LISBON
SCHOOL OF
ECONOMICS &
MANAGEMENT
UNIVERSIDADE DE LISBOA

MASTER

Master in Management Information Systems

FINAL MASTER PROJECT

Dissertation

KPIs as a measure for quality in Master Data

Inês Brás de Moura Duarte Miranda

September-2017

LISBON SCHOOL OF ECONOMICS & MANAGEMENT
UNIVERSIDADE DE LISBOA

MASTER

Master in Management Information Systems

FINAL MASTER PROJECT

Dissertation

KPIs as a measure for quality in Master Data

Inês Brás de Moura Duarte Miranda

Advisor:

Prof. Engª. Ana Maria Marques Ribeiro Dos Santos Lucas

September-2017

Abstract

In a time where companies are giving more and more importance to the implementation of information systems to support their businesses, the term Master Data is becoming more usual since it concerns the core information of a company (e.g. customer and employee data). Maintaining the highest quality for this data is nevertheless a challenge that needs to be measured through performance measures (for example: Key Performance Indicators).

The present case study has the purpose of investigating the definition, calculation, divulgation and use of Key Performance Indicators within a multinational company. To this end, a training of 6 months was provided by the company to teach participants how to calculate these values and also how to obtain all the necessary information regarding these indicators within the Human Resources department.

The analysis showed that, even though the existing KPIs are well defined and calculated, they are not enough to include all existing master data classes and are also too wide-ranging, making it almost impossible to find the root of the HR problems within the company.

Key words: Key Performance Indicators, Data Quality, Human Resources Master Data, Master Data, Performance Measures

Resumo

Numa época em que as empresas estão a dar mais e mais importância à implementação de sistemas de informação para suportar o seu negócio, o termo Master Data está a tornar-se mais usual uma vez que concerne a informação mais importante de uma empresa (p.e. dados de clientes e colaboradores). Manter níveis elevados de qualidade para estes dados é um desafio que precisa de ser medido através de indicadores de performance (p.e. Key Performance Indicators).

O presente estudo tem o objetivo de investigar a definição, cálculo, divulgação e uso de Key Performance Indicators numa empresa multinacional. Para este propósito, uma formação de 6 meses foi providenciada pela empresa para explicar como calcular estes valores e como obter toda a informação necessária relativamente a estes indicadores dentro do departamento de Recursos Humanos.

A análise mostrou que, apesar dos KPIs existentes estarem bem definidos e serem bem calculados, não são suficientes para incluir todas as classes de master data e são também muito abrangentes, tornando quase impossível que seja encontrada a raíz do problema dos RH na empresa.

Palavras Chave: Key Performance Indicators, Qualidade dos Dados, Master Data de Recursos Humanos, Master Data, Indicadores de Performance

Acknowledgements

I would like to thank my professor and advisor Prof. Engª. Ana Maria Marques Ribeiro Dos Santos Lucas for all the support and knowledge shared throughout the development of this master thesis.

Also to my dearest friend Paula Fonseca for making my English even better.

List of Tables

Appendix:

List of Images

Appendix:

Abbreviations

KPI – Key Performance Indicators

HR – Human Resources

MDF – Master Data Officer

MDO – Master Data Owner

HRBP – Human Resources Business Partner

OM – Organizational Management

HRS – Human Resources Service

Table of Contents

1. Introduction

1.1 Main purpose of the study

The main objective of this research consists of, using a unique case study, investigating the importance of the Key Performance Indicators (KPIs) as a measure of the data quality of Master Data in a company. In order to do that, the study will rely on the calculation of some specific KPIs of a multinational company that operates in four business sectors - Industrial Technology, Mobility Solutions, Consumer Goods and Energy and Building Technology - and compare the results obtained in 2015 and 2016.

Therefore, it is possible to separate the study into the topics below:

- Creation of a reference guide related to Data Quality and performance indicators (KPIs) and their relationship with Master Data through the literature review;
- The importance of data quality in an organization;
- The importance of good maintenance of data quality for Master Data;
- Clarification of the steps followed to calculate the KPIs;
- Demonstration and comparison of the KPIs results;
- The difficulties in calculating the KPIs;
- Conclusion about the importance of the KPIs as a method for the evaluation of Master Data quality.

1.2 Context of the study

In a time where companies are giving more and more importance to the implementation of information systems to support their businesses, the term Master Data is becoming more usual. Master Data is the core information of a company, related to customers, products, suppliers and accounts (Dreibelbis, et al., 2008).

As such, it is important not to forget that Master Data needs to be of good quality in order to provide the company with the best and most accurate results.

In the early stages, digitalization was about the automation of processes and the invention of the internet but nowadays is about Cloud, Big Data and the Internet of Things (IoT). Nevertheless, there is something that has been key throughout: the data quality behind all this evolution (Scheuber, 2015).

1.3 Research Question

The study has the purpose of answering the question below:

- How does the **definition, calculation**, **divulgation** and the **use** of the correct measures of KPIs impact on the quality of Master Data?

1.4 Relevance of the Study

This study is relevant since it allows a better understanding of the importance organizations give to the quality of their data, in this case a particular multinational company.

If organizations become more aware of the difference that good data quality makes to their business goals, then they will begin to take better care of their data. This will also influence other companies that will realize that, in order for them to grow professionally, they need to focus on much more than sales and revenue.

This is why the study is so important, not only for multinational companies but also for small businesses to understand: the high quality of data is key for business improvement and for better achievements.

The present study is divided into five chapters: the first being an introduction to the context and relevance of the study; the second dedicated to the literature review; the third to the definition of the research method used and how the company measures its data quality; the fourth concerning an analysis of the results and, in the fifth, the conclusions as well as the study limitations and leads to future investigations in the area.

2. Literature Review

It is fundamental for this dissertation, as well as any other scientific and academic project, to begin with a literature review (Webster & Watson, 2002).

Master Data is, as per referred below (Table 1), the core information of a company, related to customers, products and even associates (Dreibelbis, et al., 2008). Loshin (2010) agrees and even adds that this data defines the fundamental business objects used in a company across different applications, which are measured and also used in reporting. Microsoft (2006) specifies a bit more about how this data is maintained within companies:

> *(...) the critical nouns of a business and falls generally into four groupings: people, things, places, and concepts. Further categorizations within those groupings are called subject areas, domain areas, or entity types. For example, within people, there are customer, employee, and salesperson. Within things, there are product, part, store, and asset. Within concepts, there are things like contract, warrantee, and licenses. Finally, within places, there are office locations and geographic divisions. Some of these domain areas may be further divided. Customer may be further segmented, based on incentives and history. A company may have normal customers, as well as premiere and executive customers. Product may be further segmented by sector and industry.*

This definition and specification of Microsoft is in accordance with Radcliffe (2009).

Nevertheless, there is a problem with this data: it is still not as important for business as it should be (Ravn & Hoedholt, 2008), especially in what concerns the maintenance

and monitoring of quality data - that is still treated as second and third priority (Scheuber, 2015) despite its importance for ensuring data quality business rules (Loshin, 2010).

Table I - Master Data Definition

Author	Master Data definition
Loshin (2010, p. 6)	"Master Data objects are those core business objects used in the different applications across the organization, along with their associates metadata, attributes, definitions, roles, connections, and taxonomies. Master data objects are those key "things" that matter the most (…)"
Dreibelbis et al., (2008, p. 2)	"Master Data is some of the most valuable information that a business owns. It represents core information about the business – such as customer, suppliers, products, and accounts – and the relationship between them. Each of these domains of master data represents information that is needed across different business processes, across organizational units, and between operational systems and decision support systems. In essence, master data defines an enterprise."
Loser et al., (2004, p. 1)	"In the context of business data processing, master data denote a company's essential basic data which remain unchanged over a specific period of time. These will include, for example, customer, material, employee and supplier data."
Radcliffe (2009, p. 32)	"Master data is the consistent and uniform set of identifiers and extended attributes that describe the core entities of the enterprise and are used across multiple business processes. Examples of core entities are parties (for example, customers, prospects, people, citizens, employees, vendors,

	suppliers and trading partners); places (including locations, offices, regional alignments and geographies) and things (such as accounts, assets, policies, products and services).”

In the table below, there are several definitions, from different authors, for the concept of Data Quality so it is easier to understand its importance in daily and future business for the companies. One thing is clear: all the authors agree that good maintenance of the quality of the data is a key factor for the future of a company as a means to differentiate itself from its competitors.

