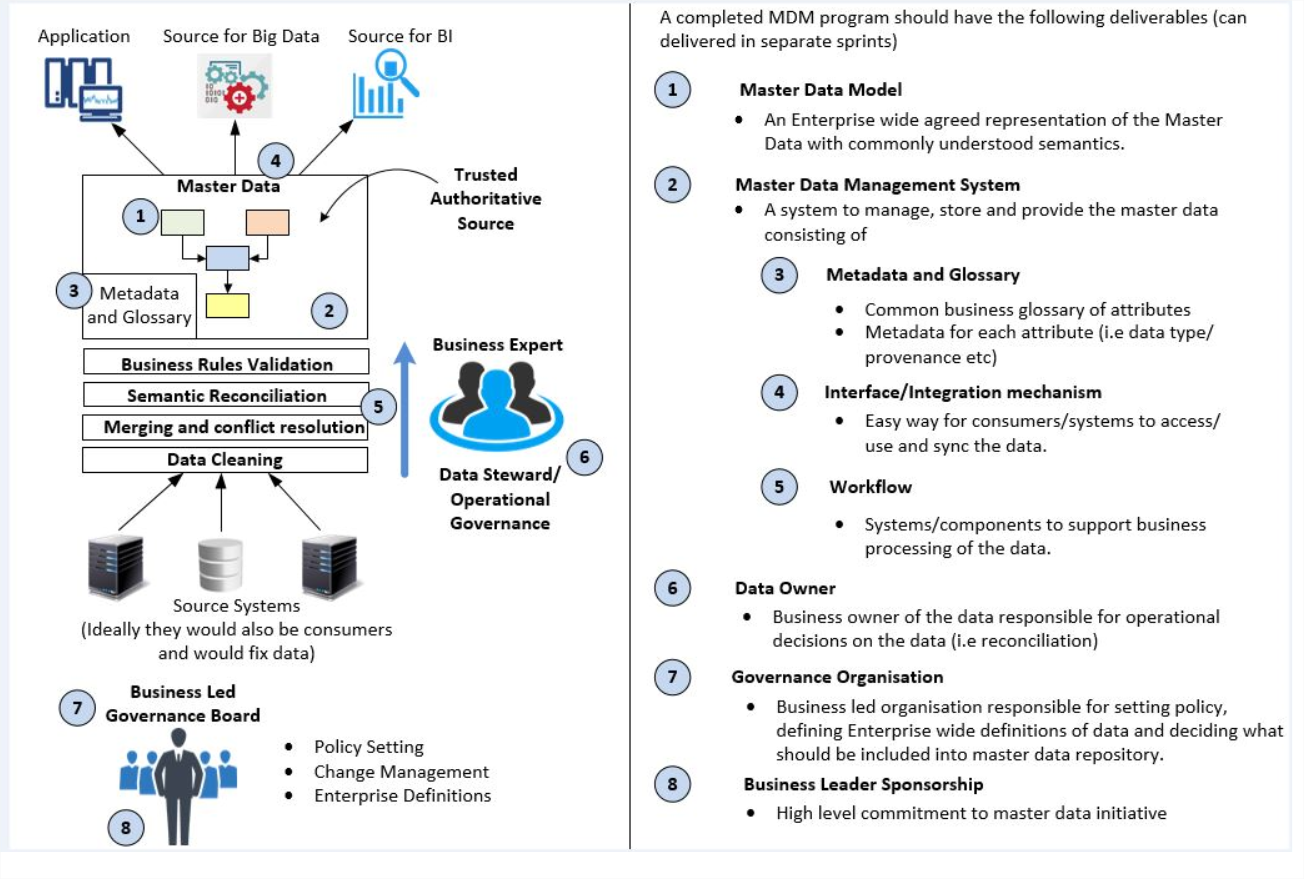


Figure : Complete MDM program.

### 3.1. Master Data Model.

At point one is the first difficulty, an enterprise-wide, agreed representation of the Master Data.   
It needs to be worded in such a way, the entire company can understand.

### 3.2. Master Data Management System.

An MDM system needs to manage, store, and provide the Master Data consisting of:

#### 3.2.3. Metadata and Glossary.

- Making a list of business terms/attributes and their explanation

- Metadata for each attribute: author, date, data type, etc.

#### 3.2.4. Interface/integration mechanism.

Easy access to work with, access and sync the data, like a single, comprehensive User Interface that takes the necessary data out of the correct modules.

#### 3.2.5. Workflow.

Systems/components to support business processing of the data. In short, systems to support the people working on processing the Master Data (like Excel, etc.)

### 3.3. Governance organization.

As stated earlier when defining the Data Owner, a person in a Senior capacity should take this role, within their field of expertise.   
This should indeed be business led, as their experience tells them which data is relevant and important. The decisions on what to include into the Master Data repository should be theirs.

All these owners, in their own fields, will make up the Business Led Governance Board.

### 3.4. Business Leader Sponsorship.

The commitment for clear Master Data Management is very obvious at TPCE.   
They started the Atlas-project, which will help in implementing a consistent system that will support all departments, referring to the mythical figure of Atlas, who carried the world on his back.

## 4. Materials and methods.

This information was gathered from PowerPoints supplied by TPCE and supplemented by website research using trusted websites in the field, like Gartner, etc.  
Then the information was used to write a comprehensive text, detailing all the components needed to understand the subject matter.  
This information is the starting point to solve the current problem TPCE is dealing with.

# Chapter 3: MDM implementation at TPCE: Introduction

## 1. Current situation at TPCE.

Currently TPCE uses a mainframe for all its data, called New Parts Application (NPA) and a Warehouse Management System (WMS) called AS400. NPA automatically interfaces with AS400 at the end of the workday. Which means the next day AS400 can access all the new data.

These systems are currently falling short of the rising demands. As stated in the introduction, the operations at TPCE have expanded a great deal and they want to give their dealers, as mentioned before, a platform for easily managing their own orders such as the Dealer Parts Management Module (DPMM).

Where operations worked fine on only NPA and AS400 before, the expansion has caused a lot of extra information that needed to be processed. This information got supplemented by templates that were made in Excel and Access databases. Of course, this makes for a very unclear structure. Which data is to be found where? Oftentimes these templates and databases were only to be found on a specialist’s computer. That’s why the need arose to transition to a new comprehensive and, most importantly, central system.

The Transportation Department (TRD) are transitioning to their own system to streamline operations, called Oracle Transport Management (OTM).

So, in order to achieve all this, a comprehensive, singular set of Master Data is required to support all these platforms.   
The departments are in varying stages of determining this set of Master Data. There is much discussion of the definition of this Master Data per department and about determining ‘Roles and Responsibilities’.

The main focus in this thesis will be on OTM, as they are launching the quickest. TRD is not a standalone entity within TPCE. They interact with Finance & Accounting (F&A), Warehouse, Purchasing and various hubs and depots.

In order to establish which data is most relevant and needed, visualizations need to be made. By mapping the processes, gaps and necessities become apparent. To map all the data and processes into one graphic would be very confusing. So the OTM data is split into the following pillars: Forwarders, Suppliers and Customers. This research is focused mainly on fleshing out the Forwarders pillar.

This work will provide a template which can be used to tackle these pillars and help other departments when their time comes to compile their Master Data. If this part of the groundwork is done well, it should cut down the time needed to transition between systems significantly for future departments.

The first step is to visualize the ideal, AS-IS and TO-BE situations in managing the data concerning Forwarders. These visualizations will serve as presentation materials, to convey the process to other departments that will go through the transition later.

## 2. Ideal situation.

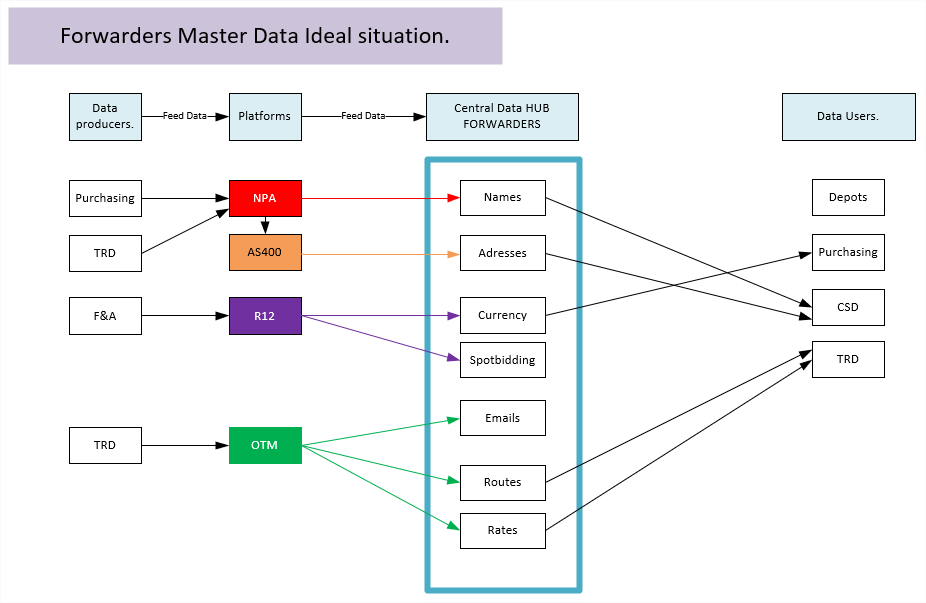
It may not seem logical to start with the ideal situation, rather than the current situation, but this is the Toyota Way of dealing with problems. It gives the team a very clear objective to work towards.

