Decoding the data economy: a literature review of its impact on business, society and digital transformation

Andrea Sestino[a], Adham Kahlawi[b], and Andrea De Mauro[c]
[a] Università degli Studi di Bari Aldo Moro, Bari, Italy; [b] Università di Firenze, Florence, Italy;
[c] Università di Roma Tor Vergata, Rome, Italy

Abstract

Purpose - The Data Economy, emerging from the current hyper-technological landscape, is a global digital ecosystem where data is gathered, organized, and exchanged to create economic value. This paper aims to shed light on the interplay of the different topics involved in the Data Economy, as found in the literature. Our research provides a comprehensive understanding of the opportunities, challenges, and implications of the Data Economy for businesses, governments, individuals, and society at large, while investigating its impact on business value creation, knowledge, and digital business transformation.

Design/methodology/approach – We conducted a literature review that generated a conceptual map of the Data Economy by analyzing a corpus of research papers through a combination of machine learning algorithms, text mining techniques, and a qualitative research approach.

Findings – Our findings revealed eight topics that collectively represent the essential features of Data Economy in the current literature, namely 1) Data Security; 2) Technology Enablers; 3) Business Implications; 4) Social Implications; 5) Political Framework; 6) Legal Enablers; 7) Privacy Concerns; and 8) Data Marketplace. Our resulting model may help researchers and practitioners develop the concept of Data Economy in a structured way and provide a subset of specific areas that require further research exploration.

Originality - This is the first paper exploring the Data Economy opportunity for business value creation from a critical perspective.

Practical implications - Practically, our paper offers managers and marketers valuable insights to comprehend how to manage the opportunities deriving from a constantly changing competitive arena whose value is today also generated by the Data Economy.

Social implications - Socially, we also reveal insights explaining how the Data Economy features may be exploited to build a better society.

Keywords: Data Economy; Digital Transformation; Data Security; Literature Review; Privacy Concern; Data Marketplace; Business Strategy

This is the **preprint** version of the article published on August 2nd, 2023 in European Journal of Innovation Management © Emerald Publishing Limited. To cite this article: Sestino, A., Kahlawi, A. and De Mauro, A. (2023), "Decoding the data economy: a literature review of its impact on business, society and digital transformation", European Journal of Innovation Management, Vol. ahead-of-print No. ahead-of-print. doi.org/10.1108/EJIM-01-2023-0078

1. Introduction

The tremendous technological, industrial, and economic changes that occurred since the ’90s have overhauled the economic world (Yannopoulus, 2011), forcing businesses to question the process of strategy implementation and the approach to the business (Dahlman et al., 2016). The increasing

integration of technologies both in individuals' daily life and business environment radically changes the competitive context through the deep digitalization of traditional activities, ultimately resulting in a hyper-technological world and a novel digitized normalcy (Ardito *et al.*, 2022). Data, as the main output of any kind of digital technology, has been considered a major source of business value (Marjanovic, 2022). Indeed, the major technological integrations, from the simplest as management software (e.g., as for the ERPs), to the most complex such as those belonging to the world of intelligent objects, return a huge amount of data due to their communication and their constant interconnection on the Internet (Erevelles et al., 2016; De Mauro et al., 2018; Sestino et al., 2020). Not surprisingly, by considering the great amount of value that can derive from data content exploitation, researchers and practitioners recognized it as the new black gold of the twenty-first century (Wired, 2021). The importance acquired by data (Ansari et al., 2015) and the intuition that it could be a valuable source for various uses (Davenport et al., 2012) has led research towards an in-depth study of data storage, use and manipulation techniques (Trabucchi & Buganza, 2019). The type of communication generated by such hyper-connectivity creates a massive amount of data, also called Big Data (see De Mauro et al., 2019). However, from a managerial perspective, in the digital economy, Big Data acquires a fundamental strategic value, and - in a simplistic way - together with additional technologies they constitute a new "ecosystem" also known as the Data Economy. This is because the data are then analyzed and transformed into information, which can be used to solve business problems that in the past could not be addressed, but importantly they are able to create economic value and important financial returns (Zillner *et al.*, 2021).

The term Data Economy refers to a correlated set of activities aimed at the process of extracting value and valorizing different data, by means of varied but integrated processes of generation, collection, processing, analysis, automation, and exploitation of data made possible by digital technologies (European Commission, 2020). More concisely, the Data Economy denotes the ability of companies, individuals and institutions to manage the growing amount of digital information, according to the meaning of an increasingly data-driven society, by leveraging the great information exchanges that take place today via the Internet, and activated by the key enabling factors in terms of Internet of Things (IoT), Big Data, and Cloud (Lammi and Pantzar, 2019).

Indeed, Data economy is based on the digital ecosystems favouring the digital transformation, consisting of the current enabling technologies in terms of 1) Cloud, as the primary tool able to give life to digital ecosystems, on which the Data economy is based, because technically represented by the technological platforms connected to the network capable of hosting and allowing the dialogue between devices, things, people (see Giessmann & Legner, 2016); 2) the Internet of Things (IoT) technologies, concerning both the set of connected online devices able to intercommunicate with each other as an integral part of the life of individuals (e.g. in the case of smartphones, tablets, computers, laptops), and smart devices connected to apparatuses industrial plants, machinery, also known as smart objects (see Sestino et al., 2020); and 3) Big Data Analysis, referring to the process of data analysis, management, and exploitation directed to extract and capture the value hidden in those collected data. Such a complex process of examining Big Data to uncover information may finally help businesses and, generally, societies make informed decisions (De Mauro et al., 2018; Sivarajah et al., 2017).

Considering the data market as a place where digital data is exchanged as products or services, the European data market has witnessed growth from 2016 to 2020, increasing in value from less than 60 million euros to 80 million (UNCTAD Digital Economy Report, 2021). In 2020 Data Economy's contribution to the Italian Gross Domestic Product (GDP) was 34,3 billion euros, almost 25% of the entire country's GDP (European Commission Report on Digital Economy, 2021). In this ever-changing scenario, technological, industrial, and economic changes make it necessary to rethink the contribution to the creation of value of assets (Grimaldi & Cricelli, 2009), shifting the focus to the intangible ones.