Table II - Data Quality Definition

Author	Data Quality definition
Knolmayer et al., (2006, p. 362-363)	“(…) a process which should ensure that all data is entered and approved with respect to business rules, and that every user and every system should receive new or updated master data as soon as needed.”
Brackett et al., (2009, p. 291)	“Data Quality is synonymous with information quality, since poor data quality results in inaccurate information and poor business performance. (…) A more rigorous data quality program is necessary to provide an economic solution to improve quality and integrity.”
Davenport et al., (2007)	“A data quality effort therefore aims to ensure data is correct, complete, current, and consistent, while data management ensures it is in context and access to it is controlled.”
Dreibelbis et al., (2008)	“For master data, data quality is of utmost concern. If a customer cannot be identified unambiguously due to poor data quality or if the bill is sent

	consistently to the wrong address, this is something that might turn customers away. (…) Governance of master data must support business and technical controls to ensure that master data is “timely, relevant, complete, valid, accurate, and consistent”.”
Wang et al., (1996)	“(…) we define "data quality" as data that are fit for use by data consumers. In addition, we define a "data quality dimension" as a set of data quality attributes that represent a single aspect or construct of data quality.”

The market is becoming more competitive and companies need to improve their processes in order to stay in the fight. As such, companies are now realizing the importance of data quality processes and their impact throughout the organization (Masayna et al. 2009). After implementing a data quality process, it is also imperative to measure the quality of the performance of the business and, in order to do that, a company needs to identify their main advantages and disadvantages or, in other terms, their high-value and high-risk data quality issues (Masayna et al. 2009).

Ravn & Hoedholt (2008) mention four situations that can be identified when implementing a data quality process:

- *Trends in data quality;*
- *Data quality issues before they impact critical business processes;*
- *Areas where process improvement is needed;*
- *A structured and methodological approach to measuring and monitoring the quality of data should be part of a larger master data management of information management strategy*

Social and economic impacts can come from poor data quality (Wang et al, 1996). There are several examples of the impact poor quality can have, but the worst is a company losing its reputation. It is important, however, to define what exactly poor and high quality is.

Garvin (1987) defines high quality as:

> *To achieve quality gains, I believe, managers need a new way of thinking, a conceptual bridge to the consumer's vantage point. Obviously, market studies acquire a new importance in this context...One thing is certain: high quality means pleasing the consumer, not just protecting them from annoyances.*

For that reason, it is important that customers have a unique source of data within a company that they can trust and on which quality measures can be based. Maintaining a single record of data is not that easy nowadays due to the company's growth through integration or even acquisitions of other companies that lead to the introduction of more and different data. The problem is not the amount of this data but the fact that they don't merge them, allowing several sets of the same data to be maintained as if they were from different customers/products/associates (the so-called duplicates).

Microsoft (2006) considers the following four advantages of the use of a single set of this type of data:

- *A single, consolidated bill saves money and improves customer satisfaction;*
- *Sending the same marketing literature to a customer from multiple customer lists wastes money and irritates the customer;*

- *Before you turn a customer account over to a collection agency, it would be good to know if they owe other parts of your company money or, more importantly, that they are another division's biggest customer;*
- *Stocking the same item under different part numbers is not only a waste of money and shelf space, but can potentially lead to artificial shortages.*

To achieve this high quality, it is important to first find a way to measure it and to improve it (Laranjeiro et al., 2015). The dimensions, as properties of data quality, are one of the ways to examine and understand the quality of data - such as: validity, integrity, completeness, consistency, timeliness and accuracy (Ravn & Hoedholt 2008; Laranjeiro et al., 2015; Scannapieco et al., 2005). Four dimensions stand out:

- Completeness: is a way to make sure that all relevant data is available to the business;
- Consistency: in this case the creation of duplicates in the system is a break of this dimension. A KPI that checks this dimension will help not only to prevent this from happening, but also to measure and present the existing duplicate problems within a company;
- Timeliness: considers the time that the master data takes to be available for use. If data is available when it is actually needed;
- Accuracy: make sure that the master data used is correct (if it comes from a trustworthy source) before starting to use it, especially in what concerns

sensitive and confidential data. This applies not only to customers but also to partners of a company.

There are several performance measures for quality whose major objective is not only to evaluate the actual situation of the organization, but also how to improve the business and how to prevent future performance problems (Ghalayini et al., 1996), never disregarding that it is a way to fight the competition (Gabcanova, 2012).

In what concerns Human Resources (HR), Gabcanova (2012) highlights that one of its main advantages is human capital and it is clear to executives that the associates are a key factor in a company, but for them the HR are not yet essential on their thrive to success (Becker et al. 2001). So this can only mean that it is not possible, for most companies, to measure the impact of the HR in their business (Becker et al. 2001). On the other hand, the main problem of the HR data is the data quality (Becker et al. 2001) so, as soon as managers realize that their main problem can also be their main improvement, they will see the importance HR data really has and how important it is to define ways to measure their performance as a way to get better results and to become more competitive.

But what exactly are performance measures? Several authors have written about frameworks (e.g. balance scorecard), business excellence model, individual performance measures, measurement systems and how each of them are better or not for a company, depending on its goals and mission. Neely et al. (2004) define them as a way to quantify the efficiency and effectiveness of an action, always taking into consideration the satisfaction of the customers. Keegan et al. (1989) add that performance measures derive from business.

Within these performance measures it is possible to highlight the Key Performance Indicators (KPIs) which are measures used to define whether the business is going well or not, and quantify the business goals using, usually, percentages (0%-100%), with 100% indicating the highest level of quality (Masayna et al. 2007; Ravn & Hoedholt 2008). Parmenter (2015) enhances that these indicators *"focus on the aspects of organizational performance that are the most critical for the current and future success of the organization".* KPIs are also the measure indicators mostly adopted by companies (Masayna et al. 2007).

For these indicators to have a positive impact on the measurement of the quality of Master Data, it is crucial that they are well defined focusing on improving performance (Fiksel, 2002). And, in order for that to happen, there are several steps that should be followed. According to Ravn & Hoedholt (2009) there are four basic steps:

1. Define master data objects of importance (e.g., customer data);

2. For each master data object, define a set of data quality KPIs;

3. For each KPI, define measure details;

4. Define procedures for follow-up on data quality issues.

Parmenter (2015) came to the conclusion that there are seven foundation stones for the implementation of KPIs:

1. Partnership with staff, unions, and third parties;

2. Transfer of power to the front line;

3. Measure and report only what matters;

4. Source KPIs from the critical success factors;

5. Abandon processes that do not deliver;

6. Appointment of a home-grown chief measurement officer;

7. Organization-wide understanding of the winning KPIs definition.

Masayna et al. (2007) add that six principles must also be taken into consideration when defining those indicators:

An indicator must motivate the right behavior;

A KPI must be measurable;

A KPI must be affordable;

The objective set for a KPI must be attainable;

Factors affecting the indicator must be controllable by the service provider;

A KPI must be meaningful to all parties.

After the definition of the KPIs, taking into consideration the stakeholders' and business needs, it is important to monitor them, check for useful improvements and, in the end, redefine them if necessary (Masayna et al., 2007).

Consequently, three statements arise from the present study whose purpose is to understand the relationship between KPIs and Data Quality.

S1: Well-defined KPIs have a positive impact on Data Quality;

S2: Good divulgation of the KPIs within the company will improve Data Quality;

S3: The correct use of the KPIs will improve Data Quality.

3. Methodology

There were several options regarding which research method to choose for the master thesis: experiment, survey, a case study (qualitative), archival analysis and history (Yin, 2013). The conclusion was that an exploratory case study would be the best option since its main objective is to define statements which lead to the need for more investigation. In order to do that, the research question had to be "Why" or "How" related (Yin, 2013).

According to Yin (2013), one of the possible definitions for a case study is: *"The essence of a case study, the central tendency among all types of case study, is that it tries to illuminate a decision or set of decisions: why they were taken, how they were implemented, and with what result. (Schramm, 1971, emphasis added)"* and that it can be considered an *"…empirical inquiry that investigates a contemporary phenomenon (the "case") in depth and within its real-world context, especially when the boundaries between phenomenon and context may not be clearly evident."*. These definitions clearly explain why this was the method chosen, as it is a way to analyze a real and current situation and to draw conclusions about future impacts not only on the company, which the case study is going to focus on, but also on the future of the companies.

3.1 Company under study

In order to shed light on the importance of the KPIs in measuring Master Data, the master thesis presents a case study done in a multinational company (with more than 389.000 employees and a revenue of around 73 billion euros) that operates in four business sectors: Industrial Technology, Mobility Solutions, Consumer Goods and

Energy and Building Technology. The reason for choosing this company is not only because it is a company with great impact on the IoT/Smart Cities/Automotive industry all around the world, but also because of the importance they give to their associates and more specifically HR. Having a contact in the HR-IT department was also a reason in favor of choosing this company as it provided an easier access to all the information, documentations and interviews. As mentioned before, most companies don't see the importance HR have in their company but the company this case study focuses on, sees it. They hold HR in such a high regard even though it is not their core business because they know that if they guarantee its master data quality, then they will improve greatly in their relationship with target systems, consumer systems and, in the end, they will create a big impact on the way their company is seen by the world.

