

UNIVERSITY OF TARTU

Faculty of Social Sciences

School of Economics and Business Administration

Anna Shamritskaya

CROSS-COUNTRY CHALLENGES IN THE IMPLEMENTATION OF BIG DATA
SOLUTIONS IN BANKING: THE CASE OF OTP BANK GROUP

Bachelor thesis

Supervisor: assistant Mark Kantšukov

Tartu 2020

This paper conforms to the requirements for the Analysis paper

.....

(signature of the supervisor)

Admitted for defense “ “..... (date)

I have written this Analysis paper independently. Any ideas or data taken from other authors or other sources have been fully referenced

.....

(signature of the author and date)

Table of Contents

Introduction.....	4
1.Cross country introduction and applications of Big Data Solutions in banking.....	6
1.1. Benefits and challenges of Big Data solutions in business	6
1.2. Applications of Big Data Solutions in Banking	11
1.3. Country-specific features of Big Data implementation in banking sector	16
2. Challenges in Big Data implementation in Bulgaria, Croatia, Hungary, Romania, Russia and Serbia.....	20
2.1. Overview of OTP Group.....	20
2.2. Data and Methodology	25
2.3. Results and discussions	29
Conclusion.....	34
References.....	37
Appendix A.....	42
Appendix B	43

Introduction

The amount of data we produce daily is huge and it became one of the most valuable resources. International Data Corporation (IDC) states that data is “the lifeblood of our rapidly growing digital existence” and it makes our digital existence “visible” (Reinsel, Gantz & Rydning, 2018, p.2).

The IDC Report states that less in 10 years the amount of data that are created, captured, or replicated (this is usually called Global Datasphere) will increase to 175 zettabytes which is five times more than in 2018. To understand how big 175 zettabytes is, one could imagine downloading the Global Datasphere at an average of 25 Mb/s and it would take 1.8 billion years to do so. (Reinsel, Gantz & Rydning, 2018)

Over the past 30 years, data has been changing most aspects of our life into a new way we experience the world around us. However, the main concern about the data generated is what to do with it so that life quality could be increased. Primary expectations are to allow new and unique user experiences and to discover more business opportunities (Reinsel, Gantz & Rydning, 2018). However, such advantages as greater decision-making and understanding of people’s needs and discoveries for the economic community can only result from the implementation of Big Data solutions (Almoqren & Altayar, 2016).

Many areas are already getting benefits from the insights that Big Data Solutions provide them and one of those areas is the banking sector. However, applications and implementation processes of Big Data (BD) differ from country to country due to the regulations, cultural aspects, and technical development of a place. Some countries are doing better than others; this makes a room to learn from the more experienced. To keep up with the growth of data banks should implement innovative strategies so that the adaptation process of BD could go smoothly and bring stability and prosperity. In order to do so, author of this paper believes that it is important to study challenges that occur in different countries.

The thesis aims to clarify what country-specific obstacles may emerge when introducing Big Data solutions in the bank: the case of OTP Bank Group. The OTP Group was selected for the study as it is one of the largest financial institutions in Eastern and Central Europe. Moreover, the Group is developing in terms of smart usage of data in the financial industry.

To achieve the aim of the thesis, the following tasks were set up:

- To discuss the benefits and challenges of Big Data approach in business,
- To review applications of Big Data solutions in the banking sector,
- To provide examples of country-specific obstacles of Big Data implementation in the banking sector,
- To provide an overview of the OTP Banking Group and countries of research,
- To conduct interviews with employees of Data Science departments of regional branches,
- To analyze and discuss results.

The paper consists of two parts, theoretical and empirical. The first part of this work will focus on explaining the concept of Big Data and on assigning its principles to the banks' functioning. Previously done studies will be analyzed to gather information about country-specific challenges that can be faced when implementing BD solutions in different countries. The empirical part will include a brief description of OTP Banking Group, analysis of countries where the branches, are and interviews with experts from IT, management, and data quality fields.

At last author shall thank OTP Bank Group which agreed to participate in the research and her family for everlasting support during all stages of done work.

Keywords: Big Data, financial sector, banking, analytics, data quality.

1. Cross country introduction and applications of Big Data Solutions in banking

1.1. Benefits and challenges of Big Data solutions in business

Big data concept has been widely discussed during the last decade not only in the IT sector but in all other fields as well and its definition is constantly evolving. In this section, the author of the paper aims to show how the concept has been changing over the last decade and its relevance to the business sector.

The most common and known definition for Big Data is “data sets whose size is beyond the ability of typical database software tools to capture, store, manage, and analyze.” (Manyika et al., 2011, p. 1). However, BD is not just about the size, many other aspects should characterize data, so it is called Big. Nevertheless, size is one of the main characteristics.

The first modern definition was given by Gartner analyst Doug Laney (2001). He presented three main aspects of the concept or so-called 3Vs – volume, velocity, and variety that characterize BD. The following definitions of 3Vs are provided for a better understanding of the concept (Hu Shuijing, 2016):

- Volume indicates a very big amount of data. The main consideration that it has to be hundreds of Terabytes or even more.
- Velocity specifies the speed for data processing in terms of waiting time.
- Variety indicates the diversity of data that is collected. The data can be structured (highly organized, predictable), semi-structured (not predictable, has hierarchy) and unstructured (not organized, unpredictable).

The most recent studies see data as a more complex concept than it was before; so that in 2014 Khan, Uddin and Gupta introduced 7Vs. The change happened because today

unstructured data is more involved, and it brings a lot of uncertainty. However, with modern tools and algorithms, it is possible to “clean” it and receive meaningful results.

Additional 4Vs are the following (Khan, Uddin & Gupta, 2014):

- “Veracity” – simply meaning truthfulness of the data. The concern is if the data can be trusted.
- “Validity” - which is pretty similar to veracity but still different. Meaning correctness and accuracy of data concerning the intended usage. So, in one case the data set can be valid but, in another case, it is not.
- “Volatility” – since the data storage is quite expensive and, in some cases, it does not have to be stored forever, the data has to be volatile. Moreover, some retention policies restrict enterprises to store the data for the limitless amount of time.
- “Value” – all businesses are always eager to have the value for the company, so is data. It has to have a greater value than the risks and costs that it brings. The value of the data is the desired outcome one wants to get after processing it.

All in all, definitions of the term can slightly vary. However, they do not contradict each other but only contribute to a better understanding of this complex concept. Every author has its own perception and understanding of BD, but all of them either agree with the definition given by Doug Laney or elaborate it in such a way that assumes new technological changes. The following Table 1 represents the most popular definitions and gives different perspectives on it.

Table 1

Definitions of Big Data

Author(s) and year	Definition
Doug Laney, 2001	3Vs- volume, velocity variety
Manyika, Chui, Bughin, Dobbs, Roxburgh, and Byers, 2011	“Data sets whose size is beyond the ability of typical database software tools to capture, store, manage, and analyze”
Khan, Yaqoob, Hashem, Inayat, Ali, Alam, Shiraz and Gani, 2014	“Big Data is characterized by three aspects: (a) the data are numerous, (b) the data cannot be categorized into regular relational databases, and (c) data are generated, captured, and processed very quickly.
Kuo and Li, 2014	“...when the data becomes too large and complex, impossible when handled by traditional methods”
Khan, Uddin and Gupta, 2014	7Vs - volume, velocity variety, veracity, validity, volatility, value
Zhu, Yu, Wang, Ning and Tang, 2018	“It represents large and complex data sets obtained from all kinds of sources”

Source: compiled by author

As any new technology BD has its pros and cons. In the following section, the author of the paper will try to discuss the main benefits and challenges of BD usage in the business sector. Under the business sector, the author understands any industry that generates money and is not included in the public sector.

