KAPITEL 1 / *CHAPTER 1*
THE BIG DATA DESIGN - A MODERN CHALLENGE IN DEVELOPING AN INNOVATIVE BUSINESS MODEL
АРХИТЕКТУРА БОЛЬШИХ ДАННЫХ - СОВРЕМЕННЫЙ ВЫЗОВ В РАЗРАБОТКЕ ИННОВАЦИОННОЙ БИЗНЕС-МОДЕЛИ

DOI: 10.30890/2709-2313.2021-06-05-027

Introduction

This paper tackles a relevant, and contemporary topic that, since 2010, has been one of the top technologies suggested, to help develop the international business environment, increase the competitiveness of companies through Business Analysis tools. The paper explains and analyzes various methods and analytical tools that have been applied to Big Data, in developing digital marketing strategies.

The objectives of this research are as following:

● Explaining the notion of Big Data and its operating principles at the enterprise level

● Elucidation of areas that already widely apply the interpretation of BD for strategic development and adaptation to the international business environment

● Analysis of the level of training (if any) of local companies, and implementation perspectives and challenges at the country level.

From corporate leaders to municipal planners and academics, big data has been the focus of attention , at the same time uncovering a new realm of studies. The sudden rise of big data has revealed the unpreparedness of society, entrepreneurs in their use and new areas of research.

New technological developments have appeared for the first time in technical and academic publications. Knowledge and synthesis subsequently infiltrated other ways of mobilizing knowledge, including books. The rapid evolution of big data technologies and the rapid acceptance of the concept by the public and private sectors did not leave too much time for the discourse to develop and mature in the academic field.

The leap of big data discourse to more popular media implies that a coherent understanding of the concept and its nomenclature is in its infancy. For example, there is little consensus on the fundamental question of how big the data should be to qualify as "big data". Thus, there is a need to analyze the rapid evolution of big data concepts and technologies.

A study by Singh[1] mapped the area of comics and noted the interdisciplinary nature of this field and the growth rate of the number of publications, authors, disciplines and countries involved in its development. A more recent study investigated collaboration interdisciplinary in BD research and found that the main contributors to BD research are Informatics, Engineering and Business and Economics as well as research communities are being formed on this topic. Another recent study covering seven years of BD publications showed the strong dynamics of journals and even more so, of conference publications in the field, attributing much of

[1] Scientometric Mapping of Big Data Research, Sumit Kumar Banshal, Khushboo Singhal & Ashraf Uddin

the activity to the broad interest in BD in various fields of research, as well as the strong interest of big and strong countries like China and the USA.

The modern web search tools allow us to view historical information that has been automatically extracted from various sources, and to present it in this interesting analysis of those terms that refer to the analysis of data itself and Big Data. Another notion of BI was chosen as a reference, with tangents to this topic. Business Intelligence (BI) is a technology-based process for analyzing data and providing practical information that helps managers, managers and employees make informed business decisions. Its purpose is to improve the functioning of the business and to offer better products and services to the public, while reducing the costs associated with business. Many of the tools described in this book will be used as part of business intelligence. Business intelligence is a general theme, in which the meaning and final direction of all the data that a business has, can be integrated into a coherent entity. As part of the BI process, organizations collect data from internal IT systems and external sources, prepare them for analysis, run data queries, and create data views, BI dashboards, and reports to make analytical results available to business users. for making operational decisions. implementation and strategic planning.[2] The ultimate goal of BI initiatives is to drive better business decisions that enable organizations to increase revenue, improve operational efficiency, and gain competitive advantage over business rivals. To achieve this goal, BI incorporates a combination of analysis, data management and reporting tools, plus various methodologies for data management and analysis.

1.1 The Big Data Paradigm

Before considering the definitions of Big Data, it is important to distinguish Big Data from two related concepts: information and communication technology (ICT) and "open data". For Osterwalder, "ICT encompasses all the technology that facilitates the processing, transfer and exchange of information and communication services"[3]. Therefore, ICT for Development (ICT4D) refers to how ICT "can be used to help poor and marginalized people and communities make a change in their lives"[4] with data exchange and more about the value that data brings. On the other hand, open data is accessible, digital and without restrictions on use or redistribution. Big data is not always open and sometimes will not be accessible without special skills or software. The term "big data" has its origins in the late 1990s, but became widespread only in 2009. Definitions of "big data" vary across industries such as information technology (IT), computing, marketing, social media, communication, storage data, analysis and statistics. People use the term to mean a wide variety of things that fit

[2] Hu, J. and Zhang, Y. (2017). Discovering the interdisciplinary nature of Big Data research through the analysis and visualization of social networks

[3] Schmidt, R., Möhring, M., Maier, S., Pietsch, J. and Ralf-Christian Härting (2014), "Big data as strategic facilitators-suggestions from Central European enterprises", in the International Conference on Information Systems business, 50-60 Springer,Cham

[4] George, G., Osinga, E.C., Lavie, D. and Scott, B.A. (2016), "Big data and data science methods for management research", Academy of Management Journal, Vol. 59 No. 5, pp. 1493-1507

their purposes. It was used to describe data tools, datasets, questions, problems, and answers.

In the past, processing power has been the limiting factor in the analysis of large data sets. This has changed radically since the beginning of computation. From the proliferation of transistors in the 1950s, to the creation of the first dynamic random access memory (DRAM) chip in 1966, to the spread of "cloud computing" today, advances in computing power have been reduced to massive improvements in four key areas. processing speed, data storage capacity (hardware), data analysis (using software) and connections between data sources and processors via the Internet.

● Processing speed: Although mechanical computing machines have existed since 1642, the invention of electronic components made it possible for modern computers to perform complex tasks quickly and accurately, dividing them into smaller ones.

● Today, IBM's "Watson " supercomputer can process 500 gigabytes of data per second, can be accessed from anywhere, and can be up to six centimeters high, or in the cloud (a remote server). Between 1952 and 1956, IBM developed the first hard drive, which could store five megabytes of data (a three-minute audio file by current standards) and was the size of a refrigerator. Technological advances in the 1970s and 1980s rapidly reduced unit costs and size, while increasing storage capacity.

Thanks to much lower storage costs, today's companies and governments are able to store large amounts of data without necessarily having a specific purpose in this regard. This was simply not possible in the relatively recent past.

● Data analysis: New and flexible software has been specially developed to manage large and non-uniform data sets. Analyzing the data used to require spreadsheets and databases can now take any number of structured or unstructured forms. An increasing number of companies specialize in this field, providing customized data analysis for companies, governments and non-governmental organizations (NGOs).

● The information ecosystem: Increasing our ability to share data and the value we can get from computer networks can be summed up by Gilder's Law and Metcalfe's Law, respectively. In 1997, Gilder proposed tripling the internet bandwidth each year for the next 25 years. Metcalfe, responsible for inventing Ethernet communication, suggested that the value of any network is proportional to the square of the number of connections that make it up.

Big Data can be examined on two levels. At a fundamental level, it is just another collection of data that can be analyzed and used for the benefit of the business. At another level, it is a special type of data that presents unique challenges and offers unique benefits. This is the level on which this book will focus. At the business level, data generated by business operations can be analyzed to generate information that can help the business make better decisions. This makes the business unstructured. Each type of Big Data is structured differently and therefore must be treated in its own way. There are huge opportunities for technology providers to innovate and grow management and generate even more data, and the cycle continues.

At another level, Big Data is different from traditional data in all respects: space,

time and function. The amount of Big Data is 1000 times higher than that of traditional data. The speed of data generation and transmission is 1,000 times faster. Big Data forms and functions are 10 times more diverse: from numbers to text, images, audio, videos, web logs, car data, and more. There are many more data sources, from individuals to organizations and governments, using a wide range of devices, from mobile phones to computers and industrial machines. Not all Big Data is of the same quality and value. Big Data is largely, over 90%, unstructured data.

Big data marks the beginning of a major transformation. New techniques for collecting and analyzing huge bodies of data will help us understand our world in ways we are just beginning to appreciate. In this book we are not so much the big data evangelists, but only his messengers. And again, the real revolution is not in the machines that calculate the data, but in the data itself and how we use it. In order to assess the extent to which an information revolution is already under way, trends across the whole spectrum of society must be taken into account. Our digital universe is constantly expanding.

Big Data includes all the data, about all the activities, anywhere. Thus, it can potentially transform our perspective on life and the universe. It can bring new perspectives in real time, it can make life happier and it can make the world more productive. Big Data can still be dangerous - in terms of breaches of confidentiality and social and economic disruption. There are three main categories of data sources: human to human communication, human-machine communication and machine-machine communication.

● human to human communication: People and corporations communicate more and more instantly through electronic networks. Everyone communicates by phone and email. For example, high-resolution cameras in mobile phones allow people to record pictures and videos and instantly share them with friends and family. All these communications are stored in the facilities of many intermediaries, such as telecommunications and internet service providers. Social media is a relatively new and transformative type of human-to-human communications. Social networks, social networking platforms such as Facebook, Twitter, LinkedIn, YouTube, Tumblr, Skype, Snapchat and others have become an increasingly intimate part of modern life. People instantly share messages and images via social networks like Facebook and YouTube. All of these data streams are part of Big Data and can be monitored and analyzed to understand many phenomena, such as communication patterns and the essence of conversations.

●Human-machine communication in this case is achieved even through the graphical interface. The person (operator of the machine) enters data that the software transmits to the machine in the form of commands, which it executes. Thus, the results are obtained. The graphical interface is the one that mediates the human-machine connection.

●Machine to Machine (M2M) Communications: M2M communications are also sometimes referred to as the Internet of Things (IoT). One trillion devices are connected to the internet and communicate with each other or with some main machines. All this data can be accessed and used by the manufacturers and owners of these machines. Machines and equipment have many types of sensors to measure

certain environmental parameters, which can be broadcast to communicate their status.

Due to constantly evolving data sources and increasing amounts of data generated, companies face serious problems in achieving the integration of high quality data. These challenges can also be called "The 4 Vs of Big Data". They are the veracity of the data, the volume, the variety and the speed.

Data veracity

Truthfulness refers to the veracity, credibility and quality of data. Big Data is messy. Reasons for poor data quality can range from technical errors to human errors to malicious intent. It also refers to incomplete data with the presence of errors and outliers. Such data creates a great challenge for companies, as it must turn them into a consistent, consolidated, united source of information and business intelligence.