It is important to clarify that KPIs are calculated throughout the company in several segments, but this master thesis will focus on Human Resources KPIs (the ones that refer to the measure of quality within the HR departments) and its relationship with data quality of the Master Data.

3.2 Concepts clarification

First, it is important to clarify several concepts used within the company.

There are two major roles, the Master Data Owner (MDO) and the Master Data Officer (MDF). The MDOs have the responsibility of making decisions regarding the business needs as well as the definition of the use and management of Master Data. There are one or two MDFs per country and they are responsible for assuring the quality of the

master data within their country. There is also a Master Data Quality Team located in Portugal responsible for contacting the MDFs every time that a data quality issue is found (reports and other data quality checks are done every day by this team in order to assure the best quality possible).

For Human Resources, the company uses an SAP software program called COMPAS (Corporate Organizational Master Data and Personnel Administration Service), which is a human resources system customized for the company's business needs, developed on an SAP platform. Its purpose is to deliver integrated and consistent master data for the four Master Data classes (appendix figure 1): Organizational, Position, Person and Communication data (figure 1).

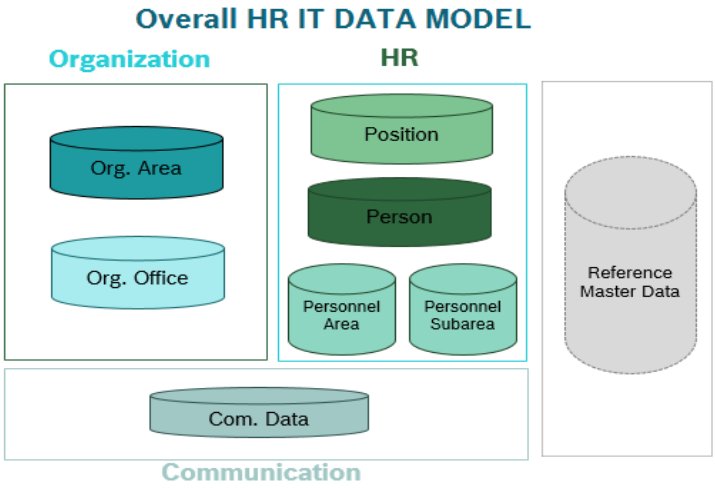


Figure 1 - Overall HR IT Data Model

Source: *Own*

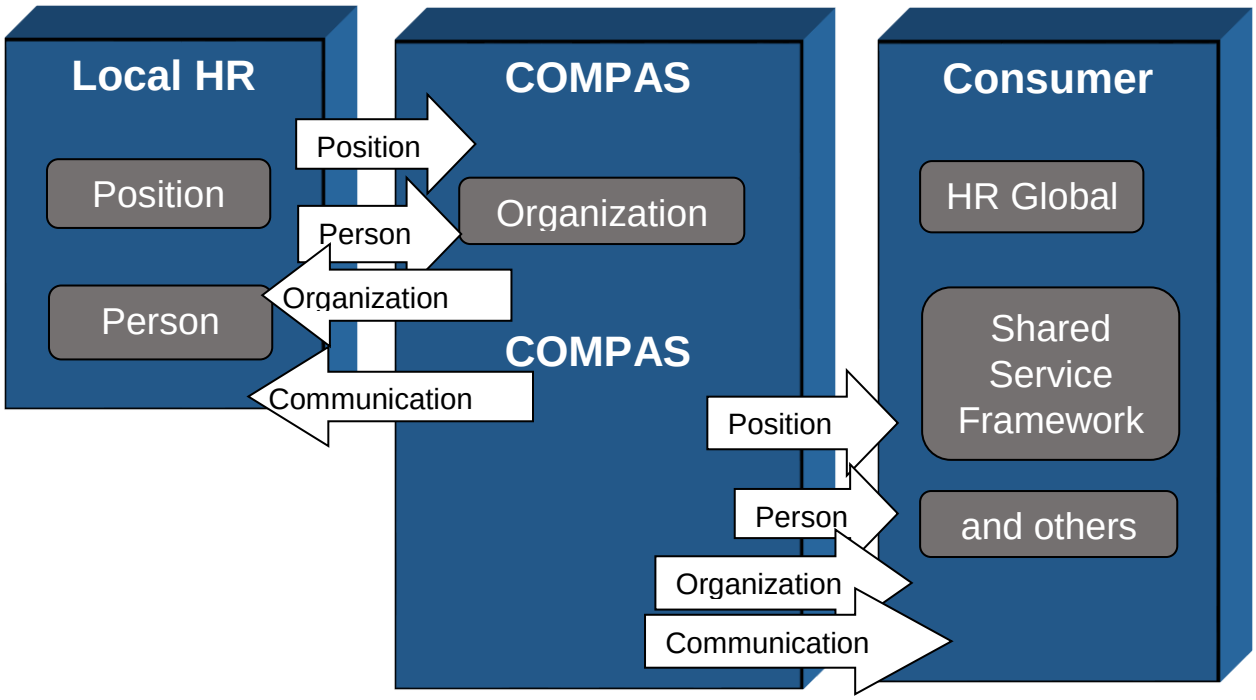


Figure 2 - Data Model process flow
Source: *Own*

On the structure above (figure 2), it is possible to understand how the four Master Data classes flow within the systems: SAP HCM systems with payroll and time management for several countries, known as Local HRs, send the information regarding Person and Position to COMPAS (Master Data database) which sends this information to all consumer systems. On the other hand, the organization data is created directly in COMPAS and from there it is sent to the Local HRs and consumer systems. Tables I to IV in the appendix detail all the objects that belong to the previous classes. Bad quality in any of these classes will have an enormous impact on all the systems. In the appendix, figures 2 to 4, the relational model for these classes is presented so it is worthwhile understanding how the classes relate to each other.

COMPAS was designed to provide HR, OM (Organizational Management) and user data of excellent quality. Training opportunities were established to safeguard this quality.

In addition, technical check mechanisms were introduced for when a new associate is hired (such as providing a table, where the name and date of birth of the new associate are inserted, hereby detecting any possible duplicates – associates with the same name and date of birth that already exist in the system. Thus, it prevents the creation of duplicates). However, the company still faces some quality challenges that cannot be prevented by the existing safeguarding measures.

3.3 The KPIs

The KPIs were then introduced in 2014 by the MDOs (though the target values – a value that has to be achieved in order for the KPIs' results to be considered as "Excellent" - were defined by the steering committee) to close the existing gap in Master Data quality that cannot be prevented by the already existing technical measures, especially as more and more follow up processes and other HR and non-HR rely on a good COMPAS data quality. These values are calculated every two months by two people (one in Portugal that calculates the HR KPIs, and one in Germany that calculates the OM KPIs) using the SAP HCM software to extract some reports which are filtered and analyzed through Excel. In the end, a separation is made by country in order to get a clearer picture of the results per country, which will then be provided to the respective MDF.

There are 8 KPIs calculated for HR and OM but, even though this thesis will focus only on the HR, a brief reference will be made concerning all the KPIs:

Note that KPIs 1-4 concern OM and KPIs 5-8 concern HR data.

OM KPIs	HR KPIs
1 – Managed offices	5 – On time data availability
2 – Abandoned offices	6 – Correct target and disciplinary manager
3 – Average office size	7 – No duplicates
4 – Official office ratio	8 – No person with more than one position

Table III - The different KPIs (OM and PA)

Source: *Own*

KPI 1: Managed offices

Data is grouped for one region.

The aim of this KPI is to evaluate the number of official offices in COMPAS without a manager assigned to it. An absence of managers can lead to problems such as a lack of people responsible for audits, a lack of people with the authority to release approvals, etc.

The KPI measures the quotient of all existing offices compared with the unoccupied offices (%) – insofar as having a manager or not is concerned.

Summary KPI1:

- Database: All official offices (region) in COMPAS underlying period;
- A residuum of approx. 5% vacant manager positions is expected.

The defined target value is **95%.**

KPI2: Abandoned Offices

The aim of this KPI is to eliminate dead knots within the organizational tree as they dilute the focus on vital parts on the structure (e.g. by worsening KPI1) and might be used by mistake (leading to data privacy and security problems).

The KPI measures the quotient of units (including non-official ones) without people and manager with all units (including non-official ones).

Summary KPI2:

- A residuum of approx. 5% abandoned is acceptable; however, the figure will be reduced next year.

The defined target value is **5%.**

KPI 3: Average office size

Data is grouped by country.

The aim is to avoid and prevent atomistic small or single-person offices from being created.

