Data-driven begins with DATA; Potential of data assets

H. Hannila[a]*, R. Silvola[a], J. Harkonen[a] and H. Haapasalo[a]

[a]*Industrial Engineering and Management, University of Oulu, Oulu, Finland;*

hhannila@student.oulu.fi (*corresponding author); rmsilvola@gmail.com; janne.harkonen@oulu.fi and harri.haapasalo@oulu.fi

Mr. Hannu Hannila is a doctoral student since Dec 2016 in parallel with a head of PLM role in the industry sector having 20+ years' work experience from the international industrial world. His doctoral thesis is currently in pre-examination phase. He has M.Sc. degree in Information Processing Science from the University of Oulu and B.Sc. degree in Information Technology from Oulu University of Applied Sciences. His research interests are data-driven, fact-based product portfolio management, product lifecycle management and product data management.

Dr Risto Silvola Risto Silvola received his Doctor of Science 2018 in Industrial engineering and management from the University of Oulu, Finland. He has 25 years career in leading global manufacturers at several different positions both in business and in IT organisations. During his career, he has participated in new product programs, global business change programs in program management positions as well as leading departments. Currently his research relates to data driven organisations and business models.

Dr Janne Harkonen received his Bachelor's degree (1st Class Honours) in Engineering Business Management from University of Greenwich in the UK and both his M.Sc. in Process Engineering and Dr (Tech) in Industrial Engineering and Management, from the University of Oulu, Finland. He has also studied in the University of North Carolina at Wilmington, USA. He has experience working in the IT and environmental technology industries. Currently, he is a senior research fellow at the University of Oulu in Finland. His research interests include product management, service management, productization, product portfolio management, and product data management. Dr Harkonen has authored and co-authored over seventy journal articles, and also a number of other publications.

Dr Harri Haapasalo is Professor in Industrial Engineering and Management at the University of Oulu, Finland. He has a Doctors degree in Technology Management and a Master's Degree in Engineering and also Master's degree in Economics and Business Administration, all from the University of Oulu, Finland. He has research interests in planning, product development, and technology commercialization. His list of publications contains more than 200 international items.

Data-driven begins with DATA; Potential of data assets

Abstract. The objective of this study is to analyse the potential of company data assets for data-driven, fact-based decision-making in product portfolio management (PPM). Data assets are categorised from the PPM standpoint, including (product/customer/…) master data, transactional data, and interaction data (e.g., IoT data). The study combines literature review and qualitative analysis of eight international companies. The findings underline the crucial role of corporate-widely combined and governed data model. Company business IT is adjusted against the corporate-level data model. The order of importance is data first, and the technology second. The data-driven mindset and culture creation are also important. The implications include understanding the role and potential of combined data assets that form the basis for data-driven PPM. Facts based on company data assets are essential for decision-making instead of "gut feeling" and emotions. The utilisation of the unused potential of data assets is promoted in the transformation towards data-driven PPM.

Keywords: product portfolio management; master data; transaction data; interaction data; data governance; data assets

Subject classification codes:

1. Introduction

Organisations are increasingly dependent on data and information; they are crucial in every operation, from customer insights to product development and future directions [37, 53]. According to Fisher [21], "*data can be the difference between business success and business failure.*" Regardless of the industry, companies need consistent, accurate, and reliable data to get the best out of their businesses. [21]. The pure data are not enough, but their value must be realised. The wisdom hierarchy (data-information-knowledge-wisdom, a.k.a. DIKW hierarchy) [1] is often and variously referred to in information/information systems and knowledge management literature to contextualise the transformation from data to wisdom [55].

The amount of digital data has been growing exponentially for years. According to Reinsel et al. [54], a Zettabyte (ZB) era is underway; the size of global digital data is estimated to reach 175 ZB[1] by 2025 being 33 ZB in 2018. The amount of data is doubling in size every year, 90% of it created during the past two years and 98% of it stored in digital form [18]. Concurrently, 80% of organisational data are either redundant, obsolete or trivial, which is an obvious risk for poorly leveraged organisational resources [3]. The roles of data and analytics are increasingly becoming mission-critical in every industry [13, 37, 53]. Thousands of data scientist roles are currently involved in both start-ups and well-established companies since companies have an enormous and growing amount of data in volume and in variety [17]. However, the extant literature provides quite little support for how data should be utilised for data-driven decision-making.

All company transactions in business IT solutions rely on master data that relate to products, customers, or suppliers, setting high-quality goals for consistent master data throughout the entire product lifecycle [28, 57, 58, 59]. Nevertheless, the reality often means inconsistencies in data definitions, data formats and values causing negative impacts and inefficiencies in organisations [4, 23, 59, 60]. According to Walker and Moran [71] "*new business opportunities will be missed without an expanded focus to connect datasets to master data – in order to realise the 360-degree view*". Based on Porter and Heppelman [51], a whole new customer relationship mindset is needed with smart and connected products, which gain a wide range of customer data and insights from product usage for sophisticated market segmentation as well as product and

---

[1] One ZB = Billion Terabytes (TB) = Trillion Gigabytes (GB) = 1,000,000,000,000,000,000,000 bytes

service tailoring and thereby for alternative pricing models. The extant literature focuses much on technology and less on how to turn data assets to a competitive advantage.

Product portfolio management is vital for companies as it has a role in tackling challenges that relate to product management, and the range of offering by attempting to optimise and allocate the scarce resources between projects to reach objectives that relate to products and technology [14]. However, the earlier narrow approach on research and development fails to address the challenges with the existing active products and their management, which has been realised by few studies [35, 67, 68]. Nevertheless, the data-driven approach has not been considered adequately in the context.

This study focuses on the potential of company data assets – master data, transactional data, and interactional data – in data-driven product portfolio management (PPM) decision-making. Focus is also on how data should be governed to realise their maximum benefits. In this study, the context and the role of data assets are approached as a combined source of providing value and meaning in the form of information and knowledge for PPM decision-making, i.e. to present how the value of data is generated and realised in a company's decision-making. The following research questions frame this study:

- RQ1: What is the role of company data assets for data-driven PPM?
- RQ2: How are data assets modelled, owned, and utilised in companies?
- RQ3: How should the company data assets be organised to support data-driven and fact-based PPM?

