



TAMPEREEN TEKNILLINEN YLIOPISTO  
TAMPERE UNIVERSITY OF TECHNOLOGY

JONNA FRED  
DATA MONETIZATION – HOW AN ORGANIZATION CAN GENER-  
ATE REVENUE WITH DATA?

Master's thesis

Examiner: Professor Samuli Pek-  
kola. Examiner and topic approved  
by the Board of Business and Built  
Environment Faculty on  
30th of January 2017

## ABSTRACT

**JONNA FRED:** Data Monetization – How an Organization Can Generate Revenue with Data?

Tampere University of Technology  
Master of Science Thesis, 61 pages  
March 2017

Master's Degree Programme in Information and Knowledge Management  
Major: Information and Knowledge Management  
Examiner: Professor Samuli Pekkola

Key words: Data Monetization, Big Data, Data analytics, Digitalization

While digitalization evolves and distinct technologies are developed further, the role of data and data analytics has grown and become more important in the eyes of organizations. Simultaneously data is not considered anymore as insignificant raw material but an important ingredient in developing business activities and enabling innovation.

Nowadays, due to the enhanced information technologies, an organization does not need to create data or its more refined forms itself but data can be sold or purchased in the same way as tangible goods or services. Despite an idea of business focused on selling data is rather novel and it has not been yet researched extensively. This thesis work studies Data Monetization phenomenon which refers to business built on data and furthermore revenue generated with data and its derivatives. The terminology related to Data Monetization has not stabilized yet and no unambiguous definition was found from the scientific literature. Hence this study aspires to clarify the phenomenon, its definition and terminology.

The main research question of this study is following: “What kind of factors are behind of and affect Data Monetization?” In order to answer the previously described question, the definition of Data Monetization is studied as well as the distinct options and measures an organization may take to enable revenue generation with data. Furthermore this study pursues to discover and identify other phenomena that associate with Data Monetization and moreover to recognize different strategic options to implement Data Monetization business. The thesis work was executed as a systematic literature review and literature sources were searched from scientific libraries and databases, such as Scopus and Google Scholar. Since it seems that this topic is not yet studied extensively and hence only few relevant pieces of literature were found, the literature sample was not confined too much.

As a result of this study an unambiguous and justifiable definition of Data Monetization was established. Such definition could not be found from the pieces of literature utilized in the systematic literature review. Additionally distinct components and aspects of Data Monetization business were recognized as well as a variety of different business and revenue generation models. From the perspective of Data Monetization it is essential to identify the valuable data and to be able to, if needed, refine and develop it further in order to enable the business transaction which in turn generates revenue. Since Data Monetization can be either indirect or direct and furthermore the organization's main or supplementary offering, Data Monetization is a multidimensional, diverse and complex phenomenon and form of business. Hence Data Monetization can be executed in variety of ways and it may offer distinct strategic purposes for the organizations.

## TIIVISTELMÄ

**JONNA FRED:** Data Monetization – Miten organisaatio voi tuottaa liikevaihtoa datan avulla?

Tampereen teknillinen yliopisto

Diplomityö, 61 sivua

Maaliskuu 2017

Tietojohdamisen Diplomi-Insinöörin tutkinto-ohjelma

Pääaine: Tiedon ja osaamisen hallinta

Tarkastaja: Professori Samuli Pekkola

Avainsanat: Data Monetization, Big Data, Data analytiikka, Digitalisaatio

Digitalisaation edetessä ja teknologioiden kehittyessä datan ja data analytiikan rooli nykyorganisaatioissa on muuttunut entistä merkittävämmäksi. Samaan aikaan dataa ei enää ajatella vain laskelmien raaka-aineena tai yksittäisinä tiedonjyväsina, vaan datan ja informaation on ymmärretty olevan keskeisessä roolissa yritysten liiketoiminnan kehittämisen ja innovoinnin näkökulmasta. Uusi tietopääoma mahdollistaa kehittymisen ja kehittämisen tietoon perustuen ongelmakohtien tunnistamisen jälkeen.

Tietotekniikan kehittymisen myötä dataa tai sen jalostetumpia muotoja ei tarvitse välttämättä kehittää itse, vaan dataa voi myydä tai ostaa siinä missä fyysisiä tuotteita tai palveluita. Silti liiketoiminnan harjoittaminen datan ja sen jalosteiden myymiseen liittyen on varsin uusi ja vähän tutkittu aihealue. Tämä diplomityö tutkii Data Monetization-ilmiota, jolla viitataan dataan perustuvaan liiketoimintaan, jossa liikevaihtoa syntyy datan tai sen johdannaisten avulla. Termille ei ole vakiintunutta tai virallista suomennosta, joka osaltaan vahvistaa käsitystä siitä, että alaa ja ilmiötä ei ole toistaiseksi juuri tutkittu. Data Monetization:ille ei myöskään löytynyt yksikäsitteistä määritelmää, joten tutkimuksessa pyritään lisäksi selkeyttämään ilmiötä ja siihen liittyvää termistöä.

Työn päätutkimuskysymys kuuluu: ”Minkälaiset tekijät muodostavat Data Monetization-ilmion ja vaikuttavat siihen?” Päätutkimuskysymykseen pyritään vastaamaan tutkimalla termin määritelmää sekä sitä, miten organisaatio voi synnyttää liikevaihtoa datan avulla. Lisäksi tunnistetaan muita ilmiöitä, jotka liittyvät Data Monetization:iin sekä erilaisia vaihtoehtoja liiketoiminnan toteuttamisessa. Työ toteutettiin systemaattisena kirjallisuuskatsauksena ja aineistoa etsittiin tieteellisistä kirjastoista ja tietokannoista, kuten Scopuksesta ja Google Scholarista. Koska aiheeseen liittyvää aiempaa tutkimusta ei juuri löytynyt, ei aineistoa koettu mielekkääksi rajata liikaa.

Tutkimuksen tuloksena luotiin kattava, perusteltu ja yksikäsitteinen määritelmä Data Monetization:ille, jota ei kirjallisuuskatsauksen avulla löydetty tähänastisista tutkimuksista. Lisäksi tunnistettiin liiketoimintaan liittyviä komponentteja ja osasia, jotka mahdollistavat liikevaihdon syntymisen datan avulla, sekä erilaisia liiketoiminta- ja ansaintamalleja. Keskeistä on arvokkaan datan tunnistaminen sekä mahdollinen jalostus informaatioksi tai jopa tuotteeksi tai palveluksi, jotta liikevaihdon synnyttävä transaktio voi tapahtua. Koska kyse voi olla ydinliiketoiminnasta tai esimerkiksi lisämyyntimahdollisuudesta, Data Monetization on monimuotoinen ja kompleksinen ilmiö sekä liiketoiminnan muoto. Sitä voidaan toteuttaa hyvin eri tavoin ja sillä voi olla erilaisia strategisia tarkoituksia organisaatioille.

## PREFACE

Data and data analytics as well as business have always fascinated me and hence the topic of Data Monetization was a natural choice for my thesis work. Data Monetization is a rather new phenomenon and it has not been yet studied much which also motivated me to dig deep into this topic.

At first the main focus of this study was on the different strategic approaches and point of views on Data Monetization but quite quickly I realized that the definition and terminology of this phenomenon were ambiguous. Therefore I had to shift the focus towards the fundamentals and basics of Data Monetization in order to be able to later on discuss about the strategic choices and options related to Data Monetization business. This, on the other hand, proved the necessity of this study since the fundamentals were not studied extensively hitherto, or at least I was not able to find such pieces of scientific literature. I believe that Data Monetization is a growing and intensifying phenomenon and form of business and hence it is likely to evolve even further. Additionally it is already a diverse and complex business approach present in the markets. Therefore this subject has a lot to offer for the organizations in the future and furthermore a lot to study for the researchers.

I would like to thank my colleague Jukka Laitinen from Avanade for introducing Data Monetization to me in the first place. I also want to thank Professor Samuli Pekkola for supervising my work and, especially, for great advises and ideas which helped me to enhance my Master's thesis work.

In Helsinki, 25.3.2017

Jonna Fred

## CONTENT

1.	INTRODUCTION .....	1
1.1	Research question.....	2
1.2	Restrictions and limitations of study.....	2
1.3	Structure of study .....	3
2.	RESEARCH METHOD.....	4
3.	FRAMEWORK OF CREATING VALUE WITH DATA .....	11
4.	DEFINING DATA MONETIZATION .....	14
4.1	What is data? .....	14
4.2	Big Data.....	15
4.3	Monetization.....	17
4.3.1	Generation of money flow, revenue and profit.....	18
4.3.2	Value .....	20
4.4	Data Monetization .....	21
4.4.1	Scope of Data Monetization.....	23
4.5	Definition of Data Monetization in this study.....	24
5.	DATA MONETIZATION, STRATEGIES AND BUSINESS MODELS .....	26
5.1	Valuable data.....	26
5.1.1	What is valuable data? .....	27
5.1.2	Refinements and derivations of data.....	29
5.1.3	Packaging and wrapping of data .....	30
5.2	IT and Data Monetization .....	31
5.2.1	Data Privacy and Information Security in Data Monetization.....	33
5.3	Investments in Data Monetization business .....	34
5.4	Business models of Data Monetization.....	35
5.4.1	Direct Data Monetization.....	36
5.4.2	Indirect Data Monetization .....	39
6.	ISSUES RELATED TO DATA MONETIZATION .....	43
6.1	Ethics, data privacy and social networks .....	43
6.2	Cultures and eligibility globally.....	46
6.3	Ownership of data .....	47
7.	CONCLUSION .....	50
7.1	Findings.....	50
7.2	Discussion and evaluation.....	54
7.3	Further research on Data Monetization.....	55
	BIBLIOGRAPHY .....	57

# 1. INTRODUCTION

Nowadays Data Monetization is a growing and developing phenomenon that concerns majority of the organizations globally according to Gartner and some major consulting companies (eg. Gartner 2015; Accenture 2016; Deloitte 2016). Whereas in the past collecting data was considered mainly as a mandatory process to enable the core business, in 21<sup>st</sup> century and especially 2010s organizations have understood the value and revenue potential related to data. Many businesses are now competing with data and analytics in otherwise relatively even markets, to be able to defeat their competitors. Although Big Data, Internet of Things and other popular concepts have praised the value and importance of data, giving an impression that the more data organization has the better, not all data is valuable nor does it yield profit (Wang & Strong 1996; Davenport 2006; Davenport 2012).

In these days it is relatively easy to collect and store data. The organizations have understood the value of data-driven decision making and thus Decision Support Systems (DSS) and Business Intelligence (BI) are utilized to enhance organizational decision making and to help streamlining business operations. Additionally Business Intelligence and data analytics may reveal new business opportunities and increase the performance of business and processes. Although data may be a crucial dynamic capability in the business world, enabling organization's competitiveness in the market, several organizations face the common challenge of data overload (Gelle & Karhu 2003; Davenport 2006; Davenport 2012). This means that the organization has more data than it can handle, understand, interpret and rationalize. Furthermore the organizations may hesitate when valuing their intellectual capital and their upcoming decisions related to data and their business – it may be very challenging to observe the valuable data from raw data mass of high volume. (Gelle & Karhu 2003) And despite data is stored and refined accordingly, the organizations may not know whether certain data is valuable, useful and a potential source of revenue or not.

Data Monetization is a complicated concept – whereas large volumes of data may be required to attain a well representative and better quality sample, thus higher value, large volumes usually lead to higher costs related to maintenance and storage. Maximizing revenue gained from data necessitates careful comparison of different views and strategies related to, for instance, revenue models, business plans, processes and data handling. Current and future technological competencies and investments will also steer organization's Data Monetization possibilities and strategies. Depending on data in question organizations may need to consider issues related to data privacy, information security and even legislation associated with aforementioned.

## 1.1 Research question

This study will explore the complex and yet rather unknown field of Data Monetization considering the technical, financial and strategic aspects of it. The main purpose and objective is to recognize the factors of Data Monetization, and furthermore to investigate how data can be monetized. The main research question of this study is “**What kind of factors are behind of and affect Data Monetization?**”. The following sub research questions will help to answer the main research question:

- What does Data Monetization mean?
- How may an organization generate revenue out of data?
- What other phenomena are related to Data Monetization?
- How can an organization discover and recognize its valuable data?
- What kind of elements organization should consider when planning Data Monetization business?
- What kind of strategic directions organization may take in Data Monetization?

These sub research questions were chosen in order to be able to understand better what Data Monetization is and how an organization can do Data Monetization business. As some major consulting companies, such as Accenture and Deloitte, have depicted, recognition and discovery of organization’s valuable data is a prominent step in Data Monetization (Accenture 2016; Deloitte 2016). Therefore this step was considered as an important sub question since without suitable data there exists no Data Monetization. Furthermore these sub research questions help to understand Data Monetization as a phenomenon, its distinct aspects as well as the factors behind and affecting Data Monetization.

## 1.2 Restrictions and limitations of study

The main focus of this study is to find out how data can generate revenue to an organization and what kind of factors are behind and affecting this revenue generation. Thus the study concentrates on the transactional, strategic and business aspects related to data and Data Monetization. This can be noticed from the sub research questions as well since they highlight the business orientation of this Master’s thesis work as well. The technical details related to data, data processing, warehousing and analytics are mainly excluded from this study. Still it should be noted that the technical aspects are inevitably tightly related to data and they may affect greatly Data Monetization as a phenomenon and a form of business. Thus this kind of technical point of views may be considered in a general level in this study, if they are discovered to affect Data Monetization, hence being a factor behind and affect it.

### **1.3 Structure of study**

The study proceeds as follows: the research method is introduced in the chapter 2 and the framework of value creation developed to support this study is described in the chapter 3. The phenomenon of Data Monetization is introduced and defined in the chapter 4 and after this in the chapter 5 Data Monetization, different aspects of phenomenon and factors affecting it are studied further. In the penultimate chapter 6 distinct issues and problems associated with Data Monetization are discussed. Eventually the chapter 7 will conclude this study depicting the findings and offering an evaluation. Additionally the ultimate chapter discusses the necessity of further research on the topic and proposes ideas and aspects on the prospective studies.



## 2. RESEARCH METHOD

This study is conducted as a systematic literature review, giving a comprehensive view on the current research about Data Monetization and the phenomenon of how an organization may create business and generate revenue with data. At first the idea was to focus on the Data Monetization strategies but soon enough it was found out that Data Monetization as a phenomenon and term is rather new and thus it has not been studied much. Hence this study took a step back to examine the nature, artifacts and definitions of Data Monetization to attain a comprehensive understanding of this phenomenon. This way the study builds a steady basis for Data Monetization phenomenon, thus giving an opportunity to further study the basic aspects and factors of revenue generation with and out of data. Furthermore the distinct aspects and elements related to and affecting Data Monetization are investigated in order to be able to answer the main research question: “What kind of factors are behind of and affect Data Monetization?”.

As mentioned previously, Data Monetization is a rather new term and phenomenon and moreover there is not a great deal of research conducted about this topic yet. When searching term “Data Monetization” on the Internet it is evident, that this term has been established and used continuously by IT and consulting companies, such as Accenture and Deloitte, and IT and business magazines, such as Forbes, Gartner and CIO (eg. Gartner 2015; Accenture 2016; Deloitte 2016). This in turn strengthens the assumption that Data Monetization is a new, hot and still evolving phenomenon that is noticed by the top actors in the IT and business fields globally, thus making it an interesting and notable subject.

As Webster and Watson (2002) suggested, the first step of a literature review was to search and discover relevant literature and contribution. Thus scientific literature was searched mainly from Scopus, Science Port and Google Scholar, first with the keyword “Data Monetization”. Still, although “Data Monetization” gives several straight hits from Google Scholar, only a few pieces of scientific literature handle Data Monetization as it is. Therefore the next step was to search for scientific literature related to associated aspects as well, such as Big Data, Business Intelligence, data analytics and business strategy and models as well. Furthermore the next search was executed with distinct combinations of the following terms:

- Value creation
- Data analytics
- Big Data
- Revenue generation
- Profit generation
- Monetization

- Strategy
- Business model and
- Business Intelligence

Since this study did not concentrate on the technical aspects of data and Data Monetization, those pieces of the found literature were mainly excluded from the sample. With this search and after filtering out the irrelevant, outdated or low-quality literature the following sample was found:

- Chen, H., Chiang, R.H.L. & Storey, V.C. (2012). Business Intelligence and Analytics: From Big Data to Big Impact. *MIS Quarterly: Management Information Systems*, Vol 36 (2), pp. 1165-1188.
- Davenport, T.H., Barth, P. & Bean, R. (2012). How Big Data is Different? *MIT Sloan Management Review*, Vol. 54 (1), pp. 43-46.
- Kamal, R. & Hong, C.S. (2015). Resilient Big Data Monetization. Cornell University, arXiv:1509.04545.
- LaValle, S., Lesser, E., Shockley, R., Hopkins, M.S., Kruschwitz, N. (2011). Big Data, Analytics and the Path from Insights to Value. *MIT Sloan Management Review*, Vol. 52 (2), pp. 21-32.
- Najjar, M.S. & Kettinger, W.J. (2013). Data Monetization: Lessons from a Retailer's Journey. *MIS Quarterly Executive*, Vol 12 (4), pp. 213-225.
- Opresnik, D. & Taisch, M. (2015). The Value of Big Data in Servitization. *International Journal of Production Economics*, Vol. 165, July 2015, pp. 174-184.
- Woerner, S.L. & Wixom, B.H. (2015). Big Data: Extending the Business Strategy Toolbox. *Journal of Information Technology*, Vol. 30 (1), pp. 60-62.
- Zott, C., Amit, R. & Massa. L. (2011). The Business Model: Recent Developments and Future Research. *Journal of Management*, Vol. 37 (4), pp. 1019-1042.

