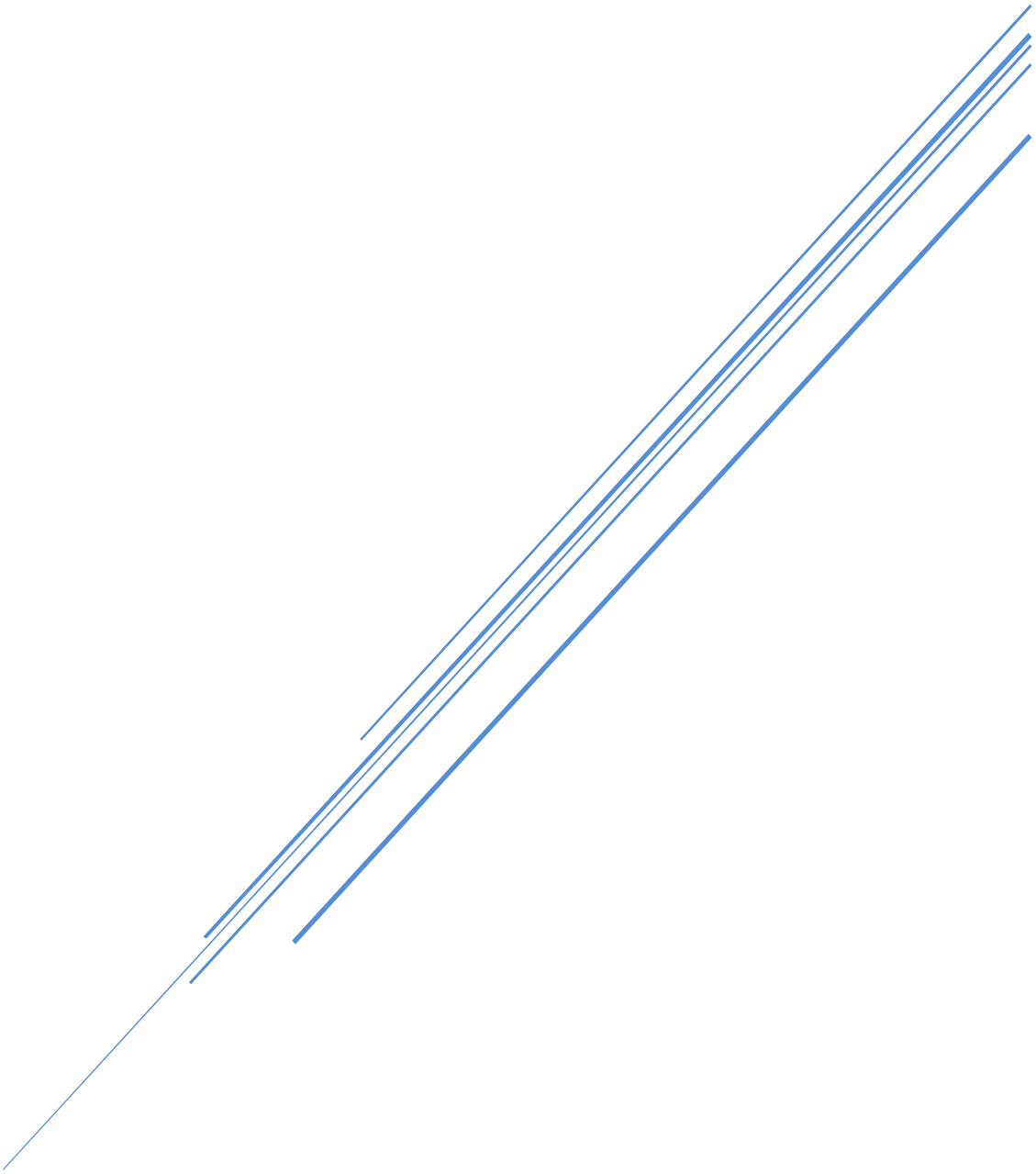
BUSINESS AND CHALLENGES IN 2024



Csiszár, Cs. M., Horváth, K. (2024). The role of data governance in corporate strategic decision-making.In: Musinszki, Zoltán; Horváth, Ágnes; Szűcsné, Markovics Klára (szerk.) Business and Challenges in 2024. Miskolc, Magyarország : MTA MAB Gazdálkodástudományi Munkabizottság (2024) 241 p. pp. 49-58. , 10 p.