Figure : Forwarders Master Data- Ideal situation.

The main data producers are TRD, F&A and Purchasing. They feed their data into the platforms that are relevant to them.  
R12 is the TME-wide accounting program used by F&A in Poland. They feed the Forwarder’s financial information into their systems, like V.A.T. (Value Added Tax) numbers, bank account numbers, etc.  
In the ideal situation all this data is stored in a single, Central Data hub, which these same departments can use, accessing this data via a uniform User Interface.

This is ideal because every data record is entered only once, eliminating the chance to make mistakes with multiple inputs. This golden record is always kept up to date by the designated Data Owner.

## 3. Current situation.

Figure : Forwarders Master Data-Current, AS-IS.

In the current situation, the data producers feed their platforms. These platforms feed their own databases, or are files on someone’s personal computer, as mentioned previously. On the data user side oftentimes the same department, but different people, in different roles, take their information from these separate databases.   
As depicted, the data records, Name, Address and Currency are found in two different databases. So the chance for mistakes is doubled.  
For instance, the address F&A uses might be a Forwarder’s old address. So the invoice will be sent to location A, while the goods will be delivered to location B, the new address.

Who is responsible for maintaining this data? Do these departments communicate changes? In order to find out, harkening back to the beginning was necessary, a flowchart was made to visualize the process of setting up a new Forwarder. How is the data gathered? Which departments are involved?

For now, TRD selects a new Forwarder to do business with. They will send two documents to the Forwarder, who has to fill them out:

* Bank account approval form gathers information like bank account number, V.A.T. number, tax regimen, etc.
* Forwarder Fact Sheet: Currently in development, it will list all details concerning a forwarder, like contact person, address(es), phone numbers, etc.

The Forwarder sends the documents back to TRD and now the data validation can start.   
F&A will verify the data on the Bank Account Approval Form and delegate the validation of certain parts (like V.A.T.) to IT&D (Information, Technology and Disabilities). Following that, Purchasing will upload the rates into OTM.  
The last verification check is for F&A in Poland. If they approve, they will generate a Forwarder code which can be used by OTM, operations can start. This whole process has a timeframe of a month.  
TRD will verify the Forwarder Fact Sheet. If ok, that data will be manually put into NPA, which will interface with AS400.  
The flowchart of this process can be found in attachment one.

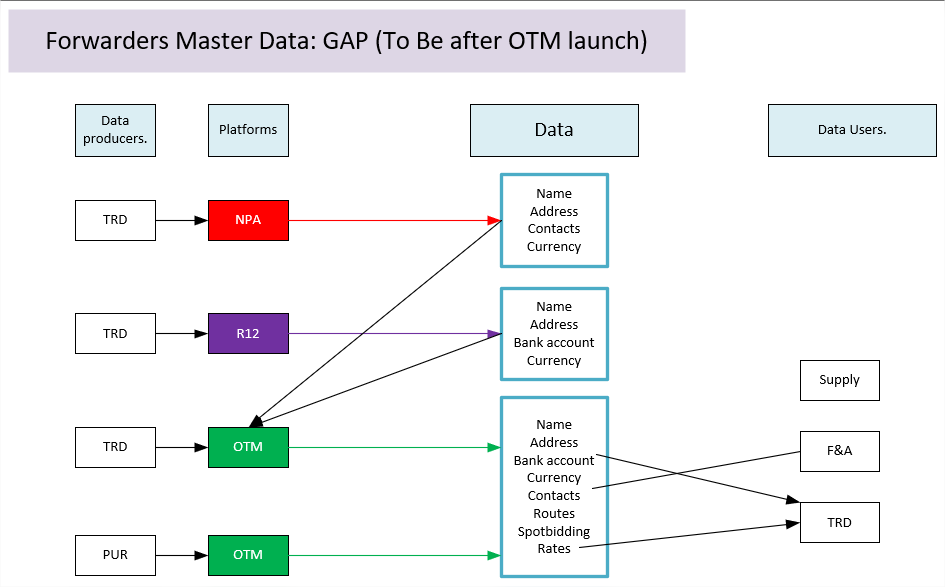
4. The gap between ideal and current situation.  


Figure : GAP, Forwarder Ideal and Current situations.

After the launch of OTM the ideal situation, as proposed by specialists within TRD, has not been reached. The Excel and Access inputs have been absorbed into OTM, which centralizes them and makes them available for use by all invested parties. Purchasing can now upload into OTM directly.  
Eventually, after all departments have made the transition, NPA will be phased out, but currently it’s still in use and its data will be used by OTM. R12 will remain in operation for the foreseeable future. So, in the best case, the NPA database will be absorbed by OTM, at least the part pertaining to transportation. This would eliminate an extra database to maintain and removes a chance for errors.   
In conclusion, this situation is already a lot better due to names, addresses, routes, rates and so on will be consolidated into OTM, eliminating the need for external Excel files and Access databases. Thusly also eliminating another avenue of error.

5. Materials and methods.  
In order to compile all this material, many meetings were held with the specialists within OTM. It was trial and error but eventually a correct representation of the Ideal, Current and Gap (TO BE) situation was reached. Trial and error, because during the mapping exercise, many issues came to light that would otherwise be forgotten. Brainstorming made certain details resurface that would otherwise have been overlooked. So, these findings needed to be incorporated until correct.   
Documents, such as OTM’s EVOLV-templates supplemented the information. EVOLV-templates are tables, made in Excel that list the most important data records that will need to be uploaded into OTM before launch and after. These will be explained further in the next chapter.

# Chapter 4: Steps in Master data implementation: A Critical Data Element (CDE) file for OTM.

The definition for Critical Data Element according to O’Reilly, a trusted online learning platform with excellent peer reviews:

“Critical data elements (CDEs) are defined as “the data that is critical to success” in a specific business area (line of business, shared service, or group function), or “the data required to get the job done.” Note that data that is critical in one business area may not be critical in another. Also note that when identifying CDEs, we often look at reports that may present values derived from underlying data, derived values”

In the context of this research, it means compiling a file with all the relevant Master Data records, pertaining to OTM.

This is a very lengthy process, broken down in the following steps:  
1. Data gathering.  
2. Data cleaning.  
3. Data consolidation.

All the necessary data is swimming around somewhere, sometimes ordered, sometimes not at all.  
So the IT departments starts extracting data from different sources and platforms.  
From NPA, AS400, Excel files and so on. Data cleaning can sometimes be skipped, if the data comes from a ‘cleaned’ source already, like an EVOLV-template, or a source OTM has no control over, like R12.

This research starts by scouring these sheets and start compiling a rough list, to start the CDE. Not just Excel sheets and databases, but since access to the actual platforms was off limits, screenshots were obtained from specialists and scoured for useful data records as well. The steps above will be shown per platform/data source used in this research, to start the data consolidation.

1. Data obtained from NPA.  
1.1. Data gathering.

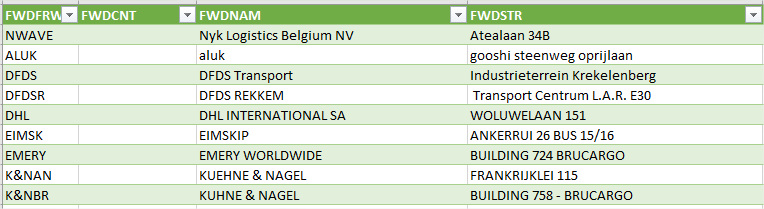
Example: Forwarders table, extracted from NPA. Note, all names of contact people have been whited out preventively, should privacy issues arise. The other used data, such as addresses and phone numbers have been checked and are freely available on the internet. Meaning, not behind a subscription paywall or business directory requiring a specific log-in, so everyone can look it up.  
Since it’s a lengthy document, it has been broken down in parts.   
Part one:

Figure : Forwarders-NPA Part one.