To begin with, Stocia (2018) in his work stated that implementation of BD solutions brings the opportunity to optimize business activities, to explore new markets, to accurately forecast, and to detect frauds. He emphasizes that one of the main advantages is the understanding “how this data provides a context based on the issue” so that the companies can correctly interpret customer behavior and preferences (Stocia, 2018, p.19). Shah, Soriano and Coutroubis (2018) stated that the adoption of new solutions offers greater productivity and profitability. This goes along with what Stocia (2018) proposed as improvements and could be a result of this adoption. However, Shah, Soriano and Coutroubis (2018) see some other factors as key ones, for instance, performance analysis, higher efficiency, new market segments, innovations and financial stability.

Sen, Ozturk and Vayvay (2016) agree with the previous statements that BD is considered as a necessary innovation. But the authors underline that with such innovation the business can expect only the generation of future opportunities. So, they stress that BD has potential but cannot promise immediate solutions. Kaisler, Armour, Espinosa and Money (2013) stated that BD supports the value creation and that there are five common ways of how it is done. They are the following: developing transparency, supporting testing and analysis of different kinds, assistance in segmentations, improving real-time analysis, and bringing innovation.

Another text that looks at how the BD application impacts the business is “Are you ready for the era of “big data”?”. This paper considers particularly the same advantages as were discussed before. However, the authors focus more on the financial part of the improvements and state that BD changes the competition between companies. They also emphasised the prosperity, lowering costs, better alignment of products with customer needs, and social impacts that such innovation brings. (Brown, Chul & Manuika, 2011)

Talking about the benefits of BD in the business most of the authors agree on the point that such innovation is needed. However, they all propose different results from it, some are more optimistic and hope for immediate results from transforming processes and altering the business strategies, but some consider it a long-term improvement resulting in higher productivity. Other difference is that some authors consider it as only a profit-centric improvement but some state that it has a social impact as well.

Like any innovation, BD solutions are challenging. and the next part will cover the main concerns. Stocia (2018) proposes five key ones: information security, high-performance architecture, misleading analysis, costs and difficulty to gather relevant information. Those come from the fact that it is quite difficult to manage large amounts of data that is mostly instructed. As Bolliers (2010) explains, the BD is not self-explanatory; it does not represent

the objective truth and the interpretation processes are debatable. It goes along with what Stocia (2018) states in his text, that the data itself and the methodology that is used for analysis, both could be deceiving.

Shah, Soriano and Coutroubis (2018) are separating disadvantages into two following groups: technological and managerial. This proposition makes it easier to address problems. Under the technical challenges, they understand the data volume, variety and all other characteristics of BD. They propose that those technical issues arise because of the nature of BD. Under managerial challenges, they rely on Portela, Lima and Santos's (2016) proposition that companies do not focus on business opportunities but only trying to solve technological issues. Kaisler et al. (2013) propose the same scheme but suggesting splitting into three categories: storage, management, and processing. The storage and processing could be sub-categories under technological issues. However, if there are only two sub-categories, there could be missing other challenges that were proposed by Shah, Soriano and Coutroubis (2018).

Brown, Chul and Manuika (2011) draw attention to the two pillars: the scarcity of talent and privacy and security issues. They focus on those because the application of BD could bring good results only if well trained and knowledgeable specialists are working and if the application follows all legislations. The challenges proposed could fall under the management issues in the previous framework

The different framework was proposed by Sen, Ozturk and Vayvay (2016, p.162), they suggest that only three following questions should be answered if the company want to utilize BD:

- Can we sense this data?
- Can we generate sensible results from it? Can we use it for better service?
- Can we convert it to a profit?

This framework could work well if the company is only starting using the BD approach; since they need to know first if the data can be used.

The challenges and issues that BD brings within its implementations depend on many various factors: how technologically advanced is the firm, how big it is, and if the management staff knows what to do with the innovation. The frameworks discussed are different from each other, but they all agree that some essential issues should be addressed: privacy, storage, and interpretation of results. The following Table 2 represents the summary of the author's findings and groups the challenges into three categories.

Table 2

Summary of benefits and challenges of Big Data application

Benefits	Challenges
<ul style="list-style-type: none"> • Optimization of business activities, • Discovery of new markets and opportunities, • Meticulous analysis, • Greater productivity and efficiency, • Lower costs, • Encouraging value creation, • Competitive advantage. 	<p>Technological perspective: volumes, velocity, variety, relevance, storage, application and processing.</p> <p>Social perspective: privacy and ethicality.</p> <p>Business perspective: conversion into profit, rapidity of solution and scarcity of talent.</p>

Source: compiled by the author

To conclude, the author of the paper can state that there are different approaches to the advantages and challenges that BD brings to the business sector. The way to address it depends on many aspects and the company should consider both benefits and issues before deciding if to follow the innovation approach.

1.2. Applications of Big Data Solutions in Banking

To start with, one should understand the importance of BD for the finance sector in general. By its nature, banks and other financial institutions are generating great amounts of data daily, so that this sector is driven by data. Moreover, the availability of data from different sources (social media, blogs, etc.) and new technologies that allow us to analyse

unstructured data gives a chance to banks to be more competitive on the market. The possibility to use modern BD solutions opens new opportunities for greater growth and sustainability in the financial sector. In some parts of the financial sector, the shift from traditional technologies (relational database management) to BD solutions happened because it is no longer possible to process and handle an exponentially growing amount of data. Another crucial point is that due to new compliance requirements, the governance and risk reporting is needed to be more transparent which is possible if using advanced software. Due to new regulations, financial institutions have to store enormous masses of data for many years and BD solutions are the way to do it according to the new regulations. (Forest, Foo, Rose & Berenzon, 2014)

For the financial industry, data is becoming one of the most important assets. The sector is using BD to alter not only business processes inside the organizations but the whole industry in general. They are using it to predict the “stock’s movement and securities in the market by developing predictive algorithm”. (Jaiswal & Bagale, 2017, p. 339)

As this paper is going to look at the country-specific aspects of a particular banking group – the following part is going to concentrate only on banks, not the whole financial industry.

There are unique ways of how BD solutions used in banking as well as are some crucial features that are pretty similar to other industries. Jaiswal & Bagale (2017) stated that one of the ways BD technologies could be used in banks is to categorize and customize customers in order to provide private services. Human-centric services are one of the key factors to survive and succeed in any business. However, in retail banking, customer loyalty is influenced by perceptions of service quality and levels of customer satisfaction while other industries have their factors (Rahman, 2005). Customer journey and satisfaction are becoming critical elements because there are countless touchpoints and channels where the

interaction between client and the firm happens; this leads to a more complex customer journey (Lemon & Verhoef, 2016). That is why the understanding, and evaluation of customer experience is seen as crucial. The research made by the Deutsche Bank (2014) extends the argument stating that to get new data banks have to be collaborative, for instance, with PayPal, Google or Amazon, because “financial institutions no longer monopolize a consumer’s financial transactions” (Forest, Foo, Rose & Berenzon, 2014, p. 12). Another common aspect is that this sector goes through the same cycle as any other industry. It all starts with the collection of data depending on the sector it could be various sources and the data by its nature could be of different types. The second stage is the analysis; the data gathered is analysed and explored for new and unique patterns, correlations and results. The way of analysing could be divided into two distinct frameworks: foundational technologies and emerging research (Chen, Chiang & Storey, 2012). Those two include more specific tools and techniques of BD solutions. The last stage is Reporting. The important aspect is that it should be in a user-friendly way and bring out new results that increase opportunities. The final common aspect is that banks provide real-time recommendations for customers about their financial statements. The real-time proposal service is the part of many industries, for example, online streaming platforms or retail industry.