Here low-quality of literature stands for literature that is not scientific or yet published. Additionally conference papers and lecture notes were excluded from the sample. Since Data Monetization, data analytics and Big Data are rather new terms and phenomena, literature published before 2010s was excluded from the sample at this point.

After identifying the pieces presented above, more literature was searched by exploring the bibliographies of the found literature. This method was suggested by Kitchenham et al. (2009) as well as Webster and Watson (2002) since it gives a deeper view on the topic and furthermore reveals the original, prior articles that give the foundation for the topic. At this point, when discovering the literature and ideas behind the previously found sample, also older literature was accepted to the final sample of this systematic literature review. The reason for this is that especially the theories related to business, economics and strategy are developed tens of year ago and thus when defining and studying the business and strategic aspects of Data Monetization older literature may be utilized as well.

With this method wide range of distinct but still relevant studies were found. Moreover this study's literature sample is presented in the table 1:

*Table 1. Final literature sample for systematic literature review*

Source literature	Found from bibliography	Found from bibliography
<b>Chen et al. (2012):</b> Business Intelligence and Analytics: From Big Data to Big Impact.	<b>Davenport (2006):</b> Competing on Analytics.	
<b>Davenport et al. (2012):</b> How Big Data is Different?		
<b>Kamal &amp; Hong (2015):</b> Resilient Big Data Monetization.		
<b>LaValle et al. (2011):</b> Big Data, Analytics and the Path from Insights to Value.		
<b>Najjar &amp; Kettinger (2013):</b> Data Monetization: Lessons from a Retailer's Journey.	<b>Granados &amp; Gupta (2013):</b> Transparency Strategy: Competing with Information in a Digital World.	<b>Granados et al. (2006):</b> The Impact of IT on Market Information and Transparency: A Unified Transparency Framework.
<b>Opresnik &amp; Taisch (2015):</b> The Value of Big Data in Servitization		
<b>Woerner &amp; Wixom (2015):</b> Big Data: Extending the Business Strategy Toolbox.	<b>Constantiou &amp; Kallinikos (2015):</b> New games, new rules: Big data and the changing context of strategy.	<b>Aaltonen &amp; Tempini (2014):</b> Everything Counts in Large Amounts: A Critical Realist Case Study on Data-based Production.

		<p><b>Cordella, A. (2006):</b> Transaction Costs and Information Systems: Does IT Add up?</p>
		<p><b>George et al. (2014):</b> Big Data and Management.</p>
		<p><b>Porter (1996):</b> What is Strategy?</p>
		<p><b>Teece (2007):</b> Explicating Dynamic Capabilities: the Nature and Microfoundations of (Sustainable) Enterprise Performance.</p>
		<p><b>Boyd &amp; Crawford (2012):</b> Critical Questions for Big Data: Provocations for a Cultural, Technological, and Scholarly Phenomenon.</p>
		<p><b>Lee et al. (2014):</b> A Cubic Framework for the Chief Data Officer: Succeeding in the World of Big Data.</p>
		<p><b>Clemons (2009):</b> Business Models for Monetizing Internet Applications and Web Sites: Experience, Theory, and Predictions.</p>
<p><b>Zott et al. (2011):</b> The Business Model: Recent Developments and Future Research.</p>		

As it can be noticed from the table 1, a great deal of seemingly relevant literature was found from Constantiou & Kallinikos (2015)'s bibliography. Although the scientific article written by Woerner & Wixom (2015) was relevant for this study, its bibliography consisted mainly of research briefings written by the writers themselves. Similar issues occurred with Davenport's (2006; 2012), LaValle et al. (2011) and Zott et al. (2011) articles as well and therefore new literature was not found from these pieces of literature.

After reading through the pieces of literature sample, some articles were excluded from this study since they did not discuss about Data Monetization or any relevant aspect of it but rather some other view on data or IT. The following pieces of literature were disqualified:

- **Aaltonen & Tempini (2014):** Everything Counts in Large Amounts: A Critical Realist Case Study on Data-based Production.
  - Study is about production of physical goods and enhancing the production practices.
- **Clemons (2009):** Business Models for Monetizing Internet Applications and Web Sites: Experience, Theory, and Predictions.
  - Studies the monetization practices of advertising on applications and web sites.
- **Granados & Gupta (2013):** Transparency Strategy: Competing with Information in a Digital World.
  - Studies strategies where all the information is shared and the market is transparent.
- **Granados et al. (2006):** The Impact of IT on Market Information and Transparency: A Unified Transparency Framework.
  - Studies strategies where all the information is shared and the market is transparent.
- **Kamal & Hong (2015):** Resilient Big Data Monetization.
  - A highly technical view on Big Data and its utilization in network communications.
- **Lee et al. (2014):** A Cubic Framework for the Chief Data Officer: Succeeding in the World of Big Data.
  - Study of CDO's work, offers advice on how CDO may succeed with Big Data.

After this the core literature sample of this study was formed and classified in accordance with the topic or theme of the research to illustrate the different views and aspects related to Data Monetization:

### **Data Monetization:**

- **Constantiou & Kallinikos (2015):** New games, new rules: Big data and the changing context of strategy.
- **Najjar & Kettinger (2013):** Data Monetization: Lessons from a Retailer's Journey
- **Opresnik & Taisch (2015):** The Value of Big Data in Servitization
- **Woerner & Wixom (2015):** Big Data: Extending the Business Strategy Toolbox

### **Big Data and Analytics:**

- **Boyd & Crawford (2012):** Critical Questions for Big Data: Provocations for a Cultural, Technological, and Scholarly Phenomenon.
- **Chen et al. (2012):** Business Intelligence and Analytics: From Big Data to Big Impact.
- **Davenport (2006):** Competing on Analytics.
- **Davenport et al. (2012):** How Big Data is Different?
- **George et al. (2014):** Big Data and Management.
- **LaValle et al. (2011):** Big Data, Analytics and the Path from Insights to Value.

### **Strategy and Business Model:**

- **Cordella, A. (2006):** Transaction Costs and Information Systems: Does IT Add up?
- **Porter (1996):** What is Strategy?
- **Teece (2007):** Explicating Dynamic Capabilities: the Nature and Microfoundations of (Sustainable) Enterprise Performance.
- **Zott et al. (2011):** The Business Model: Recent Developments and Future Research.

This sample is rather comprehensive in terms of the three topics represented above. The found literature already suggests that Data Monetization as a phenomenon and term is not yet studied very extensively, and therefore the initial literature sample regarding Data Monetization was rather small. Only four scientific pieces of literature that discuss about Data Monetization explicitly were found. Hence the newness of this field of business and its relation to Big Data were indeed supported by the sample.

It should be noted that also some other pieces of literature are used in this study to support the systematic literature review's sample and deepen the view on distinct aspects, phenomena and their relationships. Especially important are Choo's (1995; 2002) scientific articles that create the basis for a value creation framework developed and introduced in

the chapter 3. Still, this sample will have an essential role in studying Data Monetization and simultaneously attempting to answer the research and sub research questions presented initially in the chapter 1.

### 3. FRAMEWORK OF CREATING VALUE WITH DATA

In order to be able to study and investigate the topic of Data Monetization and furthermore answer the research questions with a comprehensive and concise manner, a framework of value creation process related to data is introduced and explained. This framework will be utilized in the subsequent chapters to give support and structure to this study and additionally to help to analyze the relationship between life-cycle of data and the distinct aspects and forms of Data Monetization business.

According to Choo (1995; 2002) and Thierauf (2001) data, information and particularly their refinement, knowledge, form a fundamental basis and vital condition for any organization and business. In other words, no organization or business would exist without data and its refinements since some sort of data, information or knowledge is always needed when decisions are made, innovations are created and designed as well as when an organization is trying to grasp the environment and the dynamics of markets (Choo 1995; Thierauf 2001). Hence these aspects are essential for both the organizational planning and implementation of Data Monetization business, but also for the Data Monetization business itself.

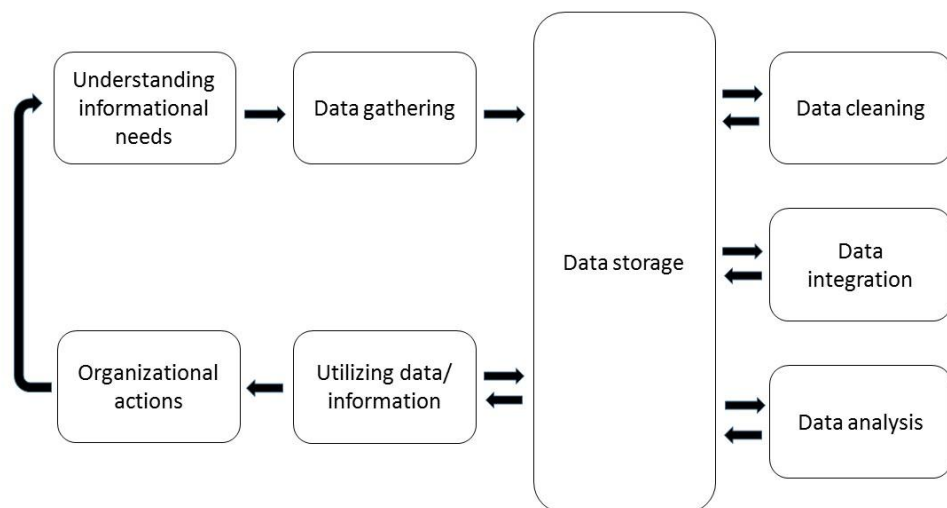
Thierauf (2001) argues that value creation process is tightly related to the knowledge creation process since data or information by themselves are not necessarily useful for an organization but serve as an essential raw material to further refined knowledge which may be used, for instance, in decision making. Hence the value of data, information or knowledge is realized when eventually utilized, for instance, in decision making. Furthermore value can be added by refining, combining and deriving the possessed data, information or knowledge, which means that simultaneously new data, information or knowledge is created (Thierauf 2001). Thus on this point of view arguably work and effort can be considered as an input, resulting an output with an increased perceived value.

Indeed Choo (1995) has formed a model to describe this vital process present in the organizations. According to this model an organization must first interpret and make sense of data or information to be able to understand what they have, what is still needed and how the possessed data or information could be turned into something beneficial. After this the interpreted information is processed by individuals and converted into knowledge that is shared and utilized in the acts of design and innovation. Ultimately the created knowledge can be used in decision making that steers and guides organization's actions and moreover its business. (Choo 1995) An important aspect of this process model is that at least part of the data or information present in the flow is external, which is comprehensible when taking into account Thierauf's (2001) and Choo's (1995) point of views



on utilizing data, information and knowledge in understanding environment and, in case of business, the market dynamics. Additionally not only tacit knowledge should be applied but also explicit knowledge and facts as a basis of knowledge creation (Choo 1995).

Later on Choo (2002) has developed his original views and reinforced the cyclical nature of the process. Indeed, an organization's informational needs are never satisfied since the organization is not functioning in a vacuum – furthermore its operational environment is continuously changing and developing, which means that the organization and its business must adapt and develop as well. Therefore when the previously described process reaches the last step, namely the decision making phase, the process will initiate due to the new questions and informational needs aroused, especially through the monitoring phase that ensures that the organization learns from its actions and is able to enhance its performance and processes in the future (Choo 2002). For this study a framework is developed based on the research conducted by Choo (1995; 2002) and Thierauf (2001) and it is illustrated in the picture 1:



**Picture 1: Value creation framework. Adapted from Choo (1995; 2002) and Thierauf (2001)**

The framework presented in the picture 1 illustrates the cyclic and continuous process of value creation based on data and information in an organization. The process usually initiates from the top left corner, from the step of understanding organization's informational needs. These needs will steer and, ultimately, are a prerequisite for the search and gathering of relevant data. When the necessary data is gathered, it will be stored to some kind of data storage, such as a database or a data warehouse. The data may be cleaned, integrated and processed as well as analyzed further, depending on the situation and need. The distinct processing actions and methods may increase the value of data, and hence

add value for the end user. After this the data or information can be distributed and shared through which it can be further utilized. While proceeding with the framework, organizational actions and changes can be applied based on the utilized data and the cyclical process will be initiated again.

It is essential to note that monitoring takes place in each phase of the process, which means that correctional actions can be evoked if needed. This type of situation may occur if, for example, the data gathered does not meet the organization's actual needs. Thus the organization may return back to the first phase, review its needs and initiate data gathering with a more comprehensive plan and approach. Hence this process is also dynamic, which means that the arrows presented in the picture 1 are not casted into iron but other, alternative approaches and paths may be taken if problems are detected while monitoring the process. Additionally the process may be cancelled or stopped at any step if the organization considers this to be a beneficial way to act.

## 4. DEFINING DATA MONETIZATION

Data Monetization, as a term, is rather new but it has been used as a synonym for generating money with data. The term ‘monetization’ can be defined as the utilization of something of value as a source of profit (Merriam-Webster 2016a). Contrary to Merriam-Webster’s rather narrow definition, scientific literature seems to offer a wider definition for ‘monetization’ stating that it may be generation of money or revenue (for instance Najjar & Kettinger 2013; Woerner & Wixom 2015). Hence Data Monetization is not necessary only about profit yielding but it can be interpreted as a wider phenomenon from the point of view of monetary transaction. This view is also supported by MOT Dictionary’s (2016) definition: “*convert or adapt to trade based on the exchange of money*”. Nevertheless, it is obvious that there exists contradiction between these definition, since there is a significant difference between terms “money”, “revenue” and “profit”. Thus, at this point, it seems that Data Monetization stands for utilizing data as a source of positive, incoming money flow.

Since currently Data Monetization is recognized and defined only in a few scientific article, it is important to study the terminology and the definitions first. The assumption is that there is not yet a scientifically accepted, uniform definition for Data Monetization, which means that the contradiction between distinct research and their point of views may occur. This, on the other hand, is not uncommon and furthermore strives the field of research. At the end of chapter 4.5 is stated the derived definition for Data Monetization that is used in this study.

### 4.1 What is data?

In order to be able to define and understand Data Monetization, it is essential to define and comprehend data. Thus at this point a brief definition of data will be given.

Data can be described as separate, qualitative fragments or bits of knowledge which do not represent anything meaningful or greatly useful by themselves (Alavi & Leidner 2001; van Belle & Ruiter 2014). An example of data would be a number series “1/4/2016” – this number series becomes useful after it can be interpreted (van Belle & Ruiter 2014). Moreover after the interpretation data becomes information. Interpreting, on the other hand, may require some previous information or knowledge and this may lead to a situation where different people interpret same data differently, depending on their intellectual capital (Rowley 1998). For instance an American would interpret the previously introduced number sequence as a date “fourth of January 2016” whereas European would say that the data is “first of April 2016”. Thus a date that consists of quantitative data and the interpretation combined together become more meaningful, qualitative information with

the help of individual's intellectual capital, namely possessed information and knowledge (Rowley 1998; van Belle & Ruiter 2014).

## 4.2 Big Data

The megatrend of digitalization has generated other trends such as Internet of Things (IoT) and Big Data. According to Woerner and Wixom (2015) and Constantiou and Kallinikos (2015) Big Data is one key factor that has been behind data analytics, especially prescriptive and predictive analytics, as well as Data Monetization. Thus a deeper look into Big Data needs to be taken in order to be able to understand Data Monetization business and furthermore to be able to answer the sub research question about how an organization can recognize and discover valuable data.

Enhanced technologies, social media, increased amount of sensors and digitalization have created an expanding and continuously growing mass of real-time, heterogeneous data. Furthermore Big Data, term standing for this enormous data mass, was developed. (Davenport et al. 2012; George et al. 2014) As depicted by Davenport et al. (2012) Big Data affects everybody and every organization nowadays, and although everybody is talking about it, Big Data is mostly used as a buzzword to describe potentially more insightful data. Despite distinct stakeholders and organizations may define Big Data little differently and the problematic nature of Big Data is commonly accepted in the scientific literature (for instance Davenport et al 2012; Chen et al. 2014; George et al. 2014; Woerner & Wixom 2015; Constantiou & Kallinikos 2015).