Too many small-sized offices lead to a lack of transparency in responsibilities and reporting lines, as well as an increase in maintenance. Offices with five associates or more are accepted as suitable.

Simultaneously, there is a form that must be filled in when an office with less than five associates is planned. However, these offices are counted as not suitable.

To evaluate KPI 3, the total number of all associates in a region in relation to all offices is measured.

Summary KPI 3:

- Database: Whole COMPAS data at a certain date. Data is grouped for one region underlying period.
- A transparent and efficient organization requires a balanced relationship between offices and associates; therefore this KPI serves as a landmark to prevent too many micro-units or even single-person-offices.
- To normalize the target value, expected first-best figure (8 associates/office) is fixed as denominator.
- Defined target value is **95%.**

KPI 4: Official Office Ratio

Data is grouped for one region.

The purpose of this KPI is to determine the relationship of non-official offices in COMPAS with all offices within the region and simultaneously prevent the use of those offices for other purposes than inactive work contracts as this leads to problems in consecutive systems.

The KPI is calculated by the quotient of all existing offices compared with the unofficial offices (%).

Summary KPI 4:

- Database: Whole COMPAS data at a certain date. Data is grouped for one region underlying period.
- The non-official organizations should only be established for inactive work contracts. If these were created for any other reason, they should be extinguished. A residuum of approx. 5% can be assumed as a realistic figure for non-official offices (designed to host inactive workforce e.g. due to partial retirement).
- Defined target value is **95%.**

From KPI 1-4, to double check the results, a request will be sent to each region on the database, which will allow for a better analysis and, subsequently, for appropriate actions to be taken.

KPI 5: On time data availability

Databases are all SAP master data actions which were changed in the corresponding KPI evaluation period.

For the interpretation of the single datasets, they compare the change date with the start date or end date. Only actions which have had an effect on the infotype (SAP specification used to save and organize information) of actions carried out at associate level are considered.

If the change date is more than 10 days later than the relevant start or end date, it is counted as "belated" and thus has a negative effect on the KPI.

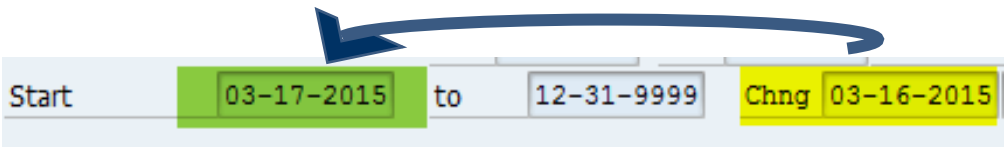
The KPI measures the quotient of all actions compared to the belated maintained actions (%).

Summary KPI 5:

- Database: All actions of the underlying period.
- "Externals" won't be considered for the KPI calculation.
- It is known that the responsibility of the on-time data maintenance depends also on the punctual delivery of the data. With this KPI, it will be easier to find the reasons for the belated data maintenance and to take suitable actions.
- Defined target value is **99,5%** (percentage of associates not affected by retroactive modifications).

$$\text{Formula} = \left(1 - \frac{\text{number of relevant retroactive modifications}}{\text{number of related actions in COMPAS while corresponding period}}\right) * 100$$

This KPI compares the “change date” with the “start date” or “end date”



If the change date is more than 10 days later than the relevant start date or end date, it is counted as “belated” and thus has a negative impact on the KPI.

The actions considered are: hiring (when an associate joins the company), exits (when an associate leaves the company) and reentry (if an associate previously worked at the company but left for some time and has now returned).

For example: An associate is hired to start on 01.08.2017 (start date) and the request is sent from the HRBP to the HRS Team. The data for this associate needs to be introduced in the system (COMPAS) by 10.08.2017 (changed date) at the latest, so it won't be negatively considered in the KPI.

If, on the other hand, the associate leaves the company on 31.08.2017 (end date), then that information needs to be introduced by 10.09.2017 (changed date) at the latest.

Employees that work as outsourcers (externals) are not considered in this calculation since the company does not pay them directly.

KPI 6: Correct target and disciplinary manager

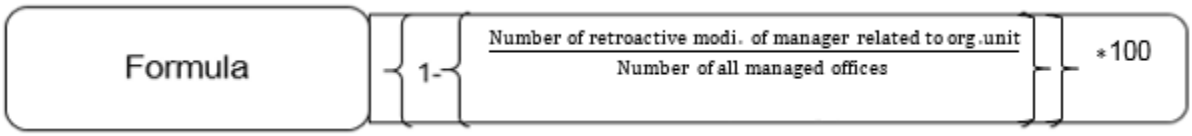
As an analogy, with the KPI 5 method, the change date is compared with the start or end date of the affected dataset.

The KPI measures the quotient of all retroactive modifications related to all managed offices (departments).

The database will receive a list of all organizational offices (departments) which were affected by retroactive modifications.

Summary KPI 6:

- Database: All retroactive modifications on manager related OM-Objects as well as all managed offices.
- Defined target value is **99,5%** (percentage of associates not affected by retroactive modifications).



This KPI, as per analogy with KPI 5, also compares the start/end date with the change date. But, in this case, it does not concern the hiring or leaving of an associate but the action of assigning a manager to a department, also called Organizational Unit. It is mandatory in the company that every time a new department is created or every time a manager leaves a department that a new one is immediately assigned. Organizational units without a manager are not allowed.

As such, this KPI checks if the assignment of the new manager to the org. unit occurs exactly after the previous manager leaves, not leaving any “blank spaces”.

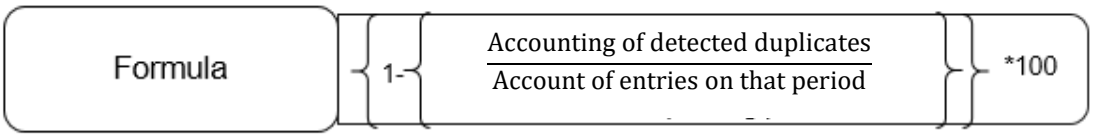
KPI 7: No duplicates

The aim of this KPI is to measure the number of double created persons. If an associate is created twice in the system, it will cause a problem in the consumer system connected to COMPAS.

The KPI calculates the relation between the number of detected duplicates and the number of associates entered in the system.

Summary KPI 7:

- Database: All COMPAS data related to the Person object.
- The responsibility for the merging process lies with the MDFs of each region (necessary in case an associate is created twice in the system by mistake).
- Defined target value is **99%** (percentage of associates with no duplicates)**.**



This KPI checks all the associates that were introduced in the system in that time period (the 2 months considered in all the KPI calculations) and how many were merged (the duplicates that were created, found and merged together to become one again). The target of 99% means that at least 99% of the entries in the period the KPI is measured should not have a duplicate.

KPI 8: No person with more than one position

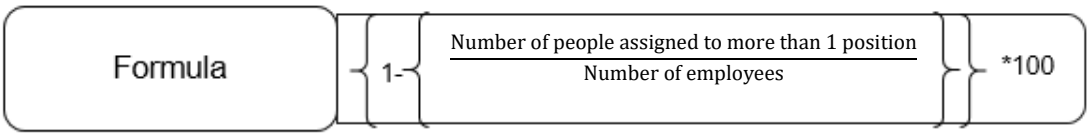
This KPI will allow to determine if the associate has more than one position assigned to him.

It is a basic principle that a person must be assigned to one position and this position must be assigned to a single office (department) - the one the HR costs are accounted to.

The existing formula will calculate the quotient between the number of people existing and the number of associates with more than one position (%).

Summary KPI 8:

- Database: All COMPAS data related to the Person object.
- One person = one position rule should be standard applied.
- Defined target value is **99%** (percentage of associates with one position)**.**



This last KPI checks how many associates, in that period, were placed in two or more positions compared with the total amount of employees in the company. The target (99%) stands for the percentage of associates that should be in only one position; 1% is the tolerance for associates with more than 1 position (even though, ideally, this should be 100% of associates in only one position).

3.4 What to do with the KPI results

As previously mentioned, the KPIs are calculated every two months and the results are sent to the MDOs as well as to the MDFs and are discussed in several meetings.

During a meeting with one of the associates, who is responsible for the data quality in the company, it was possible to get some information about the origin of these KPIs and their actual impact within the company.

As mentioned before, the KPIs were introduced in the company by the end of 2014 by the MDOs. They defined these KPIs as they focus on the areas that they felt had the most problems, and this was the most efficient way they found to actually measure the impact of the choices made in the company for these Master Data Classes (why they chose these performance measures and not others is not known). They decided then that a separation would be made by country, so it would be clearer which countries were having the worst results so that preventative and correction measures could be taken.

With every round of KPI monitoring, the company investigates improvements in key figures in many regions - important steps in the right direction. What motivates the MDOs even more is that some regions managed to exceedingly increase their figures, and the general awareness the quality topic received since they started the monitoring.