Most importantly to note here are the bright green records on top. FWDNAM, FWDSTR, etc. The meanings of these records are easily deduced from context here, but on its own, it would need a definition.

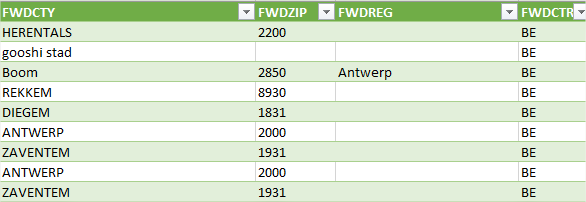
Part two:  


Figure : Forwarders-NPA Part two.

Continued from Part one, City, zip codes, regions and country of origin. Master Data records to consider here, FWDCTY, FWDZIP, FWDREG, FWDCTR.  
Begs the question, why keep these codes? Why not use the full name/definition as the name for the record?  
This has to do with database constraints and document constraints. Oftentimes there is a character limit, so it makes sense to keep these around.

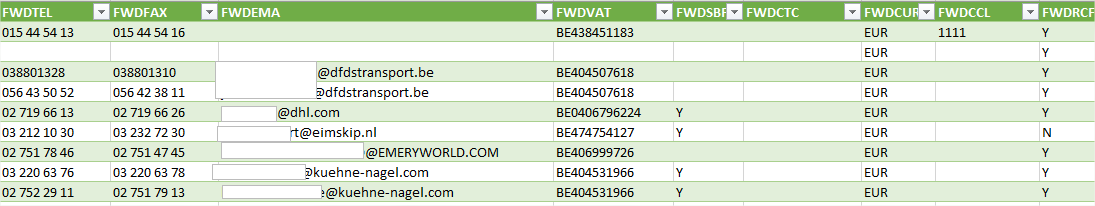
Part three:  


Figure : Forwarders-NPA Part three.

Listed here are phone numbers, fax numbers, e-mail contact, V.A.T. numbers and currency. These can still be easily deduced. The first difficulties start to appear with the following records: FWDSBF, FWDCLL and FWDRCF.

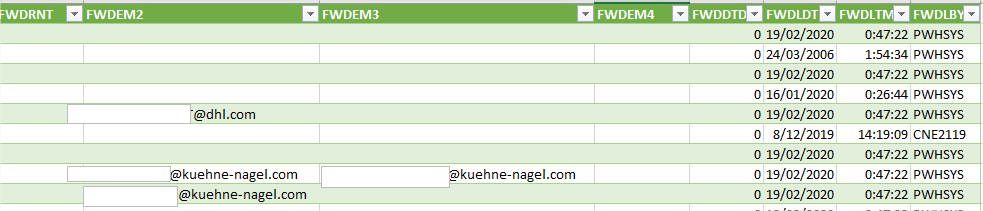
Part four:  


Figure : Forwarders-NPA Part four.

The problem becomes completely apparent in the last part. FWDEM2 and 3 can still be explained as Forwarder email two and three, but the rest doesn’t make any sense, at least not to the layman.

### 1.2. Data cleaning.

Meetings were held to determine what data is important. What seems irrelevant to one, could be important to another. There is a need to reach consensus on irrelevant data and its removal.   
Data was judged on the following criteria:

* Relevancy
* Correctness

Both criteria needed to be met in order for the data record to be kept. For instance. A fax number may have been correct but is no longer relevant to TPCE because faxes are generally no longer used. So this record is highlighted in red and will be removed in the CDE which will be discussed later.

For this research, all the Forwarder codes have been transposed into a new file.

|  |  |
| --- | --- |
| Record | Definition |
| FWDFRW |  |
| FWDCNT |  |
| FWDNAM |  |
| FWDSTR |  |
| FWDCTY |  |
| FWDZIP |  |
| FWDREG |  |
| FWDCTR |  |
| FWDTEL |  |
| FWDFAX |  |
| FWDEMA |  |
| FWDVAT |  |
| FWDSBF |  |
| FWDCTC |  |
| FWDCUR |  |
| FWDCCL |  |
| FWDRCF |  |
| FWDRNT |  |
| FWDEM2 |  |
| FWDEM3 |  |
| FWDEM4 |  |
| FWDDTD | Detention days |
| FWDLDT | Last maintenance day |
| FWDLTM | Last maintenance time |
| FWDLBY | Last maintenance by |

Table : Transposed NPA records.

As shown above, definitions were found for the previously unknown FWDDTD, FWDLDT, FWDLTM and FWDLBY, by consulting OTM specialists and NPA screenshots.

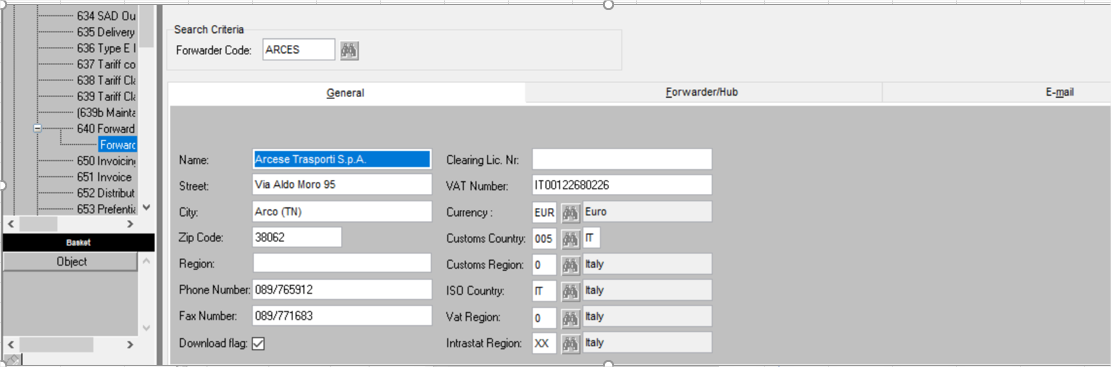
NPA screen 640 is the one that deals with Forwarders.  


Figure : NPA screen 640: Forwarders.

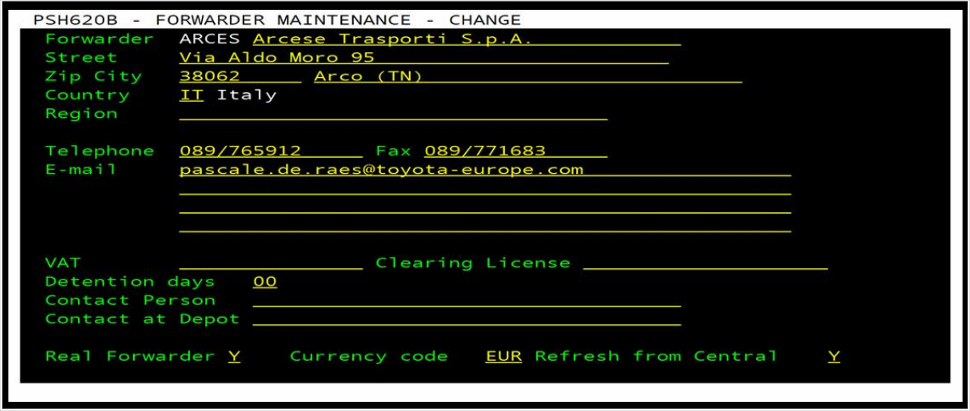
Many of the record definitions can be found in these screens. However, a few are missing.   
One of them, the FWDDTD (Detention days) record can be found in AS400, the WMS system. More specifically in AS40’s SH39 screen.  


Figure : AS400-SH39 Forwarder screenshot.

This was found out by meeting with relevant team members. This was not listed in NPA as such, because detention days is a very specific warehouse/transportation related item.   
The records, FWDLDT, FWDLTM and FWDLBY which mean Last Maintenance Day, Time and by whom it was done. These records are only important to IT. That’s two sources of data down.

2. Data obtained from R12.  
2.1. Data gathering.

R12 is an Oracle Software Program, used mainly by the Finance and Accounting department (F&A), located in Poland. They handle the majority of TME’s finances. There is a local F&A department in Zaventem, but they handle more local things, like payroll for personnel.   
F&A Poland deals with Forwarders’ bank account number, the verification thereof, V.A.T. regimes per country/region etc.   
In NPA, there is also a record for V.A.T. numbers and V.A.T. regions. So questions arise. Where was this data found first? Which platform is the lead source?