According to Merriam-Webster (2016b) Big Data can be defined as “*an accumulation of data that is too large and complex for processing by traditional database management tools*”. Indeed, as the term itself suggests, Big Data is associated with tremendous volumes of data that is usually unstructured, real-time and dynamic (Davenport et al. 2012, Chen et al. 2014). Although distinct ways to define Big Data exist, most of them seem to reflect the quite commonly used definition of three (3) V letters. Although the amount of V letters may depend on the source and depth of the definition, it seems that most of the definitions contain the following three V letters: Velocity, Variety and Volume. Thus these three terms to depict Big Data are regularly used and furthermore commonly accepted (Chen et al. 2012; Chen et al. 2014; Jagadish et al. 2014; Opresnik & Taisch 2015). Additional V letters to describe Big Data are Value and Veracity, depending on the source (such as Buhl et al. 2013; Chen et al. 2014; Opresnik & Taisch 2015) Hence some variation exists in the use, amount and naming of these additional terms in Big Data's definitions. The distinct V letters are described in the following table 2:

*Table 2: Different V letters for Big Data*

<b>Term</b>	<b>Definition</b>
Volume	Describes the scale of data. In case of Big Data, the scale is big or even enormous. According to IBM (2016a, b) 2,5 quintillion bytes of data is created every day and, for instance, The New York Stock Exchange processes 1TB of trade data during every trading session.
Velocity	New data is generated very fast and so does the already existing data evolve and become outdated. Thus dynamic Big Data stream may be real-time, requiring real-time processing techniques as well.
Variety	Different forms and types of data, thus refers usually to the unstructured nature of Big Data.
Value	According to Opresnik and Taisch (2015) the significance of Big Data relates to the value that may be unveiled when processing, analyzing and applying Big Data and its derivatives. Thus value is tightly related to other attributes of Big Data, such as veracity, since without valid, useful and good-quality Big Data not much value can be created or added.
Veracity	Describes the uncertainty, validity and accuracy of data. Hence it could be claimed that veracity stands for the quality of Big Data.

Although Big Data is a promising and prominent phenomenon these days, there are also challenges related to Big Data, its processing, analysis and utilization. According to Boyd and Crawford (2012) Big Data has aroused arguments for and against the gathering, storing, analyzing and, in some cases, publishing of these huge masses of data and its derivatives. One of the most significant concern and cause of anxiety is related to privacy

threats and tracking of people. Since the mobile devices, applications, software and even home appliances are collecting and sending data somewhere, this concern is real. (Boyd & Crawford 2012). Close to this are concerns associated with information security, objectivity and reliability of Big Data (Boyd & Crawford 2012; van Belle & Ruiter 2014; Gerlach et al. 2015). For instance when data is collected from social media, issues related to the previously mentioned concerns may appear because the party collecting the data cannot be absolutely certain of the accuracy, objectivity and reliability of data produced, intentionally or unintentionally, by another person or actor.

Due to the properties of Big Data, technical challenges exist as well. The enormous volumes of varying, unstructured data challenge the traditional database and server technologies (Constantiou & Kallinikos 2015). Diverse data formats and types require special attention on the data storage technologies, or the fast cleaning processes to enable utilization of Big Data. As one of the key features of Big Data is timeliness, it is essential to be able to store, process and analyze data in near real-time. Otherwise Big Data may not be very useful, since its dynamic nature offers a lot of benefits in terms of real-time decision making. Furthermore Big Data requires much more calculation, processing and storage capacity than smaller, uniform and structured data sets. (Constantiou & Kallinikos 2015) Thus computational power as well as algorithms used in analysis and pattern recognition require caution and attention. But nevertheless, as argued by Boyd and Crawford (2012), Big Data is not useful if it cannot be processed appropriately. They also state that although the technical aspects are taken into account, a poor quality Big Data will not give accurate, reliable and useful results and derivatives, no matter how fast and well the data mass is stored, processed, analyzed and refined. (Boyd & Crawford 2012).

### **4.3 Monetization**

It is necessary to define and study what monetization really mean to be able to understand the phenomenon of Data Monetization. As mentioned before, monetization stands for utilizing something of value as a source of money, although the definitions seem to vary. This issue with definition regarding the monetary terminology will be discussed in the chapter 4.3.1.

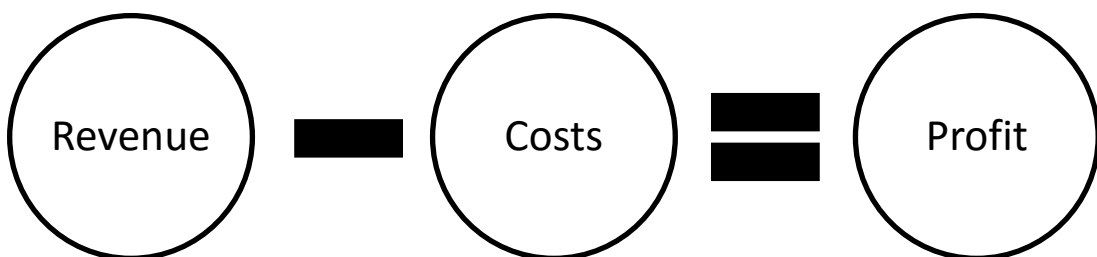
In order to be able to gain money, whether the point of view taken is on profit or revenue generation, there needs to be a transaction in which something of value is ultimately exchanged or converted into money. Furthermore the transaction requires at least two parties who are interested in executing this exchange. More pragmatic way to depict this transaction is to consider it as selling something that is worth a certain amount of money. This transaction has also one more requirement: the other party must be willing to buy this good or service and another must be willing to sell it with the equal amount of money. Thus both parties must have similar valuation of this good or service, otherwise the transaction will not occur. Consequently the laws of demand and supply are also present in the

action of monetization – when the market equilibrium is achieved, demand and supply are equal and the transaction will take place (Mossin 1966).

According to Cordella (2006) distinct costs are associated with the business transactions and thus these costs need to be considered when planning and implementing Data Monetization business. These transactional costs are an inevitable part of business actions since some groundwork and preparations must be done before a transaction may occur. Cordella (2006) has divided these costs into three groups: to the search costs, negotiation costs and enforcement costs. The search costs are associated to the costs that are required to find the opportunities for exchange whereas the negotiation costs stand for the costs related to negotiating the terms of exchange. Ultimately the enforcement costs mean the costs related to creating and enforcing the contract. (Cordella 2006) Thus it is important to consider and comprehend the transaction costs since no monetization may occur without transaction and furthermore transaction costs.

#### 4.3.1 Generation of money flow, revenue and profit

Revenue and profit are not synonyms but there is a clear difference between these two terms. According to Merriam-Webster (2016c, d) revenue can be defined as “*money that is made by or paid to a business or an organization*” whereas profit’s definition is “*money that is made in a business [...] after all the costs and expenses are paid : a financial gain*”. Thus according to these two definitions it is already evitable that profit does not equal to revenue. The crucial part of this difference is that revenue does not mean that the business is making any surplus, hence the business may have more expenses than its revenue. Whereas while considering the meaning of profit, an organization’s revenue is greater than its costs and expenses, leading to surplus and profit. The difference of revenue and profit is shown also in the picture 2.



*Picture 2: Relationship between revenue and profit*

Although the studied literature lacks a clear definition for Data Monetization, stating whether it refers to the generation of money, revenue or profit, it is essential from the

point of view of this study to rationalize the economic view on Data Monetization. Profitability is the common goal that all profit-making organizations share but it is not self-evident that an organization will be profitable. Indeed, some organizations and businesses are unprofitable, meaning that the sum of costs and expenses is greater than the incoming revenue stream. This may be intentional or unintentional – there exist also non-profit and public organizations that are not even seeking profit but their mission is rather to serve the society. (Goulet & Frank 2002) Additionally there are times when organizations may struggle with profit generation: for instance emerging businesses may face entry barriers, such as costs, that may lead to unprofitability during the first years of business (Caves & Porter 1977). Thus although seeking profit may or may not be the organization's goal, succeeding in generating profit is not self-evident or even achievable in the near future.

There are couple of ways a for-profit organization can affect its profit yielding. As it can be noted from the picture 2, profit consists of two factors: revenue and costs. Thus there are two different views for increasing or even maximizing profit. An organization may try to apply cost minimization, which means that profit is created through decreasing costs or making sure that the costs do not increase in proportion to the increase in revenue, thus leading to a greater profit. Another aspect is, naturally, to focus on increasing revenue and that way pursue higher profit. Likely the organizations will apply and keep in mind both views, since profit is all about the relation of revenue and costs.

Nevertheless, Najjar and Kettinger (2013) describe Data Monetization as an act of converting data into money by offering tangible benefits through data or by avoiding costs with data. This description is vital from the point of view of this study's definition for monetization since revenue considers only incoming money flow but does not take into account money flows inside the organization or the costs and expenses faced. Thus the definition of monetization could be considered with a wider viewpoint since data may be and is commonly utilized to cut the costs and money leaks inside the organization. The avoidance of costs is indeed an important aspect of Data Monetization but it is internal activity taken by an organization, which cannot be considered as a part of Data Monetization business.

Although the phrase of “*generating money flow*” seems to be a valid definition in monetization, it is still quite wide viewpoint and does not take into account the economic and business terminology established and stabilized in the field's literature. More commonly one discusses about profit or revenue rather than describes company's functions and operations with term ‘money’. Therefore generation of money or money flow does not seem to be an appropriate term to use in this context. Still, it should be noted that both revenue and profit restrict this study and the view on monetization.

Due to the findings related to the terminology of profit and revenue, this study will consider Data Monetization as revenue generation with and out of data. There are several reasons behind this consideration:



1. Not all organizations seek profit. This does not still mean that non-profit or public organizations could not establish Data Monetization. Thus this study does not intend to exclude non-profit or public organizations but considers Data Monetization as a phenomenon that may relate to any organization.
2. Generating revenue is prerequisite for yielding profit. Thus it would be logical to consider revenue generation in this study since it is closer to the transactions related to Data Monetization. Revenue generation may also create some other benefits than profit.
3. Although an organization is unprofitable, it does not mean that it would not be successful in Data Monetization at the moment or in the near future. Data Monetization may also be a complementing business trade that can be unprofitable at times while offering some other benefits and competitive advantage.
4. Revenue generation is considered as an appropriate level of investigation in terms of the economic point of view. Although revenue is more limited term compared to money flow, monetization's definition reinforced also the exchange of something of value to money. Thus organization's internal money flows as well as costs and expenses can be excluded from this study.

Thus, from now on, in this study the phrase of “*making money with data*” can be parallelized with revenue generation with or out of data or information. Still it is essential to bear in mind that most organizations seek profit, meaning that even though Data Monetization's primary goal is to generate revenue, in the long run the organization pursues profit yielding. In case of non-profit or public organizations, instead of seeking profit, operations should still be beneficial, not just creating expenses. Thus the ideal situation is that the revenue generated is greater than the costs and expenses faced.

### 4.3.2 Value

According to Merriam-Webster (2016e) there are couple of, rather different definitions for multidimensional matter of ‘value’. First of all, value can be described as “*the amount of money that something is worth: the price or cost of something*” and “*a fair return or equivalent in goods, services, or money for something exchanged*”. Another definition may be “*usefulness or importance*”. (Merriam-Webster 2016e) Indeed the given definitions by Merriam-Webster (2016e) seem to reflect the ambiguity of value as a term. Since value is something intangible, giving a definition is not as straightforward as with the physical matters. From Merriam-Webster's (2016e) definitions it can be pointed out that value may depend on individual's subjective perception as well as the context. This is reinforced by the last definition and the mentioned word “fair” in the second phrase. Hence value is relative to individual's experiences, perception, wants and needs. In this study value is considered from the perspective of business, thus highlighting the aspect of value creation and perceived value in transactions. Furthermore the point of view of

human dignity or values is irrelevant for this study and therefore excluded from this analysis.

Indeed value is tightly related to individual's perception and experiences. For instance Zeithaml argued already as early as 1988 that quality and value are related to each other, and that emotions, user's experiences and psychosocial consequences affect perceived value. Thus, although value may be measured by individual's perceived satisfaction and personal interests in the transaction, in the business world quantitative measures for value are preferred. Therefore value is commonly determined by the commonly accepted valuation, in most cases referred as the market price or market value. (Zeithaml 1988; Porter & Kramer 1999) This further reinforces the transactional aspect of value, which means that the final measure of value is usually achieved when an exchange of money for something worth of value has occurred (Porter & Kramer 1999; Payne et al. 2008). Hence the exchange will not take place if the valuation is not appropriate, leading to the market value (Porter & Kramer 1999). Still, the market value is not an absolute truth of something's value, but it is a highly appreciated and recognized opinion which may steer the behaviors and decisions of majority of mankind.

When considering value from the point of view of business, Ravald and Grönroos (1996) have argued that company's ability to offer superior value to the customer can be considered as one of the key elements in gaining successfully competitive advantage. Thus the ability to create and offer superior value is an essential constituent of differentiation and sustaining the competitive advantage in the long run (Ravald & Grönroos 1996). Furthermore value seems to relate very tightly to the context of value perceived by and created for customers (Ravald & Grönroos 1996; Payne et al. 2008). This perception is further reinforced with a view that value does not exist before an offering is consumed and used. From this modern point of view called service-dominant (S-D) logic, value is co-created by the supplier and the customer (Ravald & Grönroos 1996; Vargo & Lusch 2007; Payne et al. 2008) Thus the customer's experience and perception have a great role in the value creation and determination.

#### **4.4 Data Monetization**

Some of the technology research companies, such as Gartner, have noted Data Monetization as an upcoming phenomenon, created mainly by the megatrends of digitalization and Big Data. (Woerner & Wixom 2015; Gartner 2015) As discussed in the chapter 4.2, digitalization, social media and IoT with sensor technologies have increased substantially the amount of data processed and gathered by organizations. Some organizations also gather and store data that they cannot process and refine or they do not know how to utilize it. Thus the amount of data available for monetization may be enormous, but what is Data Monetization in reality?

Gartner (2015) has identified two distinct ways how an organization may generate revenue and monetary value with data: with the direct and indirect ways. The direct revenue generation means that data is traded or sold and therefore monetary value is produced in this straightforward transaction. In the indirect revenue generation data is utilized and refined to produce new information, services or products that are sold or traded. Therefore in the indirect way the key is in leveraging data and generating monetary value with refining data to something else that is valuable. (Gartner 2015) Apart from Gartner's (2015) distinction, Data Monetization can be organization's core or non-core business. Naturally Data Monetization's position in organization's portfolio of business lines is crucial when considering how much organization is willing to invest in and focus on this business activity.

Not only an organization can sell or trade data or its derivatives but Woerner and Wixom (2015) have also identified 'wrapping' as a way to monetize data. Wrapping means that an organization wraps information or data around its core product or service, mainly to differentiate it from the alternatives and competitors in the market. Thus data wrapping is associated with Data Monetization that is not organization's core business. Wrapping may make a product or service more attractive to the customers if it fulfills some kind of informational need. Furthermore wrapping may lead to greater value and thus greater revenue generation. (Woerner & Wixom 2015) Thus it seems evident that Data Monetization is rather diverse phenomenon in terms of how data can be used in business to generate revenue.

Najjar and Kettinger (2013) have studied Data Monetization and distinguished Data Monetization from data sharing. According to Najjar and Kettinger (2013) the difference is that in data sharing no price is set for the data, thus data is shared free of charge to, for instance, other members of a supply chain. Although data sharing does not necessary involve a monetary transaction or trade, it still pursues for adding value and gaining competitive advantage. Therefore it seems that Data Monetization's main target is simply financial – the aspect of adding value, gaining competitive advantage or developing business relationships may be present in Data Monetization but the core idea is in the generation of revenue. (Najjar & Kettinger 2013) Still the difference between indirect Data Monetization, as described by Gartner (2015), and data sharing may be indeterminate: as the indirect Data Monetization can mean simply leveraging data to generate revenue, data sharing may generate revenue indirectly to an organization. An example of this can be found from supply chains or networks in which sharing data may strengthen competitiveness and enhance the efficiency of an organization in the market, leading to, for instance, an increase in sales and margin as well as a greater revenue generation.

Another important point made by Najjar and Kettinger (2013) is that the avoidance of costs could also be a form of Data Monetization. This means that if an organization is able to block negative outgoing money flows with the utilization of data, Data Monetization has occurred. Thus the organization may, for example, be able to avoid costs or cut

them. Hence because of this wider view offered by Najjar and Kettinger (2013) Data Monetization may appear in varying forms and it seems that these activities and transactions executed by an organization are more common than previously assumed. Still, this aspect of Data Monetization is excluded from this study due to the business orientation and viewpoint taken.

It seems that the diversity of Data Monetization is not limited only to the phenomenon itself but also the terminology used. Opresnik and Taisch (2015) have studied Big Data's effect on the organizations' servitization where data may be used to offer added-value to the customer and new business models can be created by reusing or selling data. Servitization means, in short, the transformation from manufacturing operations towards service production as a supplementary or core business (Baines et al. 2009; Opresnik & Taisch 2015). When considering the definitions given above, servitization with the help of Big Data seems very similar to Data Monetization since data is utilized to invent and offer new supplementary services. Additionally selling data is noted as a way to create new revenue streams. Indeed, servitization seems to be a notable form of Data Monetization that is very similar to Gartner's (2015) definition of indirect Data Monetization business. Still, servitization as a term is limited to manufacturing organizations and businesses, and thus can be considered only as one, specific form of Data Monetization.