The main precondition for the lack of veracity of the data and the reason why the data become incomplete and full of abnormal values are the data silos. The existence of silos ensures the unequal distribution of data in the departments of the organization, as well as the duplication of data sets.

Data volume

Traditional data is measured in Gigabytes (GB) and Terabytes (TB), but Big Data is measured in Petabytes (PB) and Exabytes (1 Exabyte = 1 Million TB). This data is so huge that it is almost a miracle that you can find anything specific in it in a reasonable period of time. Web search was the first real Big Data application. Google has perfected the art of this application and developed many of the state-of-the-art Big Data technologies we see in use today. The main reason for the increase in data is the dramatic reduction in the cost of data storage. Data storage costs have dropped by 30-40% each year. Therefore, there is an incentive to record everything that can be observed. It's called "datafication" of the world. Data calculation and communication costs have also decreased.

Growing business entities often embrace new systems, devices, platforms, AI or IoT networks to generate new opportunities, success and optimized performance. On the other hand, the adoption of digital culture contributes to the integration of many social networks and digital channels to facilitate communication between employees, leaders, customers and departments.

However, adopting so many data channels and collecting the entire volume of data generated creates problems for organizations. Overwhelming volumes make it difficult to process and integrate data and, at the same time, require large storage, which can generate huge costs for companies. Storing too much data is not an opportunity, but a disadvantage for data-driven companies. Identifying specific data sets makes it difficult: as if you were looking for a certain red T-shirt in a huge closet with a thousand more red T-shirts. That is why organizations face the challenge of identifying the optimal storage of data, they can manage: both in terms of financial capacity and operational capacity.

Variety of data

Omniform and omnichannel data collection brings many benefits to cross-industry organizations: it identifies opportunities, measures progress, makes evidence-based decisions, strategic plans, and effectively measures risk. However,

these data create challenges for digital culture organizations.

Big Data includes all forms of data, for all types of functions, from all sources and devices. If the traditional data from invoices and registers was like a small room full of data, Big Data is like the biggest mall imaginable, that offers an unlimited variety. There are three major types of data variety. The first aspect of variety is the shape of the data. Data types range from numbers to text, graphics, map, audio, video, and more. Some of these types of data are simple, while others are very complex. There are also composite data types,that include many elements in a single file. For example, text documents include graphics and images embedded in them.

The second aspect of variety is the function of the data. There are data from human conversations, songs and movies, recordings of commercial transactions, data on the performance of machines and operations, data on the design of new products, archived data on the old, etc. Human communication data should be processed very differently from operational performance data, with completely different data expectations and objectives. Big Data technologies could be used to recognize people's faces in images; and compare voices to identify the speaker; and compare handwriting to identify the writer

Data speed

This term represents how fast data is collected, stored and processed. Often, when the volume of data, as well as its variety, are large, they have a significant impact on speed.

Big Data is generated by billions of devices and communicated at the speed of light, via the internet. Ingesting all this data is like drinking from a fire hose. One has no control over how fast the data will come. A huge unpredictable data flow is the new metaphor for thinking about Big Data. The main reason for the increased speed of data is the increase in internet speed. The speed of internet available for homes and offices is now increasing from 10 MB / sec to 1 GB / sec (100 times faster). More people have access to high speed internet around the world.

Thus, slow data processing poses challenges for data-driven companies, as slow data speeds lead to operational and process delays, which in the long run have a negative impact on organizations' strategic planning.

Every business entity that has implemented digital culture or uses Big Data as the main tool for successful development and performance faces the contemporary challenge of ensuring the integration of high quality data. But to provide a solution to this challenge, companies should analyze the veracity, volume, variety and speed of data. This will help companies to obtain high quality, reliable, fast and accessible data assets and will motivate them to make the most of their data analysis.

Nowadays, companies in distinct industries recognize Big Data as a key asset for successful strategic decision making. And it is an organizational responsibility to ensure the quality and speed of this asset. Otherwise, the business bears the risk of using outdated information, which is incompatible with the immediate needs and requirements of the organization.

1.2 Areas of applicability of BIG DATA

Big Data is changing the very nature of business, from manufacturing to healthcare retail, agriculture and beyond. The rate at which data is,and can be collected for every conceivable activity, means that there are increasing opportunities to fine-tune procedures and operations to squeeze and use information efficiently.

Industry influencers, academics and other major players agree that Big Data has become a game changer in most, if not all, types of modern industries in recent years. As Big Data continues to penetrate our daily lives, there has been a significant shift in focus from the hype surrounding it to finding real value in its use.

While understanding the value of Big Data remains a challenge, other practical challenges, including financing and return on investment and skills, remain at the forefront for many different industries adopting Big Data. That being said, according to research and market reports, in 2017 the global Big Data market was worth $ 35 billion. dollars and by 2027 is expected to reach $ 105 billion. (Fig 1.2.1)

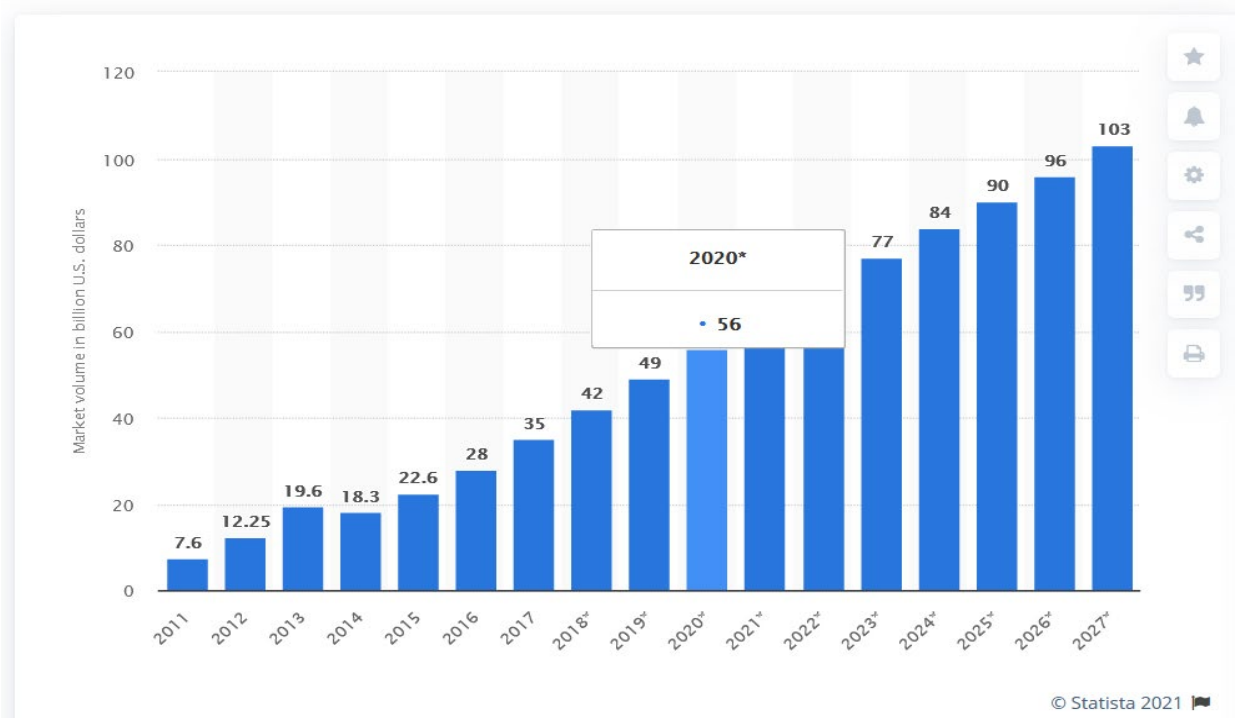

Figure 1.2.1 Global Big Data Market Size Revenue Forecasts Between 2011 and 2027 (in US $)

Sourse: https://www.statista.com/statistics/254266/global-big-data-market-forecast/

In general, most organizations have several goals for adopting Big Data projects. While the main goal for most organizations is to improve the customer experience, other goals include cost reduction, better targeted marketing and streamlining existing processes. Lately, data breaches have made increased security an important goal that Big Data projects are trying to incorporate.

Each application area has its own characteristics and requirements when it comes to data management. For the purpose of this research, the most challenging and popular application areas were selected to be analyzed in more detail: banking and fintech, healthcare, insurance, government and public sector, business and

commerce and communication and marketing. In this subchapter, an overview of the features and challenges related to big data management in those application areas will be presented, followed by an analysis of the most important challenges and issues related to database management, as well as opportunities in those areas.

The Big Data analysis in the banking market is expected to register a CAGR of 22.97%, in the period 2021-2026. The main drivers for the adoption of Big Data analysis in the banking sector are the significant increase in the amount of data generated and government regulations. As technology advances, the number of devices that consumers use to initiate transactions also proliferates (such as smartphones), which increases the number of transactions. This rapid growth of data requires better acquisition, organization, integration and analysis.

According to the Open Banking Implementation Entity (OBIE), API calls increased from one million per month in May 2018 to over 66.7 million in June 2019. PSD2 requires banks to create APIs for grouping and sharing data sets. discrete between organizations) for digital banking transactions.

JPMorgan Chase and Co. is the largest bank in the United States and the sixth largest in the world. Thanks to a large customer base of over 3 billion euros, a large volume of credit card information and other transactional data of its customers is created.

The market is also witnessing various investments from well-known investors and the government. For example, the Big Data Europe Project [57] (2015-2017) received funding from the European Union's Horizon 2020 research and innovation program. Its main goal has been to develop prototypes of Big Data applications in industries that can provide large data sets, that have an urgent need to progress towards data-driven solution approaches.

Big data analysis can help banks understand customer behavior based on contributions received from a variety of perspectives that include investment models, shopping trends, motivation to invest, and personal or financial funds. As big data analytics improve, banks can analyze market trends and make decisions about lowering or raising interest rates for different individuals in different regions. With the help of big data analytics, as the number of electronic records increases, financial services actively use it to store data, obtain business information, and improve scalability.

Financial organizations around the world lose more than 5% of their annual revenue due to fraud. While direct losses due to fraud amount to a large million dollars, the real cost is much higher in terms of lost productivity and lost customer confidence (and possible wear and tear). Various losses due to fraud remain undetected. With $ 5.7 billion in global money laundering fines issued in 2019, increasing sophistication of threats and rising compliance costs, financial institutions need advanced analysis to deter financial crime.