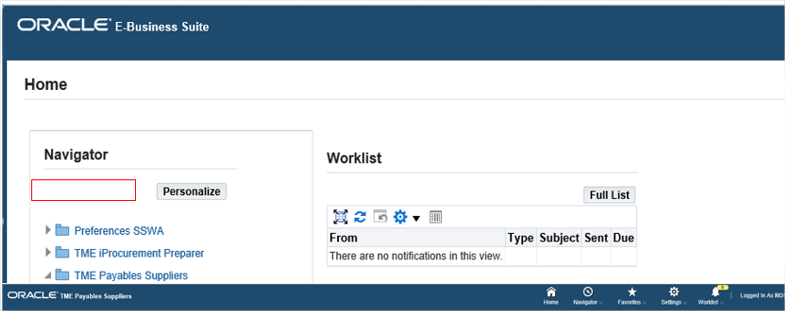
Part one: Start screen with search function.   


Figure : R12 part one.

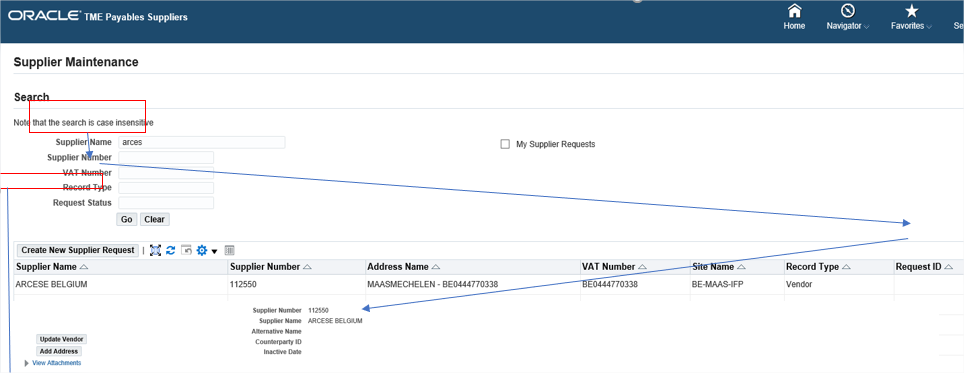
Part two: Overview of existing records.  


Figure : R12 Part two.

So there are some duplicates that are also found in NPA.

* V.A.T. numbers
* Forwarder name
* Forwarder address

There are many new records to be found too, pertaining to Forwarders (called Suppliers in R12), such as the bank account numbers, names of the banks and most striking, the forwarder code in R12 is different than the forwarder code used by TPCE. The R12 code requires a six-digit code and NPA requires a five-digit code. These numbers are still referring to the same logistics service provider.

2.2. Data consolidation.   
There are now three sources of data, NPA, AS400 and R12. Three columns are added to the working table and X marks where a certain record can be found.   
The new records found in R12 were also added. Definitions were found through meetings with OTM specialists.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Record | Definition | NPA | AS400 | R12 |
| FWDFRW | Fwd code | X | X | X |
| FWDCNT | Country | X | X |  |
| FWDNAM | Name | X | X | X |
| FWDSTR | Street | X | X | X |
| FWDCTY | City | X | X | X |
| FWDZIP | Zip code | X | X | X |
| FWDREG | Region | X | X | X |
| FWDCTR | Country code | X | X | X |
| FWDTEL | Phone | X | X | X |
| FWDFAX | Fax | X | X | X |
| FWDEMA | Email | X | X | X |
| FWDVAT | V.A.T. | X | X | X |
| FWDSBF | Real FWD? | X | X |  |
| FWDCTC | Contact at depot | X | X |  |
| FWDCUR | Currency | X | X | X |
| FWDCCL | Customs clearing license | X | X |  |
| FWDRCF | Refresh from central flag | X | X |  |
| FWDRNT | Returns notifications type | X |  |  |
| FWDEM2 | Email 2 | X |  |  |
| FWDEM3 | Email 3 | X |  |  |
| FWDEM4 | Email 4 | X |  |  |
| FWDDTD | Detention days |  | X |  |
| FWDLDT | Last maintenance day |  | X |  |
| FWDLTM | Last maintenance time |  | X |  |
| FWDLBY | Last maintenance by |  | X |  |
| R12 Code | 6-digit R12 FWD code |  |  | X |
| R12 Site | Where found in R12 |  |  | X |

Table : Updated Table including new records and data sources.

3. Data obtained for OTM.   
3.1. Data gathering.

Since the upload of data into OTM is in progress, it’s good to check which records are present on the platform itself and to see if there are any new ones.

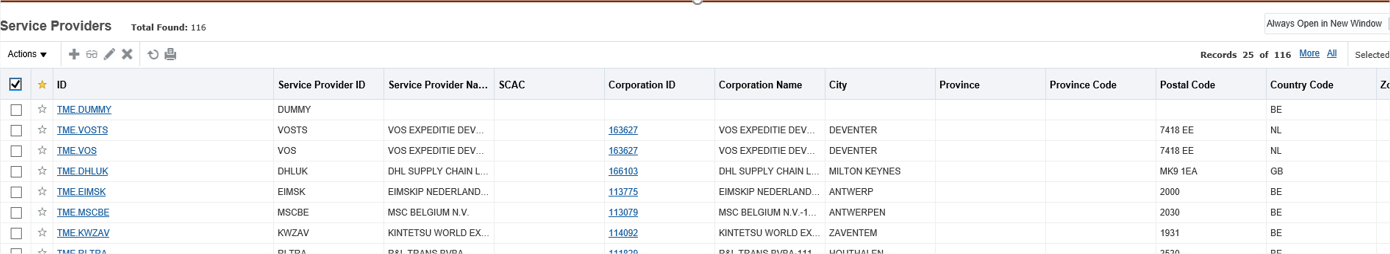
Part one:   


Figure : OTM part one.

Many records are, again, the same. It’s also apparent that OTM uses the same six-digit codes as R12.

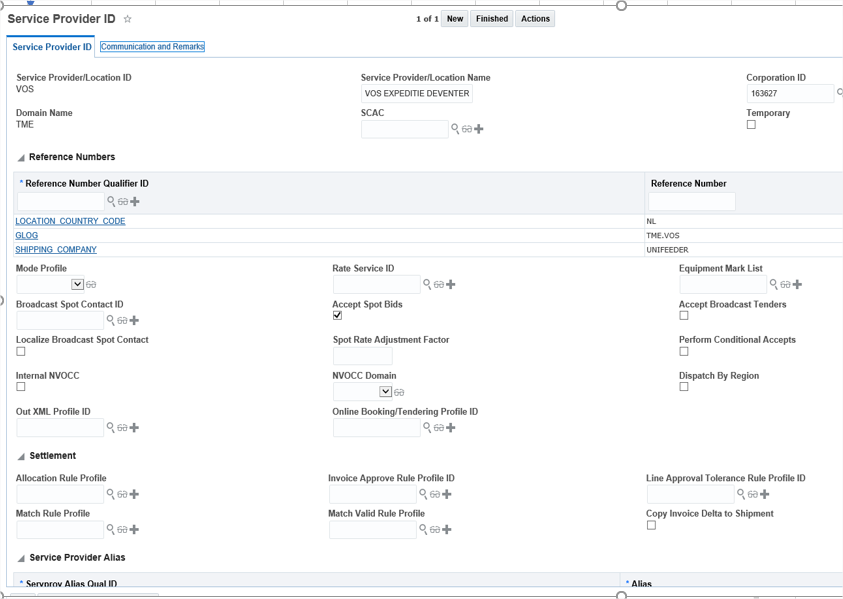
Part two: OTM screen for Forwarders continued. Looking for new records to add.  


Figure : OTM part two.

Accept spot bids is a new record found here.

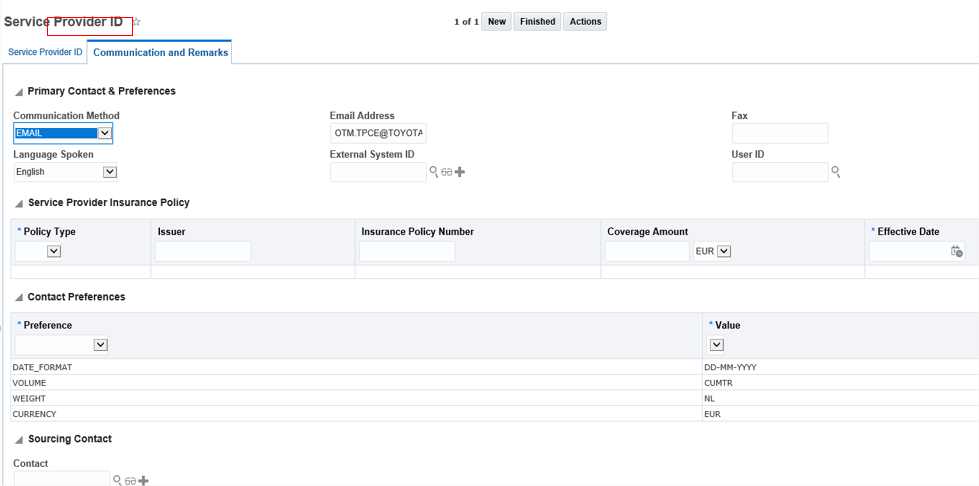
Part three:  


Figure : OTM part three.