Importantly, from an European perspective, Data Economy research efforts may be crucial because of the recent attention to this issue deriving from the EU as a whole, and each member state,

resulting the EU Data Strategy 2020 e.g., in incentivizing the creation of a Common European data spaces that finally may ensure that more data becomes available for use in the economy and society, while keeping companies and individuals who generate the data in control (European Commission, 2020): Data is thus an essential resource for economic growth, competitiveness, innovation, job creation and societal progress in general.

For instance the European Data Market (EDM) monitoring tool improved by the European Commission, providing the size and trends of the EU data market and data economy, the number of data professionals, the number of data companies and the revenues created by them, clearly show that the size of the data market has been evaluated at a growing trend, at €63.9 million in 2021, €440 billion in 2022, and expected to reach €600 billion in 2025 (IDC 4EU, 2022; Europa.eu, 2021).

By focusing on the Italian perspective, similarly to other EU countries, Italy suffers from dependence on technology providers and platforms managed by non-European providers: This can potentially make the European Union more vulnerable to external threats (e.g. with exposure to cybersecurity risks) and subject to a loss of investment and development potential by the European digital industry. However, Italy's data economy has activated an average of around €34.3 billion over the past two years (2020-2022). This value places Italy in third place in the European Union, after Germany (about 102.1 billion) and France (about 50.3 billion). However, among the top 15 European countries by value of the data economy, Italy is still the last for percentage growth in the last five years, equal to +25%, more than 10 percentage points less than Germany (+37 %) and France (+38%).

However, despite the relevance and the opportunity behind the Data Economy, because of the characteristics concerning such a phenomenon as an ecosystem built on the concept of three key enabler factors (IoT, Big Data, and Cloud), the current state-of-the-art is disorganized. More importantly, investigating the complex phenomenon of the Data Economy may be particularly interesting as suggested by previous literature (e.g., as in Akter & Wamba, 2016; Sadowski, 2019), and by considering policymakers' call for efforts in the Data Economy domain to sustain competitiveness (as in the European Data Act, unlocking the value of data from private companies in exceptional situations of high public interest, such as floods or wildfires; European Data Act, 2022) in order to systematize the concepts behind the differences in enabling technologies of the data economy (namely IoT, Big Data, and Cloud) and the opportunities deriving from the combination of these technologies for the creation of a common dataspace, with the aim of both identifying the competitive factors for companies, and contributing to a better positioning of the entire national systems as individual and then at a unitary level (e.g. at European level). Thus, in an attempt to systematize existing knowledge about the phenomenon under investigation, we implemented a literature review based on the utilization of text-mining techniques and machine learning algorithms. Similar to Sestino & De Mauro (2022), and on the basis of an original combination of two established machine learning algorithms (namely, LDA and hierarchical clustering, our review, based on the analysis of 250 academic papers, produced a human-intelligible topic structure that can serve researchers and professionals in making sense of the multifaceted nature of the Data Economy. Moreover, we also suggest a structured classification of the various streams of current research and a list of promising emerging trends.

The paper is organized as follows: in Section 2, we provide background on three fundamental enabler technologies at the basis of the Data Economy phenomenon; Section 3 offers a description of the Methodology we implemented to carry out the review; Section 4 then describes the conceptual model we identified as a result of our review; finally, in Section 5 we propose our conclusions and offer some perspectives for future research.

2. Theoretical Background: The Data Economy enabler technologies

2.1 Internet of Things

The Data Economy is built on the concept of three key enabler technologies, namely: the Internet of Things, Big Data (and the related Artificial Intelligence applications), and the Cloud (Tavazzi et al., 2021). The Internet of Things (hereafter, IoT) is a technology-based network of connections between objects, people, or between people and objects (Morgan, 2014). Whereas some scholars define IoT as a complex of specific technologies such as RFID (Berte, 2018), others qualify IoT as focusing on the result of implied technologies: a network of devices collecting and exchanging data (Gokhale, 2018). IoT now finds application in everyday life and different industrial fields as a differentiating factor in business competition (Attaran, 2017). From agriculture (Shenoy and Pingle, 2016) to healthcare (Kodali et al., 2015) and the energy sector (Motlaghet al., 2020). This importance as a key factor for economic growth shows results in the IoT industry's figures and raises debate about challenges and future opportunities. The IoT market has grown at a 20% Compound Annual Growth Rate (Al-Sarawi et al., 2020) and is estimated to contribute fourteen billion dollars to the global economy by 2030 (Constantinescu, 2019). Current literature covers challenges such as security, privacy concern, ethics, and Data quality problems are caused by inadequate domain-specific knowledge of proper measurement and connectivity (Hassija et al., 2019; Halim and Hutagalung, 2022).Furthermore, the literature covers cultural challenges like the worry about the emergence of new risks, user acceptance, and the fear of the ability to understand the new technology (Peter et al., 2023). Disadvantages such as an increased technical complexity (Ploennings et al., 2018), critical points in contraposition to Machine-to-Machine (M2M) technology (Balaji et al., 2019), and performance-related considerations (Haroon et al., 2016). Additionally, multiple opportunities have been identified concerning the Cloud (Biswas et al., 2014), Blockchain (Reyna et al., 2018), and Data Analytics (Marjani et al.,2017). Indeed, IoT is considered both an enabler (Sestino et al., 2020) and a driver (Tang, 2021) for digital transformation (Del Giudice, 2016).