RQ1 is based on a literature review of data-driven culture and decision-making, corporate data assets and their role as fact-based support for decision-making, and in relation to data governance. This is truly important since PPM decisions in companies are still made based on emotions and feelings, and the insights of the growing amount of company data are not realised and utilised for fact-based decision-making. Support is sought from the extant literature to understand how company data assets should be utilised as a raw material in fact-based decision-making. RQ2 focuses on an empirical analysis of eight international companies to figure out the current state of data governance and the companies' current ability to utilise data assets for fact-based decision-making. Analysing eight international companies and their practises should provide an adequate sample to understand the phenomenon. RQ3 proposes a framework as a practical solution on, how data assets should be combined and governed to support fact-based PPM. RQ3 aims to fulfil the gap between the best practises and the current state based on the extant literature. Also, the structural nature of the data is clarified to cover structured, semi-structured, and unstructured data.

2 Literature review

The literature review aims to compile relevant areas of the current literature to clarify the role of company data assets for data-driven product portfolio management (PPM) (RQ1). This necessitates to understand the data-driven culture and decision making (section 2.1). However, when aiming to data-driven PPM it is also necessary to understand the characteristics of corporate data assets (discussed in section 2.2), and the role of data assets in data-driven, fact-based PPM (discussed in section 2.3). Also, the discussion on data governance cannot be ignored when aiming for data-driven (discussed in section 2.4). The construction of company-widely governed data model for fact-based product portfolio management decision-making is synthesised in section 2.5.

2.1 Data-driven culture and decision-making

Data volumes are growing exponentially, and companies desire more and more data-driven these days [72]. One of the most critical foundations in the transition to data-driven is to start trusting data and establish a data-driven culture, which often requires also organisational changes [52]. Organisations have to rethink all their data management practises [2]. "*Gut feeling*" and "*experience*" can be supported and supplemented by advanced data analytics [21]. However, to get data utilised to support company targets is not easy [72]. Companies are not necessarily aware of the value of their data and information silos, and they must also understand that the (IT) technology is not an elixir that should be trusted blindly [21]; efforts are needed to get balanced with people, processes, and technology [2]. The role of technology is support by people in decision-making [72] and yet, the data must be systematically governed at a level beyond the (IT) technology [2, 3, 21, 52]. Data form an asset that should show on the balance sheet, like other assets [26 p643]; the company's strengths and weaknesses are in assets and skills compared with its competitors [50]. According to [21], "*one of the biggest mistakes that organisations make is to approach data as a technology asset, which it is not.*" Data is also often seen as a cost rather than a strategic asset [2, 21].

The single version of the truth is lost when information is managed in silos [21, 58]. The paradigm shift is necessary for the transition from "*analogue*" to "*digital*" through data and analytics. Enterprise reporting based on siloed data, business applications, data warehouses, and analytics applications is played-out – instead of this, the data must be seen enterprise-widely as a raw material for any decision. [37]. Also, massive data feeds cannot be handled by the traditional data management systems [22].

LaValle et al. [39] bring out three levels of analytics capabilities in organisations; 1) **Aspirational** companies are focusing on efficiency or automation of existing processes, searching for ways to cut costs. Not all necessary capabilities (people, processes, or tools) exists to collect, understand, incorporate, or act on analytic insights. These companies are the furthest from the analytical goals. 2) **Experienced** companies have some analytic experience, and are looking to go beyond cost management, developing better ways to optimise their organisation by analytics. 3) **Transformed** companies have a competitive advantage based on substantial experience using analytics through the organisation, and the ability to optimise people, processes, and tools. They also focus less on cutting costs compared with aspirational and experienced organisations through effective use of insights. Transformed companies are also most focused on driving customer profitability and making targeted investments in niche analytics. Higher levels of analytics adoption provide them performance advantage; transformed organisations perform more likely three times better than aspirational organisations within their industry [39].

Thusoo and Sharma [65] highlight guidelines in the transition to a data-driven company, which starts to identify, combine, and manage all sources of data, and to eliminate data silos. Secondly, analytics models are needed to predict and optimise outcomes based on all transactions and interactions necessary to get better insights. Thirdly, the culture to trust data as an enabler for better business decisions must be adopted. [65].

2.2 Characteristics of the corporate data assets

In the 1990s, most data in companies was transaction data created by business applications (e.g., ERP and CRM) in business processes, and related data infrastructures were built accordingly. Since the Internet, a growing amount of interactional data is generated by interactions between people or machines, webpages, and social media, and in a different structural format. [65].

Data assets of a company are classifiable based on their nature as structured, semi-structured, and unstructured. An online tech dictionary [8] explain these as follows: 1) **Structured data** is stored, processed, and accessed based on the data model. Storing format is typically in tables in a database and managed using Structured Query Language (SQL). 2) **Semi-structured data** is a type of structured data but lacks the strict data model structure. 3) **Unstructured data** is information that does not reside in a traditional row-column database and often includes text and multimedia content. 80-90% of enterprise data is estimated to be unstructured. [8]. Unstructured data is often associated with big data first described by Laney [37], who provided three dimensions of high V's; volume (the size of the data), velocity (changing rate of the data), and variety (different data formats and types, structured and non-structured). Afterwards, some other V dimensions have been provided, such as veracity, variability, and value [22]. Big data is also often referred to as semi-structured data, such as XML standard or corresponding or unstructured data, such as Weblogs, social media data, and real-time data, such as event data, spatial data, data generated by machines. [26 p642, 45].

In this study, we focus on the role of master data, transactional data, and interactional data in the decision-making of product portfolio management (PPM). These data assets are classified in Table 1. Master data is a glue between business IT applications and business processes [33] and ideally, remains the same and unaltered through the entire product lifecycle [58]. All transactions in business applications necessitate master data [33, 57, 58, 59, 63]. Today's products have not only got tremendously smart with sensors, microprocessors, and software, but they are getting connected with other products, production equipment, people and ecosystems providing a wide variety of usage and connectivity data about the behaviour of the product. This all generates a massive amount of data (e.g., IoT data) [11, 51], which provides whole new analytics and business intelligence opportunities for companies.

Table 1. Characteristics of the company data assets discussed in this study; master data, transaction data, and interactional data.