Big Data analysis for fraud prevention is determined by metadata; all records are exhaustively linked based on the combinations of attributes in the data. Using statistical techniques, collective entities are identified and collapsed to produce

[5] https://www.big-data-europe.eu

unique views of networked entities. Discrete, delimited networks within the data could also be generated, which helps to represent the relevant statistical groups of activities and relationships.

Another method is widely used to obtain relationships between fraudulent activities, including multiple suspicious activities in a single account or similar action patterns in different accounts. In-depth analysis looks for similarities that help indicate fraud between transactions or sets of transactions. Relationships are increasingly complex, so they often avoid simple monitoring techniques.

The process of predictive analysis will go through many stages. Like any activity within a large organization, it will start by defining the problem. The problem needs to be defined precisely, so that IT models and systems have something to address and all members of the organization understand the objectives. The scope of the effort should be more limited in order to obtain more reliable results.

Healthcare data sets consist of data on physiological and other exposure, medical imaging, disease management and the like, which are highly disorganized and distributed.[8] The source data is largely collected from clinical and administrative databases and electronic medical records. Specifically, BDA in healthcare has many potentials, such as integrating health information to discover new patterns and help physicians discover diseases in their early stages or accurately predict their trajectory. Moreover, the integration of BDA-enabled clinical data results in many benefits, such as improved health process efficiency and quality of care, and reduced health care costs. However, big data in healthcare also faces some challenges:

- the idea of applying big data in the field of healthcare is relatively young, so there is a lack of knowledge and experience in recognizing the data to use and how to use it; -lack of adequate IT infrastructure; -processing information without human supervision could lead to incorrect interpretations;
- technical challenges, such as the integration of different types of data from a variety of sources, fragmented data and inaccurate data;
- Patients' privacy concerns may arise due to the exchange of data between stakeholders.

These challenges are usually addressed by implementing Big Data governance, developing an information exchange culture, using security measures, and training key personnel to use BDA. Wang and Hajli conducted a study of 63 health care organizations to explore the relationship between BDA capabilities and business value [6]. In their research, Wang, Kung, and Byrd grouped BDA benefits into five categories[7]:

1. Benefits of IT infrastructure;
2. operational benefits;
3. organizational benefits;
4. managerial benefits;
5. strategic benefits

[6] N. Mehta and A. Pandit, “The Competition of Big Data Analysis and Healthcare: A Systematic Review.”, Int. J. Med. Inform., Vol. 114, pp. 57-65, 2018

[7] Y. Wang and N. Hajli, "Exploring the Path to Successful Health Data Analysis", J. Bus. Rez., Vol. 70, pp. 287-299, 2017.

The functional approach of predictive analysis in healthcare is quite numerous. A common problem that occurs in hospitals are secondary infections. Patients may enter with a viral or other disease and then contract bacterial pneumonia while in the hospital. This issue is ripe for machine learning, provided the data is available. Big data could be used in conjunction with machine learning to develop models that would estimate a patient's risk of developing secondary infections. The hospital may take steps to reduce the risk, such as keeping the patient's exposure to others to a minimum during their hospital stay. Healthcare staff and other health workers could also implement additional measures to protect and wash their hands. Another application in the field of healthcare is to determine who is at risk of various diseases before the onset of symptoms. For example, patients could be tested to determine their risk of developing diabetes or heart disease. Big data analysis could be used to find patterns in data for breast and prostate cancer, identifying patients at high risk of developing these diseases using previously unknown hidden patterns so that additional intervention or screening can be applied.

Both big data and predictive analytics can be used to improve hospital staff. For example, we could determine the best times or days to increase staff using predictive analytics and determine where to allocate staff.

Big Data is expected to affect insurance in several ways. The most anticipated is data analysis, although what this means is not always clear from covering this topic. The second is to subscribe and set prices, with different views on how much big data could affect them. Distributions and sales are more obvious ways, given how big data could allow for better targeting and understanding of consumer behavior. Complaints handling could be simplified using big data, and marketing could be better targeted with big data.

In the insurance sector, it is sometimes useful to distinguish between "costs" (calculation of the technical premium) and "pricing" (the actual business decision to offer a policy at a certain premium level). Insurance rates should be based on predictions rather than actual costs. Most rates (costs) are determined by the statistical analysis of past losses based on the specific variables of the insured. The variables that produce the best forecasts are the criteria according to which the first ones are established. However, in some cases, historical analysis does not provide sufficient statistical justification for selling a rate, such as earthquake insurance. In these cases, disaster modeling is sometimes used.[8]

The emergence of big data has given rise to the possibility of risk-based pricing being applied more and more, as more data could improve the predictability of policyholders' behavior or incidents. However, given the competitive landscape of insurance, insurers may not want to distinguish between policyholders at a similar level of risk, but on the basis of risk sensitivity or propensity to change. Insurance sets prices according to groups of people who have similar risk profiles, either, for example, by sex or age, for car insurance.

Big data provides new sources of information for understanding policyholders

[8] Y. Wang, L. Kung, and T. A. Byrd, "Big Data Analysis: Understanding Its Capabilities and Potential Benefits for Healthcare Organizations," Technol. Forecast. Shock. Change, vol. 126, pp. 3-13, 2018.

to fine-tune risk classification. The advantages of a higher risk classification are the setting of risk-based pricing, allowing insurers to combat adverse selection by trading low risks. Potential low-risk policyholders may not want to pay for a price that reflects the wider population of the risk fund. Risk-based prices may be much fairer than low-risk policyholders, as low-risk policyholders would usually subsidize a high-risk policyholder in a risk fund.

In the public sector, big data typically refers to the use of non-traditional data sources and data innovations to make government solutions more responsive and efficient. Governments deal not only with general issues of integrating big data from multiple sources and in different formats and costs, but also with some special challenges. The biggest is data collection; governments have a cult following because data comes not only from multiple channels (such as social networks, the web, and crowdsourcing), but also from different sources (such as countries, institutions, agencies, and departments). The transformation potential that Big Data has on the government is vast. This research focuses on big data solutions with applications in service delivery, policy making and citizen involvement - areas where big data can play a transformative role. It also looks at key initiatives in which the government needs to facilitate actions for the efficient use of data.

Big data analysis can be used by governments to improve existing services and to use new data sets to generate completely new public services.

An area of great opportunity is the use of big data algorithms to detect fraud, non-compliance and bottlenecks in government operations. Governments are big data producers, most of which are unstructured and text-intensive. Text analysis and machine learning algorithms are indispensable to analyze administrative data to obtain information. They can automate systematic, multi-level checks on insurance, procurement and tax records to signal entries that need further examination or to identify performance bottlenecks that need attention.

● Public utilities: Remote sensing data from satellites and ground sensors can provide a wealth of real-time or near-real-time information to monitor the supply and quality of public utilities, such as water and energy.

● Education: Personal data from devices, exam data, and other sources can be used in innovative ways to monitor student performance, better understand teaching practices, and help parents and students identify the best fit for a school. Big data can complement traditional sources to allow for new delivery regimes that tailor lessons to performance and monitor progress toward government and global goals.

● Public safety: Police forces rely on big data and predictive analytics to make better police decisions. Basic information, such as the type and location of the crime, can help officers make smarter patrol decisions.

Political decision makers use satellite imagery, mobile phone data and more to produce alternative economic indicators for new - and real - time - political perspectives.

The role of information in the policy-making process is a major concern in the information age. Big data is a viable source of high-frequency, granular data that can provide in-depth information on human mobility and economic behavior to better inform policy decisions. The entire globe is now imagined daily by satellites, with

quality and accessibility continually improving. Traditionally, policy is based on reports that are largely informed by the traditional survey and administrative data and statistics - tools that are slow and labor-intensive. Instead, big data can provide strategic information on a more frequent, disaggregated and cost-effective basis. Traditional surveys - such as censuses, official statistics and enterprise data - will always be needed, but robust big data proxies are likely to continue to emerge and become more integrated into policy-making processes and decisions. Only data from satellites, mobile phones and social networks can change the dynamics between information and policy making.

Satellite imagery, mobile phones, vehicle sensors, video streams and social networks are used by policy makers and planners for traffic and urban planning. Remote sensing, mobile phone data and machine learning can give policy makers and planners a much better understanding of urban mobility, land use and urban change. Policy makers, now, and in the future, can use this information to provide access to jobs and to make cities more sustainable, inclusive, productive, resilient and livable.

Government employment agencies are experimenting with big data to inform the most appropriate policies to help people return to work, such as tailoring training services to different segments of job seekers.

By applying machine learning to online and social networks, governments can be more receptive to citizens' feelings by introducing a new dimension of civic engagement.

Citizens involvement - the interaction between government and citizens to strengthen the voice of citizens - can play a full role in improving service delivery and policy-making processes, as long as there are appropriate mechanisms in place to translate citizen feedback into action. Big Data analysis can make this interaction smarter, more targeted, more personalized and more responsive. Governments have recently begun to adopt customer relationship management approaches, which are now widespread in the private sector in many countries. These systems can potentially be used to improve government-citizen interactions with timely information and to enable citizens to better manage service requests. Big Data analysis can also improve the performance of government processes, enable better decisions, and improve the transaction experience. Benefits include:

- A more informed and involved electorate: informing voters can help them assess the performance of politicians and legislative issues of interest and can increase turnout.
- Monitoring and feedback on services and policies: Big data can be a catalyst to improve service delivery and proactively listen to feedback on performance and quality, as well as to assess sentiment about policies and interventions.
- Citizens' voice and collective action: Big Data analysis helps identify real-time trends and mobilize pressure on policy makers.

In business, big data creates many opportunities, such as increasing operational efficiency, improving customer service or developing new products and services. Nowadays, many companies use (or are starting to use) big data to create valuable prospects to stay competitive in the market. However, such companies face several challenges, which must be addressed in order to make full use of big data: the lack of

analytical and managerial staff, the need to ensure the right infrastructure and the economic benefits of big data are clear to users, organization and economy in general, etc. In their report [9], the authors present several relevant information obtained from their research:

- data has become an important part of every industrial and business function;
- big data creates value in many ways (for example, by replacing / supporting human decision-making, innovating new business models, products and services);
- the use of big data will become a key basis for competition and growth for organizations;
- there will be a lack of talent needed for organizations to take advantage of big data;
- some issues will need to be addressed to capture the full potential of big data (e.g. data policies)

For the international economy and trade, the following Big Data influences were analyzed. From the in-depth analysis of the theory of international trade, we can see that, no matter what level the theory of development is based on, the main driving force for promoting economic and trade development is the comparative advantage between trading subjects. existing theories expose all the dominant sources of business topics from the perspective of technology and factor endowment. However, as a very important information carrier, data is the main source of profit and value creation. the new strategic height of each country, which builds the country's comprehensive competitiveness. From all walks of life, big data brings new development and business opportunities for businesses and, at the same time, improves the core competitiveness of businesses. All this will become the new impetus for the international economy and trade and thus promote the process of world trade.