2.2 Big Data and Artificial Intelligence

The definition of Big Data (hereafter BD) ranges from the technological capacity of managing data along their value chain to the cultural change progressively invading business and society (De Mauro, 2015). As a set of information primarily characterized by volume, velocity, variety (De Mauro *et al.*, 2016), veracity, value, and variability (Sestino et al., 2020), BD are constantly being analyzed and transformed into value for businesses and institutions. BD has been recognized as one of the drivers of digital transformation (Kostakis and Kargas, 2021). However, literature has identified a mix of benefits and barriers related to the utilization of BD. On the one hand, sales and marketing activities have benefited from the application of data analytics capabilities; on the other hand, low maturity, weak business support, difficulties in embracing new design paradigms (Russom, 2013) and privacy have put value creation at risk. Indeed, a privacy concern has arisen along the whole BD lifecycle making it necessary to work on a framework for its regulation (Cumbley and Church, 2013). Capitalizing on BD implies deep changes in business, such as: an increased focus on data flows in the place of data stocks, new professional roles such as data scientists, and novel IT roles designed to build a dynamic information ecosystem (Davenport et al., 2012). Furthermore, taking into account the attribute of the veracity of data, extensive research has been carried out: as a matter of fact, high-quality data is the requirement for analyzing and using BD in a productive way (Cai and Zhu, 2015). All in all, successful implementation of BD can have a major impact on firm performance, Industry 4.0, and innovation – but these benefits can only be realized with effective planning, positive workforce attitudes, and alignment with strategic goals. Overcoming organizational obstacles in these areas is crucial to unlocking the full potential of BD (Malik and Rybkowska, 2023). Artificial Intelligence, referred to as AI, is a technology which enables machines and software to reason, learn and interact intelligently (Surya, 2015), establishing connections with BD in predictive and descriptive data analysis, and improving the process of value creation (Surya, 2015). Their combined application shows opportunities in a wide variety of fields: precision agriculture to face up social challenges (Bahat and Huang, 2021), food industry (Misra et al., 2020), education (Luan et al., 2020), ICT for smart cities (Mashiko, 2020), and even to tackle the recent pandemic (Pham et al., 2021) who

also resulted in an acceleration of current societies digital transition (Sestino et al., 2021; Tosheva, 2020). In brief, the extensive utilization of AI in various applications necessitates conformity with standard regulations and new realities during implementation (Bui and Nguyen, 2023).

2.3 Cloud

Cloud computing refers to a variety of technologies (Malik et al., 2018) where both the applications delivered as a service via the Internet and the hardware and software as providers of these services merge together (Armbrust et al., 2010). Cloud platforms provide access and storage infrastructure through web services, ensuring scalability, reliability, performance, and tailored configurability (Wu et al., 2010). Cloud as a cutting-edge technology has innovated the ICT world and even broadly and profoundly transformed business fields such as the manufacturing industry and related business models, enabling the alignment of product innovation with business strategy and creating intelligent networks (Xu, 2012). Along with AI, Big Data profound cross-interconnections have been built (Chen, 2020), triggering a change in knowledge and networks (Zhuge and Sun, 2018). Together with challenges, some issues are raised, such as security threats (Shaikh and Haider, 2011) for both consumers and providers (Singh and Chatterjee, 2017), even in specific fields such as health (Lohr et al., 2010). More specifically, the Cloud is the technological platform on which the Data Economy is based (Harikrishna & Kiran, 2020). Thus, Cloud technologies are now fundamental for the ongoing digital transformation process that includes businesses, individuals, and institutions at all levels. Cloud technologies can have several functions that improve the performance and functionality of underlying technologies, from simple data storage to the creation of complex IT infrastructures (Alzakholi *et al.*, 2020). In this sense, it is now accepted that Cloud technologies are one of the primary keys to accelerate the digital transformation process thanks to the scalability, reproducibility of the solutions, the ease of implementation, and the high levels of performance and security that can be reached (Antonopoulos & Gillam, 2010). Moreover, Cloud technologies may enable the creation of innovative "business environments" based on collaboration between different subjects through information exchange (Strømmen-Bakhtiar, 2019). All in all, The proliferation of data, along with the progress in data science, has resulted in a requirement for a different type of cloud platform that can cater to the restrictions of limited hardware capabilities at the IoT device layer. The platform must be versatile, offer stringent security features, and be able to connect effortlessly with edge devices and supercomputers via high-performance networks (Suzumura et al., 2022; Volpert et al., 2023). A summary related to the contents present in this Section, is resumed in the synoptical Table 1.

Data Economy key enablers	**Related relevant literature**	**Main contents**
Internet of Things	(Ben-Daya *et al.*, 2019; Chen *et al.*, 2015; Dijkman *et al.*, 2015; Fleisch *et al.*, 2015; Del Giudice, 2016; Krotov, 2017; Sestino et al., 2020).	- IoT represents all those intelligent objects interconnected to the Internet that can interact with other objects and humans, based on intelligent algorithms. - IoT contributes to restructuring production and service delivery processes. - IoT may be exploited as an activator of business digitalization strategies. - IoT allows capturing a significant amount of data useful to decision-making and strategic management.

Big Data & Artificial Intelligence	(Allam and Dhunny, 2019; Duan *et al.*, 2019; Günther *et al.*, 2017; Sestino and De Mauro, 2022) Chong & Shi, 2015; Erevelles et al., 2016; Sestino & De Mauro, 2021; De Mauro et al., 2019; De Mauro et al., 2020	- BD represents the vast amount of data (5Vs) gatherable from ERPs, IoT, and each Internet-based interaction able to contribute to data generation; BD contributes to more conscious and informed decision-making. - BD helps to optimize the current offer to increase margins, expand the offer of new products and services, reduce time to market, and enrich customer engagement. - Analysis of BD models and their applications (e.g., AI, ML) helps to investigate and anticipate consumers' and competitors' behaviours, through prediction and recognition models.
Cloud	(Aleem and Ryan Sprott, 2013; Berman *et al.*, 2012; DaSilva *et al.*, 2013; Ho *et al.*, 2017; Marjanovic, 2022; Quinn *et al.*, 2014; Zhao and Zhou, 2022)	- Cloud represents the technological platform on which digital technologies are based, as a virtual infrastructure. - Cloud activates new digital business environments. - Cloud may increase Flexibility, no hardware costs, collaboration, data protection and security, sustainability, interoperability of information systems

Table 1. Synoptical table regarding the three Data Economy KETs

3. Methodology

In this Section, we introduce topic modeling techniques and describe each of the five steps we adopted to describe the essential concepts related to Data Economy by means of our literature analysis. Figure 1 provides a graphical summary of the various steps of the methodology utilized in this study, namely: 1) Document acquisition; 2) Text preprocessing; 3) LDA modeling; 4) Model selection; and 5) Topic discussion.