In the paper “Survey on Big Data in Financial Sector” (2017) the authors present four ways of how the banks can exploit BD technologies. Two of them that are common for most of the business industries are target marketing to increase competitiveness and real-time recommendation service. A study by Tiana, Hana, Wang, Lua and Zhan (2015) sees BD as extremely beneficial information that could provide such financial institutions as banks with a possibility to get a comprehensive view of both market and clients. They define at least three ways of how banks could utilise BD analytics precise customer analytics, risk analysis,

and fraud detection. These approaches are similar to any other business and can lead to smarter trading and providing customized service, thus, to reach a competitive advantage.

Nevertheless, other ways are pretty specific for the sector. One of which is cost controlling. It is important to identify unproductive factors and schemes that the bank is using because they could cause losses. (Jaiswal & Bagale 2017) The research made by Deutsche Bank in 2014 claims that due to new regulations the liquidity planning and asset and liability management have to be reconsidered. This goes along with the cost controlling.

Other peculiarity is the identification of frauds in an early possible time. Authors of two different papers Gutierrez (2014) and Jaiswal and Bagale (2017) agree that with BD it is possible to prevent most types of risk, for instance, identify high-risk borrowers, frauds in transactions and outliers. Moreover, the study by Deutsche Bank (2014) which is concerned with fraud detection looks at it differently. They proposed that fraud identification lies with the new regulations that are external and not only for internal usage. Another difference is that Deutsche Bank emphasises the use of external information in order to combat credit fraud and security breaches. They insist that unstructured information from internet channels and GPS data is crucial and that around 85% of financial data is unstructured (Forest, Foo, Rose & Berenzon, 2014).

The further point is that a transformation from face to face to online banking happened. This occurred due to new consumers' behaviour and because of the affordability of the technology that allows making transactions online safe and quick. As stated by Deutsche Bank (2014, p.13) BD plays a crucial role in "enabling customer centricity", the transition that happened recently brings with it not only a competitive advantage but safety, regulation, and many other concerns. According to Indriasari, Gaol and Matsuo (2019), the main role of BD in the banking sector is to develop digital banking innovations. The following examples could be seen as innovations (Indriasari, Gaol & Matsuo, 2019):

- Analyse the customer experience through each channel (mobile banking, e-Banking, direct interactions on the office),
- Find patterns that lead to the generation of sales,
- Analyse performance of banking digital channels,
- Optimization of the innovation strategy,
- Drive customer satisfaction, loyalty and recommendation, thus, improve the financial performance of the bank.

In regard to the transition of the bank to the digital era, in literature, Mobile banking has been described as one of the transition milestones. Mobile banking arose due to the integration of internet technologies and mobile networks. It offers similar services and practically has the same advantages as internet banking. But it represents a new paradigm as anywhere anytime service. (McGovern, Lambert & Verrecchia, 2019)

Moreover, Mobile banking was identified as a primary point of contact between banks and their clients. While there are no official standards of how mobile banking services should be like, banks are already in competition with ApplePay, Android Pay, etc. But most banks integrate into their mobile apps complex business ecosystem in order to keep up with the ongoing changes. (Mullan, Bradley, & Loane, 2017)

The survey by PwC (2020), there were defined six priorities for retail banks in the new digital era. They are the following (PwC, 2020):

- Developing a customer-centric business model,
- Optimization distribution,
- Simplifying business and operating models,
- Obtaining an information advantage,
- Enabling innovation, and the capabilities required to foster it,

- Proactively managing risk, regulations, and capital.

These goals are directly related to technological innovation adoption. Having a clear, unique and creative strategy should allow banks to be proactive in terms of innovation adoption which will bring benefits in the evolving and uncertain future.

All in all, the author of the paper concludes that like any other non-public sector the target of the bank is to increase profits. One of the great solutions is to use BD because it gives many options and has potential. However, due to high the level of security and legislation, banks have their unique challenges on the way of innovation introduction. And it is very important to know beforehand those obstacles; Table 3 is the summary of the author's findings on this topic.

Table 3

Summary of common and industry relevant aspects of Big Data applications

Specific	Common
<ul style="list-style-type: none"> • high risk borrowers • fraud detection • mobile banking • Electronic banking • liquidity planning • assets and liabilities management 	<ul style="list-style-type: none"> • customer-centric • high quality service • Big Data cycle • Real-time services

Source: compiled by the author

1.3. Country-specific features of Big Data implementation in banking sector

To find studies that discuss the aspects of BD implementation, three well-known scientific databases were employed (JSTOR, ScienceDirect, and EBSCO Discovery). The used keywords were “Big Data”, “Implementation of Big Data”, “country-specific aspects”; also, the criteria by which the author narrowed down the search were the year of publication and the focus of research. As expected, the topic of discussing the aspects of BD implementation in the banking was studied before. Most of the studies found by the author generalize the conclusions, examine one country only, do not compare the results to other

cases and do not focus on country-specific challenges. The papers that involved challenges and benefits of BD implementation in banking industry were chosen for the analysis. An important note is that those works' main goal was not to study the distinct aspects but to study other points that are the consequences of those differences. Even from the international perspective, to the author's best knowledge, there are no works on studying the country-specific challenges and comparing the results from several countries that take into account the banking sector only and rely on a particular bank group. This might be because the field is yet unexplored, and it has been changing rapidly over the past decade.

The novelty of the author's topic is to acquire information on differences that are faced when introducing BD solutions in the branches of a particular banking group. So that the differences that are coming from, for instance, the cultural background, political regime or legislations, could be compared. This research attempts to bridge the gap in understanding that different country-specific aspects can influence the implementation of BD based solutions in commercial banks, and there is a possibility to develop a universal strategy that could help to be quick on changes and bring stability to the banking sector.

The study that is a special project of IDC Russia (2015) investigates the perspectives of global and Russian markets for BD technologies in the financial sector. It outlines factors that influenced the development of BD analytics in the Russian banking sector and gives suggestions for the future. The challenges mentioned in the paper are consistent with those noticed by others (Brown, Chul & Manuika, 2011; Shah, Soriano & Coutroubis, 2018). However, there is one peculiarity that was not noticed before, the study sees a large number of BD solutions' providers as an obstacle for the Russian market since the providers are foreign and do not exactly understand the specificity of Russian bank systems and the market.

Kshetri (2016) in his paper sets as one of the objectives the investigation of how formal and informal institutions support and limit the use of BD in the Chinese financial

sector including commercial banks. The paper states that the main challenge is that banks and financial institutions do not have enough data and the desired level of competency to provide financial services to low-income families and micro-enterprises in China. However, BD could become a game-changer and help financial providers to overcome two key problems: information opacity and transaction costs. The first challenge appears due to the nature of data and was mentioned as an obstacle in other papers (Shah, Soriano & Coutroubis, 2018; Bolliers,2010; Stocia,2018); the second challenge could be described as business-related and as well stated by other authors (Brown, Chul & Manuika, 2011; Kaisler, Armour, Espinosa and Money,2013).

Almoqren and Altayar (2016) attempt to describe the motivational factors that led the banking sector in Saudi Arabia to use BD and Data Mining tools. Moreover, it highlights some issues on the way of adoption that are consistent with the literature studied (Shah, Soriano & Coutroubis, 2018; Bolliers,2010; Stocia,2018).

Saxena and Al-Tamimi (2017) look at BD and Internet of Things (IoT) technologies in Omani banks. The study takes into account four different types of banks: national, international, Islamic and specialized, and discusses the integration of BD analytics in Omani banks. Moreover, the study gives a clear framework of what should be improved in order to provide better services and maintain better customer relationships.

Srivastava, Singh, Tanwar and Tyagi (2017) investigate the suitability of Big Data Analytics (BDA) technologies in the Indian banking sector. It emphasises the challenges of BDA adoption for Indian banks and states that the sector is still in “explore and experiment state” (Srivastava, Singh, Tanwar & Tyagi, 2017, p.4).