Master Data	Enterprise-widely integrated critical business information of product, vendor, customer.	Das and Mishra (2011)
	Primary master data (customers, suppliers, materials) interlinked with all company transactions.	Knapp and Hasibether (2011)
	The shared language of the company.	Myung (2016)
	The organisations most important data assets.	Vilminko-Heikkinen and Pekkola (2017)
	Product master data: item codes, item names, product structure, life cycle status, product classification (HW, SW, service, or combination).	Kropsu-Vehkapera (2012)
	Data and attributes connecting business processes via IT systems, and ideally remaining consistent and unaltered through the entire product lifecycle	Silvola et al (2011b); Kropsu-Vehkapera (2012); Allen and Cervo (2015); Waddington (2008); Das and Mishra (2011)
Transactional Data	Structured data generated by business processes and business applications, such as ERP and CRM.	Thusoo and Sharma (2017)
	Orders, invoices, payments, deliveries, storage records.	Haug et al. (2013); Chaki (2015)
Interactional Data	Data generated by interactions between people and between machines (e.g., IoT data).	Thusoo and Sharma (2017)

Internet of things (IoT) is currently spread everywhere; consumer electronics, domestic appliances, wearables, cars, aeroplanes, autonomous vehicles, and manufacturing devices to mention a few, are all equipped with a variety of sensor technology connected to the Internet. From a manufacturer perspective, this provides an opportunity to close the loop back to product-related business data and product lifecycle phase information and activities [27]. The combination of Big Data, IoT, cloud capacity, and distributed processing techniques provide an enormous potential for decision support throughout the entire product lifecycle [40]. According to Economist Intelligence Unit [19], IoT technology has created innovations due to data, which gives whole new insights and new revenue opportunities from products and services and is enabler towards whole new markets and industries. Companies can create and capture customer value in a different way than in the past. [19]. Smart, connected products are reshaping competition within industries, and have the potential to expand the industries to a whole new level [51].

2.3 The role of data assets in data-driven, fact-based product portfolio management

This chapter focuses on the role of data assets in product portfolio management (PPM). The data assets included are master data, transactional data, and interactional data. The actuality in many organisations is that data on business objects is siloed in several systems, inconsistent with conflicting versions and definitions, and with incorrect values [10, 25, 38, 53] which make data and related information unreliable [21, 42, 61]. Without timely, relevant, and trustworthy data, decisions cannot be made based on facts [5].

All company transactions are performed against consistent master data [6]. Transaction data, such as orders, invoices, payments, deliveries, storage records [24, 13] are necessary for almost all business processes and interlinked with master data, and for that reason the quality of master data is in a vital role [28, 57, 59]. Several issues have been reported because of poor quality of master data, such as inconsistencies [4, 60], and missing data management strategy resulting in poorly utilised data assets [21, 46].

Internet of Things (IoT) has brought manufacturers very close to the end-users increasing product data exchanges [20]. Operational characteristics of products can be widely utilised for design, marketing, customer care and whole new sales opportunities [11, 51]. The business concept and business requirements reflect uniquely on the company data structure, making the data a strategic asset [4], where master data represents the DNA of the company’s business [9, 59], and thus must be managed at strategic, tactical, and operational levels [30, 59, 47].

PPM process is defined as the top management's strategic analysis and decision-making tool on how to manage and renew product offering through the entire product lifecycle to avoid product portfolio explosions and cannibalisation between existing and new products. [66, 67]. From the practical standpoint in companies PPM relates to the question of how to strategically and financially renew and balance the product offering. This is; how to design, manufacture, market and sell, deliver, care, and finally ramp-down company products to resonate with the tree cornerstones of PPM; strategic fit, value maximisation, and portfolio balance [14, 15, 56, 67]. As an example, when a company launches a new product, they must consider which product(s) need to be removed from the product portfolio to avoid cannibalisation and product portfolio explosion as well as to keep the product offering strategically, technically and financially balanced. The foundation for PPM is created by the general commercial and technical product structures of company products through the product portfolio aligned with the company strategy and accordingly standardised, corporate-widely integrated product master data (PMD). Examples of PMD are such as item codes, item names, product structure, life-cycle status, and product classification by its nature for hardware (HW), software (SW), or service, or any combination. [30, 31, 32]. The role of PMD is to connect and synchronise data, business IT solutions and related processes. [23, 33, 58]. The majority of PMD is created during the new product development process [58], but regardless of the life-cycle phase, the product-related data is interlinked with the product structure [16, 35].

The data-driven, fact-based PPM necessitates reliable data and information from all data sources. This is inevitable since PPM decisions are typically made based on emotions and gut feelings [14, 41, 43], and today's decisions determine the business performance in the coming several years [14], whereas wrong decisions can result in destructive consequences [15]. From the company's financial perspective, this is inevitable since 20% of company products bring typically 80% of sales volume [68].

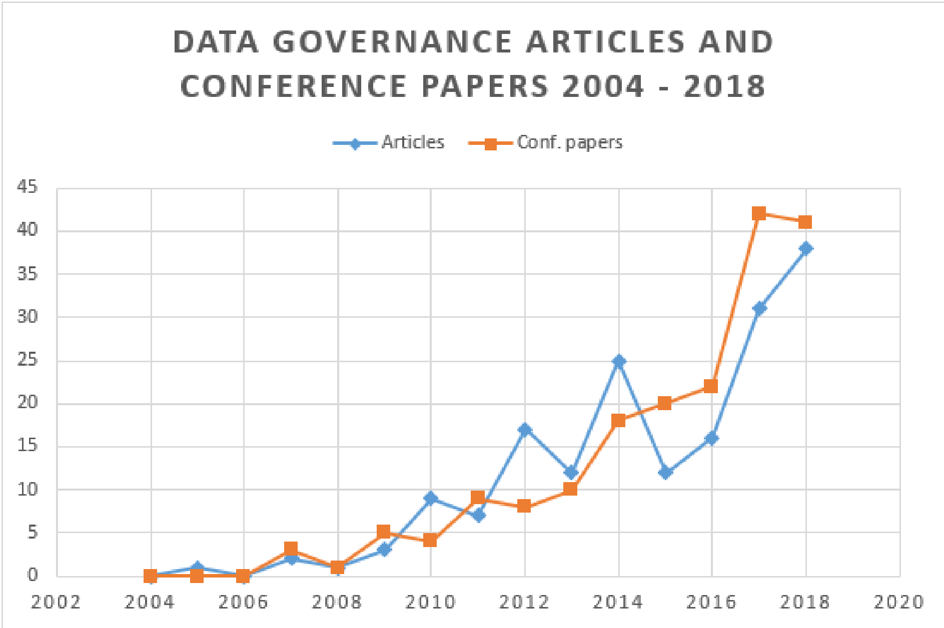
2.4 Data governance

Real-time business intelligence has become a competitive advantage for companies. [65]. Nevertheless, the literature has been focusing much on IT systems and algorithms, while data analytics and data governance have gained less attention despite growing data volumes in companies [12, 34]. Surviving in complex environments requires well-integrated processes, disciplined data architecture, and consistent data and information management [13].