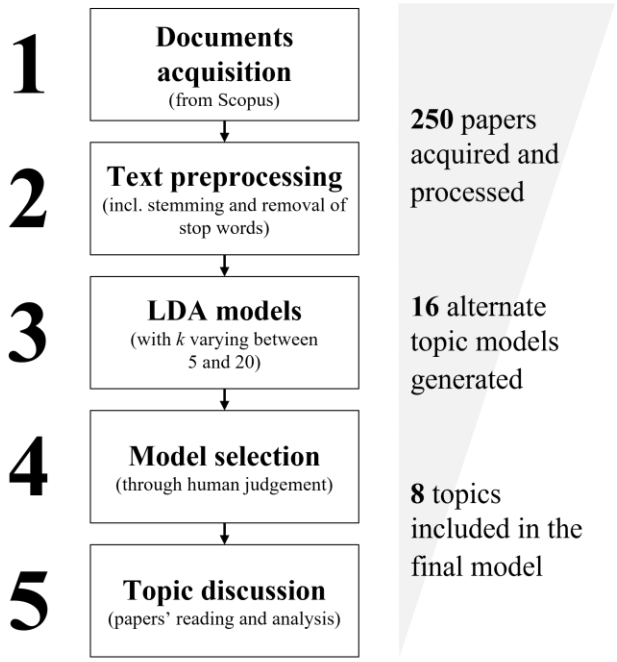

Figure 1. The five phases related to the proposed research design and the quantifications of entities involved at each step.

Literature reviews are a typical and crucial component of research projects to give a thorough overview of the body of knowledge already available on a given subject and spot patterns and trends in the literature. Starting an academic study with a comprehensive literature evaluation has numerous advantages. Firstly, it makes it possible to recognize the fundamental contributions to the field's advancement in science—those that most obviously influenced later investigation. Secondly, by categorizing earlier works, researchers can better comprehend the thematic organization of the field and decide where to concentrate their research efforts. Thirdly, historical trends of particular themes can be discovered through a systematic assessment of earlier publications spanning a considerable amount of time. Finally, it enables researchers to identify the most promising emerging trends and project the future prominence of each field of study.

The introduction of the Internet has changed how scholarly research is carried out and communicated. Researchers now have more access than ever to a wider variety of literature thanks to the growing availability of electronic versions of research articles (Borgman, 2007, 2015). The process of conducting academic research, including how academics find, access, and evaluate research literature, has undergone major modifications as a result. One of the most significant developments brought about by the availability of electronic versions of publications is the move toward more thorough and effective literature reviews. Although accessing single publications has become easier, the development of a framework that captures the entirety of a subject without omitting any significant contributions is still a complex and time-consuming task. Scientific literature reviews that take into account a vast number of pertinent papers are necessary to make sense of a particular topic. A common method for doing systematic reviews is based on the adoption of text-mining techniques. Such techniques take an unstructured data set (such as the body text of academic papers, their abstracts, or just their titles) and extract a more structured description of the subjects it contains.

The output of a traditional text mining method for literature reviews is a set of *N* clusters of

documents, each of which represents a topic covered in the literature. This method applies clustering algorithms over a matrix that describes the relative presence of popular keywords in a corpus of documents (document-term matrix), (Delen and Crossland, 2008; Fung *et al.*, n.d.; Moro *et al.*, 2014; Sunikka and Bragge, 2012). This method has the benefit of being very straightforward and equipped with processes that have been proven to be effective in determining the right amount of clusters to use, such as those outlined by Milligan and Cooper (1985). The most significant disadvantage of using clustering to extract topics is that each document can only be connected with a single cluster (i.e. a single topic). However, making links across other fields, often unexpected ones, is undoubtedly the most inventive and fascinating aspect of a scientific effort. The structural complexity of a scientific domain cannot be demonstrated by a topic model that does not allow the fact that articles usually discuss numerous topics at once. This necessitates the use of mixed membership models, which violate the notion that a unit must belong to a single cluster (Airoldi *et al.*, 2008, 2015). In fact, mixed membership models presume that individual units may be part of several population categories at the same time.

The Latent Dirichlet Allocation (LDA) is one of the most frequently utilized mixed-membership models for text mining, (Blei, 2012). LDA uses bayesian Estimation Techniques to estimate the number of parameters, including the vector indicating the topic proportion of each element (document) in each group (topic). The distribution of its words can also describe each topic, and the list of a topic's most likely words (keywords or top terms) can reveal the issue's essence. A human user can determine a relevant description of the issue by looking at the topic keywords and considering those documents treating the topic to the highest degree. The main result of a literature review will be a description of the subjects and the connections between them; in fact, these components are able to provide a clear and organized image of the topical material of the issue being studied.

The corpus of documents to be studied and the quantity k of themes to be extracted are the inputs for LDA. We choose to use the paper titles as documents, as done by (Vo and Ock, 2015). There is no universally accepted method for selecting the best value for k. A study of methods for determining the ideal number of topics was put forth by (Ponweiser, 2012). One method is to identify the harmonic mean of likelihood's greatest value on a set of samples produced by the Gibbs sampler, as was done by (Griffiths and Steyvers, 2004). As suggested by (Taddy, 2012) and implemented in the R package maptpx, calculating the topic model's marginal likelihood is an alternate strategy.

However, Chang et al., (2009) conducted tests to demonstrate how well these model optimization methods matched the objective of "human-interpretable" corpus breakdown, concluding that “topic models which perform better on held-out likelihood may infer less semantically meaningful topics” (p. 1). Instead of maximizing likelihood-based measurements, the authors advise topic model developers to embrace evaluation approaches that rely more on the "real-world" task that the work eventually wants to achieve. Instead of maximizing likelihood-based measurements, the authors advise topic model developers to embrace evaluation approaches that are more reliant on the “real-world task that the work eventually wants to achieve. This position was also corroborated by (Blei, 2012), who stated that there is no technical basis for assuming that the model’s accuracy would translate into a more effective organization or a simpler interpretation of the same. In a more recent survey, several performance criteria were assessed across multiple topic models (Abdelrazek *et al.*, 2023). The authors concluded that evaluation criteria are competing with each other and suggested assessing the nature of the application to select the most suitable model.