The role of data governance in corporate strategic decision-making

Csilla Margit, Csiszár, csilla.csiszar@uni-miskolc.hu
Kata Horváth, kata.horvath@uni-miskolc.hu

Abstract: In the context of the 21st century's rapid globalization and digital transformation, organizations face increasing pressure to adopt strategic and resource-efficient data management practices. Data has emerged as a critical asset, fundamentally influencing business performance and competitiveness. This study explores the multifaceted dimensions of digitalization and the data-driven economy, underscoring the strategic significance of effective data management within modern enterprises. Drawing on an extensive literature review, the research examines the roles and responsibilities linked to data as a corporate asset and develops a theoretical framework for understanding data governance – a domain that remains underexplored both in Hungarian and global academic spheres. The findings aim to enhance corporate data governance practices, offering insights to strengthen strategic decision-making processes.

Keywords: data governance, data management, data-driven economy, digitalization, information
JEL code: M10, M15, M21, L20

Introduction

Thanks to various scientific and technological advancements, the data-driven digital economy has developed with remarkable intensity. By 2020, its estimated added value exceeded USD 32.6 trillion, accounting for 43.7% of the global GDP (Guo et al., 2022). Digitalisation has led to a huge explosion of data, while improving the infrastructure needed to exploit data abundance by supporting technological innovations and broadening access to data as a resource. The digital economy, conceptualized by American researcher Don Tapscott in 1996, is a network-based economy that can be understood as a collection of economic activities based on information technology (Zimmerman, 2015; Pan et al., 2022; Carlsson, 2004). The digital economy is driven by information digitalisation and the internet (Sidorov – Senchenko, 2020). In his 2021 dissertation, Tóth divides the digital economy into two subdivisions. One segment is the digital sector, which basically includes tools for managing digital information. Another significant aspect of the digital economy is the data economy, and the intersection of the two areas can be identified as the so-called attention market.
The transformation and spread of the digital economy also facilitates the development of traditional industries, and today the entire modern economy can be considered to be digital in the sense that there is no economic sector that does not use digital data in its daily operations (Li et al., 2021; Tóth, 2021). Thus, the development of technology driven by digitalization results in the datafication of many aspects of human life and economy (Ciuriak – Ptashkina, 2020). The simultaneous development of digitalization and technology enables not only an increase in data volume but also its intensive extraction. With the help of technologies such as artificial intelligence, data volumes that were previously incomprehensible become manageable, allowing the extraction of hidden, valuable information. The vertical and horizontal expansion of the available data and the intensive development of technological capacity mark the inflection point of the curve of economic development. This is also the point that represents the transition between the knowledge-based economy and the data-driven economy (Ciuriak, 2018).
In the following chapters, we will first review the key milestones in the emergence of data as a crucial resource and interpret the concepts of business information and data. Next, we will illustrate various aspects of evaluating information as capital, and after synthesizing the

available literature, we will outline the diversified conceptual framework of data goverance and data management.

1. Literature Review

1.1. Data is the New Oil

Technological development and digital transformation have fundamentally changed our perception of the world, society and economy (Ciuriak – Ptashkina, 2020). As a result of the digitalization of the global economy, information, data, and knowledge have become key production factors, significantly influencing global competition and transforming into economically relevant values and resources (Tóth, 2021). The amount of globally available data is growing exponentially year by year, the volume of data doubles approximately every two years (Ciuriak, 2018), which is partly due to the actors of the global economy as well as the members of the consumer society (Ződi, 2017).
Over the years, data has become a critical resource, affecting many elements of the value chain as well as countless industries, both in production and resource allocation and distribution (Mayer-Scönberger – Cukier; 2013). Based on all this, it does not seem at all an exaggeration that so much literature refers to the data as the new oil. The primary basis of comparison is to be found in the economic impact and role of resources, but the analogy is also apt from the point of view that, similarly to oil, data does not represent value in itself, in its "raw form", in order to exploit and realize its full potential, it must be refined and shaped according to business and use purposes. As gasoline or PVC can be extracted from oil through transformation, raw data is converted into customer information compiled for the purpose of supporting corporate management decisions.
However, there are some significant differences that we must emphasize. First, data is a non-competing good, meaning that one person's use of data does not limit its availability to other users. Secondly, data cannot be used up, its amount does not decrease depending on the usage, and even simultaneous access to the data is provided. Finally, the inexhaustible source of data, in contrast to the limited extraction of oil, does not result in a narrow supply of resources, but supports constant expansion (Haucap, 2019; Hirsch, 2014; Shepherd – Popovich, 2010).