Data governance emerged as a topic in the early 2000s (Figure 1). The previous topical issue a decade earlier was enterprise information management (EIM) in connection with structured data management in data warehouses and data mart technologies. In the 1990s, structured data was easy to model, store, and transfer by custom-written stored procedures in relational database management systems (such as Oracle, SQL Server, and DB2) and later automated by extract-transfer-load (ETL) process. [13]. Today, when 80% of organisations' data is semi-structured or unstructured [45], the role of information management has changed, and whole new technologies and processes are required to govern data and to ensure the quality of data [13].

The explosion of data has resulted in a need for data governance [12] which has started to gain attention also in the literature over the past decade (Figure 1). Waddington [70] and Fisher [21] define data governance as a process to standardise business data and related metrics through the organisation, including data definition, propagation, ownership, and quality. In other words, it is to ensure consistent, high-quality, and relevant data through the organisation [9].

Figure 1. Yearly published articles and conference papers including the term “data governance” in Scopus.

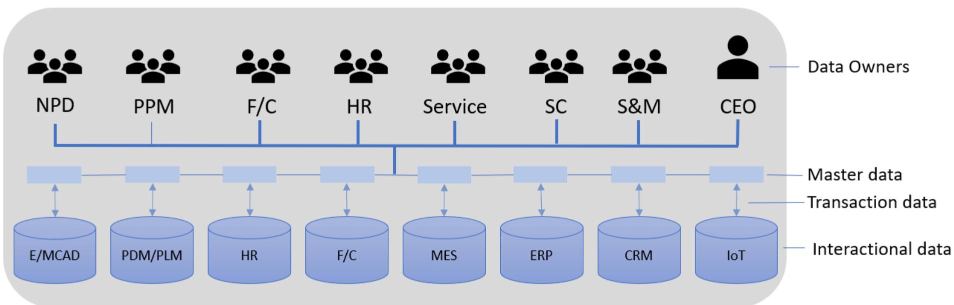


Historically business applications (e.g., Enterprise Resource Planning (ERP), and Customer Relationship Management (CRM) systems) have been responsible for the data quality to support application-specific transactions, which will, however, not work for analytics because of whole different requirements for the data and related attributes [64 p535]. The final quality of the data is a sum of each step from data capture to analytics and decision-making [62]. Data governance is a company-specific activity based on the uniqueness of the company data [12]. Several scholars have proposed a company-level data management practises [22, 30, 48, 57, 59] to differentiate [4] and survive with competitors [56, 61, 63].

2.5 The construction of company-widely governed data model for fact-based product portfolio management decision-making

This study analyses the potential of company data assets (master data, transaction data, and interaction data) in data-driven product portfolio management (PPM) decision-making, and the role of company-wide data governance and data model in this context. The framework in Figure 2 suggests a conceptual, corporate-level data model for fact-based product portfolio management combining data owners, master data, transaction data, and interactional data. Master data, whether they are the product, customer, or supplier master data, connect business applications and reflect the DNA of a company's business, and is ideally created only once to ensure one single source of the truth [28, 33, 58, 59]. All company transactions in operations are created against the master data [33, 57, 63]. The fundamental objective of this model is to connect all data assets from different data sources (e.g., from Electrical/Mechanical Computer-Aided Design (E/MCAD), Product Data Management / Product Lifecycle Management (PDM/PLM), Human Resources (HR), Finance & Control (F/C), Manufacturing Execution System (MES), ERP, CRM, Internet of Things (IoT) to the same backbone via master data. This is an enabler to realise further the full potential of company data assets for data-driven decision-making, where reporting and analytics can be adjusted based on, and through several data sources, concurrently.

Figure 2. A company-widely governed conceptual data model for fact-based PPM to system independently connect data producers and data consumers of a company to the related business IT systems.



Data owners, whether they are data producers (like for example new product development, NPD, where the majority of product master data is created) or data consumers that are thus operating with the same data assets. The other data owners in Figure 2 are representatives from different functions of the company, such as ~~product portfolio management (~~PPM~~)~~, ~~finance and control (~~F/C~~)~~, ~~human resources (~~HR~~)~~, service, supply chain (SC), sales and marketing (S&M), and the company CEO. Since companies' organisational structures vary, and the data model of a company is unique [12], a one-size-fits-all framework is not appropriate, but the framework can be established by the elements presented in this framework in Figure 2.

3 Research process

Companies are having a vast and rapidly growing amount of data and information but lack the ability to analyse and govern them, which weakens their ability in the transformation to data-driven decision-making and culture. Data assets are one of the most unused assets in companies. This study examines the potential of company data assets and data governance from a practical perspective supplemented by the existing literature aiming to ensure the scientific and managerial applicability.

3.1. Research objective and strategy

The research objective of this study is to reduce a managerial gap and to complement earlier research related to data management, precisely focusing on the role of data assets and data governance for product portfolio management (PPM) decision-making. Three research questions (RQ) frame the research. The RQ1 is based on extant literature and aims to answer what are the role of company data assets for data-driven decision-making in PPM. The RQ2 - how data assets are modelled, owned, and utilised in companies – is answered through an empirical analysis of eight international companies. The RQ3 aims to provide baselines on how the company data assets should to be organised to support data-driven and fact-based PPM.

The research is a qualitative multiple case study aiming to bring out relevant awareness from several cases. To get strong and reliable evidence, and to understand the similarities and differences between several cases, multiple case study was used [7]. This study is exploratory by nature, examining the current state of data management and proposing a framework, which combines data owners, master data, transactional data, and IoT data in organisations to gain data-driven and fact-based support for PPM decision-making. The existing literature forms the basis for company data assets and data management regarding data assets and data governance.

3.2 Data collection

Semi-structured in-depth interviews [44] complemented by company internal and publicly available materials to gather empirical data were used to obtain a comprehensive understanding of the phenomenon. Informants represented theoretical and practical expertise in the subject to guarantee information-rich in-depth understanding and insights instead of an empirical generalisation. [49 p230]. Company internal quality and process documentation were utilised for triangulation.