Indriasari, Gaol and Matsuo (2019) research how BD influences the digital transformation in the Indonesian banking sector. Their paper tries to investigate the relationship between the application of Artificial Intelligence (AI) and BDA in banking, and

their influence on customer experience. Even though the study's main goal is not to examine the country-specific aspects, it presents challenges and opportunities for BD application in Indonesian banks. The obstacles brought up by the authors align with the framework in Table 2. However, their study and investigation of the Chinese financial sector (Kshetri, 2016) mention several legal or/and political aspects that were not discussed in previous papers. Due to this fact, the author of this paper concludes that there is another group of challenges that could be titled as Legal.

The author concludes that the results from all mentioned above studies show some patterns (see Table 4). The main issues in most countries are the lack of specialists, adoption costs and information quality. Moreover, some studies (IDC Russia, 2015; Kshetri, 2016; Indriasari, Gaol & Matsuo, 2019) pointed out several peculiarities that could be described as country-specific challenges.

Table 4

Studies on implementation Big Data technologies in a specific country

Country & Year	Author(s)	Methodology	Results
Russia, 2015	IDC Russia	Qualitative research based on previous studies	<ul style="list-style-type: none"> • Large number of solution providers/ outsourcing; • Lack of specialists; • Issues with personal data.
China, 2016	Kshetri, N.	Case studies	<ul style="list-style-type: none"> • Huge support from the government; • Less ethical issues since the “a lower degree of concern about privacy”; • Financial industry is under the transformation – new regulations.
Saudi Arabia, 2016	Almoqren., N. and Altayar, M.	Quantitative approach	<ul style="list-style-type: none"> • System quality (system integration, usefulness, security, performance and effectiveness); • Information quality (visualization, personalization, reliability and accuracy); • Basel III regulations.

Oman, 2017	Saxena, S. and Al-Tamimi, T. A. S. M.	Studies based on social media	<ul style="list-style-type: none"> • No transformation accordingly to IT changes and innovations; • Institutional barriers such as regulations and standards; • Issues with trust and security to new technologies; • Technical challenge such as telecommunication infrastructure.
India, 2017	Srivastava, A., Singh, S.K., Tanwar, S. and Tyagi, S.	-	<ul style="list-style-type: none"> • Shift to “Data Culture”; • Lack of skilled workers; • Adoption cost; • Emerging of new technologies.
Indonesia, 2019	Indriasari, E., Gaol, F., L. and Matsuo, T.	In-depth interviews with CIOs or IT executives and Qualitative Data Analytic	<ul style="list-style-type: none"> • Infrastructure limitations; • Data accuracy; • Legacy technologies lacking integration capabilities; • Regulatory constraints; • Inflexible business process and teams; • Lack of properly skilled team; • Cost of adopting.

Source: compiled by the author

All in all, there are obstacles that should be taken into account when introducing BD based solution in a new country. This study in the following parts will try to find out new patterns and other challenges by researching more countries in another region.

2. Challenges in Big Data implementation in Bulgaria, Croatia, Hungary, Romania, Russia and Serbia

2.1. Overview of OTP Group

The following part will cover three main points: the historical background of the OTP group, including the current situation, financial highlights, and world ratings; overview of departments participated in the research and external environment the banks operate in.

The National Saving Bank of Hungary was established in 1949; it was state-owned and domestic bank providing retail deposits and loans. In the following years, the bank has

widened its range of activities and responsibilities. So that by 1989 the bank has operated as a multi-functional commercial bank and in 1990 when the NSB became a public company and was renamed to National Savings and Commercial Bank. The activities of the bank were separated into non-banking along with supplementary units. (History, n.d)

The privatization of the bank began in 1990. It was a result of the introduction to the Budapest Stock Exchange, Firstly, the state's ownership of the bank decreased to a single voting preference, shortly after that the state's ownership was fully eliminated. Today OTP Group described as having free-float shareholder structure around 68,61% and dispersed ownership of private and institutional investors. (OTP Bank, 2019)

After the end of privatization, the main focus of the Groups was Central and Eastern Europe which provides huge opportunities for acquisition and growth. By now the bank has already done several successful acquisitions and became the key player in the region.

Following is the list of subsidiaries' names and countries of operation (History, n.d):

- DSK Bank, Expressbank in Bulgaria,
- OTP banka d.d in Croatia,
- OTP banka Srbija and Vojvodanska banka in Serbia,
- OTP Banka Slovensko in Slovakia,
- CJSC OTP Bank in Ukraine,
- Crnogorska komercijalna banka and Podgoricka banka in Montenegro,
- Mobiasbanca in Moldova,
- SKB banka in Slovenia,
- and OAO OTP Bank in Russia.

In order to give an overview of the financial performance of the OTP Bank Group, the key financial ratios such as revenue, net income, market capitalization, and P/E ratio are presented in Table 5.

Table 5

Consolidated financial highlights in millions EUR

Indicators	2018	2019
Total income	2,764.03	3,316.08
Net income	997.875	1,269.48
Market Cap (billion)	9.8	13.1
P/E ratio*	9.7	10.3

Notes: *P/E ratio (trailing, from adjusted net earnings)

Source: compiled by the author based on the Summary of the full-year 2019 results.

Positive growth in all indicators is due to the “strong domestic demand and the rapidly expanding investment activity” (Summary, 2020). Moreover, the investors’ confidence in Hungary considerably strengthened, so that both S&P Global and Fitch’s Ratings increased their credit rating by one notch (to ‘BBB’). This leads to the conclusion that the banking group is enhancing its positions and becoming financially more stable each year. With steady growth and reliability that the banking group maintains, they are turning into the key player not only in the domestic terms but in the Central and Eastern Europe region as well.

In total six subsidiaries took part in the research: Bulgarian, Croatian, Hungarian, Romanian, Russian and Serbian. They were chosen according to the recommendation of Head of OTP Data Science department in Hungary. The key reasons are that these subsidiaries have sufficient volumes of data and relevant expertise. The information in the following Table 6 will give a reader a brief overview of each subsidiary. Moreover, it is important to state that in terms of Data Science departments and innovation management, the subsidiaries have the privilege to decide on their own the way to manage the data acquired.

Table 6

Overview of subsidiaries.

Subsidiary	Established	# of clients	# of employees	BD implementation
Bulgaria	2003	>0.5mln	>3,000	-
Croatia	2005	>0.5mln	2,000	-
Hungary	1990	-	>9,000	2010 & 2017*
Romania	2004	>0.35mln	1,364	-
Russia	1994	>3.7mln	>10,000	2016 & 2017*
Serbia	2006	-	1,327	2019

Notes * in Business and in IT correspondently

Source: compiled by the author based on Overview of Expressbank. (n.d), Overview of OTP banka Srbija.(n.d), Overview of OTP Bank (n.d), OTP in Russia (n.d), OTP Banka (n.d), Annual Report Romania (2019).

The author of the paper believes that in order to have a better understanding of challenges banks in the chosen countries have, it is important to look at the external environment they are operating in. In the previous part of the paper, four main dimensions of challenges have appeared: technological, legal, social and business. For each dimension, the author of this paper has chosen indexes that rate countries in different terms. For the technological assessment of countries, Global Innovation Index was chosen because it shows the level of innovation in a country that corresponds to the level of technological development (Dutta, Lanvin, & Wunsch-Vincent, 2019). In terms of legal dimension, the author believes that it is important to notice that the subsidiaries are operating in different legal zones. While Hungary, Bulgaria, Romania, Croatia are European Union (EU) member states, Serbia has the status of candidate country to the EU and Russia is not a part of the EU. For the social dimension the Digital Evolution index is employed; one of the aspects it measures is a digital trust which is important to study because it shows the attitude across countries to interactions with the digital world (The Digital Evolution Index, 2017). For the Business dimension, The Global Competitiveness Report 2019 (World Economic Forum)

was chosen as an index. It rates countries based on the Global Competitiveness Index (GCI) which combines 103 indicators that captures concepts matter for productivity and show the overall state of a country in terms of global competitiveness which is ‘the attributes and qualities of an economy, that allow for a more efficient use of factors of production’ (Schwab, 2019, p.2).