1.2. Business Information

Pásztor and Popovics (2010) tackle information as "interpreted data", which means the aggregation, grouping and processing of the available data(s) according to purpose. Starting from this approach, we can also get closer to understanding the concept of data, since the definition mentioned above identifies data as the core of information**.** According to Bács (2023), data is a factor devoid of knowledge surplus value (information value), which only becomes information (or knowledge surplus) when a change occurs in this lack of value. In other words, we can say that data is a fact without added value, which becomes information depending on the given context, and its value is determined by an externally defined system of relations.

The different types of business information are illustrated in Figure 1 below.

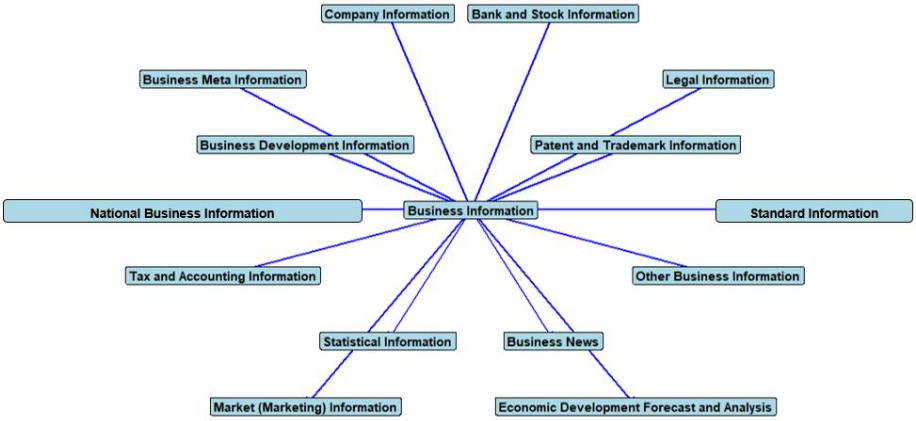


Figure 1: Different types of business information
Source: own edition based on Kiszl (2004)

1.3. Evaluation of information as capital

As Nordhaus and Samuelson (2016), Allen (1990), Chaudhary (2023) point out, information has economic value and, as such, can be subject to market exchange. However, it is extremely expensive to produce and relatively easy to reproduce (Coiera, 2000), which increases the risk of market failure. For this reason, specific state regulation of information goods and resources is needed, such as legislative practice aimed at protecting intellectual property (Nordhaus – Samuelson, 2016). However, determining the economic value of information is not an easy task, moreover, according to Birchler – Bütler (2010), it is only possible ex ante. In their view, the explicit estimation of the value of information is a paradox, since news can only be sold before its information content becomes known. Thus, it is actually the news, the source of information (newspaper, internet access, etc.) that is sold, not the actual news. In this sense, the value of information is actually determined by the value of the source of the information and not by the intrinsic value of specific information. The value of the source of information is primarily given by the possibility of reaction, which also determines the expected utility. Thus, in the interpretation of Birchler and Bütler (2010), the value of information (which is actually the value of the source of information) is determined by the change (increase) in the expected benefit expected from obtaining the information, and the results of the reaction and behaviour change generated by it.

A deeper understanding of the problem of determining the value of information is facilitated by looking at information as an intangible asset and analysing the issue by extending the validity of the limits defined by Kaplan – Norton (1996). In his 1996 work, Kaplan – Norton identifies the following factors as obstacles to the valuation of intangible assets:

1. the fact of indirect value creation;
2. dependence, which links information to a specific context, situation, degree of consistency with corporate strategy;
3. relativity of value, variability of its potential;
4. the close interconnection of intangible assets, which makes it impossible to determine value independent of corporate strategy or other organizational contexts.