#### **4.4.1 Scope of Data Monetization**

As already noted, definitions of monetization or Data Monetization are not truly unambiguous, although the phenomenon may seem very simple at the first glance. Whereas the key part of the definition, that is, the aspiration to generate revenue with data, is easy to understand, the reality seem to contain a great amount of distinct ways to implement Data Monetization. Thus defining Data Monetization precisely is challenging and highlights already the complexity and multidimensionality of this phenomenon.

One found ambiguity of Data Monetization relates to data itself. Although the discovered definitions of Data Monetization discuss about generating revenue or profit with data, it seems that in the reality organizations may be dealing with information instead. For instance LaValle et al. (2011) mention both data and information when describing how an organization may generate more value out of their information, using data analytics and Big Data. Thus data and information are sometimes mixed up, which may indicate that these two terms are hard to actually separate from each other. Naturally ignorance can affect as well, but since the article of a recognized, scientific magazine uses these terms side by side, this cannot be omitted when considering the definition of Data Monetization. Furthermore, as mentioned before, Data Monetization can be indirect, which means that instead of selling data, an organization may sell information. Thus information is tightly related to Data Monetization, although the term itself does not suggest this.

On the other hand Wixom (2014) has given a broader yet more precise definition for Data Monetization, stating that “[Data monetization is] the act of exchanging information-based products and services for legal tender or something of perceived equivalent value”. It should be noted that in this definition Data Monetization is not limited to data but revenue may be generated with information-based products or services. As Woerner and Wixom (2015) argue in their later studies, the term “*information-based products or services*” may refer to a wide range of distinct offerings with varying level of complexity. Thus an information-based product or service can be, for instance, raw data, enhanced data, a derivative or result of analytics, process design or even process execution. (Woerner & Wixom 2015) Therefore it seems that Data Monetization is not, in reality, limited to transactions related to only data, but it is rather a hypernym for all the actions which purpose is to generate revenue with or out of data or information-based products or services. This is indeed supported by Gartner’s (2015) partition of indirect and direct Data Monetization as well.

This broader definition of Data Monetization given by Wixom (2014) does not clearly suggest that Data Monetization would necessitate money in the exchange data or information-based products or services for something of comparable value. Therefore Data Monetization might mean that an organization could exchange data to, for instance, a tangible product. Although this might not be a common act taken by companies, it is still a possible scenario. Despite the definition given for monetization excludes this kind of action since the transaction should generate revenue. This means that depending on the gained value and its form as well as the country’s legislation some tangible and intangible products or services can be considered as revenue in the company’s accounting.

#### **4.5 Definition of Data Monetization in this study**

In this study Data Monetization is thus thought as the revenue generation with and out of data and data-derived and information-based products and services. Thus the definition and point of view on Data Monetization is expanded from data to also its derivatives, which means that an organization can, for instance, derive and refine its data to other products that are then, in turn, sold forward, generating revenue for the organization. This indeed expands the variety of distinct business and monetization models that can be exploited in Data Monetization – data or refined information-based products and services can be sold, leased, rented or, as pointed out by Woerner and Wixom (2015), they can be wrapped to another products or services.

This study does not limit the business models related to Data Monetization or the actual origin of data utilized and possibly derived. Indeed the only limitation is related to the aspect on monetization – since this study focuses on examining Data Monetization as a business, monetization is defined here as revenue generation. This means that Data Monetization is considered here similarly as Wixom (2014) determined the phenomenon. Therefore the important but here still irrelevant observation made by Najjar and Kettinger

(2013) about the cost avoidance dimension of Data Monetization is excluded from this study's scope.

The reason for this study's rather wide viewpoint is the lack of previous studies related to the subject – in order to be able to determine and closely examine this phenomenon and its constituent, Data Monetization needs to be considered as a fairly unrestricted phenomenon. And nevertheless a wider point of view is necessary in order to answer the main research question of this study: “What kind of factors are behind of and affect Data Monetization?”

## 5. DATA MONETIZATION, STRATEGIES AND BUSINESS MODELS

As distinguished by Gartner (2015), Data Monetization can be indirect or direct. Thus both forms of Data Monetization business are studied and the associated business models and strategies will be identified. As Data Monetization is a multidimensional and varying form of business, it is crucial to understand the different phases and factors of Data Monetization before proceeding in this study.

From the literature of Najjar and Kettinger (2013), Woerner and Wixom (2015) and Constantiou and Kallinikos (2015) it may be pointed out that Data Monetization is likely to be usually a non-core business for organizations that have potentially valuable data due to their other business activities and thus existing information technology, infrastructure, marketing and other business capabilities to support and enable the launch of Data Monetization business. This is an important observation since if an organization considers Data Monetization as a non-core business activity, the Data Monetization strategy and business model are likely to be reinforced by the core business and thus by the enterprise-wide strategy and vision. Hence Data Monetization cannot always be considered as a sole business activity but rather as a business line amongst a wider business portfolio. Still, there exist organizations focusing on Data Monetization which means that both possible emphases must be considered in this study.

An organization has distinct options on how to implement and execute direct or indirect Data Monetization business. Furthermore these ways and options are considered as business models. The business model is supported by the organization's strategy and objectives and thus create the basis for Data Monetization, stating the goals and measures taken to achieve this goal (Zott et al. 2011). After planning the business model and strategy as well as the operations and operation execution, needed IT must be procured. After this Data Monetization business can be launched and the operation execution will be started.

### 5.1 Valuable data

One of the sub research questions is determined as follows: How can an organization discover and recognize its valuable data? In order to be able to monetize data, the valuable data that could be used in Data Monetization business and that could generate revenue must be discovered and identified. This is the case with both indirect and direct Data Monetization. In this chapter term "data" is used instead of mentioning both data and data derived information-based products and services for sake of simplicity.

The discovery and identification of valuable data and its derivatives can be seen as an important part of the beginning of Data Monetization. This view is reinforced by Accenture and Deloitte, two successful and appreciated consulting companies offering Data Monetization consulting services to companies. These offerings both highlight the essential aspect of identification of valuable data, promising to help the customer to discover and determine the valuable data and its worth (Accenture 2016; Deloitte 2016). Similarly the framework of value creation introduced in the chapter 3 suggests that the data processing and sharing phases are very dependent on the phase of understanding the information needs and furthermore the step of data gathering. If these two steps are underrated, it is likely that this underestimation affects greatly the subsequent phases of this framework.

The identification of valuable data is not always easy, as argued by Gelle and Karhu (2003). They highlight the common issue regarding the expanding volumes of data, leading to a situation where the ‘data overload’ may confuse an organization and thus disturb the discovery and identification of the valuable data. Another potential issue related to the identification and discovery of valuable data arises from the value creation framework introduced in the chapter 3. As argued especially by Thierauf (2001), value and value created are realized only when they are experienced or utilized and therefore data may not be implicitly valuable since a certain degree of uncertainty is associated with it. The perceived value is, after all, dependent on the utilization of the data and hence the alleged value may not resemble the realized value.

Nevertheless, as argued by Zeithaml (1988), Porter and Kramer (1999) and Payne et al. (2008), transaction may succeed and occur only if both parties have an equivalent perception on the value present in the exchange. Thus in Data Monetization, the data or information-based product or service must be worth some value – otherwise no one would purchase it.

### **5.1.1 What is valuable data?**

According to Zeithaml (1988), value and quality are tightly related to each other. This view is supported by a great deal of researcher, for instance, Strong et al. (1997), Gelle and Karhu (2003) and Setia et al. (2013). Hence value and quality are both difficult aspects to measure because of their subjective nature – both quality and value are dependent on the perceiver’s experiences, needs and view on the subject. Thus both value and quality are reliant of one’s subjective opinion and the experience on how well, for instance, a tangible product fulfills one’s needs, wants and expectations (eg. Zeithaml 1988; Setia et al. 2013). Although the valuable data in this study is considered mainly as data that is worth money and may generate revenue, it is still essential to notice the subjective aspects on determination of value and quality.



In order to see data as something valuable, meaning that someone is willing to pay for it, data must fulfill a certain level of quality and constitute potentially interesting and useful meaning and content. This is an important view regarding Data Monetization business as well. Furthermore someone must value data enough to be willing to pay for it, thus data must be somehow useful and meet clients' requirements and needs. This approach on value is supported by the framework introduced in the chapter 3 that is based on Choo's (1995; 2002) and Thierauf's (2001) studies. Although different companies and people may value data a little differently, it is clear that the features and quality of data weigh in one's valuations. Still technically high quality data is not necessarily valuable from the point of view of Data Monetization business. This is true if, for example, data is commonly accessible and available, outdated or otherwise its content and context are not found interesting and useful. Nevertheless, one man's trash may be another man's treasure meaning that the right customer must be found since people and businesses value and appreciate matters differently. Furthermore some organizations may value volume over quality – in some cases the sample size is more important than the absolute quality of the sample, thus meaning that one would rather choose a large mass of data with poorer quality than a small mass of data with high quality.

As concluded previously data quality affects the value of data. Thus it is necessary to investigate data quality further. In general quality can stand for superiority or adequacy (Golder et al. 2012). However this definition is rather narrow and measuring of quality is challenging. Golder et al. (2012) argue that quality is difficult to define and there exist several distinct definitions but, similarly to value, quality is dependent on the expectations and experiences. Setia et al. (2013) offer a different view on quality, suggesting that data quality consists of variety of different features, such as consistency, completeness, integrity, accuracy, validity, reliability and timeliness. This view is supported by Wixom and Todd (2005) who also argue that the validity of data is not a synonym to reliable and objective data – although a data set may be valid and accurate, the users of data may not be able or want to rely on this data. Thus valid and accurate data set is not necessarily reliable or objective, something to be worth users' trust.

Data quality is also dependent on the tools, information systems, processing methods, interfaces, databases and data warehouses utilized. Thus originally high quality data may be deteriorated with poorly selected practices or tools. To ensure the quality, validity and usability of data, data processing and maintenance must be performed. (Strong et al. 1997; Setia et al. 2013) In some cases data utilization depends heavily on data processing since the original data format may be incompatible with the tools and software used or otherwise useless without some processing. Thus in this kind of situation the value of data cannot be monetized without first investing in some time and work. An example of such situation would be when the data is not in any widely supported format that someone else could utilize easily – this might mean that no one would be willing to buy this data since it cannot be utilized easily and without work or even investments to technologies that

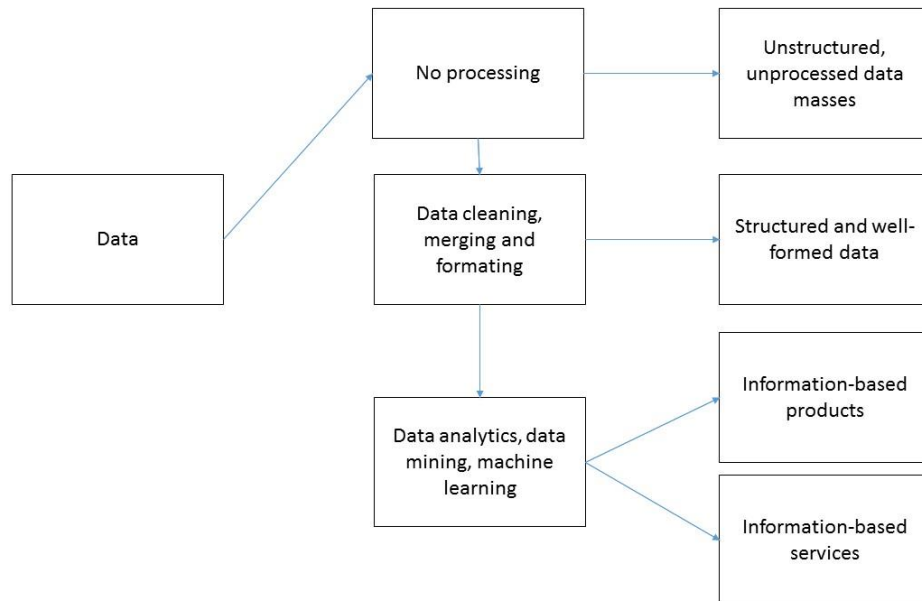
would enable the needed conversions. Also the maintenance of data quality and data is important – otherwise the previously valuable data might become invaluable due to quality issues.

### **5.1.2 Refinements and derivations of data**

In Data Monetization the data sold can be anything between the unstructured raw data and highly processed and refined information-based products or services. In order to be able to utilize, refine and even sell the data it first needs to be gathered and stored. As described in the value creation framework presented in the chapter 3, the process may or may not contain distinct activities associated with data, for example data integration or data analysis. Hence the next phases of data processing depend on the strategy taken by the organization, the data source or sources used and the type of data stored. Here the distinct Data Monetization strategies play an important role since there is a variety of different ways to handle and process data, depending on the data or data-derived information-based products and services.

As described by Wixom (2014) Data Monetization is a wider term for selling practically any kind of information-based products or services. Thus distinct levels of refinement, processing and derivation can be imagined when raw data is first refined into a more structured form of data, further into information and even further to information-based products and services. Simultaneously value is added to the end product or service offered to the customers at every step described. Hence, logically, the more processing is done, the more value is added and the higher costs of processing, the higher expected revenue.

From the studies of Najjar and Kettinger (2013), Wixom (2014), Woerner and Wixom (2015) and Constantiou and Kallinikos (2015) it is possible to distinguish three different layers of data refinement and derivation that refer to the distinct phases and steps that can be taken in terms of data processing and refinement. An essential point here is that these steps or layers are consequent – in order to be able to do data analytics, the data must be first cleaned and otherwise processed. These 3 layers are rather generic, and it may be possible to determine many other, more detailed states. At this point the resolution of three layers is considered to be adequate to introduce the varying possibilities that an organization may have regarding data and data monetization. The different layers are presented in the picture 3:



**Picture 3: Three layers of data refinement and derivation**

As presented in the picture 3, raw data can be refined into well-formed and structured data and furthermore derived into information-based products and services. For some organizations the starting point of Data Monetization may be already the refined form of data – for instance a sample of structured and well-formed data. Thus the organization’s current processes, information systems and data warehousing as well as analysis methods and technologies may steer the Data Monetization business. This means that the organization does not necessarily sell raw data but either structured data or information-based products or services.

### 5.1.3 Packaging and wrapping of data

Besides monetizing data or information-based products or services as they are, according to Najjar and Kettinger (2013) an important value-creating element in Data Monetization is the packaging of data. Additionally Woerner and Wixom (2015) have identified ‘wrapping’ as an essential element in Data Monetization.

These terms differ from each other slightly: in packaging data of different sources, formats or contexts is combined as a package, creating a greater sample or collection of samples that may, for instance, complement each other or enable distinct aspects to the same subject. Furthermore one data product, meaning one package of data, is sold. (Najjar & Kettinger 2013) Najjar and Kettinger (2013) do not determine whether the package contains raw data or instead structured and processed data. Hence it seems that the package may contain data of varying quality or a mix of unstructured and structured data. For instance the same data could be offered in different formats and the package could also contain the original raw data.

The data wrapping, mentioned by Woerner and Wixom (2015), means that data is wrapped and sold with organization's other products or services, usually with its core offering. Woerner and Wixom (2015) argue that wrapping is usually done because of three reasons:

1. Differentiation
2. Increased attractiveness
3. Greater value-creation

Obviously in data wrapping the main product or service sold is not data or information-based product or service wrapped. But rather the wrapped data or derivative is a complementing element that is somehow related to the core product or service and adding value in the eyes of the customers. (Woerner & Wixom 2015) An example of data wrapping would be a software that is complemented with relevant, actual data that may be utilized in this software by the customer. The added value would be created when the customer can utilize the data as a benchmark or complementing customer's own data.

## **5.2 IT and Data Monetization**

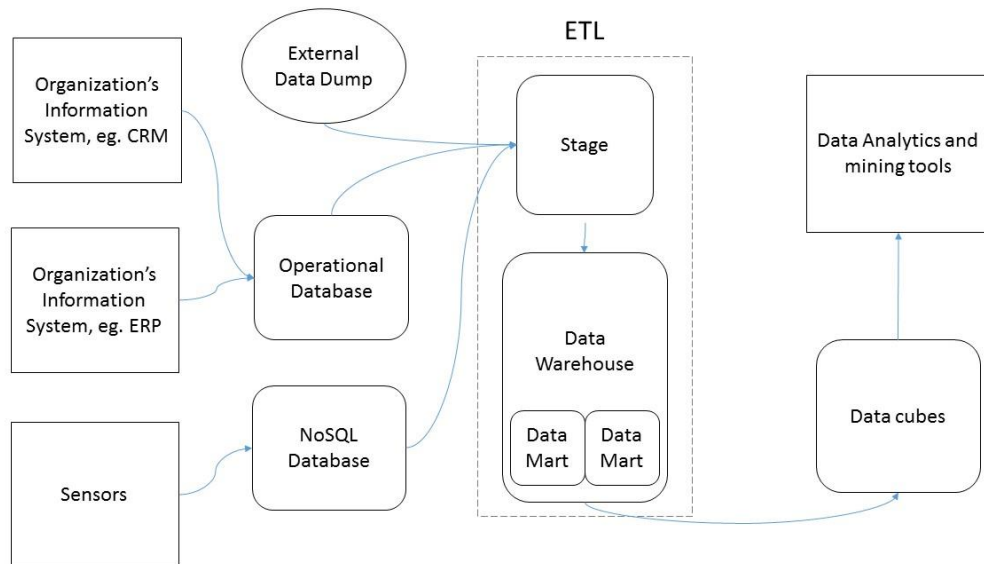
As mentioned already in the chapter 5.1, organization's original choices and investments regarding IT most likely affect greatly on how an organization can plan and execute Data Monetization now and in the future. Although it could be argued that the organizations may make new investments on IT that would enable and support Data Monetization in an ideal way, several studies have shown that IT projects, especially comprehensive ones, are usually costly, time-consuming and require compromises.