Advances in information technology and technology have promoted international trade in recent decades, and the high pace of its development is mainly due to the application and development of Big Data. Among them, Big Data optimizes logistics. in the commercial market, few countries have end-to-end production solutions before selling goods. They usually sell related semi-finished products first to other countries, and then these companies implement value-added processes before exporting to customers in other regions. take faster steps to use big data to improve logistics.

In the past, international trade mainly involved flow factors, trade in goods and trade in services, and so on. The international trade placed these three on the same level and studied them superficially. In the big data era, digital commerce and information commerce have become increasingly important for the flow of information. Digital commerce and derivative information commerce place the analysis of the market environment and the supply relationship in a more important position than physical goods. Enterprises and thus promoting the development of industrialization.

Since the beginning of World War II, the model of international economic and

[9] J. Manyika et al., "Big data: the next frontier for innovation, competition and productivity", 2011.

trade interests has taken the form of north-south polarization. Developed countries have long held a certain dominant position in international trade by virtue of their advanced technological advantages. Over the years, many developing countries have carried out reforms in various aspects, improving trade conditions and promoting increased trade, but the effect is not significant. The emergence of big data has offered new opportunities to developing countries, the application of changing the traditional way of trade is widespread, they are not involved in the absolute advantage of the international trade country, but in big data the development and application of information technology in the transmission mechanism, The transmission mechanism for achieving foreign trade growth provides the theoretical basis. It can help developing countries to use information on supply and demand in the international market and reasonable positioning and estimates that reduce the imbalance between supply and demand, reduce costs, optimize resource allocation. so as to improve the international competitiveness of developing countries, change the global economic and trade model, promote the diversified development of the world economy.

Big data plays an important role in digital marketing. Every day, digitally shared information grows significantly. In this digital age, "Data is King" and is irreplaceable for every organization and business for their digital marketing strategies. With the help of big data, marketers can analyze every action of the consumer. It provides better marketing information and helps marketers develop more accurate and advanced marketing strategies.

The application of predictive analytics to sales, advertising and marketing is huge and can be applied in several ways. For example, Facebook, which is essentially an example of using machine learning, big data and predictive analytics together. The system, in this case, examines the different characteristics of Facebook users and uses them to predict which customers will purchase a particular product. Predictive analysis can also be used for cross-selling and up-selling. Big data can be collected about customers who have responded or not to the cross-selling of an associated product. It can then be analyzed using machine learning techniques to find hidden trends, patterns and correlations in the data. Once they are known, they can predict which customers might respond.

Another strategy would be split testing of different marketing and advertising methods, up to copy variations to generate large data sets that can be analyzed. The system can then predict which customers will respond to what message or type of marketing. This will allow a company to develop more personalized and personalized marketing campaigns. They will probably be much more efficient and more satisfied with their customers, because they will receive more of what they want than being bombarded with irrelevant advertising.

The great advantage of predictive analytics is that it eliminates assumptions. Having to guess or use intuitions in the critical points of the company's operations has always been a weakness, but now that weakness is eliminated. At first glance, this seems to help only large corporations, but this is not true. As you probably know, data is for sale everywhere. This means that small businesses can use the data collected by large companies to improve their operations and prospects.

The great advantage of predictive analytics is that it eliminates assumptions. This helps to avoid guessing or intuitive use in the critical points of the company's operations has always been a weakness, but now that weakness can be eliminated. At first glance it seems that this only helps large corporations, but this is not true, because data is for sale everywhere. This means that small businesses can use the data collected by large companies to improve their operations and prospects.

1.3. The Use of BIG DATA in Modern Business

With the help of big data, companies aim to provide improved services to customers, which can help increase profits. Improved customer experience is the main goal of most companies. Other goals include better target marketing, reduced costs and improved efficiency of existing processes.

Big data technologies help companies store large volumes of data, while providing significant cost benefits. Such technologies include cloud-based analytics and Hadoop. They help companies analyze information and improve decision-making. In addition, data breaches represent the need for increased security, which the technological application can solve.

Big data has the potential to bring social and economic benefits to businesses. Therefore, several government agencies have formulated policies to promote the development of Big Data.

Businesses in all industries are looking at ways to use big data in business. Its uses are ready to improve productivity, identify customer needs, provide a competitive advantage and provide opportunities for sustainable economic development.

Different industries have responded to the trend of using and implementing analytics in different ways. Retail and sales seek to collect as much information about their customers' lives as possible so that they can meet their changing needs more effectively.

Production seeks to streamline operations. Equipment calibration settings can be recorded and refined, and product storage media can be monitored to determine the optimal conditions that lead to minimal damage and waste.

For global companies, this can mean collecting and analyzing data from factories around the world, allowing them to study minor variations and understand their results.

The efficiency of each machine - and the people - involved in the manufacturing process can be recorded so that companies know what works and can make improvements where they are needed. And in agriculture, data analysis helps the industry meet the challenge of increasing global food production by 60%, as forecasters have said will be needed by 2050 because of the growing population. Of course, in business, once a product has been grown or manufactured, it must be sold and distributed. Petabytes of customer data, already collected by large retailers, tell them who will want to buy what, where and when. Amazon, for example, uses its S3 system to keep track of millions of stock items across dozens of warehouses and

distribution centers around the globe. Operators can track deliveries in real time to see where and where they should go. At the point of sale, retailers can use the data to determine where the stock should be displayed, which stores will sell most of certain products, and will track customer movements around stores. Loyalty cards are not new, but an increasingly sophisticated analysis of customer habits will lead to an increase in the ability of retailers to predict what will be bought. This has advanced to the point where Amazon believes it will soon be able to predict what you will buy accurately enough to send it to you before you even buy it.

Every industry is learning to reap the benefits of Big Data analysis and it seems certain that finding innovative methods of data collection, recording and analysis will play an important role in business in the foreseeable future. Even something as subjective and "human" as Human Resources is transformed by Big Data and analysis. Finding and keeping the right people is a major issue for most companies. Talent management is full of challenges, and the cost of failed management and leadership is enormous. The average cost of executive failure is estimated at $ 2.7 million. their role.

Software is used to predict a number of things, including how long an employee is likely to stay in the job. Although this type of Big Data analysis currently focuses on customer-oriented roles, it is only a matter of time before it reaches higher levels of management. Certainly, improving the performance of top executives has a "disproportionate effect on the company", so Big Data solutions are sure to be considered.

Once the progress of big data development made it possible to analyze big data sets at high speed, big data began to influence the internal processes of many companies. It can determine how decisions are made, strategies are created and customer relationships are sustained.

In this subchapter, we discuss the impact that big data has on businesses today and we present the main preconditions for big data to benefit. For many businesses, big data is not a choice, but a natural reality, because the amount of structured and unstructured data grows exponentially, along with a wide network of IoT devices that capture it.

The main business opportunities presented by big data for any industry include:

● Automation: Data-based IT infrastructures allow companies to automate time-consuming processes, such as data collection and analysis.

● Trends and perspectives: Big data reveals hidden opportunities and patterns that can be used to tailor products and services to end-user needs or to increase operational efficiency.

● Data-driven decision making: machines learn about big data to allow predictive analysis and to make informed decisions.

● Cost reduction: Big data statistics can be used to streamline business processes in order to eliminate unnecessary costs and increase productivity.

Online and offline retail competitors are biting each other trying to win customers. In order to stay at a safe distance, companies need to provide a comparatively better customer experience. Big data offers retailers new ways to stay ahead and innovate.

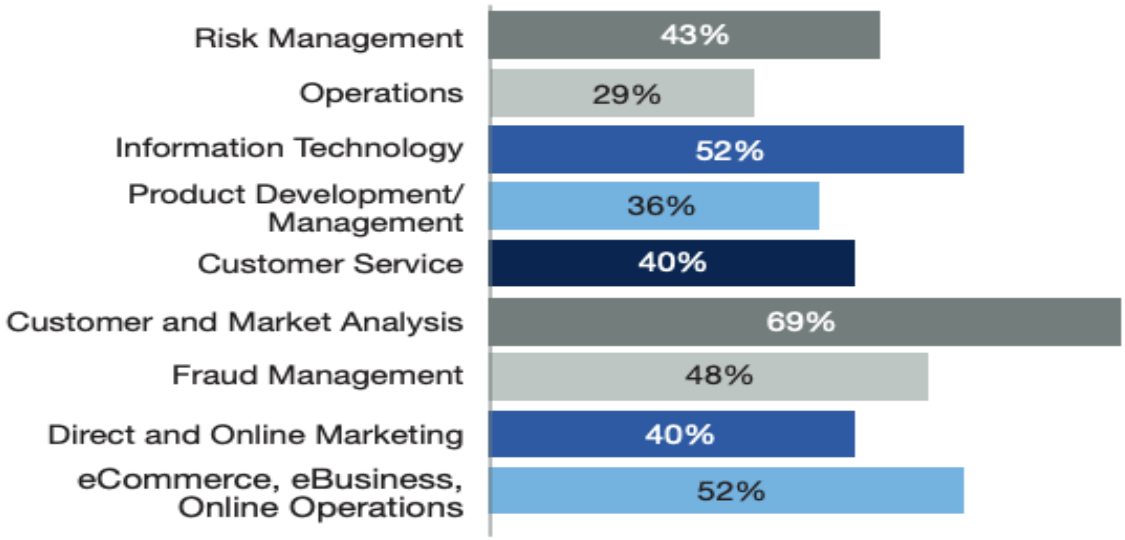

Figure 2.1.1 The Use of Big Data în Business

Sursa: https://info.idc.com/digital-data.html?utm_medium=idc-tagged&utm_sources

Big data is now one of the critical factors in the omnichannel race of customer involvement. By collecting data from multiple channels, such as social networks, call logs, store visits, browsing history, and more, retailers can get a complete view of their customers and adjust their operations accordingly, from customer service marketing.