In addition, the precise determination of the value of information is made difficult by the fact that, due to its easy reproducibility, traditional supply-demand relationships based on scarcity cannot be applied to a significant proportion of informational goods (Coiera, 2000). Finally, we must also take into account that the value of information is equally incomprehensible without its "packaging", i.e. without determining and managing the costs of the information system used

for production (Pásztor – Popovics, 2010). On the cost side, transport costs do not facilitate the evaluation process either, given that bundles of communication networks allow almost free transfer of information (Coiera, 2000). Overall, it can be stated that evaluating information as corporate capital is an extremely complex process that requires careful consideration and necessitates consistent and conscious data asset management by companies.

2. Establishing a conceptual framework for data governance

The previously described global processes, the digital transformation and the gradual reorganization of the general order of the world's economy require a conscious, resource-focused data approach of companies. Data and its instrumental management have become a determining factor and condition of corporate performance and market competition both among market participants and public authorities. At the same time, the roles and responsibilities related to data assets, as well as the theoretical and conceptual framework of data governance, can be considered rather poorly discussed in the literature. For these reasons, we considered it is necessary to review and synthesize all theories based on the available publications, which can bring us closer to a more differentiated understanding of the key processes of data governance and data management.

2.1. Data as a corporate asset

The economic significance of data and the possibility of its evaluation have already been discussed earlier, but the interpretation of the asset-like nature of data has remained unclear. For the economic definition of data, we consider the interpretation operationalized by the European Commission (2022) as guiding, which identifies data as information—primarily numerical and factual—that has been collected with the aim of supporting business processes and decision-making. In this way, even simple Excel spreadsheets compiled by companies can be considered data sources of economic importance – depending, of course, on their relevance to corporate activities.

The conscious corporate utilization of data as an input provides significant economic benefits, which further expands the possible dimensions of the scope of interpretation of data assets. The use of data inputs supports both quantitative and qualitative (assortment-based) growth of corporate activity. One of the significant advantages can be realized by the company along with increasing scale returns. Given that data is a non-competing good, its accumulation contributes to the improvement of data-based production and services, which generates even more data volume due to increasing demand and expanding user base. In the end, compared to the invested resources, the system generates much higher output and revenue. A similar advantage can be gained by the company by consciously managing assortment-based data sets. By diversifying the supply mix, companies can realize economic benefits by eliminating data silos (hermetically sealed data sets) belonging to certain subfields, i.e. through data linking and contextualization (OECD, 2015).

Data assets are also a phenomenon that can be interpreted at the level of the national economy, national data assets are a concept defined by law. According to Act CI of 2023, national data assets are *'all public data, documents and cultural public data managed by bodies performing public tasks, as well as other personal and protected data processed by them, regardless of their manifestation'.*

2.2. Data governance

There is no single literature proposal for a conceptual definition of data governance (or data asset management) (Naguib et al., 2024; Al-Ruithe et al., 2019; Mohanta – Jamdagni; 2024), so in our view, the most adequate conceptualization can be ensured by reviewing and synthesising the different interpretations and different approaches.

In most cases, research and academic papers written on the topic address data governance either through the description of relating decision-making areas or roles, or provide a superficial overview solely aimed at supporting the research, without any deeper interpretive intent (Abraham, 2019). For example, a study published by the OECD (2015) identifies data governance as a set of decision areas such as:

1. access to and reuse of data,
2. portability and interoperability,
3. connectivity and integration,
4. quality and care,
5. ownership and control,
6. value and pricing.

Weber et al. (2009) approach the concept of data governance through data quality management. In their interpretation, data governance refers to the area of corporate activity that focuses on defining data-related roles, then assigns responsibilities to these roles depending on the decision-making areas, and on the other hand, defines standards and guidelines related to data quality management (harmonized with corporate strategy, the data policy and legal provisions in force) for the organization as a whole.