The new records found here are volume and weight. Currency is a duplicate.

An extra data source here were something called EVOLV-templates. As mentioned before, these are Excel spreadsheets with pertinent data that needs to be fed into OTM. These are still a work in progress so there was still a need to double check the data with the OTM screenshots. There are EVOLV-templates for many different groups like suppliers, supplier currency, calendars etc. The one relevant to this research is, of course, the Forwarders one.

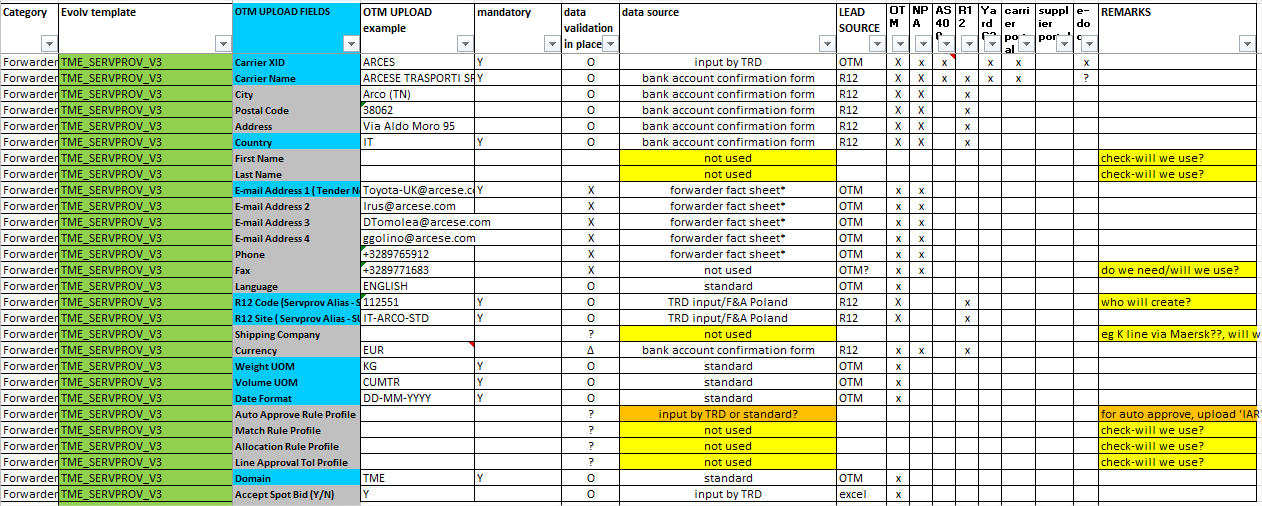


Figure : EVOLV Template: Forwarders.

There is a lot to take in here, but upon closer inspection, all the same records are present here. Even with an OTM check if these will be used or not. Meeting with OTM specialists gave conclusive information about this. A good thing to note here are the lead sources and the data source. Meaning, before any information was uploaded, there had to be a form to fill out.

The EVOLV-templates are a great upload tool, with a lot of technical information for the OTM people doing the uploading. It’s also a great springboard to make a visualization CDE.

### 3.2. Data consolidation.

There are a few things that need to be added to the table, such as a legend with L, which stands for Lead Source and X stands for ‘data is found on platform’.   
New records:

* Accept spot bid: Y/N
* Weight UOM
* Volume UOM
* Date format
* Domain
* Time Zone

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Legend: | L= Lead source |  |  |  |  |
|  | X= data is found on platform | |  |  |  |
| Record | Definition | NPA | AS400 | R12 | OTM |
| FWDFRW | Fwd code | X | X | X |  |
| FWDCNT | Country | X | X |  |  |
| FWDNAM | Name | X | X | X |  |
| FWDSTR | Street | X | X | X |  |
| FWDCTY | City | X | X | X |  |
| FWDZIP | Zip code | X | X | X |  |
| FWDREG | Region | X | X | X |  |
| FWDCTR | Country code | X | X | X |  |
| FWDTEL | Phone | X | X | X |  |
| FWDFAX | Fax | X | X | X |  |
| FWDEMA | Email | X | X | X |  |
| FWDVAT | V.A.T. | X | X | X |  |
| FWDSBF | Real FWD? | X | X |  |  |
| FWDCTC | Contact at depot | X | X |  |  |
| FWDCUR | Currency | X | X | X |  |
| FWDCCL | Customs clearing license | X | X |  |  |
| FWDRCF | Refresh from central flag | X | X |  |  |
| FWDRNT | Returns notifications type | X |  |  |  |
| FWDEM2 | Email 2 | X |  |  |  |
| FWDEM3 | Email 3 | X |  |  |  |
| FWDEM4 | Email 4 | X |  |  |  |
| FWDDTD | Detention days |  | X |  |  |
| FWDLDT | Last maintenance day |  | X |  |  |
| FWDLTM | Last maintenance time |  | X |  |  |
| FWDLBY | Last maintenance by |  | X |  |  |
| R12 Code | 6-digit R12 FWD code |  |  | X |  |
| R12 Site | Where found in R12 |  |  | X |  |
| Accept spot bid Y/N | Fwd does spot bidding? |  |  |  | L X |
| Weight UOM | Unit of measurement |  |  |  | L X |
| Volume UOM | Unit of measurement |  |  |  | L X |
| Date Format | How is date formatted |  |  |  | L X |
| Domain | Location |  |  |  | L X |
| Time zone | Time zone. |  |  | L X |  |

Table : Updated Table, draft for CDE.

4. Data obtained from additional, smaller sources.  
4.1. Data gathering.

As seen in the current situation (AS-IS) graphics there were additional sources. For Forwarders specifically there are a few additional sources:

* Carrier Portal
* Supplier Portal
* Excel files
* The Yard C3 software program
* E-doc

The Carrier and Supplier Portals are designed to gather and dispense data easily. The Forwarder Fact Sheet, for instance, is in development. This will centralize all the relevant data pertaining to Carriers (also known as Forwarders) and Suppliers.

4.2. Data consolidation.  
These records had a fairly low impact for now. Data was obtained by meeting with the relevant people, showing them the Excel table that was being worked on. Four more columns were added and X’s were put where the relevant data record was found per platform.  
Since there are no new records added, the table will only be shown partly, just to reflect the new columns added.

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Record | Definition | NPA | AS400 | R12 | OTM | Car. Portal | Sup. Portal | Yard C3 | E-doc | Excel | Access |
| FWDFRW | FWD code | X | X | X |  |  |  | X | X |  |  |
| FWDCNT | Country | X | X |  |  |  |  |  |  |  |  |
| FWDNAM | Name | X | X | X |  |  |  | X | X |  |  |
| FWDSTR | Street | X | X | X |  |  |  |  |  |  |  |
| FWDCTY | City | X | X | X |  |  |  |  |  |  |  |
| FWDZIP | Zip code | X | X | X |  |  |  |  |  |  |  |
| FWDREG | Region | X | X | X |  |  |  |  |  |  |  |
| FWDCTR | Country code | X | X | X |  |  |  |  |  |  |  |
| FWDTEL | Phone | X | X | X |  |  |  |  |  |  |  |

Table : Updated with extra platforms.

## 5. Making the CDE based on all previous work.

The latest iteration of the worktable became the basis for the Critical Data Element (CDE). Through meetings however, a few things were found, that needed to be added.  
Where does the lead source get its data from, meaning the original data source? This is good to know, in case there is a mistake. In that way they can track down easily where the information came from and the responsible person can be addressed.

As an example of this Original Data Source, an excerpt of the Bank Account Approval Form. Something new Forwarders are required to fill out.