Data sharing platforms could be the answer to the challenge of data collection. By anonymously sharing data on a central platform, smaller retailers could build a large set of data from which useful information can be obtained for each retailer.

In product development, customer demand can be anticipated using big data to discover trends such as seasonal demand, successful promotion in different customer segments, popular complementary products and more. Using this information, it is possible to create templates for launching new products and services.

By feeding big data to referral engines, retailers can deliver personalized recommendations even to anonymous surfers. This way, visitors are more likely to find exactly what they want, as well as buy more.

Instead of giving way to online stores, brick-and-mortar retailers can use IoT and big data to deliver unparalleled in-store experiences.

Offline companies can analyze data from mobile apps, online and offline sales, customer locations and in-store behavior, and use the information to enhance offline experiences, optimize store design and marketing, and encourage repeat purchases.

Manufacturers use big data to improve operational efficiency and improve business processes, with the ultimate goal of increasing profits.

To ensure operational efficiency, manufacturers need to analyze production processes in order to react to abnormal events that can disrupt production life cycles and lead to customer dissatisfaction.

Big data helps manufacturing companies keep more processes under control, for example, by correlating downtime with other events to understand why shutdown

occurs.

Big data can also be used to prevent or predict when another shutdown may occur in the production process and reduce the amount of scrap by up to 27%. Manufacturers use structured data, such as equipment launch date, make and model, as well as unstructured data, such as sensor data and error logs, to keep them proactive and save costs.

Moreover, big data helps plan equipment shutdown. It is also used to predict that certain equipment may not operate in accordance with the specifications, identifying voltage caused by overload or defective parts. With the help of big data, manufacturers understand exactly how items move through their production lines. For example, they can reveal a blockage that causes an increased production time, thus optimizing processes.

Regardless of the company's field of activity, the use and processing of Big Data increases from department to department.

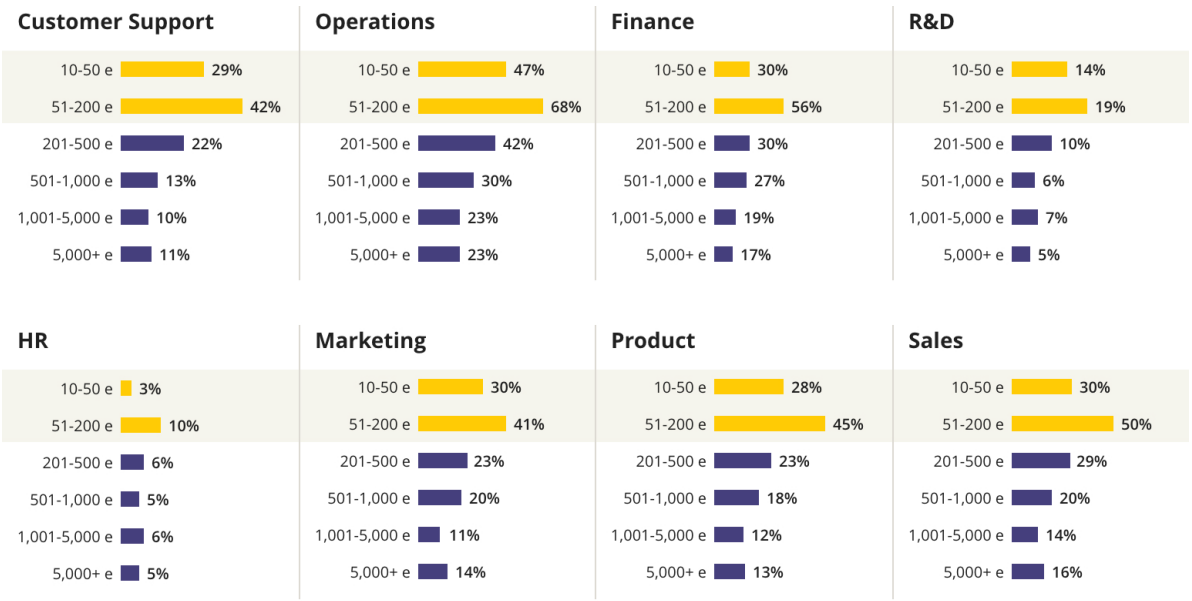

Figure 2.1.2: Expenditures for the implementation of BD processing technologies in US companies

Source: https://research.aimultiple.com/big-data-stats/

While large enterprises are struggling to stay agile in the face of more analytical BI solutions and data sources, many small companies have grown due to the cloud infrastructure in place. In addition, they are still small enough to use a single centralized BI solution. Therefore, it can be seen that companies with between 10 and 200 employees reported the highest use of analysis.

Companies with 51 to 200 employees led the largest enterprises in the use of analysis in each department, 68% of small enterprises using analysis in operations, 56% in finance, 50% in sales and 45% in product.

As we explore the new use cases of companies targeting crisis management, there remains a strong correlation between increasing analysis in these departments and the types of problems companies are solving in the COVID-19 era.

Companies are using data to overcome new challenges. Not only has the

pandemic affected the frequency of use and budgets for data analysis, but it has also influenced the expansion of new cases of data use. Improving efficiency, supporting customers, predicting change and results are the three fastest growing use cases, and these use cases extend regardless of the size of the company.

These use cases aim to address some of the recent challenges of COVID-19, as low revenues require companies to become more efficient, changing customer habits requires companies to adapt to their new needs, and companies of all sizes are looking to the future.

In particular, businesses are more likely to see cases of extensive data usage in terms of cost reductions and cost reductions compared to other new use cases, probably indicating greater pressure and less short-term flexibility. Smaller businesses report that they have started using data to improve efficiency, predict changes and results, and support their customers. New cases of data usage leading to cloud migration have also been identified. Respondents from all industries who reported using data for the following new purposes are also most likely to consider retraining.

● Embedding analytics: 35% of those who have started embedding analytics in their products are likely or very likely to move to the cloud.

● Cost reduction: 31% of those who use cost-cutting analytics are likely or very likely to move to the cloud.

● Sales funnel optimization: 31% of those who started tracking the sales funnel using analytics are likely or very likely to move to the cloud.

There are many ways in which big data can be used to fit your business model, especially when it comes to advertising. Below will be a mini case study on how the online retail giant uses big data today.

Specifying the products that customers might want to buy is the most common upselling tactic in e-comm. A very relevant example is the way in which Amazon uses big data to find out exactly the type of products that the consumer will want to buy in the future. which reached a net value of $ 1 trillion at the end of 2018 - provides customers with an understanding of what factors determine these recommended products. In this context, Amazon cites a variety of data points to find out what its customers want.

These factors include:

● When customers shop

● How customers evaluate their purchases

● What customers buy with similar buying habits

Obviously, the last factor is the most important in terms of big data. Amazon is able to determine exactly what kind of products you want to buy based on customers with similar buying habits to yours.

Similarly, any business can use this type of data to make predictions for its customers. When there will be an increase in sales, there will also be legitimacy, in trends. For example, Amazon has noticed that people who buy TVs also tend to buy a TV stand - which the retailer has started selling in the hope that customers will buy them together.

Operational risk is particularly high in financial institutions. Scammers are

constantly trying to develop schemes to take advantage of both people and companies. As big data has evolved, however, financial institutes have realized that they can use this information to stop misleading scams.

Banks, for example, now use big data to monitor their transactions on a front-to-back line of business to help eliminate fraud at all levels. They look at information about who sends / receives money, how often these people engage in this behavior, where they live and how much money they send.

This type of technology can be useful for any business, not just for banks. As the data is collected, trends appear and anything that deviates from "business as usual" triggers a sticky digital note on the transaction. This makes it easier for companies to identify fraud when it occurs and to keep their operational risks to a minimum.

Big data is absolutely vital to learning how customers can be made to make important decisions when they reach the company's website. Businesses use this data to find out the behavioral patterns of their customers and to direct them to a sale or conversion.

Specifically, the data analyzes every action a customer takes when landing on the business website. You can see the customer's keystrokes and how the mouse moves; it can predict the actions that will be taken further.

As more customers come to the website, it will be easier to find out the most common behaviors and plan ahead. The marketing plan will bring a lot of favorable results and there will be increases in sales, as more data is gathered about what customers want and need.

1.4. The SaaS business model for interpreting metrics within companies

The same data that is channeled into referral systems, predictive algorithms, and customer segments can be reused and packaged to inform management of the successes or failures of its operations. At the same time, the skill stack that defines data engineers, data analysts, and data scientists is predominant in the development and delivery of KPIs by data strategists or business intelligence analysts.

An important distinction to be made in the case of KPIs is in their design. Instead of being largely determined by experimental methods, such as statistical analysis supported by data scientists, KPIs are largely designed by leaders who communicate to data engineers, data analysts, and data strategists accurately. what they want to see. These requests are then translated into indicators (performance indicators) that are created in the data and serve to properly measure the performance of leadership-based goals and priorities. KPIs can also provide aggregated or detailed insights into customer behavior, giving product creation and management teams oversight of customer interactions with products and services, especially as these products and services are tailored to changing needs.

The vital role of the data strategist lies in his ability to best understand the needs of management (or even a customer) and to communicate effectively with each other. The data strategist - responsible for developing KPIs and bringing them into reality - is positioned to understand the management perspective, while anticipating the

intention behind the KPIs. Data strategies often need to help design effective indicators because they have a clearer sense of the technical constraints and opportunities in the data. Therefore, while KPIs are designed by management, they are generally a collaborative effort. It positions data strategies between both data and enterprises, providing ample opportunities for growth and development.

The development of KPIs requires a strong technical understanding of both data and engineering methods. Once a design has been established, a KPI is developed in the same way as many other database-based products. Although some performance indicators fail to indicate performance measures due to limited data, poor performance indicators, or poor quality data, to name a few of the potential problems, there is a natural process of trial and error around indicators before a KPI to be established.

One of the most popular data sources behind KPIs is customer, digital or user-generated data. Nowadays it seems that the business cannot collect enough information about the customers and their activities. On the one hand, companies risk violating the privacy and piracy of personal information data, on the other hand, failing to adapt to consumer preferences by collecting enough data to create the products and services they want, need or need. which can be connected.

If properly designed, frequently updated and clearly developed, KPIs provide management with a valuable perspective on the business - a potentially decisive factor in the success or failure of a team, product or company.

Following the right values is an important part of running a successful business. Big data is crucial to help organizations achieve this goal.