In the case of the present study, choosing a large value of k (e.g., as for those values of $k \gg 20$) could conflict with the objective of providing a short and concise explanation of the essential

literary topics within the Data Economy domain. Conversely, a low number of *k* would preclude the model from understanding the diverse range of fundamental issues that have been covered in the literature. In particular, during our initial runs of the LDA model, we discovered that utilizing $k < 5$ would cause us to ignore important areas of the literature and paint a simplistic picture of the Data Economy multi-faceted landscape. Therefore, we decided on the number of topics by choosing the model that produced the most understandable output in the authors' eyes, as was done in earlier literature (Almgerbi *et al.*, 2022; Delen and Crossland, 2008; Sestino and De Mauro, 2022).

We leveraged Scopus, a literature database containing more than 87 million documents from about 7000 international publishers (Elsevier, 2022), as a source for our collection of papers. We exported from Scopus 250 conference and journal articles containing the full term 'Data Economy' in either the title, the abstract, or the keywords. Then, we performed the following pre-processing operations on the input abstractions before running the LDA model on the entire corpus: (1) Punctuation and white spaces were all eliminated. Therefore, only single words were taken into account for the next several steps. Compound words with intra-word dashes were the only ones that were an exception to this criterion and were kept. (2) Lower caps have been used in place of all caps. (3) The Porter's approach was used to stem the corpus (1980). This returned each word's stem back without the suffixes. (4) The abstracts were cleaned up by eliminating common stop words in English (such as articles and conjunctions) and other irrelevant phrases (like years and copyright information). The terms "data" and "economy," which were contained by definition in all publications, were also eliminated (see Figure 1 above).

Afterward, we applied LDA to the corpus of 250 documents multiple times with the number of topics, *k*, ranging between 5 and 15, inclusive. By means of a collaborative review session among all the authors, we concluded that the model with $k=8$ produced the most intelligible results.

Lastly, we have utilized the output of the selected LDA model (most recurring keywords and topic presence for each document) to synthesize a description of the essential concepts within each topic. In doing so, we have also carefully analyzed the papers included in the review that displayed a high degree of presence of each topic. The resulting description of each topic and the resulting conceptual model are summarized in Section 4 of this paper.

4. Findings

We have identified eight topics by applying the previous Section's methodology. In this Section, we will present each topic's definition, derived from some of the most influential papers we encountered in our literature review. In addition, we will present the literature review most closely related to each topic to provide illustrations of the essential concepts connected to each topic.

4.1. Topic 1: Data Security

Data security refers to the measures and protocols put in place to protect sensitive and confidential data collected, stored, and shared by organizations from cyber-attacks that are becoming more complex and raising complicated security problems. This can include measures to prevent unauthorized access to data by individuals or systems. In addition, this can consist of measures to ensure that the original information is not shared and that only reports are communicated, as well as measures to ensure that data remains confidential and is used in a way that is compliant with relevant laws and regulations. Data security is crucial in the data economy, as organizations rely on the collection and analysis of data to drive innovation. In our literature review, we encountered multiple proposals of frameworks able to ensure high-level security of personal data. For example, Murtagh et al. (2012) discuss the emergence of a new data economy in contemporary bioscience and the ethical concerns surrounding the sharing of individual data. The authors present DataSHIELD, a technology that enables shared analysis of data without actually sharing any individual-level data, as a potential

solution to these concerns. They present an ethnographic study of a workshop testing DataSHIELD and demonstrate that it is practical, flexible, and fulfils the requirements of ethics committees. The authors also highlight that the success of DataSHIELD depends on more than just the technology itself, as it was based on social practices and scientific and ethical motivations. Johannis et al. (2020) discuss the importance of point-of-care testing (POCT) in hospitals and the need for hospital information technology (IT) security regulations to be adapted to the specificities of POCT.

The authors describe the results of a meeting held in 2019 that addressed these issues and established consensus in the areas of user, data, and update management, as well as network connections and user-friendliness. They recommend optimizing user management by connecting to a directory service, encrypting patient data and data transmission, using secure communication protocols, and establishing an updated management plan with protocol-based testing for remote services. The authors also emphasize the importance of providing an organizational structure for POCT-IT security, continuous training and awareness for this topic, and a focus on usability. We found that many papers covering data security are also dealing with privacy concerns and related regulations. For example, Reinhardt (2022) discusses the implications of the protection of personal data as a fundamental right, as outlined in Article 8 of the EU Charter of Fundamental Rights, for a digitized society. The author addresses specific issues related to the understanding of this article, including the concept of data protection as a fundamental right, the object of protection, and the meaning and scope of the article with respect to private parties. The author also explores the concept of the "horizontal effect" of Article 8, which refers to the way in which the provisions of the Charter apply to private parties, and the regulatory implications of this concept for the data economy. Notably, the author highlights the need to clarify the possible effects of Article 8 between private parties and the doctrinal grounds on which these effects can be constructed.

4.2. Topic 2: Technology Enablers

Technology enablers refer to the various technologies and tools that enable organizations to collect, store, and analyze data that can maintain the stability of data economy systems. We found that papers covering this topic discuss hardware and software tools such as data storage and processing systems, analytics platforms, and data visualization tools. Technology enablers also include technologies such as artificial intelligence and machine learning (De Mauro et al., 2022), which can be used to extract insights and value from large and complex data sets. These technology enablers are essential in helping organisations manage and utilize data to drive innovation and growth effectively. For example, Chuang et al. (2018) discuss the concept of the IoT data economy and propose the Reliable IoT Data Economic System (RIDES) as a way to provide a reliable, automatic, and intelligent trading environment for IoT data. RIDES uses smart contracts and blockchain technology to enable automated trading without human intervention and protect against tampering. It also includes a Smart Decision Making Module to help users develop pricing strategies and a trade dispute procedure to allow buyers to request refunds in the event of malicious sellers. The authors discuss the use of edge computing to alleviate the burdens on IoT devices and present a case study showing that the RIDES system can correctly perform transactions on IoT data. Lammi and Pantzar (2019) discuss the role of the citizen-consumer in the modern, digitalized data economy. The authors examine how the role of the citizen-consumer has evolved over time and how technological advancements have influenced their everyday lives. They focus on the commercial side of consumer citizenship and compare two periods in the history of technology: the 1930s-40s, when the US car industry promoted the mobile citizen-consumer, and the post-1990s, when the Internet and cell phones introduced a greater sense of mobility. The authors discuss how the digital turn has provided citizen-consumers with new channels of operation. Pennino et al. (2022) discuss the use of blockchain technology to support economic transactions in the IoT-based economy, increasing the overall business value (Sestino et al., 2022). The authors review several real-world IoT projects and applications that may use blockchain technology to facilitate transactions and consider several architectural choices in light of the requirements of these applications. They also survey additional financial tools that blockchain

technology can potentially bring to an IoT ecosystem and discuss the architectural impact of these tools. The authors note that there are few examples of IoT projects that fully exploit the potential of blockchain technology and discuss open problems and future research directions for making blockchain adoption easier and more effective for supporting an IoT economy.