The results that are presented in Table 7 suggest that in terms of competitiveness (Business dimension) all six countries are pretty similar, their scores do not differ much, they all are in the middle of the rating system. For the social dimension, there was only data for three countries; Bulgaria has the biggest score meaning that among these countries, the Bulgarian population has greater digital trust. In the legal context, all the EU member countries have to follow the General Data Protection Rules while other countries have their local laws that are less strict than GDPR but still employ the same legal framework. In terms of technological development, Hungary has the highest score while Serbia has the lowest. The two indices GII and GCI have stated Serbia as the least innovative and competitive country and ranked Hungary 1st in GII and 2nd in GCI in this group.

Table 7

Values of indexes across countries in each dimension

Dimension	Hungary	Bulgaria	Serbia	Russia	Croatia	Romania
Business*	47/65	49/65	72/61	43/67	63/62	51/64
Legal	GDPR	GDPR	Local law	Local law	GDPR	GDPR
Social	2.25	2.75	-	2.2	-	-
Technological	33	40	57	46	44	50

Notes: * the first number is the place in rating, the second is the actual score out of 100

Source: Dutta, Lanvin, & Wunsch-Vincent (2019), Schwab (2019), The Digital Evolution Index (2017).

All in all, OTP Banking Group is suitable for the research because it is a significant player at the national level as well as in Central and Eastern Europe region; and it shows the

fastest growing ratings as a financial group with acquiring new subsidiaries and developing in new countries (Deloitte, 2020). The fact that departments are in one geographical region and scored pretty similar in rankings provided but in different stages of economic development and have different legal systems makes it not only interesting but important to research; since new and exciting patterns could be found. However, these similarities could result in similar or the same outcome.

2.2. Data and Methodology

In the following section, the author of the paper is going to introduce the way the interview plan was crafted, what method of collecting the data is used to fulfil the research aim, the framework for the analysis of acquired data, and the primary data overview.

Firstly, the deductive approach was adopted. The hypothesis that is derived from the literature review states that there are some challenges of BD application in the banking sector that could be described as country-specific. There are already certain patterns with common issues in all countries studied, nevertheless, there is a sense that some new and not studied yet patterns could appear. As was stated before four main groups of challenges emerged, they could be categorized under one broad topic such as external factors.

Secondly, the method of gathering the data chosen is web-based one-to-one interviews. However, due to the pandemic situation and high load of some experts involved, it was decided to have not only web-based one-to-one interviews but allow respondents to answer the questions in a written but detailed form. The interview method is chosen as the most convenient for the respondents and because interview is the most common method of researching this topic; some of the previous studies (Russia, 2015; Indonesia, 2019; Saudi Arabia, 2016) were conducted under the same approach. The experience and expertise of actual data experts are of great demand when it comes to analysing the details the bank must take into account when applying BD in a new region. Qualitative method dominates in this

kind of research since the challenges that are connected to BD implementation in a new region usually include detailed data and specific issues. In conclusion, the semi-structured interview makes it possible to collect more details and accurate answers (Wilson, 2014).

Finally, the sampling method that is in use is purposive sampling. It means that the participants of the research were chosen strategically, so that those people interviewed are relevant to the research questions being posed, have enough expertise and knowledge to make adequate and relevant conclusions. However, this kind of sampling does not allow to generalize to a population. (Bryman, A. & Bell, E., 2011) IT managers, Data scientists, Data quality experts, and CRM managers who have experience with BD or work projects that involved BD-based solutions, usually represent the sample of such interviews and this research is not an exception.

The way the OTP Bank Group became the focus of the research could be described in a few steps. First, the author of the paper chose the region. The choice has fallen on the Central and Eastern Europe (CEE) because the region has not been researched yet, has relatively stable development in terms of macroeconomic trends as well as in the banking sector and the most importantly the region has a trend of growing digital transformation (Deloitte, 2018). Secondly, to learn about the country-specific features there was a need for a banking group that operates in different countries. Deloitte report (2018) has defined the 15 leading groups in the CEE region and the author of the paper has contacted them all. The list of the banking groups contacted and the reasons for rejection could be found in Appendix A. In most cases the main reason was strict security requirements and rules in terms of data protection. Finally, the OTP Bank in Hungary accepted the cooperation and discussed the interview details with the author via the Email communication channel. In total, it was decided to arrange six interviews, having 1 interview in each country with IT managers, Data quality experts and/or CRM managers. Nevertheless, there were conducted seven interviews

online and two experts gave their explicit answers in written form. The following Table 8 is the representation of the sample.

Table 8

Overview of study participants and interviews

Interviewee	Position	Country	Length & format	Experience (years)
1	Head of CRM department	Bulgaria	38min; online	15
2	Senior marketing research analyst	Bulgaria	30min; online	3
3	Senior director of Sales and Marketing department	Croatia	Written format	10
4	Head of Data science and Business solutions department	Hungary	Written format	10
5	CRM department manager	Romania	44min; online	15
6	Director of Retail Risk advanced modelling & Analysis	Romania	44min; online	15
7	Head of CRM department	Russia	32min; online	10
8	Modelling and BD analysis department manager	Russia	33min; online	9
9	CRM and Data science manager	Serbia	48min; online	12

Source: compiled by the author

In order to conduct interviews, the plan was crafted. It has four core themes that are derived from the literature review part: definition of BD, how BD is used in banking, challenges in implementing and managerial perspective (see Appendix B). Challenges in implementing theme, which is of the most interest for this paper, has been divided into four sub-themes that were derived from the literature interview: Legal, Technological, Business and Social aspects. The division into sub-groups was done in order to collect more structured and precise data to fulfil the research aim. Findings from the literature review (see Table 9) were the basis for making up interview questions; the initial version of the interview was given to the Head of OTP Data Science department in Hungary who has relevant experience and expertise for the research.

Table 9

Examples of the main obstacles of Big Data introduction in each dimension

Business	Legal	Social	Technology
<ul style="list-style-type: none"> • Skilled team • Adoption costs • Readiness level • Scarcity of resources • Conversion into profit 	<ul style="list-style-type: none"> • Undeveloped regulations • Institutional barriers • Support from the government 	<ul style="list-style-type: none"> • Trust and ethicality 	<ul style="list-style-type: none"> • Data accuracy • Emerge of new technology • Storage, application and processing

Source: compiled by the author

His review and comments on the survey prompted the author of the paper to re-think the structure of the interview and questions included. Moreover, where not relevant questions were eliminated and some changes to the structure were made. The interview questions were distributed in two languages: English and Russian. It was translated in the following manner: English-Russian-English, to discern whether the translation did not distort the meaning of questions. Each interview started with a short introduction to the topic and an example of country-specific challenges of BD introduction. The first part of the interview included questions about the interviewees' understanding of the BD concept and what solutions are used in the banking industry in their country. These parts are further analysed as Theme 1 and Theme 2 correspondently. The second part of the core questions was split between the four categories for research: Technological, Legal, Social and Business aspects that could be described as external challenges. For each part the author has identified keywords for the content analysis that would lead to identifying the external challenges faced in particular countries; this part will be referred to as Theme 3. Besides the challenges found in the literature review, the author of the paper came up with the questions that identify the manager's perspective, the way people perceive the necessity, importance, and complexity of

BD implementation. This group of questions referred to as Theme 4. The closing part of the interview was a checking point of any missing or novel challenges that were not addressed in the interview.

2.3. Results and discussions

This subchapter will talk about the results from the interviews linking them together with findings from the literature review. Subchapter will follow the Themes that were mentioned before and analysis of the findings.