3.5 Statements

S1: Well-defined KPIs have a positive impact on Data Quality.

As noted by Ravn & Hoedholt (2009), for the KPIs to be considered “well defined” they need to be related to important master data objects. Considering that the KPIs for this company measure customer data, the first step is fulfilled in order to achieve “well-defined” KPIs.

Each master data object should have a set of data quality KPIs (Ravn & Hoedholt, 2009) which, in the case of the company in study, applies to master data classes and not objects. The step is then not fully followed.

A measure should be defined for each KPI. This is also not being done since the measures, for this company, are applied in general for all of them and not for each individually.

Future procedures for follow-up are being applied and lessons learned are being taken into consideration for the next steps.

According to Parmenter (2015), the KPIs should only measure what matters and their source should be the critical success factors which, in the case of this company, apply even though there are not enough KPIs for all critical factors.

The KPIs must be measurable, affordable and attainable (Masayna et al., 2017). What leads to more discussion in the previous sentence is that the target values for each KPI are not attainable for the majority of the countries involved because not all the factors are controllable by the company; they may diverge because of law restrictions or process definition within each country. This shows acceptance regarding the statement because, as defined by the three authors above, it is imperative that the performance measures (KPIs) are well defined. Only then they are effective. And, in the case of the HR KPIs presented above, they follow the steps defined by the authors above. This doesn't mean that if they are well defined then they will increase data quality, but it is indeed a step in order to achieve a positive impact on Data Quality.

S2: Good divulgation of the KPIs within the company will improve Data Quality.

The MDFs were contacted in order to find out what the main reasons for the target values not being achieved in their country could be. This revealed that their main concerns were that the maintenance teams and the managers responsible for requesting the data to be maintained were not aware of the impact that delays in requests could have on KPIs because most of the countries do not understand the importance that these quality KPIs could have on their daily business and on the overall company goals. When talking with the MDFs to better understand how they could improve their results, it was clear that, on the following month's KPIs, some improvements had taken place (even though they were not enough).

This shows acceptance regarding the statement because, without knowing its importance and its impact, there is certainly no reason to strive to better.

S3: The correct use of the KPIs will improve Data Quality.

"More than the correct use of the KPIs, there is the understanding of what quality means" – this was a reference made in the interview with a company's associate responsible for DQ - "Taking the correct measures after getting the results is crucial but understanding what we are measuring is even more important". This was definitely not their strength in the past, as they just took the numbers defined from the top management without asking for what the organization really needed.

Now they are trying to improve and that is why they created a Master Data Quality Team in Portugal. With this and the several reports they extract every day, it was possible for

them to start defining what quality is. It is also clear to them that the processes within the company for every country need to be harmonized in order for the KPI goals to be achieved. In this case they defined “correct use” as a global use, as for all the countries to have the same rules and to be coherent with the processes, without allowing a lot of exceptions (this is also a big problem they are facing nowadays: every country or every location requires exceptions for everything, and allowing this will only separate the company more). Once again this doesn’t mean that the correct use of the KPIs improve data quality directly but it will indirectly, as the correct and better use of Master Data (especially regarding its quality) will have a positive impact on the KPI results.

4. Result Analysis

With the training and documentation provided, it was possible to calculate the KPI results for the year 2016 and compare them with the results already available from 2015.

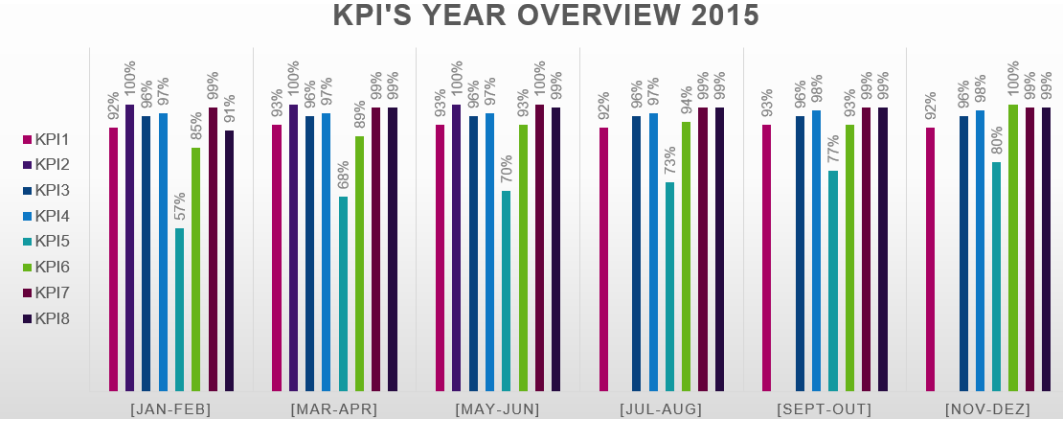


Figure 3 - KPI Overview for 2015
Source: *Own*

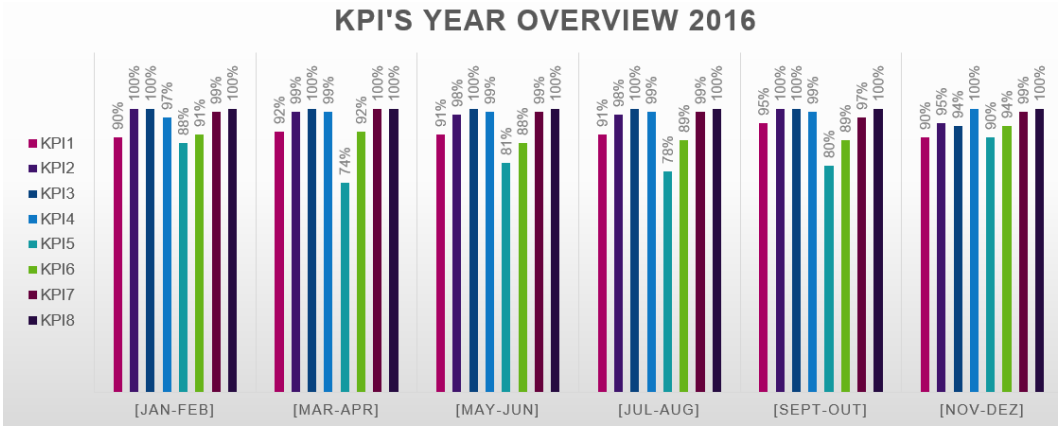


Figure 4 - KPI Overview for 2016
Source: *Own*

A deeper analysis into KPIs 5, 6, 7 and 8 showed that, on average, KPI 5 and 6 were clearly the ones with worse results (figure 5 and figure 6), even though KPI 5 had a slight increase from 2015 to 2016. KPI 6, on the other hand, decreased sharply in March-April from 2015 to 2016, with a slight increase afterwards.