The roles and responsibilities of the informants represented business domains, such as sales and marketing, supply chain and manufacturing operations, product development (R&D), product management, product data management, product lifecycle management, and finance. A variety of roles were included to gain perspectives comprehensively, to reduce bias, and to get the best knowledge in the subject within the companies. Companies were from several industrial sectors, headquartered in Finland, from a few million to several billion in euros measured by company turnover. Half of the companies design, manufacture and sell configurable, modular smart products involving HW, SW and service, for varying industrial sectors. Two of the companies' products are preconfigured HW sales items including SW for business-to-business

(B2B) customers and business-to-customer (B2C) consumers. From PPM perspective their challenge is to find a balance with the optimum number of sales items to maximise the number of pieces sold and to avoid low selling sales items. One company has a niche, competitive technology providing highly customised HW products and services to their original equipment manufacturer (OEM) customers. Their PPM challenge relates to productising their service and technology offering instead of always starting from scratch. One company provides a wide variety of HW and SW (i.e. SW updates) spare parts in parallel with quite standardised after sales services for highly complicated products with a long lifecycle. The overall PPM challenge for all the companies is to productise and standardise their offerings to make possible to strategically renew their product offering, to maximise the value of the product portfolio and to optimise the number of sales items and configurable components as a part of master product. The service is described in this study as a combination of processes, e.g. service program including the various stages of maintenance steps. Overall, 47 informants were interviewed in 14 interview sessions. The size of the company and its organisational structure influenced the size of the interview group, varying from 2 to 13 informants.

The informants were interviewed by a group of two to four researchers. Each interview session was documented by participating researchers. Interviews were recorded when allowed, which provided support for the data analysis in the later phase. Informants got interview questions before the interviews. Informants were split into organisational groups they represented, or in some smaller companies, the whole group was interviewed in one session.

3.3 Data analysis

The data collected was first summarised and analysed by each researcher individually followed by group analysis by four researchers to obtain a consensus and to achieve a deeper understanding of the phenomenon in companies. The following methods were carried out in data analysis:

- Inductive thematic analysis
- The second source documentation (i.e., the company internal and external documentation) for triangulation to ensure the trustworthiness
- Familiarisation to recognise themes and patterns.

After the interview sessions in participating companies, a validation workshop was arranged to discuss the findings. The research group presented the findings, and both oral and written feedback was collected from the participants. The target of the validation workshop was to ensure that company representatives agreed with the findings.

4 Results and analysis

The companies were asked about the company data model, data ownership, and whether they utilise IoT technology in their products and if they have found ways to make use of the data the IoT technology provides. The aim was also to recognise how data is governed in companies. This study explores the status quo of the existence of company level data model related to business IT solutions and whether the IoT data is connected to the same model or not. Data ownership goes hand in hand with data model, so the definition of data ownership in companies was also explored. The characteristics of companies and key findings are summarised in Table 2.

None of the companies had a consistent, corporate-level data model implemented, half of them did not have data model(s) at all. The other half had defined their data models on the IT solution level, typically for PDM/PLM, ERP, or PIM system, or a combination. Companies that did not have a company data model defined did not have defined data ownership either. Similarly, companies that had data model defined for some IT solution seemed to be considering data ownership as well.

The type of products in all companies consists of HW, SW and service, except one company which provides only a combination of HW and service. IoT technology was a part of the product in seven out of eight companies. IoT is utilised, e.g., to get information from product installations, monitor the use of products, for predictive analysis related to maintenance, and to gain insights about how products are commonly used. One of the companies had monetised the IoT data to sell it to the customers. Some of the companies were in early phases seeking for opportunities to utilise the identified future potential of IoT data in their business. However, none of the companies currently connect IoT data to other data assets – such as master data or transactional data – in terms of PPM. However, combining data was recognised to have the potential to provide new business opportunities, by one of the companies.

Common for all companies, regardless of whether they had defined the data model and data ownership or not, the company-level data governance was not recognised clearly if at all. Separate business units were operating in their silos, without any consideration of relation to master data, transaction data, or IoT data. Very much time was used by all the companies for more or less manual, time-consuming data analysis and reporting, especially for cost calculation of products.

Table 2. Characteristics of companies and the summary of key findings.

Case company	Type of Business	Company Data Model	Data Ownership	IoT data
A	B2B	No data model, but corporate-level driver exists to do it. No actions ongoing.	Not data ownership. "*Somebody, somewhere*."	Utilized to capture information from product installations. The need recognized to connect with product master data.
B	B2B	Customer, and product master exists, done on IT solution level.	Data owners defined in separate business domains, no company level coverage, the need recognized and development ongoing	Widely utilized, but separately from other corporate data assets.
C	B2B, B2C, OEM	Data model on IT solution level in PDM and ERP.	Ownership for PDM and ERP data partly exist	Widely utilized, but separately from other corporate data assets.
D	B2B, OEM	Data model on IT solution level in PDM. The need recognized and work initiated.	Clear data ownership in R&D, no other.	Partly utilized for product failure and maintenance analysis and reporting. Separately from other corporate data assets
E	B2B, B2C	Data model on IT solution level model in PDM and in PIM.	Data ownership model documented into PPT	First initiatives to utilize IoT in product monitoring. Separately from other corporate data assets
F	B2B	No data model	No data ownership	Partly utilized for product analytics, data is sold to some customers. Separately from other corporate data assets. A vast future potential recognized.
G	OEM	No data model	No data ownership	N/A
H	B2B	No data model	Business unit level data ownership under concideration	N/A. Future potential identified. Data security as a challenge.

As a tangible example of the potential of better utilising the company data assets, case companies have been considering the potential of IoT data in several ways. In one company the potential of IoT data was considered as a part of service delivery to enhance quality and accuracy of the service and close the loop for the business data to optimise the cost of service delivered. In another company, the IoT data is already widely utilised to simulate the product usage and predictive maintenance needs, and the future potential was seen to lie in connecting this data to the related business data of the service. Some companies were considering utilising an electronic serial number of the product for the marketing support, or to recognise additional sales potential as a part of sales channel analyses based on product or device registrations and insights they provide from the field. Almost all companies that had IoT as a part of their products had explored the possibility for new business opportunities by monetising their IoT data one way or another; either by selling the data to the customers or monetising the insight the data provides, not only for the company itself but also for customers or third parties. Overall, the companies found connecting IoT data to other data assets useful and supportive from the business perspective. The interviewees saw vast potential for variety of opportunities in the future.

5 Discussion

A data-driven approach to make fact-based decisions on company products and the entire product portfolio, instead of individual opinions, emotions or gut feeling, begins with data. The potential of company data assets, master data, transactional data, and interactional data are currently underutilised. The current focus emphasises too much the technology (Information Systems) over the possibilities of converting the data assets to a competitive advantage. The order of importance should be seen as data first, and the technology second. The data-driven mindset and culture creation are also vital.