Here an important aspect is that the organization may have other forms of business, for example core products, which do not yet involve Data Monetization business. These previously established and already functioning operations may have steered the organization's investments and development in terms of IT – therefore different organizations have varying starting points for Data Monetization. This is one of the key factors affecting and guiding Data Monetization business since the foundation of IT and data management for Data Monetization may have been built on very different vision and strategy. Thus these choices of information systems, interfaces, infrastructure, databases, data warehouses and data gathering also determine the possibilities of Data Monetization.

When an organization starts a new business, establishing new investments may be an initial and necessary step to enable the business in the first place. Although Data Monetization is considered as the generation of revenue with selling data or information-based products or services, it should be kept in mind that none of the world's organizations strives to make operational loss, which means to lose money while conducting Data Monetization. This is true also for the non-profit organizations, which means that Data Monetization business should be beneficial although these organizations do not seek profit.

Therefore the investments made to enable Data Monetization must be reasonable in respect to the expected revenue generation and other benefits. Otherwise Data Monetization would not be economically sensible option. Thus Data Monetization is rather heavily dependent on the existing IT solutions in the organization. These solutions also determine what kind of data or information-based products or services already exist and can be refined and derived.

As discussed previously, data must be stored somewhere and it must be maintained in order to tend the data quality. Additionally if the organization wants to process the data further, data cleaning, transformation, data analysis and data mining technologies may be needed. Hence Data Monetization may require one or more of the components presented in the picture 4:



**Picture 4: Different IT components in Data Monetization**

The particular components shown in the picture 4 are an example of the essential components needed in Data Monetization since there are several distinct technologies which to choose from. For example with Big Data distributed processing and storage systems as well as stream processing are preferred which means that for example Apache Hadoop and its tools and services such as Apache Hive, HBase and Spark are preferred over the traditional options, such as relational databases, physical servers and batch processing. (Chen et al. 2012) This view is also supported by Davenport et al. (2012), who argues that IT automation is necessary in the future due to the increasing volume, speed and complexity of data.

Still the basic idea is similar than in the picture 4 – data is gathered from somewhere, combined, cleaned, transformed and loaded to another storage. Depending on the data and purposes, data could be processed even further, for example with the methods of data

analytics or machine learning. At this stage the organization would be monetizing information-based products or services rather than just data. Hence if the organization wants to offer highly processed information-based products or services, highly sophisticated IT components and aforementioned methods are also needed.

The extent and need of IT components depends on the organization's data and its Data Monetization business' products or services offered to the market. Indeed, if the organization wants to monetize raw data, for example sensor data, it does not necessarily need ETL process, data cubes or any processing tools. Thus only a data storage for this particular data is needed, for example a NoSQL database. This is an important observation highlighting the diversity of practices and means that may be associated with Data Monetization. Moreover the organization, its IT capabilities and strategy prominently affect the starting point and furthermore the future's opportunities related to Data Monetization in terms of the technical aspects of data, data processing and the end product or service offered.

Additionally some channel for the purchasing and sharing must be offered, otherwise no sale transaction would occur. Thus the data must be stored into a suitable storage from which it can be shared or copied to the customers efficiently. As, for example, Davenport et al. (2012) notes, with Big Data the cloud-based platforms are tightly related to the processing and handling of data. Moreover the offered data or information-based products or services and organization's data storages and platforms used are essential factors affecting the availability and the choice of distribution channel in Data Monetization. The organization may, for instance, share the data through its platform, offering simultaneously platform-services to the customers. Another option would be to offer data in batches that can be simply downloaded to customers' own environments. As already mentioned by Najjar and Kettinger (2013) as well as Woerner and Wixom (2015), packaging and wrapping are rather commonly used methods in Data Monetization. In spite of distinct methods related to and forms of Data Monetization, the IT capabilities and infrastructure constructing the basis for purchasing data or information-based products or services must be also thought through, invested in and implemented.

### **5.2.1 Data Privacy and Information Security in Data Monetization**

When discussing about sharing, analyzing and mining data, data privacy and information security are topics that raise questions and for a good reason. The organizations' data storages may contain sensitive data that must be secured appropriately, not only because of ethical reasons but in several countries the legislation reinforces data privacy and confidentiality. Thus organizations that deal with Data Monetization must be aware of and take responsibility for data privacy and information security. Information security's two important features are the following:

1. Prevent unauthorized access to the data, databases, computers and websites (Culnan & Williams 2009)
2. Protect from corruption and destructive forces
3. Ensure integrity and availability (Zissis & Lekkas 2012)

Data privacy and information security have two important aspects regarding Data Monetization. First of them is that certain sensitive data may be illegal to sell or share and furthermore the actual definition of sensitive or private data that cannot be distributed is determined by the existing legislation. This kind of sensitive data could be, for instance, personal data containing one's name, address and social security number. Another aspect is that if some sensitive data exists in organization's databases, organization must protect this data and ensure its privacy and security. (Culnan & Williams 2009; Gerlach et al. 2015) According to Zissis and Lekkas (2012) this concerns also situations where sensitive data can be shared with permission of the person in question, which means that the process of data sharing must be secured. Thus the platforms, infrastructure and the actual transfer in question must be appropriately secured as well.

### **5.3 Investments in Data Monetization business**

As discussed previously in the chapters 4.3 and 4.4, organizations are pursuing to generate revenue and, in the end, profit or other benefits from Data Monetization. In other words, Data Monetization business has to be beneficial to the organization, creating and offering value to it – otherwise organization would not jump into Data Monetization business. Due to the basic surmise of and motive for establishing business, the investments planned and made must be in balance with the expected and realized generation of revenue. Therefore the expenses and investments faced may also steer the decisions related to the offering, business model and strategy in Data Monetization.

For some organizations Data Monetization does not oblige great investments as they may already have the necessary IT in place and monetizing data might only require some re-allocation of current workers. In this kind of situation Data Monetization is not definitely organization's first business and it has had other products or services previously. For some other organizations Data Monetization may be the first business attempt ever. It is also possible that the Data Monetization business may require attention and investments regarding data gathering and storing, data processing capabilities and ability to initiate data wrapping, for instance. Thus these organizations described above have very different starting points for Data Monetization business, reflecting also the way how an individual organization wants to execute Data Monetization.

The barriers of entry is one widely recognized force or factor in the economics field of market dynamics that is described by, for example, Porter (2008). According to Porter (2008) the requirement of capital is one of the sources of barriers to entry and indeed investments lie in this category. Furthermore investments may be an obstacle to be able

to start Data Monetization, especially in case where the organization's core businesses and already existing technologies do not support the initiation of Data Monetization business. When considering the fundamental process that concludes to the revenue generation, support in terms of investments may be needed in several fields, such as IT, R&D, marketing, distribution and sales. Probably IT is the major factor for two reasons: first of all Data Monetization is IT intensive business and secondly IT investments are usually enormous compared to the investments related to, for example, sales operations. And as described in the chapter 5.2, several IT components may be needed in Data Monetization.

Depending on the organization, Data Monetization may be the core or non-core business and this business orientation is likely to be reflected on the magnitude of investments made for Data Monetization. Therefore if considering the core business or an organization, the investments made for this trade are probably greater than those for the non-core business activities. After all the investments must be covered by the business transactions. As already described in the beginning of this chapter, Data Monetization business is likely to be, at least initially, a non-core business for organizations that gain, produce and refine data as a secondary product from the core business activities (Najjar & Kettinger 2013; Constantiou & Kallinikos 2015; Woerner & Wixom 2015) Thus the original core business is playing a key role in terms of the initial capabilities and starting point of Data Monetization for a majority of organizations, offering the information systems, infrastructure and technologies to initiate this new business and hence giving a great advantage compared to the emerging organizations with Data Monetization as their core business.

## **5.4 Business models of Data Monetization**

The business models and strategies that may be applied in Data Monetization are studied and furthermore the different paths taken to generate revenue with data are considered. Additionally in this chapter the diverse phenomenon of Data Monetization is present which means that no single, definite and unambiguous business model or strategy can be pointed out. Since there are a great deal of distinct ways to execute and establish Data Monetization business, the distinct business models that can be implemented are studied and considered on a general level. Clearly the ultimate goal of Data Monetization is to generate revenue to the organization, thus setting the objectives for Data Monetization business.

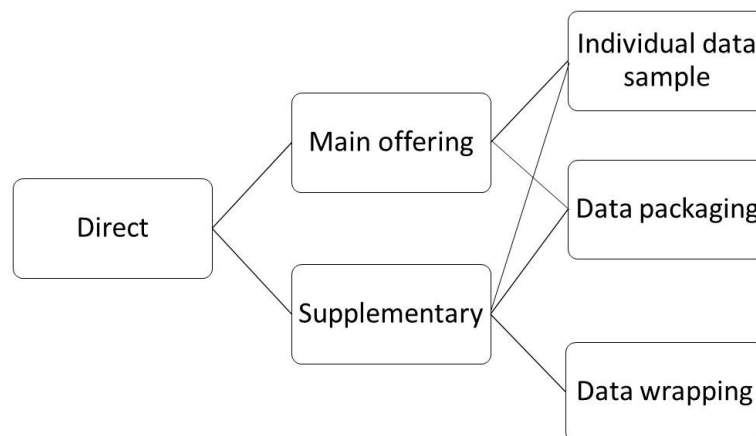
Gartner's (2015) view on dividing Data Monetization into direct and indirect way of execution is used here to study further the distinct possibilities to generate revenue with data. Hence it is essential to distinguish direct and indirect Data Monetization from each other once again: direct Data Monetization means selling directly data or refined data products to the customers whereas indirect Data Monetization stands for selling products or services of some value that are derived from data. (Gartner 2015) Despite the definition of Data Monetization takes into account information-based products (Woerner & Wixom 2015), this form of Data Monetization does not seem to fall into the category of direct or



indirect Data Monetization unambiguously. Whereas information-based services are clearly derived from data and cannot be considered as direct Data Monetization since something else valuable besides solely data is sold, information-based products are a different story. As defined by Woerner and Wixom (2015), information-based product can be a set of raw data or result of analysis but it may be also a design. When considering raw or enhanced data, information-based product falls into direct Data Monetization, but for instance a physical product that is designed based on data is already indirect Data Monetization. Thus this distinction is essential when considering Data Monetization models and strategies through Gartner's (2015) partitioning of direct and indirect Data Monetization business.

### 5.4.1 Direct Data Monetization

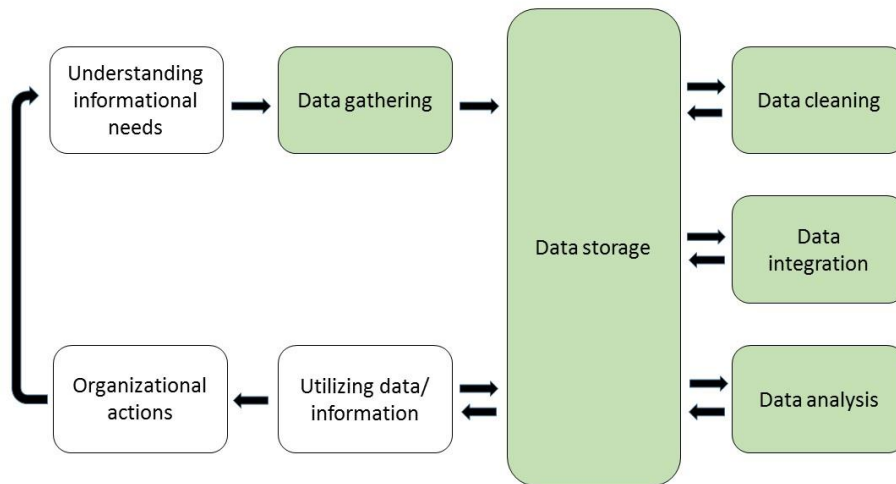
As described by Gartner (2015) the direct Data Monetization is a rather straightforward approach on how an organization can generate revenue simply by selling data. Data wrapping and data packaging are also methods of direct Data Monetization. Still, an organization may also sell just unstructured or structured data to the customers without supplementary value-adding products. In this kind of business data is playing the key role, whereas in data wrapping and in some cases of data packaging data may be the value-adding supplementary (Najjar & Kettinger 2013; Woerner & Wixom 2015). The picture 5 depicts the different offering options of direct Data Monetization in a rather general level:



***Picture 5: Distinct products of direct Data Monetization***

As shown in the picture 5, direct Data Monetization can be established by offering data as a main product or a supplementary product. The type of data offered may vary from unstructured raw data to highly sophisticated results of data analysis in all of these option

in the picture 5. Nevertheless the key is in offering data or data refinements to the customer. The different levels of derivation and refinement of data were presented in the picture 3 and discussed in the chapter 5.1.2. Thus the data sample, package or wrapping may contain anything from raw data to results of data analysis. When considering direct Data Monetization from the perspective of value creation framework introduced in the chapter 3, value is created or added in the rather beginning of the process. These steps are illustrated and highlighted in the picture 6:



**Picture 6: Value creation in the direct Data Monetization business**

As presented in the picture 6, the value is created mainly in the phases of data gathering, data storing and data processing in which data may be refined and analyzed into more sophisticated samples or result sets, but yet no indirect derivations exist. These essential mechanisms of value creation are particularly important from the point of view of generating revenue since, as previously mentioned, the value of product or service is usually measured by its market price. This means that the more valuable and non-imitable an information-based product is, the greater amount of revenue it will likely generate (Grant 1996).

In the direct Data Monetization the value is highly dependent on the data or information sold to the customers as a primary or a supplementary product since, after all, there is not much else to offer for the customer. Hence all the value adding or value creation aspects relate to the data, data's content and context as well as organization's data processing and storing capabilities. This view can be parallelized with a tangible product, to elucidate the essence of this perception. A tangible product, for instance a chair, is valuable when it is usable, beautiful, comfortable and durable. Moreover if the chair is broken, uncomfortably designed or otherwise inappropriate, a customer might not appreciate and value it, nor would have purchased it.

As described previously in chapter 5.2., part of Data Monetization business is to be able to distribute and share sold data to the customer. This aspect is the first step of the value creation framework that is not considered as a value adding or creating activity. The reason for this is that the distribution channels must exist to enable the business and the customers expect to get the product or service through some channel. Therefore distribution or sharing of the information-based product is not considered as a value adding activity in direct Data Monetization and hence the value creation activities are specifically limited to data and data processing, not to the other aspects and operations related to Data Monetization transactions.

One option to enable the distribution and sharing of data is to offer a platform as a supplementary service that the customers can use to handle data and from which they may download all or part of the data if they wish to. This kind of model is called PaaS (Platform as a Service) (Davenport et al. 2012). Some sort of platform must exist in any case to enable Data Monetization in terms of data warehousing and, in some cases, data processing. Such platforms exist already and several large vendors offer PaaS as part of their business portfolio: for example Microsoft Azure or IBM Bluemix could be utilized in Data Monetization. When considering the picture 5, PaaS could be combined with all the main and supplementary offerings depicted. Still, PaaS is not the only way to execute and establish Data Monetization, but one choice of infrastructure and an option of business model.

Partially because of the type of offering in the direct Data Monetization, the continuity of business can be seen as an essential factor affecting the strategy and business model taken. A logical conclusion would be that once a customer purchases a set of data, the customer will not purchase the same data set again. Therefore the direct Data Monetization is more likely to be a supplementary business. This is also supported by Woerner and Wixom's (2015) view on how the amount of companies selling tangible products and engaging in Data Monetization is increasing. Otherwise the organization must be able to develop and create continuously new and interesting data and data packages that customers are willing to buy.