4.3. Topic 3: Business Implications

Business implications are how data and technologies impact an organization's operations, strategy, and bottom line. This can include the potential benefits of using data to drive innovation and growth in the organization, such as increased efficiency and improved decision-making. Therefore, in the data economy, it is crucial for organizations to understand the potential business implications of their data-related activities and to implement appropriate measures to manage and mitigate the potential risks, such as data breaches or legal and regulatory challenges. The literature review includes works that debate the topic from different perspectives. Vertesi and Dourish (2011) researched how scientific collaborations share data and how their norms of data-sharing are closely tied to how they produce and acquire data. The authors draw on ethnographic work with two robotic space exploration teams to illustrate how these norms are best understood in the context of data production. They argue that digital artefacts are embedded in a broader data economy and that this perspective has implications for understanding data in its interactional context and introducing systems or policies that may conflict with the value of data in its context of production. Fosso Wamba et al. (2019) discuss the role of information quality in big data analytics (BDA) and how it impacts business value, user satisfaction, and firm performance. The authors propose a theory on information quality dynamics based on the appraisal-emotional response-coping framework and test it using data from 302 business analysts across various organizations in France and the USA. They find that information quality in BDA has four significant dimensions: completeness, currency, format, and accuracy. They also find that overall information quality has a significant, positive impact on firm performance, which is mediated by business value and user satisfaction. The authors propose an REBUS-PLS algorithm to automatically detect groups of users with similar behaviors when determining their perceptions of BDA information quality. The study offers managers a set of determinants for information quality and business value in BDA projects in order to enhance user satisfaction and firm performance. Jeon et al. (2022) discuss the emerging trend of open government data (OGD), through which public information is available for free online, and explores the value of the OGD service as a stated willingness to pay (WTP). The authors collaborated with the Seoul Metropolitan Government, which operates one of the largest OGD portals in the world, to measure the WTP of three stakeholder groups: people using OGD, people not using OGD, and city officials. The estimates suggest that people using OGD are willing to pay an average of $0.8 per month for the service, but this value is not significantly different from that of people not using OGD ($0.6). The authors also find that a higher proportion of people using OGD refuse to pay any amount for the service compared to people not using OGD, while a higher proportion of city officials give true zero responses compared to people not using OGD. The authors discuss the merits and limitations of quantifying intangible digitized public goods and propose future directions for research.

4.4. Topic 4: Social Implications

The social implications refer to the ways in which the collection, use, and sharing of data can impact society. In the data economy, vast amounts of data rapidly grow and increase through various sources, including social media, sensors, and other digital technologies. This data is often used to improve society through the improvement of the efficiency and effectiveness of various systems, such as transportation and healthcare. However, the collection and utilization of data can also raise concerns regarding privacy, security, and other social issues, which may negatively impact society. To assess the nature and scope of these impacts on society, it is crucial to consider factors such as the parties having access to this information and the ways in which it is being used. Gaining a clear understanding of the social implications of the data economy is vital for ensuring responsible and ethical use of data.

We found in our review several papers dealing with this topic. For instance, Cuquet and Fensel (2018) discuss the impact of big data on society and the need for a comprehensive strategy in Europe to optimize the use of data for societal benefit and increase innovation and competitiveness. The authors propose a research roadmap to capture the economic, social, ethical, legal, and political benefits associated with using big data in Europe. The roadmap considers the positive and negative externalities of big data, maps research and innovation topics to these externalities, and provides a time frame for addressing these topics in order to deliver social impact, skills development, and standardization. It also identifies which sectors will benefit most from each research effort. The goal of the roadmap is to guide European research efforts in developing a socially responsible big data economy and to allow stakeholders to identify and meet big data challenges while considering the societal impact, positive and negative externalities, and concrete problems worth investigating in future programs. Benthall and Goldenfein (2021) argue that the regulation and ethical analysis of technology and businesses doing data processing, which are based on liberalism and focused on the privacy and autonomy of individuals and the ordering of a public market, are largely incompatible with the techno-political and techno-economic dimensions of artificial intelligence (AI). The authors analyze liberal regulatory solutions in the form of privacy and data protection, regulation of public markets, and fairness in AI and show how the data economy and AI have surpassed liberal legal imagination. They argue that organizations use AI to exceed the bounded rationality of individuals and each other, leading to the private consolidation of markets and an unequal hierarchy of control that operates mainly for shareholder value. The authors propose the use of data intermediaries or sociotechnical systems composed of individuals working towards collectively pursued purposes, as an alternative to AI. They discuss the attention cooperative, a model of a social system that prioritizes its incoming and outgoing data flows and can form and maintain its own autonomous purpose, as an example. Schneiders et al. (2022) discuss the potential impacts and considerations of peer-to-peer (P2P) electricity trading, a new data-driven business model being trialed in the energy sector. The authors examine under what circumstances P2P electricity trading can be considered part of the sharing economy, drawing comparisons to the sharing economy in other sectors. They argue that a policy approach promoting dialogue and innovation, such as using regulatory sandboxes, is necessary to reap the full societal benefits of P2P electricity trading while avoiding risks to infrastructure and individuals. The authors suggest that using such a tool can help prevent the breakdown of trust between policymakers and platform companies that has been observed in other sectors.