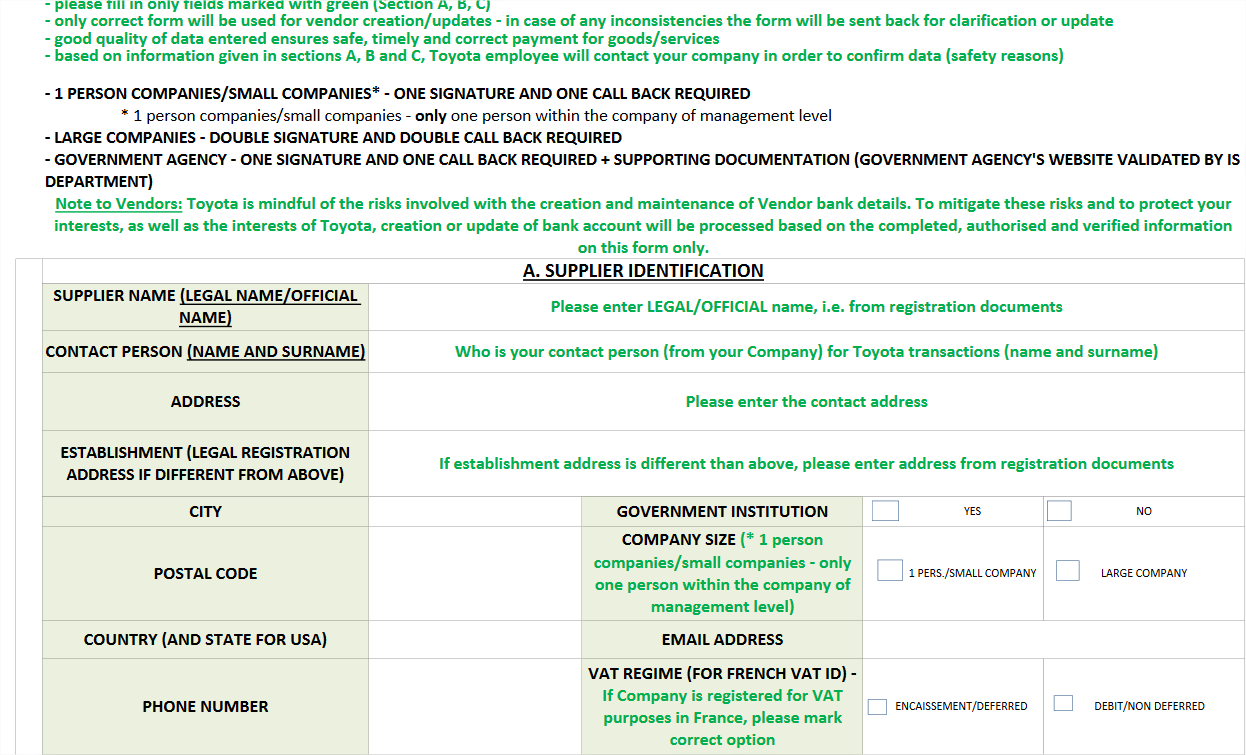


Figure : Bank Account Approval form excerpt.

The Lead source per record can be found in the EVOLV-template. What about the other platforms? To continue the line of addressing the responsible person, it was vital to find out if the information was put in manually or as a result of an automatic interface.  
These two factors were added to the legend and many specialists were consulted per platform to establish the mode of input.

Next steps to take in making a suitable lay-out:

A) Take the earlier working table and make a better layout. Add colors per platform.

B) Add definitions, explaining the codes. Rudimentary, from the perspective of TRD, meaning self-explanatory terms for TRD are skipped for now.

C) Add all platforms TRD comes into contact with. Codes were established.  
**L**= Lead source and first manual input.  
**IF**= Interface, automatic input  
**M**= Manual input

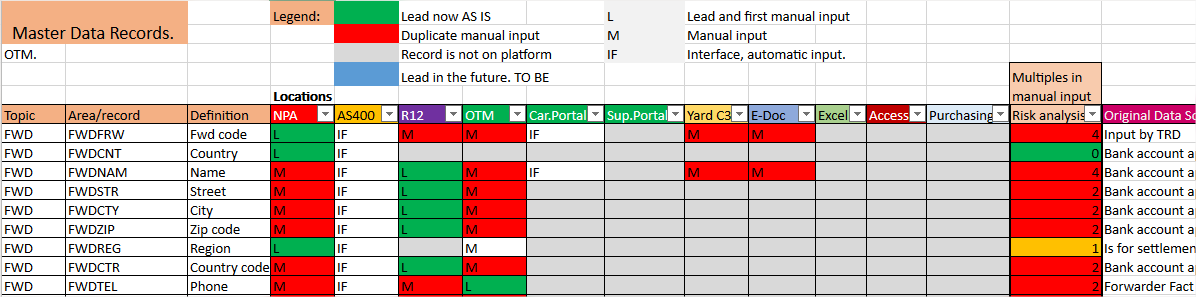
‘Lead in the future’ was added to the Legend, so it could be added in the future. Currently there is no consensus yet, on who should be the ‘Lead in the future’, so it’s skipped for now.  
This results in the following lay-out.  


Figure : CDE lay-out example.

Denoting the presence of Data with an ‘X’ was fairly nondescript however, so the X’s got changed into L, IF or M, depending on the input mode.  
Those input modes were found out by speaking to many different TRD/OTM specialists. It was established which record existed on which platform and what mode of input (automatic/manual) was used.  
There are specialists that only work with Yard, E-doc (a platform that offers transportation documents to carriers), etc. These suggested people were contacted and they offered their help and expertise.

In the case of R12, it was not possible to speak to someone from F&A Poland. So, screenshots were requested from the R12 screens accessible by OTM people. By examining these screenshots, it became apparent which records were available in R12, as explained previously. The information about the type of input was supplied by OTM specialists.

It’s a good way to easily convey which data records are used by whom, this also helps to narrow down possible sources of mistakes.

Conclusion: The two most important things are:

* Determining on which platforms the data is available,
* Determining the mode of input, manual or interfaced/automatic.

### 5.1. Risk analysis.

Many manual inputs means there is a lot of room for mistakes. In order to track this, a risk analysis was made in Excel by using conditional formatting with determined values.   
Green: Values < 1. Aside from the Lead (a.k.a. the first manual input), there are no other manual inputs. Reducing the chance drastically, but never to zero, as there could be mistakes in the Original Data Source.

Orange: Value =1. Aside from the Lead, there is one other manual input. There is one chance for mistakes, but one extra input might be inevitable, hence the Orange annotation.

Red: Values > 1. As usual, aside from the Lead, there are several manual inputs here, which increases the risk of making mistakes even more.

This analysis serves two purposes:

* Finding out which manual inputs could be automated.
* To simply serve as a visualization to highlight problematic records.

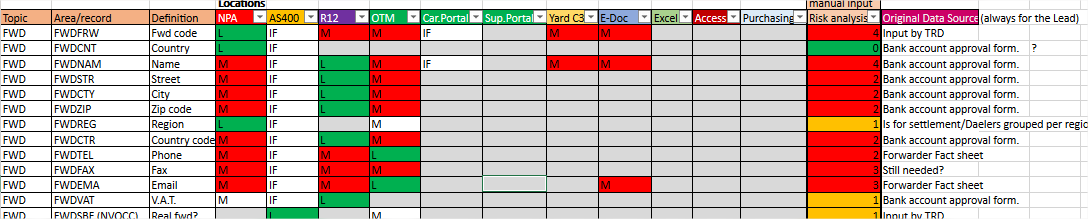
Next to the Risk analysis is the Original Data Source (always for the lead).  


Figure : CDE continued.

For instance, the FWDEMA (E-mail) data record. There has been a mistake and no one knows where to start looking for the right E-mail address. This CDE appoints the Lead source, where all the others copy or interface their data from. If they can’t help, the CDE also provide the Original Data Source. In this example, the Forwarder Fact Sheet, which a new Forwarder has filled out, providing their information. The correct record should be listed there. If not, the Forwarder can be called, as a last resort.

5.2. Basis for Rules and Responsibilities determination.  
Meetings with specialists determined the Lead source for each data record. Meaning, they are the ones that enter this data first. So when establishing Data Owners and Data Stewards, this Lead will be the most plausible candidate for this role.

## 6. Conclusion.

The CDE will serve as a basis for the Business Glossary. In an IT and Business mix, some of the information in the CDE is convenient during the transitional period, but may not be so relevant when serving as a business template. This will be discussed in the next chapter.

Conclusion: The CDE serves as a tool to initially compile, clean and visualize potential problems with the data. It offers a list of records that will definitely have to be included in OTM as Master Data.

Uses in the full Atlas project scope:

* Serve as a template/ basis of comparison for the departments that are next in line for an overhaul.
* The method of compilation can be copied and continually improved by gathering experience across the departments.