To analyze a series of metrics, the SaaS (Software as a Service) business model was proposed. As an important service model for advanced computing, SaaS uses a defined protocol that manages services and applications. The popularity of advanced calculations has reached a level that has led to the generation of large data sets. Big data is evolving with great speed, large volumes and great diversity. Such an amplification of the data called into question the existing database tools in terms of their capabilities. Software as a service (or SaaS) is a way to deliver applications over the Internet - as a service. Instead of software being installed and maintained. SaaS applications run on a SaaS provider's servers. The provider manages access to the application, including security, availability, and performance.

During the COVID-19 pandemic, companies resorted more intensively to such services, due to the fact that the general activity of the consumer passed on the Internet, which made possible a more detailed analysis of its behavior through the tools offered by SaaS applications. SaaS is now the gold standard for companies, from accounting firms from old schools to artificial intelligence blood startups and from small family stores to large multinational corporations. The use and demand for SaaS products has increased dramatically over the years and we see this trend continuing in 2020 and beyond.

According to the 2019 report [10], two key trends were revealed: the turnover of applications was higher than that of employees, 39% of typical medium-sized

[10] https://customerthink.com/?__cf_chl_jschl_tk__=5b2b797fc625420f04719ac853c4e54e2b5d14a1-1621016028-0-

companies changing applications in the period 2017-2018; and that the selection of SaaS applications has been increasingly decentralized between departments.

In fig. 2.2.1 Analyzing the 2020 report data analyzes over 10 years of data on SaaS spending and usage from over 1,000 companies. Thus it can be concluded that:

● Total costs per company for SaaS products have increased by 50%, and the number of unique applications used in each company has increased by about 30% year-on-year.

● Companies are analyzing over 30% of their applications each year at this time.

● SaaS spending in the last year for SME companies has doubled by 100%, for medium-sized companies it has increased by 66%, and for corporate companies

In conclusion, by 2020, an overview of how small, medium and large organizations use SaaS applications and the basic purchasing modalities for SaaS builders, consumers and investors has grown considerably, due to the data and metrics that generate and contribute to them. adopting the development strategy.

We selected and analyzed 4 most commonly used values / metrics in the analysis of company performance, products and sales reports.

One of the most critical values of SaaS - the average income per account or commonly known as the average income per customer (AIPC). The main reason this value matters is that it helps companies identify the profits they make from each individual customer who is connected to them.

The value of AIPC is mainly used by companies that operate on a subscription-based business model or by those that provide a certain type of service. For example, AIPC is widely used by telecommunications companies, social media companies and banks.

AIPC is a value of profitability that is not recognized by accounting standards such as GAAP or IFRS. However, companies in relevant industries report value on financial statements and devote a significant amount of time to discussing AIPC results.

Its role is to provide an overview of the company's profitability on account. In addition, it reveals the products or services of the company that generate the most and the lowest income.

Average revenue per account (AIPC) is a measure of profitability that evaluates a company's revenue per customer account.

AIPC is calculated by dividing the company's revenue for a period by the number of accounts for the same period. Key applications of the AIPC metric include the following:

● Comparison: AIPC is commonly used to compare a company's performance with that of its competitors, as well as to compare its performance over time.

● Customer segmentation: AIPC can also be used to indicate the company's revenue segments. Value can help identify which services or products generate the most / lowest revenue or which customers (new vs. existing) generate the most revenue.

● Revenue forecasting is another important application of AIPC. Value is typically used as input for calculating a company's recurring revenue (CRR or ARR).

At the same time, recurring income values that are distinguished by predictability and stability are generally used as a basis for revenue forecasting.

- AIPC formula: the value is commonly used as input for calculating a company's recurring revenue. At the same time, recurring income values that are distinguished by predictability and stability are generally used as a basis for revenue forecasting.

While sales figures may fluctuate, the generation of potential customers should remain on a steady upward trajectory. This is why the lead capture rate is one of the most important measures for tracking growth in real time.

To understand what a lead capture rate or LCR means in business, it is essential to define what qualifies as a lead. A lead in marketing is not just any potential customer. It is a potential customer or client who has already demonstrated a quantifiable interest in the company's services or products.

LCR is a key performance indicator (KPI) that tracks the growth rate of qualified leads in real time.

When it comes to marketing, sales and lead generation, LCR is undoubtedly a powerful performance. Using LCRs can be a better measure of real-time growth than sales revenue, especially when analyzing the month-to-month figure. Using LCR, a company can predict its sales growth trajectory.

Although it is used month-to-month for real-time lead tracking, the LCR is also useful for longer-term use. By comparing LCRs from year to year, a company can identify changes in strategy more precisely and can understand if they have translated into an increase in the number of customers over time.

The lead capture rate of a lead is easy to calculate, formula (2.2.2). It aims to change the percentage of the volume of qualified potential customers by comparing this number of potential customers from one month to another.

The financial value of each client in the medium / long term, - CLTV is one of the most important metrics to measure for any growing company. By measuring CLTV in relation to the cost of acquiring customers, companies can measure how long it takes to recoup the investment needed to win a new customer - such as the cost of sales and marketing.

The customer lifetime value (CLTV) is the value that indicates the total income that a company can reasonably expect from a single customer account. It takes into account the value of a customer's income and compares this number with the estimated life of the company. Companies use this value to identify significant customer segments that are most valuable to the company.

One of the most difficult changes to adapt to in the new world influenced by the evolution of Big Data is measurement and reporting on the move, from an emphasis on short-term conversion to one that reflects the impact of business engagement. The measurement of customers needs to move from blunt units to a more precise and sophisticated set of specific values, such as "increasing portfolio share", "loyalty for life / net present value" and "profitability trajectories in the customer segment".

CLTV tells companies how much revenue they can expect to generate a customer during the business relationship. The more a customer continues to buy from a company, the higher the value for life.

The historical approach is based on the amount of gross profit from purchases made by buyers of a company in the past, only data on previous acquisitions are needed. The historical approach is valid only if customers have similar preferences and stay with the company for the same period of time. Changes in customer behavior are not taken into account. So, if buyers change their interests and how they buy, this will affect the result.

A predictive approach to calculating CLTV aims to model a customer's transactional behavior and predict what they are likely to do in the future. It is more accurate than historical CLV because it applies algorithms that can predict the total value of a buyer. Together with previous acquisitions, this approach explains a customer's actions. This metric can be calculated according to the formula (2.2.3)

The Product Involvement Score (PIS) is a new value recently introduced into the SaaS business landscape. PIS provides a unique, quantitative business metric to measure the overall commitment of a product in order to quickly diagnose a product's performance. Historically, product leaders have had to rely on qualitative values, such as NPS, the net promoters score NPS uses a simple recommendation score to determine the virality of a product or service. It has always depended on customer questioning and has never provided an accurate look at the actual user behavior. While important, NPS may not be the only business value for a product. PIS is intended to fill this quantitative gap and partner with NPS, truly examining user behavior and quantifying how users interact with a product.

PIS is a way to highly assess the health and success of a product through quantitative engagement data, as a sentiment-based and highly subjective counterpart to NPS. It is applied as a business-level KPI for a product and, when further analyzed with segments and data filters, can be used to gain a more detailed understanding of the involvement of user groups over the life cycle of the product. The information can then be used to stimulate product development, in-app messaging and awareness campaigns, with the ultimate goal of increasing product involvement and user satisfaction, indicated by improved values or, as we like to say, offering excellent products on which users love.

PIS consists of another 3 metrics: adoption, attachment and growth, and in the end is the average score of each component. Possible scores for all components range from 0 to 100.

Adoption analyzes the percentage of users who use the product effectively, as indicated in the base events. Adoption may involve the use of visitors or the account for any or all major events to all visitors or active accounts in the selected time period. It is calculated according to the formula.

Attachment refers to the average percentage of users who are users with a high frequency return rate relative to the daily / weekly / monthly values of active users. The attachment component finds the ratio of the average daily or weekly active visitors to the average number of weekly or monthly active visitors in the selected time period.

The increase compares the number of active visitors or active accounts in the current time period with the previous time period and composes this growth rate for one year to determine if the product has strong growth and virality. Growth does not

consider new visitors or accounts, but only activity.

The calculation of the growth score consists of two parts, the measured growth rate and the projected annual growth rate. First, the percentage change in the total active users between the current and previous time periods, the growth rate, is calculated. Then, the projected annual growth rate is calculated to determine the score of the growth component. Possible scores range from 0, if there is no increase or negative increase, to 100, if the calculated annual growth rate is 100% or higher.

Data, values and analysis all mean different things, but they work together to support strategic goals. Values cannot be developed without data. Without values, there are no trends to analyze, so it will be harder to find the relationships in the data. Analysis is essential to make the data resources collected through the various means and digital tools they need for the development of strategic business plans.

1.5. The potential to adapt, if it exists, the local market in the Republic of Moldova to the trend generated by Big Data and IT

Digital technologies (information technology, telecommunications, robotics, AI) are reshaping economic opportunities for individuals, companies and countries globally. The improved capabilities of these technologies and their increasing global adoption may lead companies to automate certain tasks or jobs or move work to new locations. This could involve job losses or wages, but also changes in global trade patterns. These technologies also create opportunities for individuals and businesses: they reduce transaction costs, connect firms and individuals with each other and with markets, enable innovation and support improved productivity. In short, technological change will affect different businesses, individuals and economies differently, depending on their level of training.

The current situation in the Republic of Moldova was analyzed from two perspectives: how prepared the workforce is, to adapt to current digitalization conditions and the ability of SMEs to integrate digital tools and automate the work, management and production process. This report presents the results of an analysis of what Moldova could do to prepare for (and take advantage of) the growing digitalization of economic activities between sectors inside and outside Moldova. It indicates some of the directions that could be integrated into future programs, funded by development partners or considered for action by the government.

In the last 10 years, Moldova has evolved dynamically and emerged as a new destination for Western companies to outsource their IT projects, software development and other technology-related services. Chisinau, the capital of Moldova, is the place where most business activities in the IT sector take place.

Moldova has the potential to become an innovative and technological development and delivery center, where highly qualified, multilingual, knowledgeable professionals are ready to develop complex projects and help companies achieve their business objectives at better than cost prices. outsourcing media from other parts of Europe or anywhere else in the world.