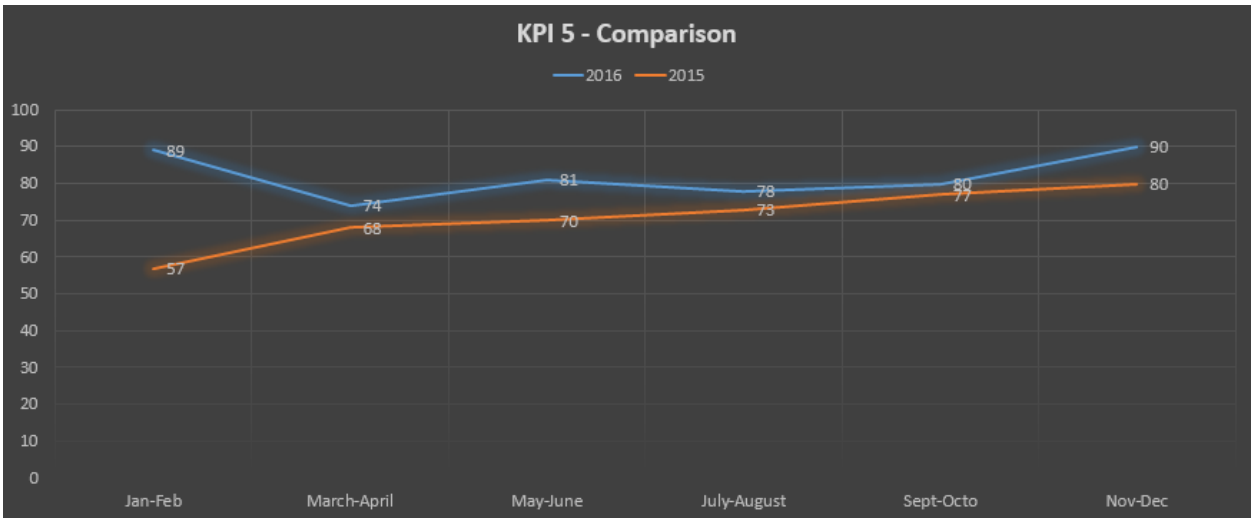


Figure 5 - KPI 5 Comparison for 2015/2016
Source: *Own*

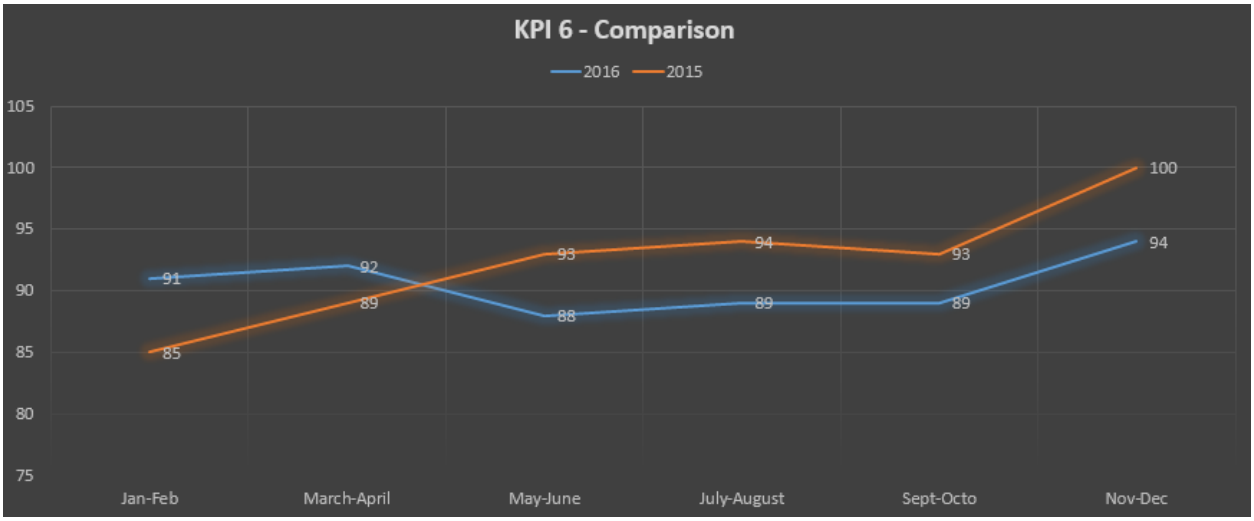


Figure 6 - KPI 6 Comparison for 2015/2016
Source: *Own*

Bearing these results in mind, the MDFs were contacted in order to understand what the problems in their countries could be that prevented them from achieving the target results.

The conclusions, at the end of 2016, are shown in figure 7:

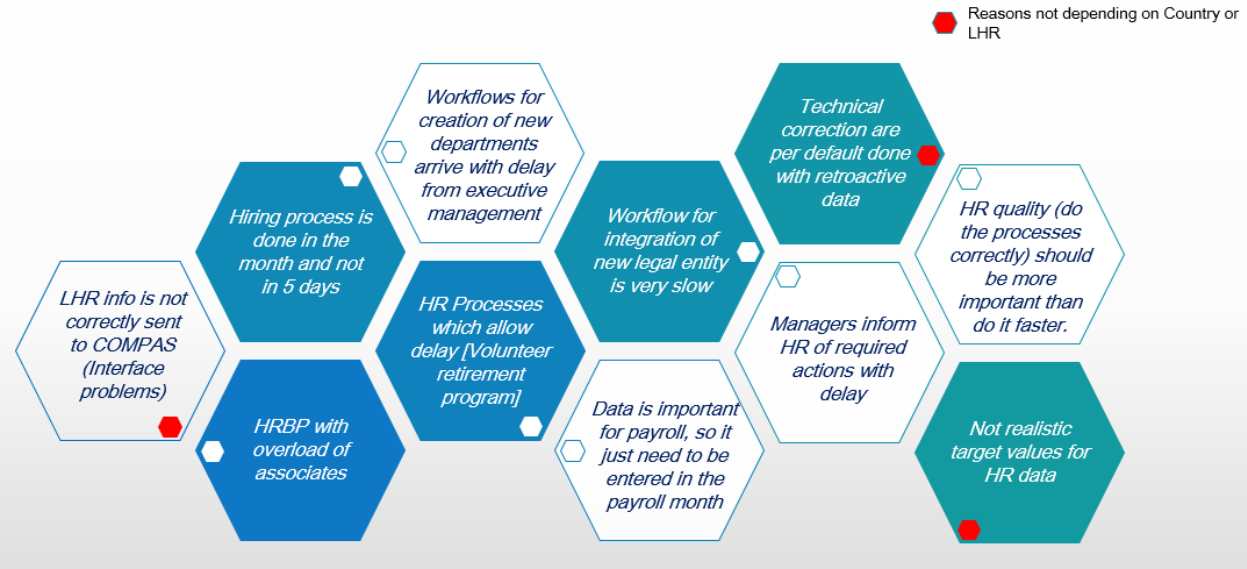


Figure 7 – Reasons why target values are not achieved (end 2016)
Source: *Own*

And some suggestions were made:

- Establish a limit of associates per HRBP;
- Awareness for managers to contact the HRBP at an earlier stage;
- Business partners support in submitting requests on time (at least 5 days before the start date).

The main problem frequently highlighted by the MDFs was that the communication between the Human Resources Business Partner (HRBP) and the HRS Teams took place with a huge delay (the HRBP is the person responsible for a group of associates per location and has the responsibility of informing the HRS Team when a hiring, exit or reentry of an associate happens. They can only process the changes when they get a request from the HRBP, until then no action is taken, even if the associate has already started to work/has left the company).

Five months later, the same question was put to the MDF in order to check if there were any improvements and the results are shown in figure 8.

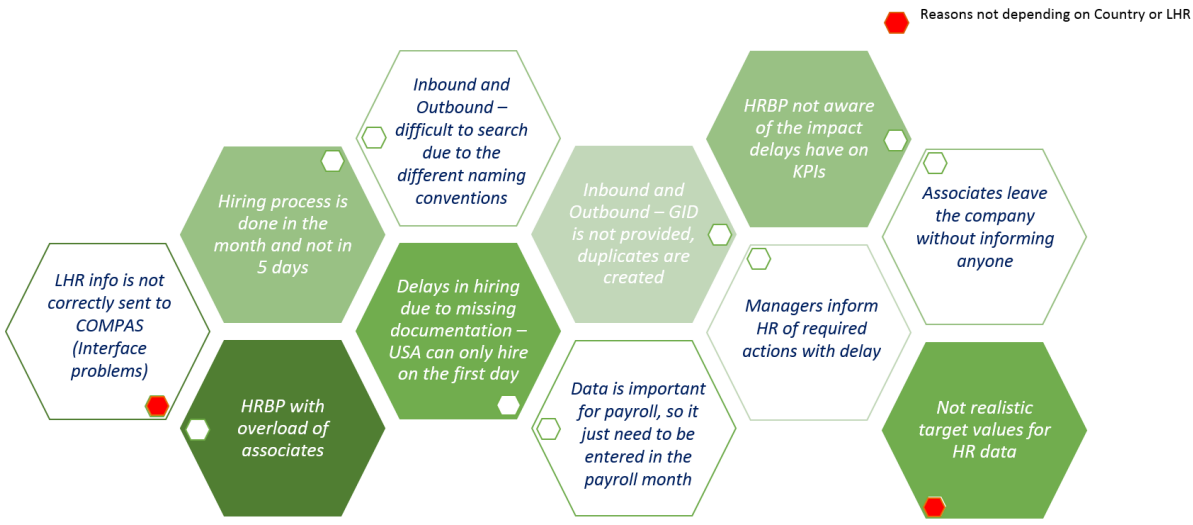


Figure 8 - Reasons why target values are not achieved (mid 2017)
Source: *Own*

The results, surprisingly for the company, were quite similar – HRBP communication to HRS is still the major complaint. Then the question that arises is: Was something done between those periods in order to improve the results? Yes. Was it enough? Clearly not. The company instructed the MDFs to establish some rules and inform the HRBPs about the maximum time that they can take to send the requests. Spain was the only country that took serious measures as they forbade delayed requests from the HRBPs. The feedback obtained from the other HRBPs was that they try, but as they have so many associates under their supervision and such a large workload, it is difficult to handle everything on time.

The conclusions taken from these results of the KPIs, of their evolution and of the actions taken to improve them are that the KPIs' results are still not improving, which means that further action from the company should be taken; the processes within the countries need to be discussed and defined so a common goal is determined; more capacity is needed since the HRBPs have too many associates to handle; communication within the company needs to improve to prevent the creation of duplicates.