The previously described seems to be a definite challenge since discovering and creating new data without other business lines that would generate the data is nearly impossible. An option would be to either buy this data initially from some other company and to refine and perhaps analyze it in order to come up with some differentiated and valuable data, or to utilize some open source data that is available to everyone and that could be potentially used as a source of revenue without neglecting the property rights associated with this particular open source data. Both options are arduous and especially risky – the data bought or gathered can be utilized and sold by the competitors and thus the original data set is by no means a factor of differentiation. This is the case with both individual data sets and packaging of data sets to larger wholes. Therefore it is more likely that the direct Data Monetization exists as a supplementary business that utilizes the data available from

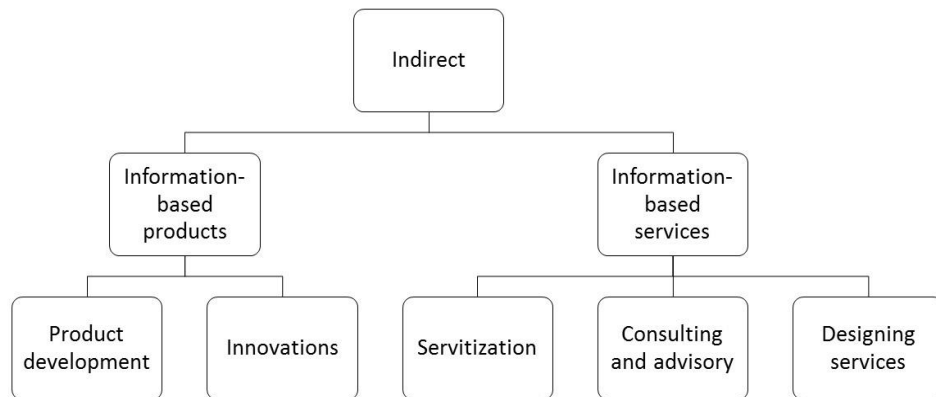
the core business lines and thus creates new revenue streams to supplement the organization's business line portfolio and profitability.

Whereas the direct Data Monetization as a core business may not be the most riskless, continual and common form of trade, the supplementary direct Data Monetization may be a great way to gain additional revenue and truly benefit from the organization's data reserves (Najjar & Kettinger 2013; Woerner & Wixom 2015). As mentioned by Woerner and Wixom (2015), for instance Data Wrapping is a rather simple method to utilize the possessed data to add value and differentiate from the competitors. For example earlier described platform as a service (PaaS) business model may benefit from data wrapping when offering data sets and samples to the users, to be utilized and perhaps analyzed on the platform.

As depicted in the picture 5, supplementary business may associate with offering individual data sets, mix of sets in data packages or wrapping data with core products or services. With all of these supplementary business lines the strategy seems to be to offer added value to the customers and simultaneously create new revenue streams to the organization. Depending on the competencies and capabilities an organization may clean, process, mine and analyze the data to develop more refined forms of information-based products. In the supplementary direct Data Monetization business continuity is not likely to be as enormous issue as it seems to be with direct Data Monetization as a core business since data is not playing the key role. Thus smaller amounts of new and differentiated data is needed and the integration of data and core products or services is already a value adding feature. Hence the supplementary direct Data Monetization is a more riskless business model that offers seemingly easier revenue generation possibilities to the organizations that possess and perhaps also create new data themselves.

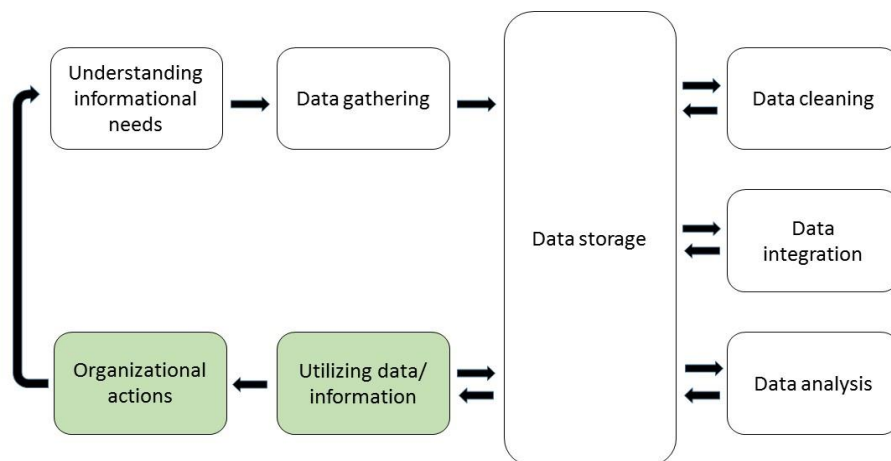
## **5.4.2 Indirect Data Monetization**

The indirect Data Monetization is probably a more common form of Data Monetization since it can be considered as a hypernym for selling a great deal of distinct information-based products and services that are derived from data. Additionally servitization with data described by Oprasnik and Taisch (2015) can be considered as one, more concrete form of indirect Data Monetization, thus including also the shift and transformation of businesses from product manufacturing towards service production. The indirect Data Monetization seems to be a rather comprehensive phenomenon since it may be associated with several distinct trades of business, whether it is manufacturing and product development or consulting and process design. The extensive and diverse nature of indirect Data Monetization business is illustrated in the picture 7:



**Picture 7: Different offerings of indirect Data Monetization**

As shown in the picture 7, indirect Data Monetization business may be divided further into the offerings of information-based products and information-based services. Information-based products and services may consist of distinct sets of offerings, such as R&D or consulting services. Thus the original data that was utilized to create the information-based product or service may be still visible, as in case of R&D, or it may be very distant from the derivatives. When considering indirect Data Monetization through the value creation framework, an organization adds and creates value in the latter phases of the process, presented in the picture 8:



**Picture 8: Value creation in the indirect Data Monetization business**

Since in the indirect Data Monetization the organization's offerings consist of the refined and derived forms of information-based products and services, the value is created and offered to the customers based on the phases of utilization and organizational actions. Before these steps the Data Monetization would not be indirect, but rather direct and therefore these steps are not considered as value adding or creating actions, although the original data and information that build the basis for indirect Data Monetization business are playing an important part in terms of the success and quality of the refinements and derivatives. Here the step of utilization of information is juxtaposed with the development of the indirect information-based products and services. This would be the case, for instance, when innovating a new product based on some original data and information. The step of organizational actions can be also considered as a phase that creates and adds value in the eyes of the potential customers due to the possible development of organization's business practices and procedures. Indeed an organization may create new business models and ways to satisfy the customers, especially in the case of information-based services.

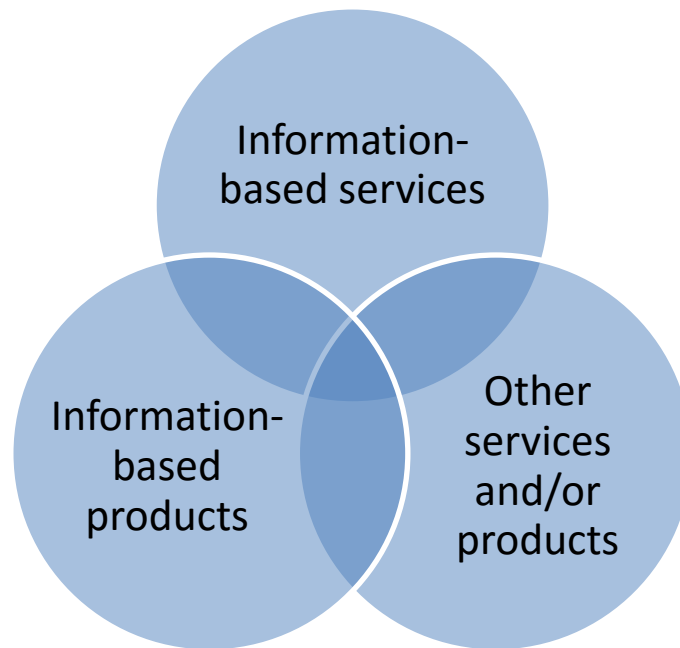
When compared to the direct Data Monetization, the indirect version of the phenomenon is more likely an organization's core business. Indeed, not many real-life companies are focused on selling mainly data, but majority of companies consider consulting or R&D as their core competence and thus they offer such services as the core business. Additionally the indirect Data Monetization consists of wide range of different ways to execute business, and unarguably there exists more variation and options in the indirect Data Monetization than in the direct Data Monetization.

Still, although the indirect Data Monetization's distinct trades of business are more suitable to be executed as core business due to the features such as easier scalability and lower risk, the indirect Data Monetization can also be established and executed as a supplementary, non-core business. An example of indirect Data Monetization as a non-core business could be the design, maintenance, development and operations services of a factory or other product manufactured by the organization. Similarly the R&D and development of new products that are furthermore sold can be seen as a form of indirect Data Monetization. Therefore the alternative of executing non-core indirect Data Monetization business seems to be very common as well. This can be supported by, for instance, Oprasnik and Taisch's (2015) views on servitization.

According to Oprasnik and Taisch (2015) as well as Woerner and Wixom (2015), a manufacturing company can expand its business rather easily with the indirect Data Monetization and information-based services. In the case of servitization the organization initiates one or more non-core service businesses based on its data and information that are intrinsically related to the organization's core business. Thus this new trade of business is rather effortless to establish and does not necessarily require a great deal of resources,

new capabilities or investment. The information-based services can be offered to the customer as a core business as well. Nowadays there exists plentiful of companies focusing on, for instance, design, consulting and advisory services.

Regarding the direct Data Monetization Najjar and Kettinger (2013) identified data packaging as an alternative business model of Data Monetization. Additionally Woerner and Wixom (2015) described a method of data wrapping that is another option to execute direct Data Monetization. Similar business models can be applied in the indirect Data Monetization as well. An organization may create packages that contain a mix of products and/or services or wrap, for instance, advisory or maintenance services around a product. As with data wrapping and packaging, applying these methods with the indirect Data Monetization business may add value and increase offering's interestingness and attractiveness, hence enhancing the organization's core and non-core business lucrativeness. The packaging or wrapping can be constituted in a varying of ways as presented in the following picture 9:



***Picture 9: Indirect Data Monetization and wrapping***

As shown in the picture 9, the wrapping or package may be a combination of information-based products and services or an information-based product and/or service combined with organization's other products or services that are usually, according to Woerner and Wixom (2015), Najjar and Kettinger (2013) as well as Opresnik and Taisch (2015), organization's core offering and business. Thus in the package an information-based product or service can be organization's core or non-core business, depending on the organization's other business trades and furthermore its overall business strategy. This also supports the view of seeing Data Monetization business and especially the indirect Data Monetization business as a diverse and complex phenomenon.

## 6. ISSUES RELATED TO DATA MONETIZATION

Although Data Monetization may offer a great deal of opportunities to the entrepreneurs and enterprises, several issues and questions related to Data Monetization exist. For instance Davenport et al. (2012) as well as Boyd and Crawford (2012) have noted such challenges or drawbacks that relate mostly to data and Big Data but are also associated with Data Monetization. Hence it is necessary to consider all the distinct aspects and qualities of Data Monetization in order to understand this phenomenon and the phenomena related to it. This is important both from the perspective of organizations that would like to establish and run Data Monetization business as well as from the perspective of this study.

Especially an organization planning and actualizing Data Monetization should be aware of both the positive and negative features of data and Data Monetization since these obstacles, challenges and issues may affect the revenue generation of organization and furthermore steer the Data Monetization strategy and practicalities. Additionally some of the issues, such as ethical challenges related to Data Monetization, may affect organization's image and hence either enhance or disrupt the organization's competitiveness in the market.

The previous chapters have already discussed the technical aspects related to data quality and data and information security. Although especially data and information security is an important aspect that should be considered carefully and extensively while planning and executing Data Monetization business, at this point of study these viewpoints will not be repeated but other issues and obstacles are discovered and introduced.

### 6.1 Ethics, data privacy and social networks

Ethics of Data Monetization is yet a rather new topic due to the novelty of Data Monetization itself. Still, it is likely to be a subject of great debate in the future, just like social media, data gathering and Big Data are nowadays (Boyd & Crawford 2012). The ethicalness of collecting, possessing, sharing, selling and transferring so called secondary data indeed unsettles people, perhaps since people are not sure what data is collected, how it is stored, who knows about it and could it been used against them. This issue is primarily focused on data related to people (Culnan & Williams 2009; Gerlach et al. 2015), and logically for example data about factory machinery is not tightly associated with ethical issues since human beings are not related to this type of data, although information security is an important aspect related to all types of data.

Indeed it seems that at least in most of the western countries the legislation steers the previously mentioned actions and especially the data privacy and information security to



some extent but considerable differences exist, for instance, between EU and United States. Moreover it is common that the particular legislation has to be obeyed only if the data is located in this particular country or region – if a person shares on one's own initiative his sensitive data to a party that is located outside this region, the legislation will not be standing. (Culnan & Williams 2009; Weber 2010) This is a very timely topic of discussion especially related to the cloud computing and platforms while several organizations offer distinct PaaS (Platform as a Service) options for their clients (Zissis & Lekkas 2012). Still, in some cases the legislation regarding Internet of Things and, for instance, sensor data is separated from the legislation steering data privacy related to personal information and social media (Weber 2010), thus leading to a more complex and diverse legislative field related to Data Monetization business.

When considering Data Monetization, Gerlach et al. (2015) have studied relationship between companies' privacy policies and users' willingness to share their data and information in the social networks. This is an essential perspective when Data Monetization business relays on monetizing the secondary data of users. More specifically the secondary data stands for data that is *“used for personal information for other purposes subsequent to the original transaction between an individual and an organization when the information was collected”*. Thus an example of data monetization with secondary data would be an application which may monetize users' pictures that have been uploaded to the service. (Gerlach et al. 2015) Weber (2010) points out another type of secondary data – nowadays and increasingly in the future IoT (Internet of Things) is a source of vast amount of data that could be potentially monetized. This data may be collected with a great amount of different sensors and gadgets, for instance with the RFID tags, microphones, movement detecting sensors, GPS and cameras. And while the amount of IoT products grows in the households, more and more data and information is collected of individuals and their habits. Still the client privacy and anonymization is an essential act and aspect related to data collected through IoT. (Weber 2010)

According to Gerlach et al. (2015) there exists a tradeoff between users' willingness to share data and content of company's privacy policy – nowadays people are more aware of privacy policies and more concerned about what may happen to their data and information. Still the privacy policies are usually hard to comprehend and furthermore accepting the policy is an only way to access the application or service behind the privacy policy agreement (Gerlach et al. 2015). Therefore it is not uncommon that many people do not read the content of policy and furthermore agree with the organizations' monetization practices with secondary data. In case of IoT and, for instance, a television that may collect data also during it is shut down, the aspect of privacy policy is much more complicated. Indeed if the gadgets are said to collect data in order to enhance the customer experience, it is explicitly challenging to determine for sure what kind of information and when is collected and transmitted by the product, how this data is processed and who may use it and for what purposes. And when an individual is purchasing such a product it may

be almost impossible to get to know these details beforehand. Furthermore, legally, the agreement is settled when the customer purchases the product and starts to use it. (Weber 2010)

Especially Facebook, Instagram and Google have been targets of discussion since, naturally, these organization collect a great deal of data from their websites and applications (Gerlach et al. 2015). It seems that people are scared of their data and personal information to be possessed by a profit-seeking company, although they may have given the information by themselves to such companies. Culnan and Williams (2009) argue that organizations should have and take a moral responsibility of individuals' data and privacy policies that would comply with law rather than to try to discover a loophole. The point of view of Culnan and Williams (2009) concentrates more on the data storage and utilization internally, which can be parallelized with data gathering, development, refinement, analyzes and derivatives that may occur in Data Monetization business. This is an important aspect on Data Monetization as well, since information security and data privacy must be taken care of while sensitive data is processed to enable the monetization transactions. Any issues with the internal data privacy practices may have as serious consequences as if one fails to follow data privacy legislation and agreed ethical guidelines in the monetization transactions. Culnan and Williams (2009) as well as Boyd and Crawford (2012) argue that previously mentioned issues may damage the reputation of organization and furthermore lead to problems with the legislative and governing bodies.

In general the ethical issues related to Data Monetization are associated mainly on the business model which focuses on the utilization of secondary data. This is rather logical since in this context secondary data stands for personal information shared by individuals in exchange for service or use of some application and thus ethics is tightly related to humanity. Gerlach et al. (2015) argue that the role of privacy policy and monetization practices with secondary data may have a significant impact on the adoption of services which collect and use secondary data. Moreover the users' perception of privacy policy may initiate a behavior of non-adopting the service or application because the "pricing-by-privacy" business model and its practices terrify potential users' more than the service appeals (Gerlach et al. 2015). Furthermore Culnan and Williams (2009) mention the societal aspects on data privacy, data collection practices and what data is ethical to monetize. Naturally each organization self-regulates its ethicalness and values to some extent, and indeed data privacy and the means of business are one aspect to be considered of.

Hence the privacy policies are indeed an important aspect and even an issue in Data Monetization business – not only due to the restrictions set by the legislation but also because these policies may affect the amount and quality of data gathered as well as the business' reputation in the market. Therefore the company's business model and data handling strategies should be adjusted carefully. But in the end, it is likely to be very challenging to balance the ethics, moral and legal aspects with the eagerness to generate revenue and to keep up with the competitors.