This study constructed a company-widely governed data model for fact-based product portfolio management decision-making. The role of data assets in this study is approached as a combined source of master data, transaction data, and interaction data to prove value and meaning in the form of information and knowledge for PPM decision-making. The data owners are also acknowledged as data quality necessitates responsibilities. The aim was to present how the value of data is generated and realised for a company's decision-making. A data model can help in addressing the adverse effect of siloed business, data and applications, that prevent the effective data-driven and fact-based approach. A company-widely governed data model enables application independent approach to master data, transaction data and interactional data. All company transactions in business IT solutions rely on master data that relate to products, customers, or suppliers, but siloed business, data and applications prevent the true data-driven approach. Necessary data consistency for fact-based decisions over the product portfolio can be reached by product structure and related master data. Also, the company business IT should be adjusted against the corporate-level data model. The business IT technology is useless in terms of data-driven approach if it does not support the efficient use of enterprise data assets. It appears that companies are currently in very early stages of utilising a data model and understanding its benefits. Also, the ownership of data assets and the effective utilisation of the data assets require efforts. Combining data; master data, transaction data, and interaction data may have the potential to provide new business opportunities aside supporting the data-driven fact-based approach and providing a guideline for organising company data assets.

5.1 Scientific implications

This study supports the earlier literature by highlighting the fundamental role of the data as a strategic asset of the company [2, 3, 21] which must be consistently governed [21] a level beyond the technology [2]. The existing PPM literature is elaborated and extended by combining all company data assets to realise the highest potential of data-driven and fact-based support for PPM decision-making.

The transition for data-driven necessitates trusting data as raw material for any decision [37] and creating a data-driven culture and organisation [52], which is in line with this study. This study is also in balance with previous studies of data management practices, which must be reconsidered to get balanced with people, processes, and technology [2]. Also, data must be governed a level beyond and separated from the technology [2, 3, 21, 52], while the role of the technology is supportive [21, 72].

This study provides a new contribution for PPM by combining master data, transaction data, and interactional data as a company-widely governed, combined data source for data-driven decision-making. This is required to remove data silos and to realise the full potential of data assets for decision-making in PPM. This supports the earlier literature suggesting consistent master data management practices to ensure high-quality transactions [6, 33, 57, 58, 59, 63]. This study provides contribution by using this basis to form a company-level data model that enables connecting interactional data to the same data model.

Based on existing literature, PPM decisions are made based on emotions and gut feelings [14, 41, 43]. These decisions have very far-reaching effects [14], while wrong decisions can result in negative consequences [15] such as product portfolio explosions and cannibalisation between existing and new products. [66, 67]. Also, since 20% of company products bring 80% of sales volume [68] the right decisions are required today. This study provides new contribution and extends current literature by providing a framework to support PPM decisions based on company data assets that are highlighted as an enabler for data-driven decision-making in PPM.

The current literature has gained marginal attention for data analytics and data governance and been much focusing on IT systems and algorithms [12, 34]. This study expands the literature by introducing the company-widely governed data model.

5.2 Managerial implications

Managers should understand that data should come first, and the technology should follow data. Even an average, medium size company's business IT costs easily rise over 10 million € as a result of poorly leveraged organisational resources, both people and business IT while both data and business IT are fragmented and siloed, and reporting is based on functional silos via spreadsheets causing a constant debate about whose Excel file contains the correct data, and how the data should be predicted. Noteworthy is that this study should not be read from a CIO perspective or from business IT technology perspective. The novelty of this paper lies in providing a data-driven approach where company executives are encouraged and justified to begin to trust their strategic data assets as a raw material for any decisions. Managers need to understand, that the transition for data-driven decision-making begins with data-driven culture creation and requires belief in the data as a raw material for decision-making. This is the first step which company executives need to adopt. This often requires changes in data management practices and organisational structures. Also, the value of data as a strategic asset must be understood. At least as necessary is to understand to get rid of data silos and begin to govern data company-widely and beyond the business IT technology – data is not a technology asset.

Today's smart, connected products provide whole new opportunities for companies based on the wide variety and a massive amount of interactional data, such as usage and connectivity data. This study provides the data-driven PPM decision-making support for company executives in the form of a company-wide data model, but this is not limited to PPM only. With right analytics and business intelligence methods, interactional data provides endless opportunities not only within industries but also create whole new industries.

5.3 Limitations and future studies.

The limitation of this study involves the fact that none of the involved companies deal with data governance exceptionally well, making this study to lack of an excellent reference. However, this deficiency is well complemented by extensive literature, and provides new research opportunities. This study was deliberately limited to cover data assets, which have a role in PPM. From this perspective, the metadata – data and information about data – was excluded as not being relevant from the data-driven PPM perspective. The technology was also deliberately excluded from the study since the framework provided is approached technology-independently. In general, the sample size could be bigger, but eight case companies provide an extensive enough coverage for the purpose of this study. Future research could study a company-wide data model as part of a larger entity, including layers of business processes, reporting and analytics, to conceptualise the entire company ecosystem, for which this study provides the basis from a data asset perspective.

6 Conclusions

Data-driven begins with the "*DATA*", data that are currently poorly utilised, but form a strategic asset for companies, and hence should be treated like other assets to gain their highest potential. This study focused on the role of *master data, transaction data,* and *interactional data* in connection with fact-based and data-driven PPM. The paradigm shift is taking place in companies in the form of growing amount of data, which requires a new mindset to deal with data and analytics in transformation towards data-driven decision-making culture.

A very topical question is why there are few or no corporate wide data models in companies. There is much literature about siloed data and interoperability challenges between business IT systems, but the literature is lacking on how to break system boundaries. Companies seem to be lacking in awareness of the value of their data assets, and understanding that the data must be governed systematically, and separately from IT technology. This paper tackles these challenges based on three research questions by 1) specifying the role of company data assets in data-driven PPM, 2) analysing the current state of data governance in companies, and 3) by providing a model on how the company data assets should be organised to support data-driven and fact-based decision-making in PPM.

During this study the authors started to discuss about "*data damagement*" as "*a combined outcome of spreadsheet effect and (data) silo effect*", i.e. how much company resources are wasted, and damage caused as a result of poor data management. Even though this came from a light debate, it was soon realized that this is the culmination of the central essence of this study, which must be taken seriously. This truly serve as a point of departure for future research. Data-driven begins with data!

References

[1] Ackoff R.L. From Data to Wisdom. J. Appl. Syst. Anal. 1989;16(1):3-9.