In their research, Ritchie and Whittard (2024) build on a much narrower understanding of data governance, according to which data governance refers to the processes and procedures that ensure ethics, fairness, and appropriateness of data use and sharing.

In its book published in 2009, the Data Management Association International refers to data governance as a practice of authority and control over data management.

Khatri and Brown (2010) identify the personnel requirements (decision-making and responsibility areas) and specific tasks of data asset management. In their study, they identify five decision-making areas, which are as follows::

1. Data principles
2. Data quality
3. Metadata
4. Data access
5. Data lifecycle

In our view, the most complex definition of data governance can be found in the study of Abraham et al. (2019). According to their interpretation, data governance aims to establish and provide a comprehensive framework for managing data as a strategic corporate tool. This refers to clarifying the corporate decision-making authority related to data, formalizing data policies, various standards, and procedures, as well as monitoring compliance. This concept highlights several important factors related to data governance. First of all, data governance is a comprehensive activity that spans several functional areas of the company, data sets and databases, which increases the efficiency of data flow by creating collaboration between individual units, eliminating data silos, and reducing duplications and contradictions. Second, data governance is a framework that structures and formalizes subfields related to data management. It further declares that the data should be treated as a strategic asset of the company. It also interprets the decision-making and responsibilities related to data, i.e. in which areas decision-making about data is justified, and who can make these decisions and how. In addition, it highlights the role of data governance in creating data policy and setting standards

and procedures. Finally, it mentions the audit function, which checks deviations from expected management practices, compliance with established data policy principles and standards, and intervenes if necessary.
Using a similar approach, the National Bank of Hungary also formulated its Recommendation No. 19/2022 (XII.1.) *"... on the design, operation, framework and control functions of the process of compiling data reporting services of credit institutions, as well as on the related data governance tasks"* (MNB, 2022). In it, the minimum activities that need to be recorded in internal regulations of relevant departments are identified:

(a) the process of preparing the data to be compiled for the reporting area,
(b) the control steps used during preparation,
(c) the systems concerned,
(d) the life cycle of data used in the field,
(e) responsibilities,
(f) responsible for data quality (MNB, 2022 p. 7).

The definition of data governance can also be approached through the interpretation of the data value chain and the corporate responsibilities identified along it. The three essential elements of the data value chain are crucial for the data economy. The process starts with data acquisition and preparation and ends with data usage and evaluation. The three pillars can be further divided into subunits and subtasks. The two tasks of the first phase of the data value chain, which are decisive for decision-making, concern data sources and data storage. In connection with data analysis, the most important tasks are the frequency of analysis, its subject matter, and the identity of those with analytical powers and responsibilities. The last pillar is the evaluation of the data and then their use according to the target area, i.e. determining the areas in which each data represents value (even in monetary terms). These three activities and the related subtasks can therefore be considered as cardinal areas of data governance and data management. From the point of view of the data value chain and the data economy, technologies and the infrastructure that enable the business use of data are of particular importance (European Commission, 2022).

2.3. Data management

Data management has been discussed several times during the discussion of the previous sections, not by chance, as the activity is closely related to data governance both conceptually and practically. The primary goal of data governance is to define the framework that ensures the efficient management and use of IT tools and resources, and to clearly designate the circle of decision-makers and responsible personas. Basically, this is the strategic level of corporate activities related to data. In contrast, data management focuses on specifics and operational activities, creates and implements the decisions and tasks prescribed by the framework (Khatri – Brown, 2010).
Several subfields of corporate data management can be identified, with numerous scholarly recommendations in this regard. According to DAMA International (2017), the key components of data management are as follows (Figure 2).
Data Architecture refers to the definition of enterprise data objects and the development of a complex data model at conceptual, logical and physical levels (Dyché – Levy, 2006). Data architecture encompasses the definition of the company's data requirements, the establishment of architectural standards and guidelines, and the designation of relevant individuals and responsible parties (EFQM, 2011).
Data Modelling and Design is practically an explicit aspect of data use, i.e. it means analysing, designing, building, testing and maintaining, i.e. the general management of data (DAMA International, 2017).