Definition of Big Data. The specialists who took part in the research agreed with the definitions provided by the author of this paper in Table 1. However, in terms of banking, the main characteristics of BD mentioned by six out of eight interviewees were that the data should be unstructured and there should not be a possibility to process it by the usual statistical means. Another crucial point pointed out was that the results from the processing should give a broader perspective on a problem. The definitions that are given by experts slightly vary but there is no contradiction. Moreover, there are several points that were mentioned by all of them: voluminous and complexity of data. This brings the author of the paper to the conclusion that even though there is no one precise definition of BD, the experts develop their own understanding of it, and it goes along with the very first definition given by Doug Laney in 2001.

Big Data implementation in banking. This theme was split into three categories after analysing the keyword from the interviews. The specialists provided the author with interesting insights into how Big data could be used in banking. It is worth mentioning that only three banks out of six presented in this research are utilizing the Big Data based solution (Russian, Hungarian and Serbian). However, other experts have related expertise and knowledge that is why they as well brought up exciting examples. The summary results of

the literature review (see Table 2) are aligned to the results of the author's findings (see Table 10)

Table 10

Big Data implementation in banking, keywords by categories

Categories	Keywords
Examples	Next best offer, compliance, risk management, transactions analysis, client segmentation, real time decisions, behaviour analysis
Necessity	Faster decision-making process, innovative way, not yet necessary, decrease employees' effort, cost efficiency.
Benefits	Address business problems, powerful calculations, flexibility, faster data processing, activity optimization.

Source: compiled by the author based on interview answers.

The most common use of BD is risk and behaviour analysis. As for the necessity of Big Data based solutions, the results were quite surprising, speakers do not believe it is yet necessary for banks. The main argument for not using BD is that banks do not generate that much data in real-time and for now banks can easily cope with processing data using a classical statistical approach. However, experts emphasised that the BD approach is innovative, and it opens many future opportunities. This fact has been proven by Sen, Ozturk and Vayvay (2016), they stressed that BD has potential but is not an immediate or necessary solution for banks yet. Brown, Chul and Manuika (2011), Kaisler, Armour, Espinosa and Money (2013) highlighted the same benefits that were mentioned by speakers.

In general, all authors and interviewees agree that BD is an innovative and exciting way to perform day to day analysis. Moreover, implementation of it in banking is not necessary yet but if the organization decides to do so it can address business problems in a more efficient way and perform better results.

Challenges of Big Data implementation. As for the technological side, the author did not find any country-specific peculiarity but the one that depends on the level of development of a bank. Departments that started implementing or that have already implemented the BD based solutions stressed that the main problem is the transformation of data with other companies. While those banks that only plan to implement highlighted three main concerns: accuracy, transformation, and infrastructure (see Table 11).

Table 11

Challenges in Big Data implementation in countries under study

Country	Business challenges	Legal challenges	Social challenges	Technological challenges
Bulgaria	Cost of adoption, future profitability	GDPR	Digital trust	accuracy, transformation and infrastructure
Croatia	Lack of specialists	GDPR	Digital trust	accuracy, transformation and infrastructure
Hungary	-	GDPR	-	IT solutions are not designed for managing these kinds of needs.
Romania	-	GDPR	Low degree of understanding	storing data; accuracy of data and transformation
Russia	Lack of specialists	No explicit interpretation of the Law	after pandemic Digital trust	transformation of information within different banks and countries
Serbia	Lack of specialists	The Law is not complete, low compliance	Digital trust	awareness of companies regarding the sharing of data

Source: compiled by the author

From the legal aspect (see Table 11), the peculiarity is different legal approaches and attitudes towards law. Countries that are in the European Union see GDPR as the main obstacle towards innovation, some experts mentioned that the law prevents companies from using BD solution because it is strict and uncompromising. Russia and Serbia share a similar

opinion about the local laws, experts from both countries see that the legislation is not complete and could be interpreted in different ways, that is misleading and could be considered as an institutional barrier. All in all, it is very important to study the legal aspect when implementing BD solutions in any organization, especially in banking because all the above-mentioned countries see it as a huge burden but in different ways.

As it is shown in Table 11, the most common issue for the most countries is that there are not enough specialists of the desired level or with the desired experience. However, in Eastern and Asian countries this problem seems to be minor or does not exist. The author of the paper believes that this challenge could be a regional peculiarity. Finally, those countries that have enough specialists consider the cost of adoption as the main problem.

Degree of concern about how the data is used and the purpose of usage is the main issue in most of the Eastern European countries from the social perspective (see Table 11). However, specialist highlighted that the public still does not fully understand how the data could benefit them. As was stated before, the main purpose is accurate marketing, so that the public can receive better offers from the bank. The author of this paper believes that this challenge could be eliminated with educational tutorials from the banks. Again, the challenge is regional, not country-specific.

Managerial perspective. Theme 4 brought up three categories: motivation of employees, flexibility and internal challenges.

Motivation. All interviewees stressed that there is higher motivation from younger specialists to work with innovations. However, they all believe that implementing such new technology as BD keeps workers up to date in terms of skills and knowledge needed on the market. This brings the author of the paper to the conclusion that BD based solutions implementation could raise greater interest from younger generations to work in banks and keeps the internal development of employees on the desired level.

Flexibility. As for now banks only see BD for the purpose of serving customers.

Nevertheless, the literature review stated that there are internal benefits for the companies: competitive advantage, simplification of business and operation models, and risk, capital and regulations management (PwC,2020; Indriasari, Gaol., & Matsuo, 2019).

Internal challenges. The complexity, lack of internal resources and evidence of profitability became the most common challenges. Moreover, in banks where BD based solutions are not yet implemented, the experts see the challenge in gathering the necessary amount of data while experts from banks that have already been using BD do not mention it as a burden. All in all, Theme 4 leads to the conclusion that countries studied have similar problems and there are no peculiarities except those that depend on whether the bank only in the beginning phase of implementing BD or has already been successfully using it.

The author of the paper thinks that because countries researched are in the same geographical zone, have pretty similar economic development and in terms of social-cultural attitudes do not differ much, they have corresponding challenges. The analysis of the findings leads to that there are regional specific challenges, not country-specific. However, the author believes that it is still important to research further, and the results cannot be generalized.

Conclusion

Having done theoretical and empirical analyses the author of the current paper was able to discover several findings of Big Data as a concept. In the first place, several definitions were presented, attributing to different viewpoints of the BD concept. The first modern definition was given by Doug Laney (2001) who assigned three key characteristics: volume, velocity, and variety contributing to the general understanding of what data could be considered as Big. However, the concept and its definition have been developing over the past decades taking into account technology change. Now, most researches consider not only three but seven characteristics of data in order to call it Big. Nevertheless, there is no one unique definition of the concept, different authors perceive the concept differently but based on the given characteristics.

At the stage of theoretical research on the basis of previous studies, the author found that there are various ways how BD can be used in business and banks in particular. Not all banks are currently utilizing BD based solutions, most of them are only at the beginning of this exciting journey. However, BD is a promising solution to the day to day business problems and can give banks more reliable information about the clients, thus lead to better business decisions and prosperity.

Besides studies that in different ways touched upon challenges and benefits of BD implementation in banks in several countries were introduced. The current bachelor thesis contains the results from the studies in terms of challenges. Findings show, that there are peculiarities in BD introduction that depend on a country: its political regime, social perception of innovations, and technological and economic development.

After deriving necessary theoretical insights for readers author has introduces the sample of nine interviews (seven in the online format and two in the written format). The sample included specialists from the IT and Marketing fields of one organisation (OTP Bank

Group) but in different countries with relevant experience and expertise thus being representative for the research. Prior to sample, the author has presented an overview of OTP Bank Group including financial highlights. Moreover, several indexes that are related to the four dimensions (business, social, legal and technological) were chosen in order to give the reader viewpoints of the current situation in countries studied. From the interviews acquired author at first has done thematic analysis through coding.