[2] Aiken P. Experience: Succeeding at data management-BigCo attempts to leverage data. J. Data and Inform. Q. 2016;7(1-2):1-35.

[3] Aiken P, Billings J. Monetizing data management. Unlocking the value in your organization's most important asset. LLC USA: Technics Publications; 2013.

[4] Allen M, Cervo D. Multi-Domain Master Data Management. Advanced MDM and Data Governance. Waltham, MA: Elsevier; 2015.

[5] Anderson C. Creating a Data-Driven Organization: Practical Advice from the Trenches. Sebastopol, CA: O'Reilly Media; 2015.

[6] Baghi E, Schlosser S, Ebner V, Otto B, Oesterle H. Towards a Decision Model for Master Data Application Architecture. IEEE Computer Society. 47th Hawaii International Conference on System Science; 2014; Waikoloa, HI, USA.

[7] Baxter P, Jack S. Qualitative Case Study Methodology: Study Design and Implementation for Novice Researchers. The Qual. Report. 2008;3(4):544-559.

[8] Beal W. Webopedia. [accessed 2019 May 15]. https://www.webopedia.com/TERM/S/structured_data.html & https://www.webopedia.com/TERM/U/unstructured_data.html.

[9] Bonnet P. Enterprise Data Governance: Reference and Master Data Management Semantic. Hoboken, NJ: John Wiley & Sons; 2010.

[10] Borgia O, Fanciullacci N, Franchi S, Tucci M. (2015). The use of product information along its entire lifecycle: A practical framework for continuous development. Iss. Ind. Syst. Eng. 2015:177-182.

[11] Borgia E. The Internet of Things vision: Key features, applications and open issues. Comp. Comm. 2014;54:1-31.

[12] Brous P, Janssen M, Vilminko-Heikkinen R. Coordinating Decision-Making in Data Management Activities: A Systematic Review of Data Governance Principles. 5th International Conference on Electronic Government and the Information Systems Perspective (EGOV); 2016; Porto, Portugal. in Lect. Notes Comp. Sci. 2016;9820:115-125.

[13] Chaki S. The expert's voice in information management. Enterprise Information Management in Practice. New York, NY: Springer; 2015.

[14] Cooper R.G, Edgett S.J, Kleinschmidt E.J. New Product Portfolio Management: Practises and Performance. J. Prod. Innov. Manage. 1999;16(4):333-351.

[15] Cooper R.G, Edgett S.J, Kleinschmidt E.J. Portfolio management for new product development: results of an industry practices study. R&D manage. 2001;31(4):361-380.

[16] Crnkovic I, Asklund U, Dahlqvist A.P. Implementing and Integrating Product Data Management and Software Configuration Management. Artech House; 2003.

[17] Davenport T.H, Patil D.J. Data Scientist: The Sexiest Job of the 21st Century. Harv.Bus. Rev. 2012;90(10):70–76.

[18] Devakunchari R. Analysis on big data over the years. Int. J. Sci and Res. Publ. 2014;4(1):1-7.

[19] EIU (2017). Economist Intelligence Unit. The Internet of Things Business Index 2017 – Transformation in motion. The Economist. [accessed 2019 May 20]. https://perspectives.eiu.com/technology-innovation/iot-business-index-2017-transformation-motion

[20] Emmanouilidis C, Beroncelj L, Bevilacqua M, Tedeschi S, Ruiz-Carcel C. Internet of Things – Enabled Visual Analytics for Linked Maintenance and Product Lifecycle Management. IFAC PapersOnLine. 2018;51(11);435-440.

[21] Fisher T. The Data Asset: How Smart Companies Govern Their Data for Business Success. Hoboken, NJ: John Wiley & Sons; 2009.

[22] Gandomi A, Haider M. Beyond the hype: Big data concepts, methods, and analytics. Int. J. Inform. Manage. 2015;35:137-144.

[23] Hannila H, Tolonen A, Harkonen J, Haapasalo H. Product and supply chain related data, processes and information systems for product portfolio management. Int. J. Prod. Lifecyc. Manage. 2019;In press.

[24] Haug A, Arlbjorn J.S, Zachariassen F, Schlichter J. Master data quality barriers: an empirical investigation. Indust. Manage. & Data Syst. 2013;113(2);234-249.

[25] Jetson J, Nelis J. Business process management: practical guidelines to successful implementations. Amsterdam: Elsevier; 2008.

[26] Kimball R. The Evolving Role of the Enterprise Data Warehouse in the Era of Big Data Analytics, in Kimball et al. (2015), The Kimball Group Reader:

Relentlessly Practical Tools for Data Warehousing and Business Intelligence. Second ed. John Wiley & Sons; 2011.

[27] Kiritsis D. Closed-loop PLM for intelligent products in the era of the Internet of things. Comp.-Aided Design. 2011;43(2011):479-501.

[28] Knapp M, Hasibether F. "Material Master Data Quality. Proceedings of the 17th International Conference on Concurrent Engineering, IEEE; 2011 Aachen, Germany:1-8.

[29] Kropsu-Vehkapera H. Enhancing understanding of company-wide product data management in ICT companies. Acta Universitatis Ouluensis C Technica 418.University of Oulu; 2012.

[30] Kropsu-Vehkapera H, Haapasalo H. Defining product data views for different stakeholders. J. Comp. Inf. Syst. 2011;52(2):61-72.

[31] Kropsu-Vehkapera H, Haapasalo H, Harkonen J, Silvola R. Product data management practices in high-tech companies. Ind. Manage. & Data Syst. 2009; 109(6):758-774.

[32] Kropsu-Vehkapera H, Haapasalo H, Jaaskelainen O, Phusavat K. Product Configuration Management in ICT Companies: The Practitioners' Perspective. Techn. and Invest. 2011; 2(4):273-285.

[33] Kumar Das T, Mishra M.R. A Study on Challenges and Opportunities in Master Data Management. Int. J. Database Manage. Syst. 2011;3(2):129-139.

[34] Kwon O, Lee N, Shin B. Data quality management, data usage experience and acquisition intention of big data analytics. Int. J. Inform. Manage. 2014;34(3):387-394.

[35] Lahtinen N, Mustonen E, Harkonen J. Commercial and Technical Productization for Fact-Based Product Portfolio Management Over Lifecycle. IEEE Trans. Eng. Manage. 2019:in press.

[36] Laney D. 3D data management: Controlling data volume, velocity and variety. META Group Research Note 6; 2001.