4.5. Topic 5: Political Frameworks

This topic deals with the legal and regulatory frameworks that govern the economy's collection, use, and dissemination of data. Such frameworks are designed to protect the rights and privacy of individuals and organizations concerning their data. Meanwhile, these frameworks ensure that data is used to promote innovation and economic growth. Therefore, political frameworks in the data economy may include laws and regulations related to data privacy, data security, and data governance, among other things. Many scientific papers have focused on this aspect. For example, Marelli et al. (2020) discuss the tension between the GDPR and the expansion of digital health technologies and big data practices within the European digital health ecosystem. It is argued that certain aspects of the GDPR, such as key data protection principles and regulatory categories, the reliance on the "notice-and-consent" model, and the narrow scope of the Regulation in terms of addressing possible harms and discriminations, are not aligned with the rapid growth of digital health. This raises concerns about the GDPR's ability to govern current developments in the field effectively and calls for appropriate policy responses. Alastalo and Helén (2021) discuss the use of personal identity numbers (PINs) as a political technology in Finland, allowing interoperability between public authorities, private businesses, and data repositories. It examines the history and uses of the PIN in the Nordic welfare state and its role in enabling interoperability and facilitating public administration and state knowledge production. The article also explores the use of the PIN in the post-welfare data economy, where it has become a national asset in the Nordic countries, particularly in the areas of biomedical

research, innovation business, and healthcare. Lai and Flensburg (2021) discuss the mobile market as an ecosystem in which powerful companies generate value by continuously collecting and distributing data. This article advances the ecological metaphor by both theoretically conceptualizing and empirically exploring the contemporary app ecosystem. It identifies the key infrastructural resources that form the foundation of app-based communication (devices, operating systems, app stores, apps, third-party services, and data access) and their ownership structures in order to examine how power is obtained, exercised, and amplified in the app ecosystem. The article also provides a critical asset analysis of Google LLC and discusses its position as a dominant player in the app ecosystem.

4.6. Topic 6: Legal Enablers

The legal enablers are laws, regulations, and other legal frameworks that enable the development and growth of the data economy while guaranteeing everyone's rights. These enablers can include various legal frameworks, such as data protection laws that guarantee the privacy and security of personal data and laws that govern the collection, use, and sharing of data. Legal enablers are essential for providing a stable and predictable legal environment where companies and individuals can operate and innovate in the data economy. They can also help protect individuals' rights and interests and ensure that the data economy is fair and transparent. Some papers reported applications, like De Hert et al. (2018), suggest a pragmatic and extensive approach to interpreting the right to data portability, with a focus on maximizing its potential to benefit the Digital Single Market and the rights of digital users. The authors propose two perspectives on the right: the minimalist approach, which sees it as a means of transferring data when an individual leaves a service, and the empowering approach, which sees it as a way to allow individuals to combine data from multiple sources. The empowering approach is preferred. Lageson et al. (2021) analyze the types and volume of personally identifiable data released on the Internet by various government agencies in the United States, including law enforcement, criminal courts, and corrections. They find that these disclosures often include valuable personal information, such as names, birthdates, addresses, and physical characteristics, and estimate the volume of data disclosed online. They also estimate that every year, over 10 million arrests, 4.5 million mug shots, and 14.7 million criminal court proceedings are digitally released online and that approximately 6.5 million current and former prisoners and 12.5 million people with a felony conviction have a record on the Internet. The authors argue that these broad disclosures reveal an imbalance between transparency in terms of state action and transparency in terms of individual privacy and suggest that the criminal legal system contributes to the distribution of Internet privacy violations and community surveillance as part of the punishment. Liddell et al. (2021) discuss the legal and policy considerations around the treatment of health information in the UK. The authors argue that health information should not be considered property due to its unique characteristics and the potential drawbacks of a property framework regarding patient self-determination and market efficiency. They recommend a regulatory model that includes strong protections for personal health data, input from relevant stakeholders, interoperability, and focusing on a health-data service model rather than a goods model.

4.7. Topic 7: Privacy Concerns

Privacy concerns refer to individual and corporate concerns about who has access to their personal information and how it is used. These concerns can be about the potential for data to be used for surveillance or other purposes that infringe on an individual's privacy. Especially since sources like social media, sensors, and other digital technologies allow the collection of vast amounts of data that can impact an individual's privacy. Addressing these privacy concerns is important for ensuring that the data economy is fair and transparent and that individuals' rights are protected. This topic has been of interest to many researchers, including Cong et al. (2021), who discuss the use of consumer-generated data in economic growth. The authors propose a model in which intermediate goods producers use data to innovate and contribute to the production of final goods, driving economic

growth. They also examine the balance between profit and potential privacy infringement for consumers, who provide data, and the dynamic, nonrival nature of data with flexible ownership. The model suggests that, while an economy can grow at the same rate as the social optimum in a decentralized system, there may be inefficiencies in the R&D sector, including underemployment of labour and data overuse. The authors suggest that subsidizing innovators rather than directly regulating data may help to address these inefficiencies. They also note that as a data economy matures, consumers' data provision may decline, leading to initial growth traps that require intervention. Eviette and Simpson (2021) explore the use of metadata in the digital data economy and the privacy concerns surrounding its collection. The authors propose a means of allowing users to control access to metadata related to them, using the Solid data decentralization project and category-based access control. The goal is to address metadata collection's implications and privacy concerns. Fainmesser et al. (2022) analyze the incentives of digital businesses to collect and protect users' data and how the business's revenue model affects this strategy. The authors argue that businesses with a more data-driven revenue model will collect more users' data and provide more data protection than more usage-driven businesses. They also suggest that current regulation in the United States, which only focuses on data protection, may harm consumer surplus and overall welfare and that a more effective policy would combine a minimal data protection requirement with a tax proportional to the amount of collected data.