As its own entity, the Risk analysis makes it stand out and it serves a purpose as a troubleshooting tool.

7. Materials and Methods.  
This information was obtained by meeting with OTM project people. The relevant data was lifted from the Data provided by IT specialists first and a lot of screenshots were provided by business specialists. A work-in-progress table was compiled, listing all the records found from the different tables, on different platforms. Which were then discussed for correctness and validity in meetings with several OTM professionals. Adjustments were made in real time during these meetings and afterwards.  
If meetings with OTM members didn’t suffice, contact was made with other specialists that OTM members suggested, in order to get a complete overview. These specialists provided additional tables, screenshots or even came by to dispense their information.

# Chapter 5: Steps in Master Data implementation: Business Glossary a.k.a. Data Catalogue.

At first glance, the Data Catalog may look a lot like the CDE-file, which is true.  
The main difference, however, is the format. While the CDE format might be more convenient to communicate with several specialists within and beyond the department, the Business Glossary template was proffered by the Enterprise Architects (EA).

The ultimate goal of the Atlas project is to have a single, clear database which displays the correct data and which has one, comprehensive interface all users from across departments can use.

This can be achieved by the use of a Data Catalog tool. One example of this, suggested by Enterprise Architecture (EA), is Collibra. Which would fit within the scope of transitioning towards Industry 4.0. standards.  
A Data Catalog tool is not just a database, but offers tools for data governance, cloud storage and security.  
It also offers analytics, which fit into the Big Data framework. These analytics will help with error detection, which segues into predictive maintenance, another Big Data strong point.

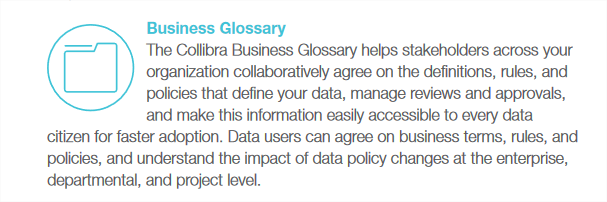
This is of course beyond the scope of this research, but it’s important to know that this system consists of many cogs and in order to run smoothly, all these cogs must be correct and in working order.   
The work builds upon itself, and one mistake early on could have a lot of consequences. Here’s what Collibra has to say about their Business Glossary.  


Figure : Collibra Business Glossary definition.

In a nutshell, this is everything the company is trying to achieve. The end goal. To have a single database, with a comprehensive User Interface that allows easy access across the company. After ‘Roles and Responsibilities’ are established, the Data Owners and Data Stewards have a good platform to upload and maintain their data with.

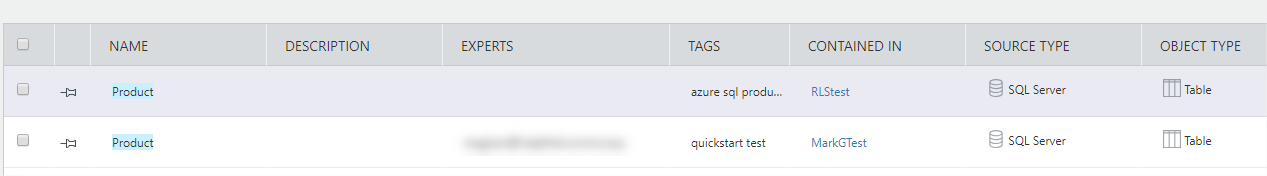
As mentioned by the people of EA, most of these Data Catalog Tools are built upon Microsoft Azure’s structure. This means data is uploaded into this tool via a specific template. This is what Microsoft Azure looks like.   


Figure : Microsoft Azure example.

This template is very general, so during many meetings, a more TPCE-specific Business Glossary template was created.

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Term Name** | **Parent Name** | **Definition** | **Description (Example)** | **Synonym** | **Owner** | **Stakeholders** | **Codes** | **Notes** | **Legend** |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |

Table : Blank Business Glossary template.

Notable differences with the CDE:

* Description
* Synonyms
* Stakeholders

The Definition alone is sometimes not enough to make a term/record completely clear, that’s why the Description column was added. Earlier it was mentioned that the definitions in the CDE served mostly to explain terms for TRD that weren’t self-explanatory. In the Business Glossary, every term needs a definition, because people outside the relevant department will read it too.  
Earlier it also became apparent that Forwarders also went by the name Carriers or Logistics Service Providers, depending on which department one would ask.   
The last, most notable difference are the Stakeholders. Where the CDE would focus on which platform the data could be found, the Business Glossary lists all departments that have use of this data. Filled out, it looks like this.

|  |  |  |  |
| --- | --- | --- | --- |
| **Term Name** | **Parent Name** | **Definition** | **Description (Example)** |
| Forwarder |  |  |  |
| Forwarder code | Forwarder | ID number for a forwarder | 5 digits, 12345 in NPA |
| Country | Forwarder | Country of origin |  |
| Name | Forwarder | Forwarder name |  |
| Street | Forwarder | Address street name |  |
| City | Forwarder | Address city |  |
| Zip code | Forwarder | Address zip code |  |
| Region | Forwarder | Region where FWD is situated |  |
| Country code | Forwarder | Official 3 letter abbreviation |  |
| Phone | Forwarder | Phone number |  |
| Fax | Forwarder | Fax number |  |
| Forwarder main contact email | Forwarder | Main email address. |  |
| V.A.T. | Forwarder | V.A.T. Number |  |
| Real forwarder | Forwarder | Real FWD or logistic service provider |  |
| Contact at depot | Forwarder | Contact person at depot |  |

Table : Filled out Business Glossary Part one.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Synonym** | **Owner** | **Stakeholders** | **Source** | **Length** | **Codes** | **Notes** | **Legend** |  |  |  |
| Logistics Partner |  |  | OTM |  |  |  | TRD | Transportation department | | |
|  | TRD | TRD, F&A | OTM | 5 | FWDFRW |  | F&A | Finance and Accounting | | |
|  | TRD | TRD, F&A | OTM | 35 | FWDCNT |  | FWD | Forwarder |  |  |
|  | TRD | TRD, F&A | OTM | 35 | FWDNAM |  | PUR | Purchasing |  |  |
|  | TRD | TRD, F&A | OTM | 40 | FWDSTR |  |  |  |  |  |
|  | TRD | TRD, F&A | OTM | 35 | FWDCTY |  |  |  |  |  |
|  | TRD | TRD, F&A | OTM | 10 | FWDZIP |  |  |  |  |  |
|  | TRD | TRD, F&A | OTM | 35 | FWDREG |  |  |  |  |  |
|  | TRD | TRD, F&A | OTM | 2 | FWDCTR |  |  |  |  |  |
|  | TRD | TRD, F&A | OTM | 15 | FWDTEL |  |  |  |  |  |
|  | TRD | TRD, F&A | OTM | 15 | FWDFAX |  |  |  |  |  |
|  | TRD | TRD, F&A | OTM | 50 | FWDEMA |  |  |  |  |  |
|  | TRD | TRD, F&A | OTM | 15 | FWDVAT |  |  |  |  |  |
|  | TRD | TRD | OTM | 1 | FWDSBF |  |  |  |  |  |
|  | TRD | TRD | OTM | 35 | FWDCTC |  |  |  |  |  |
|  | TRD | TRD, F&A | OTM | 3 | FWDCUR |  |  |  |  |  |

Table : Filled out Business Glossary Part two.

The previously made CDE was instrumental in filling out this Business Glossary. The listed platforms made it easy to determine the stakeholders, a.k.a. the departments that use that platform, are automatically the stakeholders here. Owners and source were established in the same fashion, by looking at which platform was the Lead source in the CDE. The Owner was the department responsible for this platform. The CDE that was established by many meetings, consulting specialists and constantly incorporating feedback, provided a sound basis. The groundwork done for the CDE made filling out the Business Glossary a lot easier.   
Other things, like ‘Length’ pertains to the character limit the record has. This is mainly geared towards the IT department.

Result: This Business Glossary/Data Catalogue is the uniform begin point, that other departments will continue use in order to log their most important Master Data records. Supplemented by visualized processes.  
All the research work has culminated into this and will make future implementation a lot easier.

## 1. Materials and methods

Previous materials, like the CDE, were used as a basis to fill out the Business Glossary. Meetings with the MDM team established a workable template for TPCE, starting from the generic Microsoft Azure template.  
The template was filled out by consulting specialists from OTM, to aid in finding suitable definitions and or known/synonyms.

# Conclusion.

Harkening back to the research questions stated in the very beginning, were answers found? Note, these research questions were drafted, before the situation AS-IS was surveyed at Toyota.