## 6.2 Cultures and eligibility globally

Data Monetization markets are global, especially when considering data and information-based products. Although data can be purchased from the other side of the world, it should be kept in mind that the data must be suitable and eligible on the distinct parts of globe. This issue relates to both data quality and context. When the context changes, for example when the data is bought from the other continent, it may not be applicable and useful anymore. (Strong et al. 1997) The reason may be cultural or geographical – for example car maintenance data from India could not be utilized and applied in Finnish car repair shops since the cars in India may be different due to different legislation and the environment where the cars are driven is very different as well.

In terms of quality the data format should be such that it can be utilized in different parts of world with ease. It seems to be a commonly accepted assumption that data is offered in different formats, for instance when downloading the data or saving the data on computer. Distinct data tools and software usually offer a wide range of formats, such as .csv, .json, .xml and .txt and furthermore distinct data storage as well as data source options may be supported. Although most of the tools and software support more than one option of each, for instance Microsoft's Excel can fetch data from several sources such as different databases, cloud platforms, queries and data feeds, a potential customer may still wish to have more than one option of data format and data transfer in order to be willing to commit to the transaction. This may be due to ignorance, desire to stay with a familiar technology or simply because no extra work regarding data processing is wanted.

Also other format-related issues may occur when dealing with data. For example different units of measure may vary across the globe and although the global standard of SI units would be followed, not all companies and citizens are capable or willing to utilize solely SI units. A great example of this issue is the unit of measurement of temperature: the SI unit for temperature is Kelvin (K) that is not used in everyday life in any part of the world. Moreover the most commonly used and accepted unit is Celsius (C). Fahrenheit (F) is the official unit in the United States and in some Caribbean islands (Zimmermann 2013).

The format of data may be an obstacle when considering the usefulness and usability of data, which in turn affects the potential customers' valuation of the information-based product and furthermore their willingness to buy it. As argued by, for instance, Zeitheml (1988) and Setia et al. (2013) the product or service bought is usually something that fulfills a need of an organization. The issues related to the data format will probably repel the potential customers since the information-based product or service cannot be utilized right away and some extra work regarding data processing and formatting is needed in order to fulfill the organization's needs.

Another fairly similar obstacle or issue may be the language of data. Whereas most organizations in the world are capable of and willing to use English as their working language, still some people, companies and cultures are rather opinionated and willing to utilize only their mother tongue or the official language of this particular country as the procedural language. Nevertheless most of the organizations around the world may be willing to use English as the lingua franca, information-based products and services are likely to be utilized internally and thus they may usually reflect the working language of the company. Although globalization may have affected this issue already, still all of the organization are not willing to purchase information-based products or services with a foreign language – the reason may be in lacking language skills or simply in opinionate attitude and old habits.

Offering an information-based product or service in several formats, in several languages and with several data transferring options may be appealing due to larger target group globally. But the disadvantage of such strategy is that it requires more work on developing these distinct options and furthermore increases the cost of production and storage that is very likely to eventually affect the pricing. Higher price, in turn, may damage the competitiveness of offering and the customer's willingness to make the purchase. Clearly not being able to offer a suitable product or service in a suitable form is an issue itself, but finding a balance between bare minimum and oversupply may be another challenge that an organization in Data Monetization business must seek an answer for.

### **6.3 Ownership of data**

The ownership of data has been a topic of great interest and debate especially since the origin of Big Data and cloud computing (van Belle & Ruiter 2014; Zissis & Lekkas 2015). Nowadays this debate is mainly related to loading data into the cloud, which means that this particular data is stored in one or several servers most likely out of the country (Subashini & Kavitha 2011; Zissis & Lekkas 2015). The cloud computing and social networks are indeed a field of interest in terms of data ownership since both of them are more or less on a shady area and no international standards are yet applicable. Van Belle and Ruiter (2014) have took a rather traditional viewpoint on this issue, through the legislation and copyrights related to data whereas Zissis and Lekkas (2015) have considered data ownership issue from the point of view of cloud computing deployment models.

When considering Data Monetization, data ownership is an interesting aspect at two different points of the process. First of all the data collected or bought may raise questions of the ownership of data. This kind of issues arise especially when secondary data is utilized. The second possible encounter with data ownership issue is when Data Monetization transaction occurs. If data or data derivatives are sold as a core business or in the form of data wrapping, data ownership may become a significant point of view. Especially if the customer is trying to benefit economically from Data Monetization organization's data, for example, by re-distributing it. Thus data ownership is an issue that may

not just create an obstacle for Data Monetization business but also lead to post-transactional issues.

According to Zissis and Lekkas (2015) the deployment model selected in the cloud environments is significant in terms of data governance and ownership since, depending on the contract, the governance of data and different applications may be outsourced and hence the concrete control on them is delegated to another party. When considering a public cloud, which means a shared, virtualized cloud that is accessed usually via the Internet, the most common practice seems to be that the control over data is mitigated to the owner of infrastructure (Zissis & Lekkas 2015). This, on the other hand, means that the original owner of the data must simply trust the processes and technology provided by the cloud and infrastructure owner.

Although it seems that the data governance may be outsourced to the cloud provider, more questions can arise when considering the cloud environments and PaaS in the long run. Subashini and Kavitha (2011) has discussed about the possible issues and indeed interesting questions regarding whether the original data owner is able to take the data with him if, for instance, the original data owner considers changing to another cloud provider. And if data can be taken from the cloud environment, it may be unclear whether the cloud owner still has copies of the files, documents and data as well as whether this is acceptable or not. (Subashini & Kavitha 2011) The issues described previously raise also questions about the ownership of data – if the data could not be taken completely away from the cloud environment, one may argue that the cloud owner is the owner of data. According to both Subashini and Kavitha (2011), van Belle and Ruiters (2015) as well as Zissis and Lekkas (2015) the privacy, control and ownership issues cannot be solved afterwards unambiguously but such issues can be avoided in the first place by applying tight and comprehensive Service Level Agreements (SLAs) and contracts. Another option would be to choose a private cloud in which the governance and control is not outsourced to the third party.

Similar aspects must be considered when an organization is creating plans about the environments and possible utilization of cloud computing. Additionally the ownership of data in cloud environments may be parallelized with the data privacy issues described in the chapter 6.1. since in the similar manner a user may load and generate data in the organization's environments. Still, according to van Belle and Ruiters (2014) the legislation related to data ownership and the copyrights is varying across the world and furthermore no widely accepted and agreed perception is yet formed. This may also distort the competition in the field of Data Monetization if, for instance, data privacy and ownership are interpreted differently in distinct countries and thus some organizations are able to sell information-based products that are illegal to distribute in another country. Furthermore these issues may be avoided by moving the servers to another region or favor cloud environments that are based on such areas.

At the moment it seems that most of world's legislation accepts the analysis of already published data or creative use of a subset of data which means that no separate permission is required (van Belle & Ruiter 2014). But still an enormous amount of variation is present what comes to the ownership of data and how one may use data published on the Internet. For instance commercial use of some sort of data may not be accepted in one country whereas in another country the laws regarding the copyrights are much more lax. (van Belle & Ruiter 2014) And as pointed out by van Belle and Ruiter (2014) the definition of data is also varying in eyes of the legislative bodies, thus shaping the perception on data ownership and copyrights even further. Nowadays there is not one and only truth and answer for this issue but rather a great deal of diverse opinions, options and interpretations.

Although data ownership is obviously a topic that might require a vast focus of researchers, the organizations executing and planning Data Monetization should be aware of issues related to data ownership and furthermore be aware of the significance of adequate SLAs and terms of agreement to be able to secure its own control towards the data and information-based products and services as well as to be able to protect its business against exploitation. By generating the data itself, an organization may avoid some issues and unambiguous situations but this does not necessarily protect the business of unwanted re-distribution of data or data refinements. Hence the role of awareness and preventative actions are essential in establishing a healthy and growing Data Monetization business.

## 7. CONCLUSION

This chapter will revise the findings of the study from the perspective of depicting the answers to the sub research questions and, ultimately, to the main research question. While proceeding this study's achievements will be described and the evaluation will be conducted. Finally the suggestions regarding further research on Data Monetization will be described.

### 7.1 Findings

The focus and purpose of this study was to recognize the distinct aspects and factors of and effecting on Data Monetization and ultimately to understand how an organization can monetize data. The framework of value creation with data was built in order to support and structure the study and furthermore to reflect the systematic literature review's findings against Choo's (1996; 2002) and Thierauf's (2001) fundamental views regarding the role of information in organizations and moreover the value created in the process of knowledge creation.

After initiating the systematic literature review, it was evident that no commonly accepted definition for Data Monetization existed in the researched literature but the terminology used varied and the definitions had different nuances. Ultimately some of the definitions were contradictory and vague. This study derived and discovered a rather comprehensive and rationalized definition for Data Monetization from the point of view of Data Monetization business: "*Data Monetization is revenue generation with and out of data and data-derived and information-based products and services.*" By justifying the financial aspect of considering Data Monetization as generation of revenue instead of profit enabled a wider perspective on the topic and furthermore included non-profit and public organizations into Data Monetization as well. It also turned out that Data Monetization is a hypernym for distinct offerings and ways to generate revenue which means that Data Monetization is not limited only to selling data or data sets even though the terminology might suggest so. The definition, and the chapter 4 in general, offer an answer to the first sub research of this study: "*What does Data Monetization mean?*"

As stated by the main research question, the main findings of this study focus on the distinct factors affecting and enabling Data Monetization. Since Data Monetization is, evidently, a complex and diverse phenomenon and concept, there exist different factors and aspects related to and affecting Data Monetization. These views were studied in detail in the chapter 5 and also this chapter pursued to answer the following sub research questions: "*How may an organization generate revenue out of data?*", "*How can an organization discover and recognize its valuable data?*" and "*What kind of strategic directions*

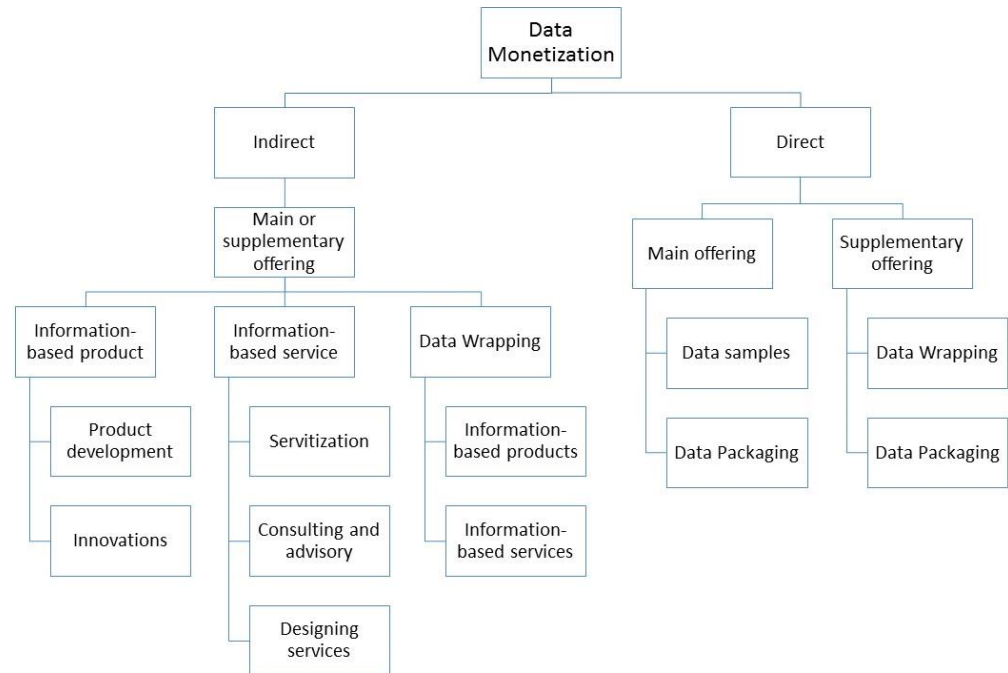
*organization may take in Data Monetization?”*. These answers are also summarized in the following verses.

The key of Data Monetization is, essentially, the data or data-derived products or services that are the source of generated revenue when they are sold or traded. Furthermore the type of information-based product or service affects significantly to the business model deployed and thus the choices of strategy and revenue generation logic. Additionally the content of data and its value affect greatly the success of Data Monetization and the potential amount of revenue generated with data since, ultimately, the higher the customers valuation is, the higher revenue will be gained. On the other hand it is important to notice that data processing and storing are rather costly operations and hence an appropriate business model and plan must be created since the investments and expenses must be balanced with the expected and realized revenue generated. The investments in terms of time, resources and money may vary across the organizations pursuing Data Monetization, depending on their approach and strategic choice. The organizations that are complementing their business portfolio with Data Monetization are not likely to face as high barriers of entry in terms of investments, IT and other competencies as the freshly founded organizations that do not have any core business on the background.

In Data Monetization business data, information-based product or service can be the main or supplementary offering. Moreover this study distinguished between the direct and indirect business models. The difference between them is that in the direct business model data and data sets are sold whereas in the indirect business model data is refined into information-based products and services that are, in turn, sold or traded. It was discovered that indirect Data Monetization is the more common form of business compared to direct Data Monetization. Furthermore indirect Data Monetization can be either organization's core or supplementary business but direct Data Monetization is more likely to be the organization's supplementary business line, supporting and creating additional value alongside the core businesses. The reason behind this may be in the risks related to establishing direct Data Monetization as a core business. Additionally the continuity of such business can be considered challenging as well as the crucial prerequisite of Data Monetization – how could an organization without other operations generate and create new data?

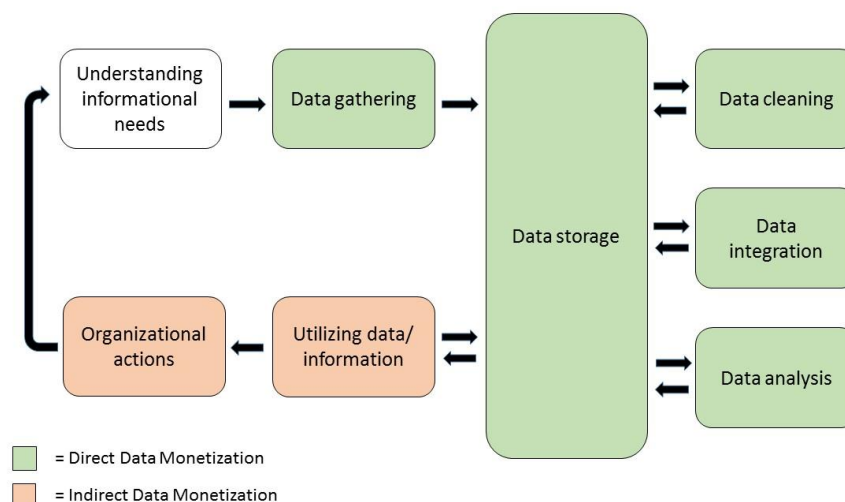
These business models described are associated with different offerings depending on the extent of business as well as the product or service offered. The indirect and direct business models and their features are revised in the following picture 10 which is constructed by combining the pictures 5 and 7 from the chapter 5.4:





**Picture 10: Conclusion of Data Monetization business**

When the indirect and direct business models were reflected against the framework of value creation with data, it appeared that in the direct business model value is created in the phases of data gathering, storing and processing whereas in the indirect business model value is mainly created in the later steps of utilizing information-based products or services and executing organizational actions:



**Picture 11: Summary of value creation in Data Monetization business**

Since the indirect Data Monetization is dependent on data from which the information-based products and services are derived, the phases of value creation present in the direct

Data Monetization appear usually prior to the steps of value creation present in the indirect Data Monetization. This also means that the indirect Data Monetization may require more investments in order to achieve the corresponding steps of value creation framework and furthermore the monetization transaction and revenue generation compared to the direct Data Monetization.

As it can be noted from the picture 11 the first step of framework is excluded from both of the business models. The reason for this is that the first step of the process model does not occur solely in Data Monetization business but it is always associated with another phase of the model. Furthermore the first step was recognized to be a crucial prerequisite for data gathering and moreover Data Monetization in general from the perspective of the organization that plans and establishes Data Monetization business. Hence the first phase is indeed an important aspect in Data Monetization but it is not a phase where the value creation is realized.

The currently known problems as well as potential issues related to Data Monetization were also discussed in the chapters 5 and 6. Simultaneously these chapters sought answers to the following sub research questions: “*What other phenomena are related to Data Monetization?*” and “*What kind of elements organization should consider when planning Data Monetization business?*”. Some serious issues and questions associated with data security and privacy, ethics and ownership of data were depicted. Additionally some challenges that Data Monetization pursuing organizations may face. Similarly the challenges related to and affecting the continuity of Data Monetization business were emphasized. Whereas the challenges related to establishing Data Monetization business may be vanquished and coped with, the current and potential issues related to topics such as ethics, data ownership and data privacy are more complicated ones. Hence it is essential that the organizations are aware of such challenges and issues and furthermore are able to adjust their actions and business when needed.