From the following empirical analysis, author was able to derive several important insights that could be used for further investigation and analysis. In the first place, not all interviewees agree that it is necessary for the banks to employ BD solutions. Nevertheless, they recognize it as an innovative and advantageous opportunity for the bank. The main reasons why not to introduce BD are that banks do not create much of real-time data and for now usual statistical approach copes with all the challenges. Secondly, technological challenges cannot be described as country-specific and depend on the innovation level in the bank. Additionally, the legal aspect imposes the most burden in terms of BD use and introduction. In the EU, the GDPR prevents the innovative approach in banking since the main mission is to protect the privacy of the population. Talking about the social perspective, it is worth mentioning that all countries have the same pattern, interviewees suggest that the public does not understand the concept thus does not have trust in BD. Finally, certain internal challenges should be addressed in banks: motivation and development of employees, internal flexibility and readiness level. Those aspects make it easier to introduce BD solutions in the bank.

All in all, the author of the paper concludes that those countries studied do not have distinctive challenges and all the obstacles mentioned by interviewees could be classified as regional. As one of the proposals for future research of the topic, the author recommends expanding the geographical region of research, the number of interviewees and conducting

quantitative research of this study. The author of the current paper investigated the obstacles that may occur when introducing BD solutions in a new department of the bank but a good continuation would be to analyse the readiness of the bank sector to employ new technologies and innovations in different regions, to study not only obstacles but benefits as well.

References

1. Almoqren., N. & Altayar, M. (2016). The Motivations for Big Data Mining Technologies Adoption in Saudi Banks. *2016 4th Saudi International Conference on Information Technology (Big Data Analysis) (KACSTIT)*, 1-8.
2. Annual Report Romania. (2019). Retrieved from <https://www.otpbank.ro/sites/default/files/assets/documents/WEB-Raport-Financiar-Anual-EN-small.pdf>
3. Bloomberg. (2020). Retrieved from <https://www.bloomberg.com/quote/OTP:GR>
4. Bol'shiye dannyye v finansovoy otrasli: obzor i otsenka perspektiv razvitiya mirovogo i rossiyskogo rynkov [Big data in the financial industry: a review and assessment of the development prospects of the global and Russian markets]. (2015). *IDC Russia*.
5. Bollier, D., (2010). *The promise and peril of big data*. The Aspen Institute.
6. Brown, B., Chul, M., Manyika, J., (2011). Are you ready for the era of “big data”? *McKinsey Quarterly*,4.
7. Bryman, A. & Bell, E. (2011) *Business Research Methods* (3rd ed.). Oxford University Press
8. Chen, H., Chiang R. H. L. & Storey, V. C., (2012). Business intelligence and analytics: from big data to big impact. *MIS Quarterly*, 36 (4), 1165-1188
9. Deloitte. (2018). CEE banking M&A dynamics remain stable. Retrieved from <https://www2.deloitte.com/content/dam/Deloitte/cz/Documents/financial-services/cz-cee-banking-mergers-and-acquisitions-study-2018.pdf>
10. Dutta, S., Lanvin, B. & Wunsch-Vincent, S. (2019). *Global Innovation Index 2019*. The Retrieved from <https://www.globalinnovationindex.org/gii-2019-report>
11. Forest, H., Foo, E., Rose, D. & Berenzon, D., (2014). Big Data how it can become a differentiator. *Deutsche Bank AG From Edge to Core*.

12. Ghasemaghaei, M. & Calicn, G., (2019). Can big data improve firm decision quality? The role of data quality and data diagnosticity. *Decision Support Systems*. 120. p.38-49.
 13. Giacalone, M. & Scippacercola, S., (2016). BIG DATA: ISSUES AND AN OVERVIEW IN SOME STRATEGIC SECTORS. *Journal of Applied Quantitative Methods*. 11. 3. P.1-17.
 14. Gutierrez, D. D., (2014). Inside BIG DATA Guide to Big Data for Finance. *Dell & Intel*.
 15. History. (n.d). Retrieved from <https://www.otpbank.hu/portal/en/AboutUs/History>
 16. IBM Marketing Cloud Report. (2017). Retrieved from <https://www.ibm.com/products/software>
 17. Indriasari, E., Gaol, F., L. & Matsuo, T. (2019). Digital Banking Transformation: Application of Artificial Intelligence and Big Data Analytics for Leveraging Customer Experience in the Indonesia Banking Sector. *2019 8th International Congress on Advanced Applied Informatics (IIAI-AAI)*.
 18. Jaiswal, A. & Bagale, P., (2017). A Survey on Big Data in Financial Sector. 2017 *International Conference on Networking and Network Applications (NaNA)*, 337-340.
 19. Kaisler. S., Armour, F., Espinosa, A. & Money, W., (2013). Big Data: Issues and Challenges Moving Forward. *2013 46th Hawaii International Conference on System Sciences*, 995-1004.
 20. Khan, M., Uddin, M. & Gupta, N., (2017). Seven V's of Big Data understanding Big Data to extract value. *IEEE*.
 21. Khan, N., Yaqoob, I., Hashem, I., Inayat, T., Ali, Z., Shiraz, W. & Gani, W., (2014). Big Data: Survey, Technologies, Opportunities, and Challenges. *The Scientific World Journal*.
-

22. Kotler, P. (1994). *Marketing Essentials*
23. Kshetri, N., (2016). Big data's role in expanding access to financial services in China. *International Journal of Information Management*, 36(3), 297-308

24. Kuper, A., Lingard, L., & Levinson, W. (2008). Qualitative Research: Critically Appraising Qualitative Research. *BMJ: British Medical Journal*, 337(7671), 687-689.
25. Laney, D., (2001). 3D data management: Controlling data volume, variety and velocity. *MetaGroup research publication*
26. Lemon, K.,N. & Verhoef, P.,C. (2016). Understanding Customer Experience Throughout the Customer Journey. *AMA/MSI Special Issue*, 80, 69-96.
27. Manyika, J., Chui, M., Bughin, B., Dobbs, J., Roxburgh, C. & Byers, A., (2011). *Big Data: The Next Frontier for Competition, the McKinsey Foundation* Retrieved from https://bigdatawg.nist.gov/pdf/MGI_big_data_full_report.pdf
28. McGovern, P., Lambert, J. & Verrecchia, M. (2019). Mobile Banking Adoption: An Exploration of The Behavioural Intention of Consumers in Ireland. *Journal of Accounting and Finance*, 19(8), 11-30.
29. Mullan, J., Bradley, L. & Loane, S. (2017). Mobile banking adoption: future channel usage, services and platforms. *Journal of Current Issues in Media and Telecommunications*, 9 (4), 369-377
30. Nguyen, T., Zhou, L., Spiegler, V., Ieromonachou, P. & Lin, Y., (2017). Big data analytics in supply chain management: A state-of-the-art literature review. *Computers & Operations Research*, 98, 254-264.