***Research Question one:* What are the specifications a system needs to qualify for Industry 4.0.?**

1. Is there a determined set of requirements to aid in the vetting process?

2. Which changes need to be made in order for the system to conform to the requirements?

The current system needed to be overhauled in order to meet Industry 4.0. standards. Specifically, the Master Data Management portion of this needed to be tackled.  
The requirements have been determined during the course of this research. There was a need for clarity, uniformity and consistency in their Master Data records, as well as a clear definition of the term Master Data itself.  
These Master Data records needed to be well-defined and presented in a single, comprehensive user interface. This, so every relevant department can make optimal use of it. The Transportation Department are the first ones to transition to the new system, OTM.   
In order to achieve this, the changes that needed to be made have been discussed in detail above. Every department needed to go through a process of Data gathering and cleaning. Every record needed to be discussed to see if it’s still relevant. For instance, the Fax numbers will not be included in OTM.   
In the CDE, duplicate data inputs were ferreted out in order to minimize the chances for mistake. This is also a good guideline to determine Roles and Responsibilities, who is the lead source and who will manage the data from now on? All this, so a single, correct record of Master Data can exist.

This Critical Data Element file formed the groundwork to help tweak the Business Glossary template and to fill it in.   
The experience/specifications gathered from this entire process will significantly help the other departments/systems to transition. The MDM team will support in this  
In essence, this research helped in determining this set of requirements needed for a smooth transition.

Research Question two: How do we create added value and optimize the supply chain by making changes?

1. How cost effective are these changes? If a change costs more than a potential gain, is it worth it?

2. What is the role Master Data can play in this process?

3. After the changes are implemented, is there an evaluation period? To see how day-to-day operations are affected?

The main research question has definitely been answered. The value added by this research culminates in the Business Glossary. This has two great benefits. The first one is filling it out the Forwarders portion of the Transportation Department, preceded by the whole process of Data gathering and cleaning. Now the department has a comprehensive list to support them in their imminent transition.  
Second, all the groundwork, all the trial and error done for the Transportation Department will make it easier for future transitioning departments to establish their list of Master Data.   
While it’s not possible to measure these changes monetarily, this groundwork will save a lot of time and effort in the near future. All the experiences are gathered and applied throughout the firm. The term Toyota uses for this is Yokoten, meaning best practices are gathered and shared in order to guarantee smooth operations in the future.

The role of Master Data here is obviously very important. Determining relevant Master Data per department and establishing Roles and Responsibilities for its maintenance are the corner stone of the entire Atlas effort.

The evaluation period is permanent. Toyota will continually monitor operations to ensure quality.

*Research Question three*: How do we tackle the implementation of those changes?

1. Will the changes be implemented gradually or radically?

2. What is the role of each individual Atlas Team member in this?

After field research, it became obvious that these changes will be implemented gradually. The Toyota way of working values every relevant employee’s input. Things get developed by a continual stream of listening to and incorporating feedback. The experiences generated by the first transitioning department, namely Transportation (for OTM) are gathered by the MDM team, which will apply them to support the future, transitioning departments within the Parts Supply Chain Group (PSCG).

The Atlas team is fairly small. Their main job is to support the departmental teams and the MDM team by providing feedback, make a roadmap of the entire project and liaise between all these teams, departments and management. As has been experienced during the twelve weeks this field research was conducted at TPCE.

So this research started with a question that came from TPCE, what is Master Data? How can we implement a proper Master Data Management, as a step to conforming to Industry 4.0. standards?  
Through meetings, visualizations and constant feedback, the CDE was made, which was the groundwork for the Business Glossary. Which will be continuously expanded upon by the next departments in line. All this has been achieved in only twelve weeks.   
This research can help other departments transition, as it lists the entire process of getting to the Business Glossary step by step. They can learn from OTM’s experience in how to efficiently gather and clean their Master Data. They can skip the trial and error period so to speak and follow the steps listed in this research.

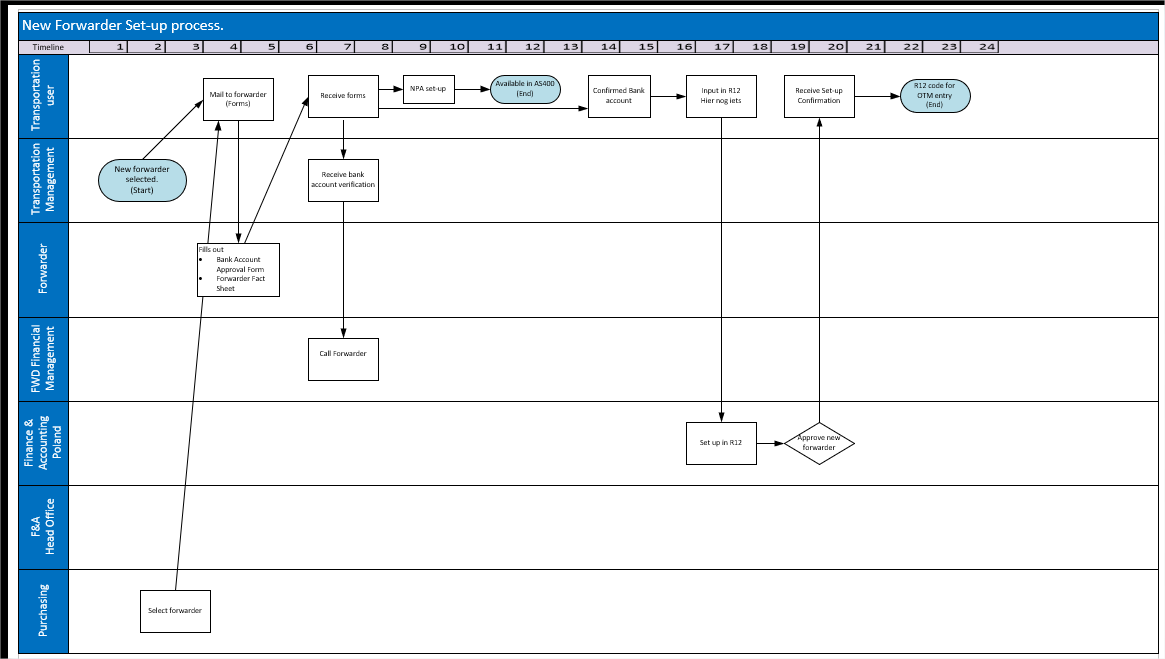
This research also discusses what Industry 4.0. is, what types of Data there are and how the general process of Master Data Management is. The departments won’t have to feel like they have been thrown into the deep end of the pool. They can slowly wade in, get familiar with the terminology and then apply the steps, tailored to their specific situation.  
To replicate the results of this research, a ‘Materials and Methods’ part has been included at the end of most chapters.

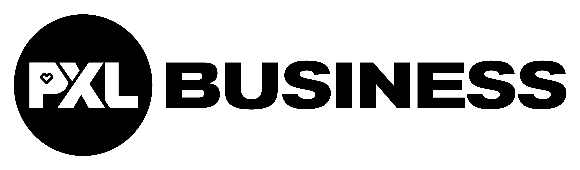
# Reference list.

## Websites.

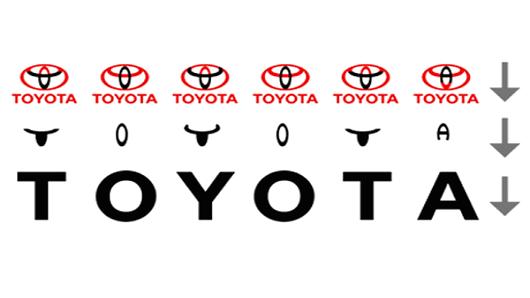
<https://www.gartner.com/en>  
<https://www.techopedia.com>   
<https://tdwi.org/Home.aspx>  
<https://azure.microsoft.com/en>  
<https://www.collibra.com>  
<https://www.ibmbigdatahub.com/infographic/four-vs-big-data>  
<https://www.whishworks.com/>   
  
  
Literature.  
Dama International (2017). Dama-Dmbok, Data management body of knowledge 2nd edition. In Dama International (Red.) *Dama-Dmbok, Data management body of knowledge 2nd edition,* pp. xx-xx. Technics Publications.

Attachments.  
Attachment one: Swimming lanes.





**Optimization of IT systems to conform to Industry 4.0. standards by using Master Data.**



**Sara De Moor 3rd year Logistics Management**

**Promotor: Nadine Mertens Academic year 2019 – 2020  
Mentor: Tom Scheepers, TPCE, Diest**