The primary task of *Data Storage and Operations* is to select hardware and software-based IT tools that provide the most efficient data management in terms of functionality, cost, reliability, complexity, capacity, scalability and manageability (Tallon et al., 2014).
Data Security is responsible, among other things, for maintaining basic security requirements related to data accessibility, authenticity, availability, reliability and integrity. The most important tasks related to data security include assessing potential risks and defining data protection roles and standards (Faria et al., 2013).
Data Integration and Interoperability is a rather complex area, incorporating tasks related to data acquisition, extraction, transformation, movement and virtualization in addition to operational support (DAMA International, 2017).
Data integration and interoperability are supported by the *Documentation and Content* component by compiling a well-structured database. These include activities related to data storage, protection, indexing and access to unstructured data sources (DAMA International, 2017).
The field of *Reference and Master Data* management refers to the management of shared data and the reduction of redundancy, which aims to ensure better data quality through standardized definition and use of data values and registers (DAMA International, 2017).
Data Warehousing and Business Intelligence refers to managing analytical data processing and providing the ability to access, report and analyse decision-supporting data (DAMA International, 2017).
Metadata is practically nothing more than a data set compiled about data, a description. Metadata summarizes and interprets individual databases, practically providing a concise and consistent description of data representations, thus helping to understand the meaning and semantics of the data (Khatri – Brown, 2010). With the help of metadata, the sensitivity level, origin and data retention periods of data can also be classified (Abraham et al., 2019).
The last area of management *is Data Quality*. As a decisive resource for corporate decision-making, adequate data quality is of paramount importance, quality errors can pose significant risks for the company. Data quality practically characterizes the compliance of data with expectations and usage requirements interpreted in a given context (Khatri – Brown, 2010). Monitoring data quality means, among other things, defining relevant metrics and constantly measuring the level of data quality (Abraham et al., 2019; Malik, 2013).
As illustrated in Figure 2, data governance brings together the various subfields of data management as a kind of framework, which are interconnected. In fact, we could say that data governance provides a holistic approach to collecting, managing, protecting and storing data that supports the achievement of specific data management territorial goals.