31. OTP Bank. (2019). Retrieved from https://en.wikipedia.org/wiki/OTP_Bank
32. OTP Banka. (n.d). Retrieved from <https://www.otpbanka.hr/en/o-nama/otp-bank>
33. OTP in Russia (n.d). Retrieved from <https://www.otpbank.ru/en/about/>

34. Overview of Expressbank. (n.d). Retrieved from
<https://www.linkedin.com/company/expressbank-bg/about/>
35. Overview of OTP Bank (n.d) Retrieved from <https://www.linkedin.com/company/otp-bank/about/>
36. Overview of OTP banka Srbija.(n.d). Retrieved from
<https://www.linkedin.com/company/otpsrbija/>
37. Portela, F., Lima, L. & Santos, M.F., (2016). Why Big Data? Towards a project assessment framework. *Procedia Computer Science*, 98, 604-609.
38. PwC. (2020). Retail Banking 2020 Evolution or Revolution? Retrieved from:
<https://www.pwc.com/gx/en/banking-capital-markets/banking-2020/assets/pwc-retail-banking-2020-evolution-or-revolution.pdf>
39. Rahman, Z. (2005). Customer experience management — A case study of an Indian bank. *Database marketing and customer strategy management*, 13 (3), 203-221
40. Reinsel, D., Gantz, J. & Rydning, J., (2018). *The Digitization of the World*
41. Saxena, S. & Al-Tamimi, T. A. S. M. (2017). Big data and Internet of Things (IoT) technologies in Omani banks: a case study". *Foresight*. 19. 4. 409-420.
42. Schwab, K. (2019). *The global Competitiveness Report 2019*. Retrieved from
http://www3.weforum.org/docs/WEF_TheGlobalCompetitivenessReport2019.pdf
43. Sen, D., Ozturk, M. & Vayvay, O., (2016). An Overview of Big Data for Growth in SMEs. *Procedia- Social and Behavioral Sciences*, 235,159-167.
44. Shah, S., Soriano, C. B. & Coutroubis, D., (2018). Is big data for everyone? the challenges of big data adoption in SMEs. *IEEE International Conference on Industrial Engineering and Engineering Management (IEEM)*,803-807.
45. Shuijing, H., (2016). Big Data Analytics: Key Technologies and Challenges. *IEEE*.

46. Srivastava, A., Singh, S.,K., Tanwar, S. & Tyagi, S. (2017). Suitability of Big Data Analytics in Indian Banking Sector to Increase Revenue and Profitability. *2017 3rd International Conference on Advances in Computing, Communication & Automation (ICACCA) (Fall)*, 1-6.
47. Stocia, L.C. (2018). Big Data in Business. *Knowledge Horizons – Economics*,10(3),15-24.
48. Summary of the full-year 2019 results. (2020). Retrieved from https://www.otpbank.hu/static/portal/sw/file/200305_OTP_20194Q_e_final.pdf
49. The Digital Evolution Index 2017. (2017). Retrieved from https://globalrisk.mastercard.com/wp-content/uploads/2017/07/Mastercard_DigitalTrust_PDFPrint_FINAL_AG.pdf
50. Tiana, X., Hana, R., Wanga, L., Lu G. & Zhan, J. (2015). Latency Critical Big Data Computing in Finance. *The Journal of Finance and Data Science*, doi: 10.1016/j.jfds.2015.07.002.
51. Wilson, C. (2014). Interview Techniques for UX Practitioners. USA, Waltham: Morgan Kaufmann
52. Zhu, L., Yu, F. R., Wang, Y., Ning, B. & Tang, T. (2018). Big Data Analytics in Intelligent Transportation Systems: A Survey. *IEEE Transactions on Intelligent Transportation Systems*, 20 (1), 383-398.

Appendix A

List and state of contacted Banking groups

List of contacted banking groups	Cooperation state/ reason for refusal
OTP Banking group	Cooperation is accepted
Erste	Strict rules and security requirements
KBC	No answer
Unicredit	Strict rules and security requirements
Raiffeisen	Strict rules and security requirements
Societe Generale	Strict rules and security requirements
Intesa Sanpaolo	No answer
Commerzbank	No answer
Swedbank	Strict rules and security requirements
SEB	Strict rules and security requirements
Citibank	Strict rules and security requirements
BigBank	Strict rules and security requirements
LHV	Strict rules and security requirements

Source: compiled by the author

Appendix B

Interview questions

Thematic block	Interview question	Theoretical ground
Introduction	In what country do you work?	Compiled by the author
	What is your job title?	Compiled by the author
	What department do you currently work in?	Compiled by the author
	For how long you have been working in this department?	Compiled by the author
	For how long have you been working in Banking industry?	Compiled by the author
Definition of BD	How would you define Big Data?	Kuo and Li (2014); Ning and Tang (2018)
	What do you consider under BD solutions? Could you please bring examples?	Tiana, Hana, Wanga, Lua and Zhan (2015)
BD in banking	Do you consider it is necessary to apply BD based solutions in banking industry? Why?	Gutierrez (2014); Jaiswal and Bagale (2017)
	Did you directly work on the project that involved either BD or BD based solutions? What was it about? How did you contribute?	Compiled by the author
	Could you please name some ways of how BD based solutions are used in Banking sector?	Deutsche Bank (2014)
	What year the branch you are working at started utilizing BD based solutions?	Compiled by the author
Challenges in implementation	Does the process of BD implementation have any kind of support/ opposition from the governmental bodies/departments/ institutions?	Kshetri (2016)
	Was the cost of adoption one of the challenges in your subsidiary? Can it be described as a major one?	Indriasari, Gaol & Matsuo (2019)
	How do you think what is the attitude towards data protection from the society perspective in your country? Is there a degree of concern? Any trust issues?	Kshetri (2016)
	Do you believe that in your region there are enough specialists of the desired level? Do you consider it as a major problem?	Indriasari, Gaol & Matsuo (2019); Srivastava, Singh, Tanwar & Tyagi (2017)
	What are the challenges in terms of technology? Infrastructure, transformation,	Saxena & Al-Tamimi (2017);

	accuracy of data? Could you bring any particular examples?	Almoqren & Altayar (2016)
	Has the process of BD implementation faced any ethical issues from legal perspective? Could you bring any particular examples?	Compiled by the author
	Do new security regulations impose more pressure on the department and the use of BD in your region?	IDC Russia (2015)
	Are there any other challenges that could be described as external and that were not mentioned before?	Compiled by the author
Managerial perspective	How would you describe internal challenges when implementing BD solutions? What are the major ones?	Compiled by the author
	What kind of issues did the organization face when started implementing BD based solutions?	Compiled by the author
	How challenging it was (implementing) at the scale from 1 to 5 where one is “not challenging at all” and five is “the most challenging so far”?	Compiled by the author
	Do BD-based solutions contribute to more flexibility in banking industry?	Compiled by the author
	What challenges do managers face when it is time to implement any new solution and how do they solve those?	Compiled by the author
	What conditions triggered the company to implement BD?	Compiled by the author
	What conditions in the organization enabled the process of BD application to go smooth?	Compiled by the author
	How motivated are employees to work with BD?	Compiled by the author
Conclusion	Do you believe it is important to research the country specific aspects of BD implementation?	Compiled by the author
	Do you think there are any country specific aspects the company must take into account when applying BD based solutions?	Compiled by the author

Source: compiled by the author

Non-exclusive licence to reproduce thesis and make thesis public

I, Anna Shamritskaya,

1. herewith grant the University of Tartu a free permit (non-exclusive licence) to reproduce, for the purpose of preservation, including for adding to the DSpace digital archives until the expiry of the term of copyright,

CROSS-COUNTRY CHALLENGES IN THE IMPLEMENTATION OF BIG DATA
SOLUTIONS IN BANKING: THE CASE OF OTP BANK GROUP

supervised by Mark Kantšukov.

2. I grant the University of Tartu a permit to make the work specified in p. 1 available to the public via the web environment of the University of Tartu, including via the DSpace digital archives, under the Creative Commons licence CC BY NC ND 3.0, which allows, by giving appropriate credit to the author, to reproduce, distribute the work and communicate it to the public, and prohibits the creation of derivative works and any commercial use of the work until the expiry of the term of copyright.

3. I am aware of the fact that the author retains the rights specified in p. 1 and 2.

4. I certify that granting the non-exclusive licence does not infringe other persons' intellectual property rights or rights arising from the personal data protection legislation.

Anna Shamritskaya

11/08/2020