The country plays the role of a bridge between the western and eastern markets

and, since 2006, Moldova has increased its ICT services and products 20 times and this growth is accelerating. Almost 80% of the IT products created in the country are destined for export. Western Europe is the largest export market for Moldovan technologies, with an export market share of 45%. It is followed by the USA at 21%, Central and Eastern Europe at 18%, the Middle East and Africa at 6%, and other regions, including Asia, represent 10%.

In 2018, the Republic of Moldova was included in the Bloomberg Innovation Index 2018 and entered the Top 15 GII economies (Global Innovation Index) 2018 in the development of mobile applications. Applications represent the global trade in fully digital goods, therefore providing an insight into how innovation, production and trade in digitized products and services are evolving in an increasingly globalized digital economy.

We place this analysis in the context of a digital future for Moldova, taking into account at least the period until 2030. This is due to the fact that the scale of digitalization of the economy - and each occupation - will probably increase only in the coming decades and that the future of the workforce is already determined.

Table 3.2.1 Mix of IT Absorption Level Indicators in the Republic of Moldova

Index	Measure Unit	Year
IT graduates / per year	2 500	2017-2018
Active IT Professionals	23 000	2017
IT Market Research Tatio	6,8%	2018
Doing Business Rank	47/190	2018
Covering 4G Fiber Optic Networks	90%	2018
Internet Speed	43,52 Mbps	2017
Internet Penetration Ratio	76%	2017
Mobile Gadgets spread Ratio	120%	2017
Network Readiness Index	77/144	2018

Source: https://www.indrivo.com/

Digital jobs include jobs in the ICT industry (created through the production of ICT such as software and IT services, telecommunications, etc.) and those created or influenced by the digitization of sectors throughout the economy (eg. in the banking sector, tourism, agriculture, Government etc.). These could be jobs in companies or they can represent the work done by individuals. These are jobs created by:

- companies in the ICT industry;

- companies in sectors that digitize
- people who carry out freelance activities in digital activities.

As a small, open economy with a declining population, Moldova should consider the steps that prepare it, the workforce - today and in the future - and its business for this future based on the digitization of various branches of the economy, as well as and e-gov systems. This report presents the results of an analysis of what Moldova could do to prepare for (and take advantage of) the growing digitalization of economic activities between sectors inside and outside Moldova. It indicates some of the directions that could be integrated into future programs, funded by development partners or considered for action by government and development agencies.

Moldova's ICT industry has evolved rapidly in the last decade. WTO data show that Moldova's net exports of telecommunications, computers and intelligence services reached $ 78 million in 2016.

According to data recently presented by the National Bureau of Statistics, in 2019 the information and communications technology industry has become one of the locomotives of economic growth in the Republic of Moldova. For the period 2015-2019, the IT sector registered an increase from 2.1 billion lei to 6.6 billion lei, which represents a more than threefold increase in annual sales revenues. Thus, interpreting the data, in 2019, the IT industry accounted for 7.1% of national GDP, which qualitatively changed the structure of exports of the Republic of Moldova.

For this period, the sales volume in the electronic communications industry amounted to 6.8 billion lei. The volume of exports of ICT products and services exceeded in 2019 the figure of 5 billion lei. Only IT exports increased three times - from 1.1 billion lei in 2015, to 3.5 billion lei in 2019, exceeding even the export of strategic products, such as wine (by over 400 million lei)

The positive evolution in the ICT sector is also accompanied by the increase in the number of companies during this period - from 1,700 to 2,300, which also offers the best salary level in the economy. With the fiscal regime of the IT Park, the over 600 resident companies registered an average salary of about 29,000 lei per month.

There are opportunities for growth. For example, thanks to the EU-Moldova Association Agreement, trade ties could increase, and the country is starting to receive attention from industry analysts worldwide. Many stakeholders (public and private) agree that there is potential for employment in the IT industry more than threefold over the next decade if infrastructure (eg IT parks), business climate and skills training programs continue. they will make progress.

Various actors have launched initiatives to increase labor demand and supply, with the hope that Moldova's digital economy, including companies producing ICT goods and services and ICT users in other sectors (eg financial services, healthcare, tourism, government), can contribute to the job creation that Moldova needs to pursue sustainable economic growth.

Positive dynamics are due to a vibrant and organized private sector and due to support from the public sector. The active association of the ICT industry in Moldova, the Association of Information and Communication Technology Companies (ATIC) has been a significant player in this space. For example, it has set up a center of excellence in the IT industry, Tekwill.

Law no. 77 of 21.04.2016[11] on the Information Technology Park establishes the manner in which this park will operate. This law aims to create the necessary premises to stimulate the development of the information technology industry, research and innovation in the field of information technology in various fields and to encourage foreign direct investment.

The park offers many benefits, including consulting services. However, the most important benefits are the tax cuts (Article 15) that allow entrepreneurs to invest this economy in development activities. The park facilitates a one-time fee of 7% of sales revenue, which is very attractive for bootstrap start-ups. Article 8 sets out what activities residents of the park should carry out in order to qualify.

Although comprehensive activities are covered, they do not provide technology-driven financial services enterprises. However, providers of pure technology solutions for the financial sector are covered.

There is already an ecosystem that supports the development and digitization of the fintech and economic sector in the Republic of Moldova,a vibrant community of supporters working together to bring about transformative changes in financial services for the benefit of e-commerce. Those with a more significant and noticeable impact are:

● Tekwill enables information technology, including the development of Fintech technical solutions. They are a center where IT professionals can connect and share ideas and resources to empower Moldova's economy. Tekwill also runs XY Accelerator for start-ups in the Republic of Moldova. Tekwill also hosted the first Fintech Conference in Moldova in 2019.

● MAICT - Moldova Association of Information and Communication Technology promotes the development of the ICT sector.

● AAFS - Association of Alternative Financial Service of Moldova, unites non-banking financial institutions. They cover Fintech, but only outside the banking sector.

● Moldova is fortunate to have many international development partners that support the government and the private sector, including: USAID, the World Bank, the European Bank for Reconstruction and Development.

Due to the activity of such organizations, in recent years there has been a rapid increase in the number of digital startups established in Moldova. The Tekwill facility is starting to become a hub for start-ups, and MAICT believes that there is a possibility that more companies will appear in the coming years around financial services ("fintech") or games, for example. Also, supporting IT professionals and business graduates in Moldova could encourage them to stay in the country.

The government has been receptive and aims to promote the growth of the IT industry through fiscal incentives, new IT parks and training programs for specialists. Parliament recently approved a new "start-up visa law", which simplifies access to the Moldovan labor market for highly qualified IT specialists. But private sector stakeholders remain uncertain about its acceptance by international workers. The Ministry of Economy and Infrastructure defines a new roadmap for the

[11] Law no. 77 of 21.04.2016, regarding the establishment and activity of IT Parks

competitiveness of the ICT industry (for the period up to 2023), and a significant focus is - in addition to improving infrastructure - capacity development, business and start-up ecosystem and investment promotion. it is also working to modernize the curriculum and ICT infrastructure in schools and is trying to develop human capital.

The number of graduates in ICT-related fields - around 6,500 per year - is also higher in the proportion of graduates compared to regional colleagues, such as Bulgaria, Hungary or Romania. However, the share is lower than the EU or OECD average. Employers and industry representatives are critical of the quality of many of these graduates and the relevance of their skills to industry requirements.

Some of the gaps both on the demand side and on the supply side. For example, there are plans to support more digital start-ups and provide them with financial and technical support to grow. Plans are also being developed to extend the Tekwill model to other cities to make the digital economy more inclusive. The use of ICT by companies in sectors. In addition to workers, technological change involves challenges for companies. If they want to sustain growth, Moldovan companies - especially exporting companies in the sectors - should invest and upgrade their technology to remain competitive. Indeed, recent research by the World Bank has shown that developing countries face an increasing possibility of losing productivity gains due to delayed innovation - including technology adoption - due to lack of managerial capacity and poor working environments.

There are several gaps, both on the demand side and on the supply side. For example, there are plans to support more digital start-ups and provide them with financial and technical support to grow. Plans are also being developed to extend the Tekwill model to other cities to make the digital economy more inclusive. The use of ICT by companies in sectors In addition to workers, technological change involves challenges for companies. If they want to sustain growth, Moldovan companies - especially exporting companies in the sectors - should invest and upgrade their technology to remain competitive.

Indeed, recent research by the World Bank has shown that developing countries face an increasing possibility of losing productivity gains due to delayed innovation - including technology adoption - due to lack of managerial capacity and poor working environments. the "horizontal" use of digital tools and technologies in other sectors[12].

For example, only about 60 percent of Moldovan companies (the survey included 360 formal companies with 5 or more employees), of those surveyed in the most recent business survey, have websites or use e-mail to trade. business. This indicator is much weaker compared to other European countries, such as the Czech Republic (over 90%) or Poland (over 80%). 19 See figure. 3.1.1

Moldova has made some progress in ICT development; however, it remains quite weak in most international comparisons. Moldova's positions do not improve considerably even compared to geographically close countries, in their historical context and in their level of economic development.

[12] (Cirera, Xavier; Maloney, William F. 2017. The paradox of innovation: the capacity of the developing country and the unfulfilled promise of technological recovery. Washington, DC: World Bank. © World Bank. Https://openknowledge.worldbank.org/handle / 10986 / 28341banking, tourism, agriculture)

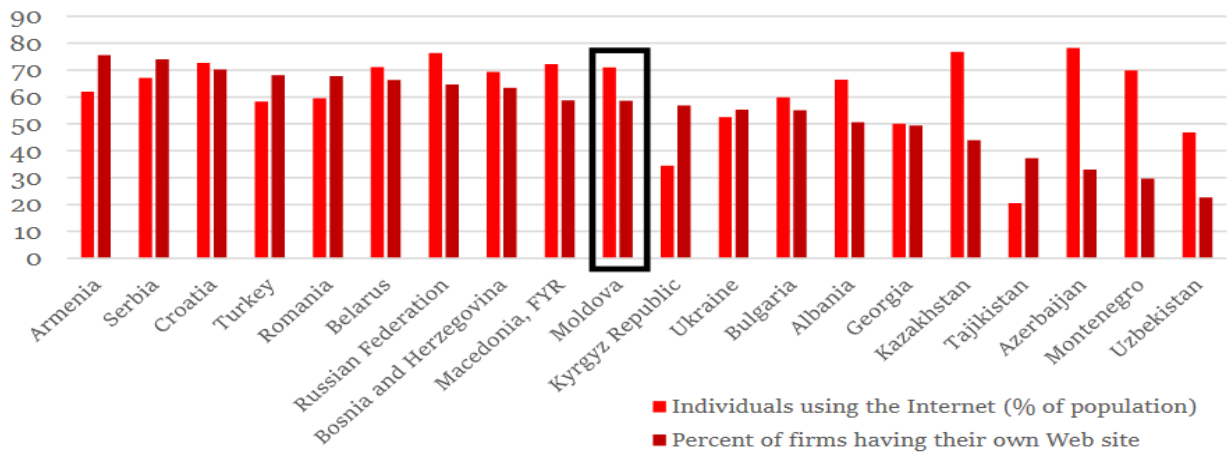

Figure 3.1.1 Indicator of the number f Internet users and companies with websites

Source: World Bank Enterprise Surveys (firm data, c. 2016), World Development Indicators (Internet usedata, c. 2016)

In its latest version of 2019, the IPR report maps the network-based training landscape of 121 economies based on their performance in four different pillars: technology, people, governance and impact. According to fig. 3.1.2, it can be seen that compared to the average of these indices, divided by pillar, the Republic of Moldova is much closer in terms of development to middle-income countries, especially in terms of technology and governance.