[37] Laney D, Jain A. (2017), 100 Data and Analytics Predictions Through 2021. Gartner report ID: G00332376; Gartner; 2017.

[38] Lans R.F. Data Virtualization for Business Intelligence Systems. Revolutionizing Data Integration for Data Warehouses. Waltham, MA: Elsevier; 2012.

[39] LaValle S, Lesser E, Shockley R, Hopkins M.S, Kruschwitz N. Big Data, Analytics and the Path from Insights to Value. MITSloan Manage. Rev. 2011;52(2):21-31.

[40] Li J, Tao F, Cheng Y, Zhao L. Big Data in product lifecycle management. Int. J. Adv. Manuf. Tech. 2015);81:667-684.

[41] Maitlis S, and Ozcelik H. Toxic Decision Processes: A Study of Emotion and Organizational Decision Making Org. Sci. 2004;15(4):375-393.

[42] Marchetta G, Mayer F, Forradellas R. A reference framework following a proactive approach for Product Lifecycle Management. Comp. in Indust. 2011; 62:672-683.

[43] McAfee A, Brynjolfsson E. Big Data: The Management Revolution. Harv. Bus. Rev. 2012;90(10):61-68.

[44] Merton R, Fiske M, Kendall P. The Focused Interview: A Manual of Problems and Procedures. 2nd ed. New York: The Free Press; 1990.

[45] Muntean M, Surcel T. Agile BI – The Future of BI. Inform. Econom. 2013;17(3):114-124.

[46] Myung S. Master Data Management in PLM for the Enterprise Scope. in A. Bouras et al. (Eds) Product Lifecycle Management in the Era of Internet of Things. PLM 2015. IFIP Adv. in Inform. and Comm. Tech. 2016;467:771-779.

[47] Ofner M.H, Straub K, Otto B, Oesterle H. Management of the master data lifecycle: a framework for analysis J. Ent. Inf. Manage. 2013;26(4):472-491.

[48] Otto B, Wende K, Schmidt A, Osl P. Towards a Framework for Corporate Data Quality Management. Proceedings of 18th Australasian Conference on Information Systems ACIS; 2007; Toowoomba, Australia: 916-926.

[49] Patton M.Q. Qualitative research and evaluation methods. 3rd ed. Thousand Oaks, CA: Sage; 2002.

[50] Porter M.E. What is a strategy. Harv. Bus. Rev. 1996;74(6):61-78.

[51] Porter M.E, Heppelmann J.E. How Smart, Connected Products Are Transforming Competition. Harv. Bus. Rev. 2014;92(11):64-88.

[52] Pugna I.B, Dutescu A, Stanila O.G. Corporate attitudes towards Big Data and its impact on performance management: A qualitative study. Sustain. (Switzerland) 2019;11(3):1-26.

[53] Redman, T.C. Data Driven: Profiting from Your Most Important Business Asset. Boston, Mass. Harvard Business Press; 2008.

[54] Reinsel D, Gantz J, Rydining J. (2018). The Digitization of the World. From Edge to Core. An IDC White Paper - #US44413318. Seagate; 2019.

[55] Rowley J.E. The wisdom hierarchy: representations of the DIKW hierarchy. J. inf. sci. 2007;33(2):163–180.

[56] Saaksvuori A, Immonen A. Product lifecycle management, 3rd edition. Springer; 2008.

[57] Schaffer T, Leyh C. Master Data Quality in the Era of Digitalization – Toward Inter-organizational Master Data Quality in Value Networks: A Problem Identification. in F. Piazolo et al. (Eds) Innovations in Enterprise Information Systems Management and Engineering. ERP Future 2016. Lecture Notes in Bus. Inf. Proc. 2017;285:99-113.

[58] Silvola R, Jaaskelainen O, Kropsu-Vehkapera H, Haapasalo H. Managing one master data – challenges and preconditions. Ind. Manage. & Data Syst. 2011;111(1):146-162.

[59] Silvola R, Tolonen A, Harkonen J, Haapasalo H. Defining one product data for a product. Int. J. Bus. Inf. Syst. 2019;30(4):489-520.

[60] Smith H.A, McKeen J.D. Developments in Practice XXX: Master Data Management: Salvation Or Snake Oil? Comm. of the Assoc. for Inf. Syst. 2008;23(4):63-72.

[61] Sriti M.F, Assouroko I, Ducellier G, Boutunaud P, Eynard B. Ontology-based approach for product information exchange. Int. J.Prod. Lifecycle Manage. 2015;8(1);1–23.

[62] Storey V.C, Dewan R.M, Freimer M. Data quality: Setting organizational policies Dec. Supp. Syst. 2012;54:434-442.

[63] Terzi S, Bouras A, Dutta D, Garetti M, Kiritsis D. Product lifecycle management – from its history to its new role. Int. J. Prod. Lifecycle Manag. 2010;4(4):360-389.

[64] Thornthwaite W. Dealing with Data Quality: Don’t Just Sit There, Do Something, in Kimball et al. (2015), The Kimball Group Reader: Relentlessly Practical Tools for Data Warehousing and Business Intelligence. Second ed. John Wiley & Sons; 2009.

[65] Thusoo A, Sarma J.S. Creating a Data-Driven Enterprise with DataOps. Insights from Facebook, Uber, LinkedIn, Twitter, and eBay. Sebastopol, CA: O’Reilly Media; 2017.

[66] Tolonen, A. Product portfolio management over horizontal and vertical portfolios. Acta Universitatis Ouluensis C Technica 574. University of Oulu; 2016.

[67] Tolonen A, Harkonen J, Verkasalo M, Haapasalo H. Product portfolio management process over horizontal and vertical portfolios. Int. J. Prod. Life Cycle Manage. 2015a;8(3):189-215.

[68] Tolonen A, Shahmarichatghieh M, Harkonen J, Haapasalo H. Product Portfolio Management – Targets and key performance indicators for product portfolio renewal over the life cycle. Int. J.Prod. Econ. 2015b;170:468-477.

[69] Vilminko-Heikkinen R. Pekkola S. Master data management and its organizational implementation: An ethnographical study within the public sector. J. Enterp. Infor. Manage. 2017;30(3):454-475.

[70] Waddington D. Adoption of Data Governance by Business, DM review 2008;18(12):32-32.

[71] Walker S, Moran M. Create Powerful Customer Experiences With a 360-Degree View of Your Products. Gartner report ID G00328027. Gartner; 2017.

[72] Wang S, Krisch U. A foundation for building a data-driven culture. Appl. Marketing Analyt. 2019;4(3):238-252.