5. Conclusion

The purpose of the present case study was to investigate if the company which this study focuses on, defines, calculates, divulges and used the correct KPIs in order to improve the quality of HR Master Data.

The calculations that were made showed that, even though the KPIs follow some of the steps in terms of how they are defined, they are not being efficient in the company since the results are not improving. This could mean that a reevaluation should be done of all the KPIs in order to understand if new ones should be created, or even if the existing ones should be changed. A separation, for example, in KPI 5, between hiring/reentry and exit could be a good way to understand what the actions that have a worse result are, if the problem is more when an associate joins the company or, if on the other hand, the problem is when the associates leave. Also, a deeper overview about what happened with KPI 6 in March-April from 2015 to 2016 would be useful to get a better overview of what happened in each country or even in the company worldwide that might have affected these so much. It is important then to review the calculation of these indicators, also to make sure that they include all business relevant information.

The divulgation of the importance and impact that these indicators have throughout the company is a matter that still needs to improve a lot because unless the associates and managers understand its importance, no action will ever be enough to change these results. Everyone should be working for a common goal and a common process needs to be developed for that to happen. Training is also required, especially for the HRS Teams, so they better understand the connection between COMPAS and the other

systems. With this it would be easier to understand the impact any change can have on the company.

Finally, the actions that are being taken after receiving the KPI results are also not enough. The feedback from the MDFs was very important as it allowed for a better understanding of the problems of their respective countries, and to understand the countries' perspective, but still no relevant action was taken by the company. It is clear that the HRBPs are overwhelmed with work so more capacity is needed; training is also required for them and for managers to better understand the changes needed to increase performance.

It is hoped that this case study contributes towards the little literature that is available concerning the relationship between Data Quality and Key Performance Indicators. It is also hoped that, taking into consideration the company under study, it is possible to understand how complex and important it can be to define, calculate and apply performance indicators, never disregarding that each company is different and a that detailed analysis is required for each case in order to maximize each one's performance and quality in their data.

The main limitation found during the execution of this case study was the fact that there was not a lot of information available concerning the relationship of this performance indicator and data quality.

For future research, it would be relevant to investigate the other performance indicators more thoroughly in order to understand which could be better for each company or department to use.

6. References

Bauer, Kent (2005). "KPIs: Avoiding the Threshold McGuffins." *Information Management* 15.4: 47.

Becker, Brian E., Mark A. Huselid, and David Ulrich (2001). *The HR scorecard: Linking people, strategy, and performance*. Harvard Business Press,: 1-2.

Brackett, Michael, and Production Susan Earley (2009). "The DAMA Guide to The Data Management Body of Knowledge (DAMA-DMBOK Guide)."

Davenport, Thomas H., and Jeanne G. Harris (2007). *Competing on analytics: The new science of winning. Harvard Business Press*.

Dreibelbis, Allen, et al. (2008). *Enterprise master data management: an SOA approach to managing core information*. Pearson Education.

Fiksel, Joseph (2002). "Substudy 5: Key Performance Indicators.".

Gabcanova, Iveta (2012). "Human resources key performance indicators." *Journal of competitiveness* 4.1: 117.

Garvin, David (1987). "Competing on the eight dimensions of quality." *Harv. Bus. Rev.*: 101-109.

Ghalayini, Alaa M., and James S. Noble (1996). "The changing basis of performance measurement." *International Journal of Operations & Production Management* 16.8: 63-80.

Herzog, Thomas N., Fritz J. Scheuren, and William E. Winkler (2007). *Data quality and record linkage techniques*. Springer Science & Business Media.

Jaaskelainen, Olli, Hanna Kropsu-Vehkapera, and Harri Haapasalo. "Lifecycle View On Product Data Quality Dimensions."

Kaplan, Robert S., and David P. Norton (1996). "Using the balanced scorecard as a strategic management system.".

Keegan, Daniel P., Robert G. Eiler, and Charles R. Jones (1989). "Are your performance measures obsolete?." *Strategic Finance* 70.12: 45-46.

Knolmayer, Gerhard F., and Michael Röthlin (2006). "Quality of material master data and its effect on the usefulness of distributed ERP systems." *International Conference on Conceptual Modeling*. Springer Berlin Heidelberg.

Laranjeiro, Nuno, Seyma Nur Soydemir, and Jorge Bernardino (2015). "A survey on data quality: classifying poor data." *Dependable Computing (PRDC), 2015 IEEE 21st Pacific Rim International Symposium on*. IEEE.

Loser, Cornel, Christine Legner, and Dimitrios Gizanis (2004). "Master data management for collaborative service processes.".

Loshin, David (2001). *Enterprise knowledge management: The data quality approach*. Morgan Kaufmann.

Loshin, David (2010). *The practitioner's guide to data quality improvement.* Elsevier: 249.

Lucas, Ana (2010). "Corporate data quality management: From theory to practice*." 5th Iberian Conference on Information Systems and Technologies*. IEEE.

Masayna, Veerawat, et al. (2007). "Data quality and KPIs: a link to be established'." *The 2nd World Congress on Engineering Asset Management (EAM) and The 4th International Conference on Condition Monitoring*: 1377-1379.

Masayna, Veerawat, Andy Koronios, and Jing Gao (2009). "A framework for the development of the business case for the introduction of data quality program linked to corporate KPIs & governance." *Cooperation and Promotion of Information Resources in Science and Technology, 2009. COINFO'09. Fourth International Conference on*. IEEE: 230.

Microsoft (2006). SOA in the Real World. Retrieved March 11, 2017, from https://msdn.microsoft.com/en-us/library/bb833022.aspx

Neely, Andy, Mike Gregory, and Ken Platts (1995). "Performance measurement system design: a literature review and research agenda." *International journal of operations & production management* 15.4: 80-116.

Parmenter, David (2015). *Key performance indicators: developing, implementing, and using winning KPIs*. John Wiley & Sons.

Pula, Edith N., Merlin Stone, and Bryan Foss (2003). "Customer data management in practice: An insurance case study." *Journal of Database Marketing & Customer Strategy Management* 10.4: 327-341.

Radcliffe, John (2009). "Magic quadrant for master data management of customer data." *G00206031, Gartner, Inc., Stamford, USA*: 32.

Ravn, Thomas and Hoedholt, Martin (2008). “Measuring and Monitoring the Quality of Master Data”. *Platon*: 1-7

Scannapieco, Monica, Paolo Missier, and Carlo Batini (2005). "Data quality at a glance." *Datenbank-Spektrum* 14.January: 6-14.

Scheuber, Ulrich. Linkedin. “Master Data Management a next step of “The digital Revolution”?”. Retrieved April 29, 2017, from https://www.linkedin.com/pulse/master-data-management-next-step-digital-revolution-ulrich-scheuber

Smith, Heather A., and James D. McKeen (2008). "Developments in practice XXX: master data management: salvation or snake oil?" *Communications of the Association for Information Systems* 23.1: 4.

Webster, Jane, and Richard T. Watson (2002). "Analyzing the past to prepare for the future: Writing a literature review." *MIS quarterly*: xiii-xxiii.

Wang, Y. Richard, Lisa M. Guarascio, and Richard Wang (1991). "Dimensions of data quality: Toward quality data by design.".

Wang, Richard Y., and Diane M. Strong (1996). "Beyond accuracy: What data quality means to data consumers." *Journal of management information systems* 12.4: 5-33.

White, Andrew, et al. (2006). "Mastering master data management." *Gartner Group, Stamford*.

Yin, Robert K (2013). *Case study research: Design and methods*. Sage publications: 19-32.