Table 3.1.2 Network readiness index 2019, Moldova

Dimension	Moldova	Lower-middle-income countries	Europe
NRI	48.93	37.18	65.20
Technology	45.02	32.15	63.08
People	40.93	30.04	57.50
Governance	55.19	46.17	73.99
Impact	54.58	40.35	66.24

Source: Network Readiness Index 2019, Moldova, by Portulans In***stitute***

Despite the remarkable growth of the ICT sector in Moldova, the decline in IPR can be explained by deteriorating competitive market conditions, relatively high prices for mobile communications and reduced broadband internet access. The large urban-rural discrepancy also reflects a large urban-rural gap in Moldova's ICT development. The Network Readiness Index is one of the leading global indices on the application and impact of information and communication technology (ICT) in economies around the world.

The general average of this indicator shown in fig. 3.1.3 of the Annex, is also influenced by the three sub-pillars of the Technology pillar, in particular the one referring to "Future Technologies", sub-item 1.3.2 "Investments in emerging technologies" and 1.3.3 "Public procurement of advanced technology products" .

These indicators refer to the level of openness of the private and public sectors, respectively, to procure, adopt and implement high-performance technologies and software in the long-term activity and development.

Digital engagement of people - which can lead to digital adoption by businesses - is also relatively low. For example, an analysis by the US International Trade Administration states that: "E-commerce is still in its infancy, with few traders offering this option in Moldova and few Moldovans accustomed to ordering online."[13] This means that Moldovan companies are losing the opportunity to access new markets through tools such as e-commerce.

Others found that about 71% of the population accessed the Internet in 2016[14.] When it comes to using social networks, since 2017, Moldova had about 20% of people who actively used social networks. This was lower than the global use, which was at 37 percent, or Romania at 49 percent and Ukraine at 36 percent. But few Moldovans seem to use the Internet for economically productive purposes. For example, less than one-fifth of Moldovans (over 15 years old) surveyed in 2017 reported using the Internet "to buy something online in the last year" and only a third "used the Internet to pay bills in the last year." ,"Internet use is also concentrated in Chisinau.

Data from a 2014 survey showed that about half of the population used the Internet at that time, and about 45% of them lived in the capital Chisinau. (At the time, 16 percent reported using internet banking services.)[15]

There are no documented studies on financial inclusion. In other words, what proportion of the population has a bank account. However, it is believed that Moldova has a low level of financial inclusion, compared to its neighbors, even though the penetration of the bank account among the population has increased in recent years. The World Bank estimates suggest that financial inclusion increased from about 18% in 2014 to about 44% in 2017.

The analyzed indicators show that the Republic of Moldova, although registering an increasing trend in the number of internet users, social networks and ICT specialists, SMEs at the national level are not willing to implement and train staff to integrate systems and software for collecting and storing data about, customer products, tracking, internal management and production.

For this reason, the ability to apply Big Data processing technology by collecting data in cloud systems is premature, although several attempts are being made at the level of government projects, such as "Digital Moldova 2020". The Digital Moldova 2020 Strategy was developed by the Ministry of Economy and Infrastructure and approved by the government in 2013, according to Hot. of Government no. 857, of October 31, 2013, in order to promote policies aimed at ensuring a sustainable growth of the ICT sector.

The document sets out the directions to ensure a systemic development of the field of information and communication technology in the coming years, for the comfort and well-being of the citizen.

[13] https://www.export.gov/article?id=Moldova-eCommerce
[14] https://freedomhouse.org/report/freedom-press/2017/moldova
[15] https://www moldova.org/en/1-7-million-moldovans-use-internet/

The strategy provides favorable conditions for the development and widespread use of ICT potential in all areas: public, private, business and the daily lives of citizens.

The vision of the strategy is to create by 2020 an advanced information society, in which the use of information and communication technology facilities, extended access to modern ICT infrastructure, rich digital content and the provision of information services will lead to economic competitiveness, good governance and therefore, to an improved welfare of the population. The strategy is structured around three pillars, each of which reflects the most important issues of the sector and develops actions in this area to achieve the objectives:

- Pillar I: Infrastructure and access - improving connectivity and network access;
- Pillar II: Digital content and electronic services - promoting the generation of digital content and services;
- Pillar III: Capacities and use - improving digital literacy and skills to enable innovation and stimulate use[16].

According to sub-item 2.3. "Important ongoing programs, initiatives and projects" were proposed to implement the following projects:

Governance: Functional Electronic Governance Center, which implements the project "e-Transformation of governance" - Strategic program for technological modernization of governance; The Open Government Action Plan for 2012-2013[17], including:

- M-Cloud Joint Government Technology Platform;
- Government Electronic Payments Service 3) Government Interoperability Platform;
- Paperless Government Initiative - SIGEDIA;
- Government e-Reporting Platform for Business;
- Government Platform for Permissive Business Authorizations;
- e-Procurement;
- e-Constructions;
- e-Justice;
- Digitization of Operational Support Systems for the Government
- Government Data Storage Infrastructure

The mCloud data storage service involves: “The IaaS service is a model for delivering IT services from the MCloud platform, in which the provider (Information Technology and Cyber Security Service) makes the requested IT resources available to the Beneficiary (APC), and the beneficiary is responsible for the administration of computer systems. When recipients request Saas services, they are actually given a virtual data center environment that provides them with computing resources (virtual processor, working memory (RAM), and disk storage space) that they manage remotely. independently through the self-service tools provided by the provider[18].

[16] Government Decision no. 857, of October 31, 2013
[17] Government Decision no. 195 of April 4, 2012
[18] https://stisc.gov md/ro/-iaas-service-mcloud

The beneficiaries being:

- public authorities,
- state institutions,
- real sector enterprises.

Moldovan citizens face a lack of local digital content and a small number of electronic services. Another important aspect is the digitization of the cultural and scientific heritage, whose National Program for computerization of the cultural field for the period 2012-2020, was approved only in 2012.

Conclusions

In this paper, some challenges of big data analysis were discussed. Finally, some real-world big data analytics applications have been introduced. Big data is a significant area that offers many potential benefits and innovations. It is a remarkable field, with a promising future, if approached correctly. The difficulty with big data comes mainly from its size, which requires proper storage, management, integration, cleaning, processing and analysis.

In this regard, Moldova has the opportunity to position itself to take advantage of technological changes and the ongoing digital transformation. To some extent, it already has some of the necessary elements: in the last decade it has accelerated the digitalisation process of its government, and its local digital industry has created jobs and increased its exports. Both the private sector and the government are aligned in their vision to realize the opportunity of the digital economy. Well-defined and implemented public policies and programs can help here, helping to create more, better and more inclusive digital jobs.

At the same time, in the regional context, the Republic of Moldova, in contrast to the neighboring country, Romania, has a culture of using digital tools and online commerce. As discussed earlier, there are several sectors in which firms - especially smaller enterprises - use digital tools and services. These range from involvement in online marketing to production automation to the use of accounting or human resource management tools. Lack of information may be a key reason for a lower rate of adoption of digital services. Businesses may have limited awareness of digital technologies relevant to their sectors, lack of knowledge about where (reliably) they source these tools or technologies, limited (digital) skills among workers, or low levels of knowledge about where to acquire them. such skills and, in some cases, poor information about the costs and benefits of such investments.

With few exceptions, rural communities remain in coverage or access to digital technologies. Even though the mobile phone has become commonplace globally, other technologies have remained within reach or inaccessible to rural workers and businesses. And the self-employed can be located anywhere, working practically for an employer located anywhere in the world. But this potential is often unrealized due to low skill levels, poor infrastructure (e.g., inadequate electricity or bad roads that increase the cost of doing business) or the limited launch of appropriate technologies due to different market failures, starting with commercial viability. the base. no exception to these trends. With a vibrant local ICT industry and wide access to accessible internet service,all through mobile networks - it may be possible for rural

residents to take advantage of digital tools. However, rural areas are leaking, most skilled workers are moving to cities or even migrating.

High exchange rates and reliance on banks and card companies to process payments are also one of the biggest contributors to the low level of cashless (or electronic) payment channels used for e-commerce.

Some recommendations to prepare and improve business and government ecosystem frameworks to become more competitive and attractive for ICT investment,are,as following :

● Continue the expansion of ongoing programs and efforts as a collaboration of the public and private sectors, supporting the IT industry to improve quality and move to higher value-added products and services; raising the awareness of younger students and their families about digital skills and jobs

● Have a long-term increase investment in education through digital literacy, incorporating digital skills at all levels of education; increasing the coverage of training programs for Moldovans outside the IT industry in digital skills

● Accelerate the digitalization of enterprises, by improving the digital skills of today's workers in various sectors, overcoming the ICT industry

● Raise awareness of non-ICT businesses about digital opportunities and the need to implement Big Data storage and analysis tools

● identify several pilot sectors to test training approaches for businesses and individuals (workers) to improve their digital skills and abilities; set up a program to connect local IT expertise with companies to identify and implement digital upgrades, potentially with a focus on SMEs and state institutions, to accumulate more data sets, to be stored and processed through cloud systems.

Acknowledgements

This manuscript is part of the project State Program"Innovative business configuration in the context of regional competition", nb. 20.800020807.42