As a conclusion it can be stated that the answer to the question “*What kind of factors are behind of and affect Data Monetization?*” is rather ambiguous. A great deal of internal and external factors were discovered in the chapters 4, 5 and 6. The next list will summarize the most significant factors that shape and steer Data Monetization business:

**Internal factors:**

- Other business trades, products and services
- Existing data; Data quality, content and context
- Data generation capabilities and competencies to refine and derive data
- Existing IT, marketing, distribution and other capabilities
- Motive, aspiration and objectives
- Ability and willingness to invest
- Strategic portfolio

**External factors:**

- Availability of external data sources
- Customers' needs and wants
- Market dynamics and competitors' offerings
- Varying legislation, globalization
- Ethical aspects, data privacy and data ownership

Altogether there are distinct business models and alternatives for an organization to choose from, depending on the organization's vision and willingness to pay attention to and invest in Data Monetization business. Hence Data Monetization can be the organization's main offering but in the same way Data Monetization can complement the organization's other business trades. Similarly the actual product or service offered can vary from simple data sets to carefully developed information-based products or services. Also the end product or service can be a combination or mix of individual information-based products, information-based services and organization's other offerings. Regardless Data Monetization's diverse and multidimensional essence it is evident that Data Monetization is a growing phenomenon and furthermore a business trade that may offer new ways to generate revenue and ultimately profit.

**7.2 Discussion and evaluation**

In general it can be stated that this study was able to give an answer to each of the presented research questions and thus overall this study succeeded and fulfilled its objectives. Still the lack of scientific literature and relevant research made it challenging to conduct a comprehensive systematic literature review on Data Monetization. Hence the extent and depth of this study may have suffered due to the narrow sample of literature and furthermore the constricted amount of reference. Nevertheless the lack of previous and relevant research on Data Monetization is an interesting finding and furthermore increased the novelty value of this study.

If considering the content of literature sample utilized in the systematic literature review it can be noted that there are some gaps and deficiencies in the studied research and scientific literature. The main gap is that the majority of research papers did not define Data Monetization and furthermore when Data Monetization was defined, the definitions were diverse and rather contradictory. Additionally the pieces of literature studied Data Monetization mainly from the technical perspectives and moreover the focus was on Big Data or analytics. This means that the transactional and strategic aspects of Data Monetization or the business models and logics applied in Data Monetization business were hardly studied in these papers.

In spite of these challenges experienced this study was able to form a structured and comprehensive definition for Data Monetization. This can be considered as a merit since the

scientific literature discovered in the systematic literature review could not give a comprehensive and justifiable definition for Data Monetization. Additionally the framework of value creation with data established in this study was successful in structuring the different phases or steps of value creation in the process of transforming data into information and eventually into knowledge and organizational actions. Moreover this framework clarified the distinct phases and ways of creating and offering value to the client in Data Monetization business.

Although this study succeeded in several ways, some aspects of findings might require more thorough consideration. As mentioned previously the sample of systematic literature review contained contradictory viewpoints on the financial aspects of Data Monetization, especially regarding whether monetization is about generating revenue or profit. This study concluded that monetization stands for revenue generation but from a critical perspective this conclusion may require further research since the currently found scientific papers have rationalized this matter in ambivalent ways. Additionally this part of definition is critical from the perspective of determining what kind of actions can be considered as Data Monetization. Hence a different approach to monetization could have given a different result.

Another aspect that might require enhancement or further research is the identification of potential issues and challenges related to Data Monetization business. The issues and pitfalls described in the chapter 6 were identified from the literature related to mainly Big Data and its challenges. In other words these issues and challenges are not directly about Data Monetization and therefore a wider study on particularly Data Monetization business' issues might be needed. Additionally even though the technical issues were discussed briefly in the chapter 5 and therefore excluded from the chapter 6 it may be necessary and useful to conduct a more thorough research on those aspects as well.

Nevertheless this study has its flaws and some improvement ideas were found but still the distinct aspects and factors of Data Monetization business were studied systematically and in a well-structured manner. The business models were studied from the point of view of value creation as well as with a more pragmatic approach. Therefore it can be argued that this study sufficiently fulfilled the expectations set in the introduction and furthermore offered something new in terms of defining thoroughly Data Monetization and researching the various aspects and factors of and affecting Data Monetization. Hence this study may have some novelty value because similar studies were not yet conducted or at least published through the channels utilized in this study.

### **7.3 Further research on Data Monetization**

As described earlier Data Monetization is a phenomenon that has not been studied extensively yet. This can be observed from the small amount of relevant literature found when the systematic literature review was conducted. Additionally the varying definitions of

Data Monetization highlighted the fact that in the scientific literature no commonly accepted definition for this term existed. Furthermore a uniform and standardized terminology regarding Data Monetization is a prerequisite for further research on this phenomenon and distinct aspects affecting and arising from it.

This study strived for deducing a uniform and comprehensive definition for Data Monetization but nevertheless the monetary and financial aspects on Data Monetization may require a further study to distinguish unambiguously between the alternatives of revenue and profit generation. Although these terms may seem rather reminiscent, they affect greatly on the definition of Data Monetization. Whereas profit generation offers a wider approach in terms of including avoidance of expenses and costs as part of Data Monetization, revenue generation excludes such aspects but allows Data Monetization for the non-profit and public organizations. Both of these views have their own advantages and disadvantages, and furthermore tradeoffs most likely occur in terms of the distinct business models and practices that are considered as Data Monetization. Since Data Monetization is likely to be a growing and evolving phenomenon, a uniform definition should be developed.

Although this study provided an overview on the distinct business models present in Data Monetization business, deeper insight on the topic is required to be able to understand better the distinct revenue generation models and logics. Furthermore an encompassing view on the revenue generation models and Data Monetization business enable expanding and deepening the study on the strategic viewpoints and aspects of Data Monetization. This sort of further research offers a wider and more comprehensive view on Data Monetization but may also be able to discover new ways to implement and design Data Monetization in the evolving and developing markets.

## BIBLIOGRAPHY

Aaltonen, A. & Tempini, N. (2014). Everything Counts in Large Amounts: A Critical Realist Case Study on Data-based Production. *Journal of Information Technology*, Vol 29 (1), pp. 97-110.

Accenture (2016). Searched 2.10.2016, <https://www.accenture.com/fi-en/insight-data-monetization-summary>

Alavi, M. & Leidner, D.E. (2001). Knowledge Management and Knowledge Management Systems: Conceptual Foundation and Research Issues, *MIS Quarterly: Management Information Systems*, Vol. 25(1), pp. 107-136.

Baines, T., Lightfoot, H., Peppard, J., Johnson, M., Tiwari, A., Shahab, E. & Swink, M. (2009). Towards an Operations Strategy for Product-centric Servitization. *International Journal of Operational & Production Management*, Vol. 29 (5), pp. 494-519.

“Big Data”. Merriam-Webster (2016b). Searched 8.10.2016 from <http://www.merriam-webster.com/dictionary/big-data>.

Boyd, D. & Crawford, K. (2012). Critical Questions for Big Data: Provocations for a Cultural, Technological, and Scholarly Phenomenon. *Information, Communication & Society*, Vol 15 (5), pp. 662-679.

Buhl, H.U., Röglinger, M., Moser, F. & Heidemann, J. (2013). Big Data – A Fashionable Topic With(out) Sustainable Relevance for Research and Practice? *Business & Information Systems Engineering*, Vol. 2, pp. 65-68.

Caves, R.E. & Porter, M.E. (1977). From Entry Barriers to Mobility Barriers: Conjectural Decisions and Contrived Deterrence to New Competition. *The Quarterly Journal of Economics*, Vol. 91 (2), pp. 241-262.

Chen, H., Chiang, R.H.L. & Storey, V.C. (2012). Business Intelligence and Analytics: From Big Data to Big Impact. *MIS Quarterly: Management Information Systems*, Vol 36 (2), pp. 1165-1188.

Chen, M., Mao, S. & Liu, Y. (2014). Big Data: A Survey. *Mobile Networks and Applications*, Vol 19 (2), pp. 171-209.

Choo, C.W. (1995). The Knowing Organization: How Organizations Use Information to Construct Meaning, Create Knowledge and Make Decisions. *International Journal of Information Management*, Vol. 16 (5), pp. 329-340.

Choo, C.W. (2002). The Knowing Organization as Learning Organization. *Education + Training*, Vol. 43 (4), pp. 197-205.

Clemons, E.K. (2009). Business Models for Monetizing Internet Applications and Web Sites: Experience, Theory, and Predictions. *Journal of Management Information Systems*, Vol 26 (2), pp. 15-72.

Constantiou, I.D. & Kallinikos J. (2015). New Games, New Rules: Big Data and the Changing Context of Strategy. *Journal of Information Technology*, Vol 30 (1), pp. 44-57.

Cordella, A. (2006). Transaction Costs and Information Systems: Does IT Add up? *Journal of Information Technology*, Vol 21 (3), pp. 195-202.

Culnan, M.J. & Williams, C.C. (2009). How Ethics Can Enhance Organizational Privacy: Lessons from the ChoicePoint and TJX Data Breaches. *MIS Quarterly*, Vol 33 (4), pp. 673-687.

"Data monetization". Merriam-Webster (2016a). Searched 11.7.2016 from <http://www.merriam-webster.com/dictionary/monetization>.

Davenport, T.H. (2006). Competing on Analytics. *Harvard Business Review*, Vol 84 (1), pp. 98 -107, 134.

Davenport, T.H., Barth, P. & Bean, R. (2012). How Big Data is Different? *MIT Sloan Management Review*, Vol. 54 (1), pp. 43-46.

Deloitte (2016). Searched 2.10.2016, <http://www2.deloitte.com/us/en/pages/deloitte-analytics/articles/is-data-your-next-revenue-stream.html>.

Gartner. (2015) How to Monetize Your Customer Data. Searched 9.8.2016, <http://www.gartner.com/smarterwithgartner/how-to-monetize-your-customer-data/>.

Gelle, E. & Karhu, K. (2003). Information quality for strategic technology planning. *Industrial Management and Data Systems*, Vol. 103(8-9), pp. 633-643.

George, G., Haas, M.R. & Pentland, A. (2014). Big Data and Management. *Academy of Management Journal*. Vol 57 (2), p. 321-326.

Gerlach, J., Widjaja, T. & Buxmann, P. (2015). Handle with Care: How Online Social Network Providers' Privacy Policies Impact Users' Information Sharing Behaviour. *Journal of Strategic Information Systems*, Vol. 24 (1), pp. 33-43.

Golder, P.N., Mitra, D. & Moorman, C. (2012). What is Quality? An Interactive Framework of Processes and States. *Journal of Marketing*, Vol. 76 (4), pp. 1-23.

Goulet, L.R. & Frank, M.L. (2002). Organizational Commitment across Three Sectors: Public, Non-profit, and For-profit. *Public Personnel Management*, Vol. 31 (2), pp. 201-209.

Grant, R.M. (1996). Toward a Knowledge-Based Theory of Firm. *Strategic Management Journal*, Vol. 17, Issue: Winter 1996, pp. 109-122.

IBM. (2016a). What is Big Data? Searched 8.10.2016 from <https://www-01.ibm.com/software/data/bigdata/what-is-big-data.html>.

IBM. (2016b). The Four V's of Big Data. Searched 8.10.2016 from [www.ibmbigdata-hub.com/infographic/four-vs-big-data](http://www.ibmbigdata-hub.com/infographic/four-vs-big-data)

Jagadish, H.V., Gehrke, J., LaBrinidis, A., Papakonstantinou, Y., Patel, J.M., Ramakrishnan, R. & Shahabi, C. (2014). Big Data and Its Technical Challenges. *Communications of the ACM*, Vol. 57 (7), pp. 86-94.

Kamal, R. & Hong, C.S. (2015). Resilient Big Data Monetization. Cornell University, arXiv:1509.04545.

Kitchenham, B., Brereton, O. P., Budgen, D., Turner, M., Bailey, J., & Linkman, S. (2009). Systematic Literature Reviews in Software Engineering – A Systematic Literature Review. *Information and Software Technology*, Vol. 51 (1), pp. 7-15.

LaValle, S., Lesser, E., Shockley, R., Hopkins, M.S., Kruschwitz, N. (2011). Big Data, Analytics and the Path from Insights to Value. *MIT Sloan Management Review*, Vol. 52 (2), pp. 21-32.

Lee, Y., Madnick, S., Wang, R., Wang, F. & Zhang, H. (2014). A Cubic Framework for the Chief Data Officer: Succeeding in the World of Big Data. *MIS Quarterly Executive*, Vol. 13 (1), pp. 1-13.

Najjar, M.S. & Kettinger, W.J. (2013). Data Monetization: Lessons from a Retailer's Journey. *MIS Quarterly Executive*, Vol 12 (4), pp. 213-225.

Mossin, J. (1966). Equilibrium in a Capital Asset Market. *Econometrica*, Vol. 34 (4), pp. 768-783.

Opresnik, D. & Taisch, M. (2015). The Value of Big Data in Servitization. *International Journal of Production Economics*, Vol. 165, July 2015, pp. 174-184.

Payne, A.F., Storbacka, K. & Frow, P. (2008). Managing the Co-creation of Value. *Journal of the Academy of Marketing Science*, Vol. 36 (1), pp. 83-96.

Porter, M.E. & Kramer, M.R. (1999). Philanthropy's New Agenda: Creating Value. *Harvard Business Review*, Vol. 77 (6), pp. 121-130, 216.

Porter, M.E. (2008). The Five Competitive Forces That Shape Strategy. *Harvard Business Review*, Vol. 86 (1), pp. 79-93, 137.



“Profit”. Merriam-Webster. (2016d). Searched 8.10.2016 from <http://www.merriam-webster.com/dictionary/profit>.

Ravald, A. & Grönroos, C. (1996). The Value Concept and Relationship Marketing. *European Journal of Marketing*, Vol. 30 (2), pp. 19-30.

“Revenue”. Merriam-Webster. (2016c). Searched 8.10.2016 from <http://www.merriam-webster.com/dictionary/revenue>.

Rowley, J. (1998). What is information? *Information Services and Use*, Vol. 18 (4), pp. 243-254.

Setia, P., Venkatesh, V. & Joglekar, S. (2013). Leveraging digital technologies: How information quality leads to localized capabilities and customer service performance, *MIS Quarterly: Management Information Systems*, Vol. 37 (2), pp. 565-590.

Strong, D.M., Lee, Y.W. & Wang, R.Y. (1997). Data quality in context, *Communications of the ACM*, Vol. 40 (5), pp. 103-110.

Subashini, S. & Kavitha V. (2011). A Survey on Security Issues in Service Delivery Models of Cloud Computing. *Journal of Network and Computer Applications*, Vol. 34 (1), pp. 1-11.

Teece, D.J. (2007). Explicating Dynamic Capabilities: the Nature and Microfoundations of (Sustainable) Enterprise Performance. *Strategic Management Journal*, Vol. 28 (13), pp. 1319-1350.

Thierauf, R.J. (2001). *Effective Business Intelligence Systems*. 1st edition, Greenwood Press, 371 p.

“Value”. Merriam-Webster. (2016e). Searched 9.10.2016 from [www.merriam-webster.com/dictionary/value](http://www.merriam-webster.com/dictionary/value).

van Belle, G. & Ruiter, L. (2014). Data and the law: Beyond the sweat of the brow: Who owns published data? And what is data? *Significance*, Vol. 11(2), pp. 28-31.

Vargo, S.L. & Lusch, R.F.J. (2008). Service-Dominant Logic: Continuing the Evolution. *Journal of the Academy of Marketing Science*, Vol. 36 (1), pp. 1-10.

Wang, R.Y. & Strong, D.M. (1996). Beyond accuracy: What data quality means to data consumers, *Journal of Management Information Systems*, Vol. 12 (4), pp. 5-34.

Weber, R.H. (2010). Internet of Things – New Security and Privacy Challenges. *Computer Law & Security Review*, Vol. 26 (1), pp. 23-30.

- Webster, J. & Watson, R.T. (2002). Analyzing the Past to Prepare for the Future: Writing a Literature Review. *MIS Quarterly*, Vol. 26 (2), pp. xiii-xxiii
- Wixom, B.H. (2014). Cashing In on Your Data. *Center for IS Research Briefing*, Vol. 14 (8).
- Woerner, S.L. & Wixom, B.H. (2015). Big Data: Extending the Business Strategy Toolbox. *Journal of Information Technology*, Vol. 30 (1), pp. 60-62.
- Zeithaml, V.A. (1988). Consumer Perception of Price, Quality, and Value: A Means-End Model and Synthesis of Evidence. *Journal of Marketing*, Vol. 52, July 1988, pp. 2-22.
- Zimmermann, K.A. (2013). Fahrenheit: Facts, History and Conversion Formulas. Available: [www.livescience.com/39916-Fahrenheit.html](http://www.livescience.com/39916-Fahrenheit.html), accessed 4.12.2016.
- Zissis, D. & Lekkas, D. (2012) Addressing Cloud Computing Security Issues. *Future Generation Computer Systems*, Vol. 28 (3), pp. 583-592.
- Zott, C., Amit, R. & Massa. L. (2011). The Business Model: Recent Developments and Future Research. *Journal of Management*, Vol. 37 (4), pp. 1019-1042.