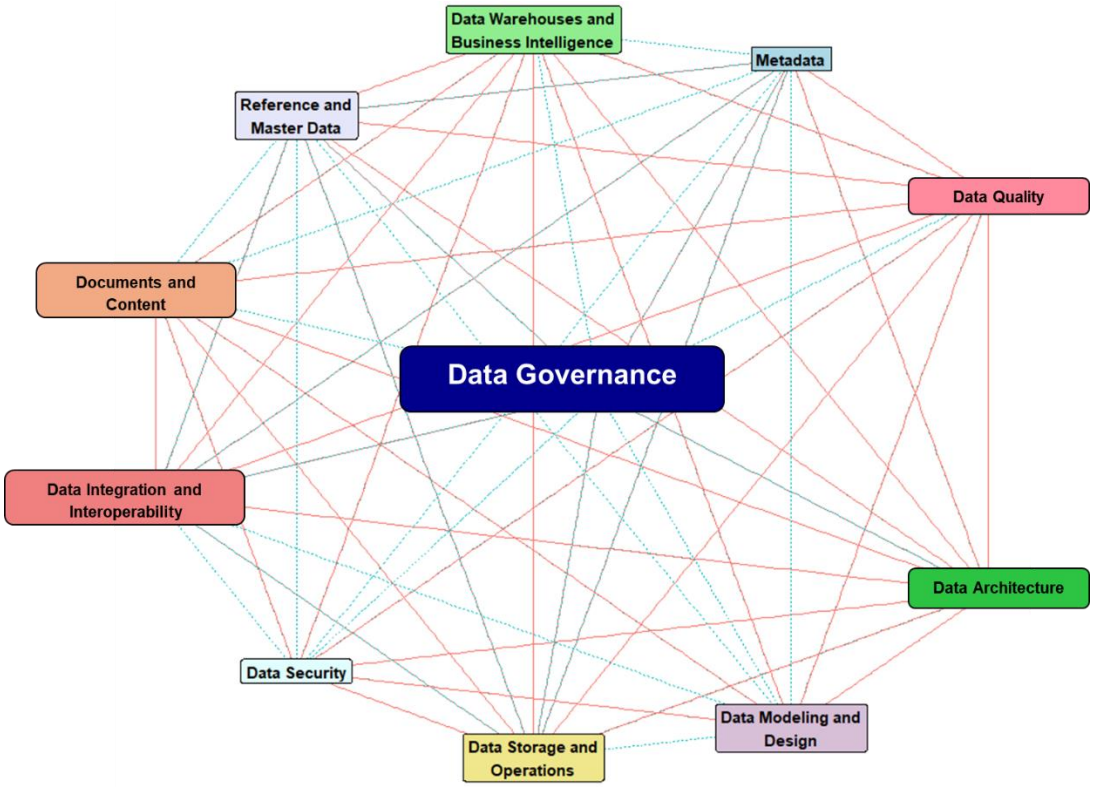


Figure 2: Main activities of data management
Source: own edition based on DAMA International (2017)

Summary

In our study, we found that digital transformation and the acceleration of global processes make it essential for companies to develop a conscious, resource-focused data approach. Treating data as an asset has become a key factor in corporate performance in the marketplace. In our thesis we examine the relationship between digitalization and the data-driven economy, emphasizing the strategic importance of corporate data sets and business information. Based on a literature review, we identified roles related to data governance and developed a theoretical framework that fills a gap in terms of scientific publications on the topic. Our findings have the potential to contribute to improving corporate data governance practices and supporting strategic decision-making.

References

Abraham, R., Schneider, J., Brocke, J. (2019): Data governance: A conceptual framework, structured review, and research agenda. *International Journal of Information Management*, 49, pp. 424–438. https://doi.org/10.1016/j.ijinfomgt.2019.07.008
Allen, B. (1990): Information as an Economic Commodity. *The American Economic Review*, 80(2), pp. 268–273.
Al-Ruithe, M., Benkhelifa, E., Hameed, K. (2019): A systematic literature review of data governance and cloud data governance. *Pers Ubiquitous Comput*, pp. 839–859.
Bács, Z. G. (2023). Gondolatok az információ szerepéről – más, egyéni szemszögből. *Nemzetbiztonsági Szemle*, 11(3), pp. 83–92.
Birchler, U. W., & Bütler, M. (2010): *Information economics* (Transferred to digital print). Routledge.