7. Appendix

7.1 Master Data Classes

Master Data Class KPI
Packaging Material
Weight for new created materials available
Organizational & HR Master Data
Managed offices[1]
Average office size
Official office ratio
On time HR & OM data availability[4]
Correct target & disciplinary manager[4]
No duplicates
No person with more than 1 position
Direct Material
Completeness of weight entry in series produc
Completeness of weight entry for components raw materials
Completeness of weight entry in shippable products[3]
Completeness of weight entry for purchased p
Assignment of data responsible department (a empty)
Assignment of data responsible department (b: valid up-to-date department name)
Evaluation of different unit of measure across systems
Customer Hierarchy BBM
Cycletime (requests finished within 24 h)
Inspection sticker for data content
Master data differences
Customer Master Data
<No KPI defined. Neither capacity nor sufficier
maturity level for KPI activities>
AMP Elements
Models master data with usage in ePS AMP
Engines master data with usage in ePS AMP
Cycletime final request handling (models & engines)
Uniform usage in business applications
Consistancy (models)
Up-to-dateness/Completeness (models & engines)
Chart of Accounts
Conformity of A-segments between PT2 and ERP-Systems (RB)
Indirect Material
Validity of material group
Acceptance of existing data sets
Roll-In (plant level)
GB internal MDF governance
Supplier Master Data
SMD process coverage (UBK-RM plants, PMD-conn.)
On time approval
Adress data completeness
Governance SMD process
Process fit
Provision of Master Data
Cost Center
Systems coverage of cost center data
Correct entries in field "ID resp. Org.office"
Validation hierarchy nodewith company code
Product Hierarchy
Connected systems to Ref-MDS
Valid product hierarchies in material master

Figure 1 - Master Data Classes

Source: *Company*

7.2 Master Data Objects for the class Position

Table I - Class Position

MD Object	**Attribute name**
Position (S)	Abbreviation
	Company Code
	Controlling Area
	Cost Center
	Description
	Direct/Indirect
	Global Employee Group
	Global Employee Subgroup
	Head of own organizational unit
	Name
	Personnel Area
	Personnel Subarea
	Position ID
	Academic title

7.3 Master Data Objects for the class Organization

Table II - Class Organization

MD Object	**Attribute name**
Org.Area	Area of operations
	Assignment UB
	BWN
	Description
	Flag “Not operative”
	Legal entity ID
	Location
	Name (Unit code)
	No Orgchart Creation
	Obj. Type
	Object abbr.
	Org Area ID
	Type of Organizational area
	Address
	DUNS-Number
Org. Office	Additional remark
	Address
	Company Code
	Controlling Area
	Cost Center
	Description
	Homepage
	ID

Org. Office	Lead BWN
	Levels
	Location
	Obj. Type
	Object abbr.
	Office suggested
	Old office ID
	Operational area
	Org Office ID
	Personnel area
	Personnel subarea
	Reason
	Role - GB-HR
	Role - HRC
	Role - HRD
	Role - HRL
	Role - HRM
	Role - HRP
	Role - HRS
	Role - Part number office
	Role – Resp. Per.
	Role - UBK Customer Responsible
	Status
	Type of Organizational office
	Unit code / Name

7.4 Master Data Objects for the class HR

Table III- Class HR

MD Object	**Attribute name**
Person (P)	Academic title
	Action Type
	Authorized Group Name
	Authorized Group Type
	Birthplace
	Communication First name
	Communication Last name
	Company Code
	Contract Type
	Cost Center
	Country of Birth
	Date of Birth
	GBHR
	Gender
	Global Employee Group
	Global Employee Subgroup
	Global ID
	Headcount
	Hidden
	HR Business Partner
	HR IT First Name
	HR IT Last Name

Person (P)	HRC
	HRL
	HRS
	Initial Entry
	Leading Person
	Legal Entity
	(Legal) First name
	(Legal) Last name
	Local Employee Group
	Local Employee Subgroup
	Local First name
	Local ID
	Local Last name
	Local PERN
	Local Salary Level I (Loc. Sal. Lev. I)
	Local Salary Level II (Loc. Sal. Lev. II)
	Location
	Nationality
	Organizational Office
	Passport First name
	Passport Last name
	Personnel area
	Personnel Capacity
	Personnel subarea
	Position
	Reason for Action

Person (P)	Start Date
	Status of Person
	Valid until
	Source System ID

7.5 Master Data Objects for the class Communication

Table IV - Class Communication Data

MD Object	**Attribute name**
Com. Data	User-ID
	Status User-ID
	email
	Phone
	Mobile
	Fax

7.6 Relationships of the Data Model

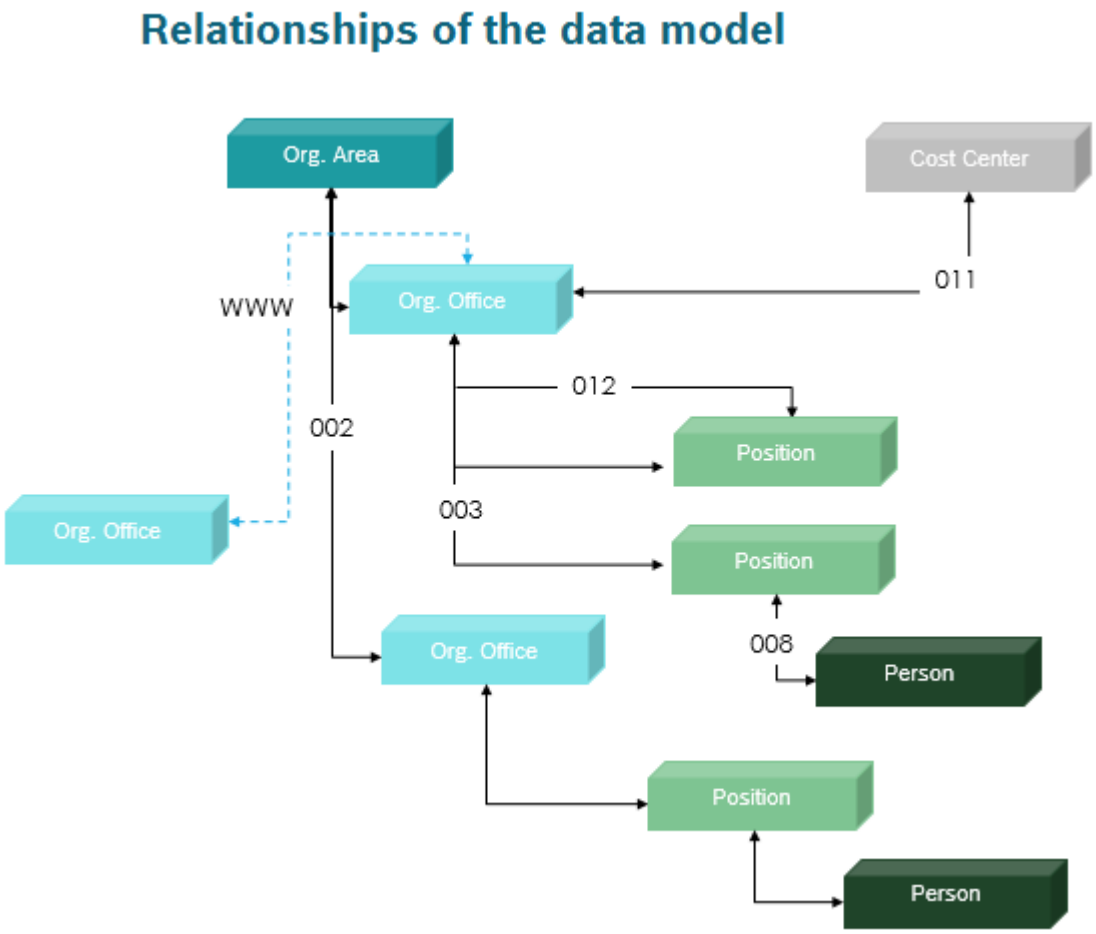


Figure 2 - Relationships of the data model

Source: *Company*

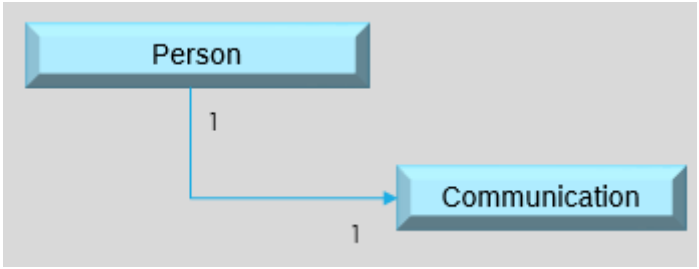


Figure 3 - Relationship between Class Person and Communication

Source: *Company*

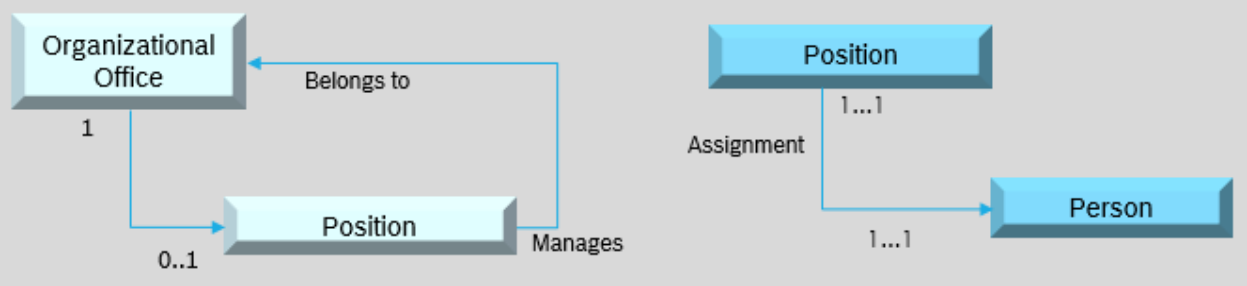


Figure 4 - Relationship with Position, Person and Organization

Source: *Company*