Carlsson, B. (2004): The Digital Economy: What is new and what is not? *Structural Change and Economic Dynamics*, 15(3), pp. 245–264. https://doi.org/10.1016/j.strueco.2004.02.001
Ciuriak, D. (2018): *The Economics of Data: Implications for the Data-Driven Economy*.
Ciuriak, D., & Ptashkina, M. (2020): *The State Also Rises: The Role of the State in the Age of Data*. Ciuriak Consulting Conference Paper.
Coiera, E. (2000). Information Economics and the Internet. *Journal of the American Medical Informatics Association*, 7(3), pp. 215–221.
Cukier, K., Mayer-Schönberger, V. (2014): The Rise of Big Data: How It's Changing the Way We Think about the World. In M. Pitici (ed.), *The Best Writing on Mathematics 2014* (pp. 20–32). Princeton University Press. https://doi.org/10.1515/9781400865307-003
DAMA International. (2017): *DAMA-DMBOK: Data Management Body of Knowledge* (Henderson D., Earley S, Eds.). Technics Publications.
Data Administration Management Association. (2009): *The DAMA guide to the data management body of knowledge: DAMA-DMBOK guide* (Mosley M., Brackett M., Earley S., Henderson D., Eds.; First edition). Technics Publications, LLC.
Dyché, J., Levy, E. (2006): *Customer data integration: Reaching a single version of the truth*. John Wiley & Sons.
Faria, F., Maçada, A. C., Kumar, K. (2013): Information Governance in the Banking Industry. *In Proceedings of the Annual Hawaii International Conference on System Sciences* (4445). https://doi.org/10.1109/HICSS.2013.270
Framework for Corporate Data Quality Management. (2014): *EFQM*.
Guo, Y., Mao, H., Ding, H., Wu, X., Liu, Y., Liu, H., & Zhou, S. (2022): Data-Driven Coordinated Development of the Digital Economy and Logistics Industry. *Sustainability*, 14(14), 8963. https://doi.org/10.3390/su14148963
Haucap, J. (2019): Competition and Competition Policy in a Data-Driven Economy. *Intereconomics*, 54(4), pp. 201–208. https://doi.org/10.1007/s10272-019-0825-0
Hirsch, D. D. (2014): The glass house effect: Big data, the new oil, and the power of analogy,. *Maine Law Review*, 66, pp. 373–395.
Kaplan, R. S., & Norton, D. P. (1996). The balanced scorecard: Translating strategy into action. Harvard Business School Press.
Khatri, V., Brown, C. (2010): Designing data governance. Commun. *ACM*, 53, pp. 148–152. https://doi.org/10.1145/1629175.1629210
Kiszl, P. (2004): Üzleti információszolgáltatás a magyarországi könyvtárakban. *TMT*, 51(9), pp. 371–378.
Li, J., Sun, W., Li, M., & Linlin Meng. (2021): Coupling coordination degree of production, living and ecological spaces and its influencing factors in the Yellow River Basin. *Journal of Cleaner Production*, 298, 126803. https://doi.org/10.1016/j.jclepro.2021.126803
Malik, P. (2013): Governing Big Data: Principles and practices. IBM Journal of Research and Development, 57, 1:1-1:13. https://doi.org/10.1147/JRD.2013.2241359
Mohanta, S., Jamdagni, A. (2024): A Survey on Taxonomy of Data Governance for Cloud-Based Services (pp. 99–109). https://doi.org/10.1007/978-981-99-8346-9_9
Naguib, H. M., Kassem, H. M., Naem, A. E. M. A. (2024): The impact of IT governance and data governance on financial and non-financial performance. *Future Business Journal*, 10.
OECD (ed.). (2015): *Data-driven innovation: Big data for growth and well-being*. Organization for Economic Co-operation and Development.
Pan, W., Xie, T., Wang, Z., & Ma, L. (2022): Digital economy: An innovation driver for total factor productivity. *Journal of Business Research*, 139, pp. 303–311. https://doi.org/10.1016/j.jbusres.2021.09.061
Pásztor, M. Z., Popovics, A. (2010): *Az információ gazdasági értéke*. pp.123–126.

Recommendation No. 19/2022 (XII.1.) of the National Bank of Hungary on the establishment, operation, framework and control functions of the process of compiling data reporting by credit institutions, as well as on the related data governance tasks. (2022). National Bank of Hungary.
Ritchie, F., Whittard, D. (2024): Using the Five Safes to structure economic evaluations of data governance. *Data & Policy*, 6. https://doi.org/10.1017/dap.2024.12
Samuelson P. A., Nordhaus W. D. (2016): Közgazdaságtan. Akadémiai Kiadó. https://doi.org/10.1556/9789630597814
Sidorov, A., Senchenko, P. (2020): Regional Digital Economy: Assessment of Development Levels. *Mathematics*, 8(12), 2143. https://doi.org/10.3390/math8122143
Survey of businesses on the data economy: Final report. (2022): Publications Office of the European Union.
Tallon, P., Ramirez, R., Short, J. (2013): The Information Artifact in IT Governance: Toward a Theory of Information Governance. *Journal of Management Information Systems*, 30, pp. 141–178. https://doi.org/10.2753/MIS0742-1222300306
Tóth, A. (2021): A tisztességes adatkereskedelmet biztosító szabályozás szükségességéről. *Állam-és Jogtudomány*, 62(3), pp. 100–121. https://doi.org/10.51783/ajt.2021.3.05
Weber, K., Otto, B., Oesterle, H. (2009): One Size Does Not Fit All—A Contingency Approach to Data Governance. *ACM Journal of Data and Information Quality*, 1, Article 4. https://doi.org/10.1145/1515693.1515696
Zimmermann, H.-D. (2000): Understanding the Digital Economy: Challenges for New Business Models. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2566095
Ződi, Z. (2017): Privacy és Big Data. *Fundamentum*, 1(7), pp. 18–30.