4.8. Topic 8: Data Marketplace

A data marketplace is a platform that allows organizations to buy and sell data. It can be a solution to many researchers, organizations and individuals to obtain data and, on the other hand, can be a solution to many data collectors to obtain additional profit from the data beyond its apparent purpose. The data marketplace can be considered a virtual marketplace where data is treated as a commodity that can be bought and sold. Organizations can use data marketplaces to find and purchase data relevant to their business, such as market research and customer or industry-specific data. Organizations can also use them to sell their own data to other organizations or individuals. Data marketplaces can be online platforms that facilitate the exchange of data between buyers and sellers, or they can be physical locations where data is bought and sold. A wide range of organizations, including businesses, research institutions, and government agencies, can use data marketplaces. Papers in the literature review have discussed the mismatch problem between data offer and data demand, so researchers like Sharma et al. (2020) propose a solution to data producers' inability to effectively monetize their data and data consumers having difficulty obtaining the data they need. The solution is to create a common data-sharing platform, or data marketplace, using blockchain technology to ensure decentralization and security. The platform would allow data producers to sell their data and data consumers to purchase it easily. The authors suggest that combining data marketplaces with blockchain technology can create fair and independent marketplaces for data. Abbas et al. (2021) conducted a systematic review of research on data marketplaces, finding that the current literature is primarily focused on technical topics such as computational pricing and architecture. They argue that for data marketplaces to move beyond the platform design stage and achieve adoption, more research is needed in non-technical areas like customer expected value and market segmentation. They suggest that this empirical research could help advance data marketplaces towards commercial viability. Maher and Khan (2022) discuss using biological and biomedical data in the healthcare industry, particularly during the COVID-19 pandemic. It notes that while electronic health records (EHR) and wearable digital devices generate a significant amount of health data, much of it goes unused, potentially due to legislative constraints, centralized data management systems, and a lack of incentives for data sharing. The authors propose the creation of a digital health data marketplace (DHDM) using blockchain and distributed ledger technologies, which could incentivize data sharing through a shared economic model. The paper reviews existing technical models and frameworks for implementing and optimizing a DHDM, as well as legal limitations and the potential socioeconomic impacts on various stakeholders.

4.9. Data Economy Conceptual Model

The eight topics identified through the LDA model application provide a conceptual map of the current literature on the Data Economy, which is visually summarized in Figure 2.

This conceptual model explores the complex interplay between the various concepts of data economy, revealing critical connections between them. In discussing the topic modeling output, we have recognized several noteworthy connections between topics that are more closely related to each other. Specifically, Privacy Concerns and Data Security are directly related, as both encompass numerous research contributions focused on protecting Personally Identifiable Information (PII) from unauthorized access (Murtagh *et al.*, 2012). Moreover, Legal Enablers and Political Frameworks represent two aspects of the broader issue of developing rules and regulations through political discourse (Koskinen *et al*., 2019). Lastly, we observed that the topics of Data Marketplace and Business Implications frequently appear together in papers, as the manner in which data is bought, sold, and exchanged in a marketplace directly influences business strategies and revenue generation models (Abbas *et al*., 2021).

In the conceptual model summary presented in Figure 2, topics with closer connections in the literature are represented as being nearer to each other.

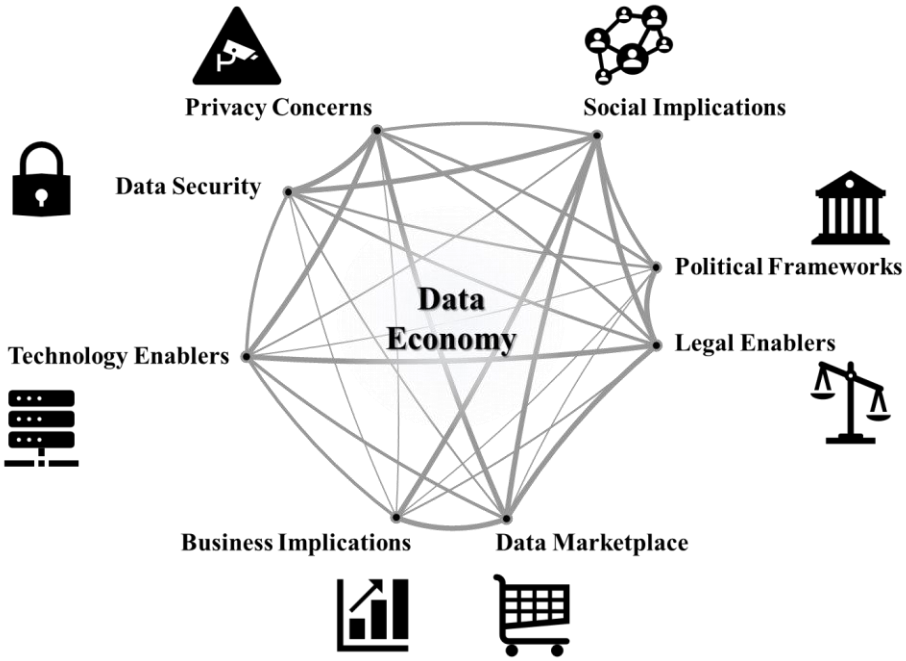

Figure 2. Data Economy Conceptual Model, displaying the eight fundamental topics identified in current literature.

5. Conclusion

This paper aims to investigate the impact of the Data Economy on business value creation and knowledge. Through a qualitative research design, a literature analysis related to theoretical and empirical studies has been performed by combining LDA topic modeling algorithm with text processing techniques. Findings shed light on eight fundamental topics dealing with the ambivalent issues and opportunities related to the Data Economy and its contribution to business value creation and knowledge, referred to: 1) Data Security; 2) Technology Enablers; 3) Business Implications; 4) Social Implications; 5) Political Framework; 6) Legal Enablers; 7) Privacy Concerns; and 8) Data Marketplace.

Findings firstly confirmed how data are an increasingly valuable asset for firms, governments, individuals, and society at large: the availability and prevalence of data have given rise to new or

significantly improved products, services, and business models and have contributed to enhancing productivity. Moreover, the technologies behind the Data Economy and the Data Economy represent a critical driver of innovation, growth, and value creation (Elia *et al.*, 2020; Marjanovic, 2022). The resulting society has transformed how businesses operate and created new opportunities for them, changing how consumers behave and giving them more power. More importantly, by considering the concept of the knowledge economy as a system of consumption and production that is based on intellectual capital, the Data Economy notably impacted digital transformation and business digitalization, affecting firms' performances as well (North and Kumta, 2020; Santoro *et al.*, 2021).

Findings reveal how the Data Economy-inspired approach may positively influence those economic activities that use digitized information and knowledge as critical factors of production. An orientation towards the Data Economy allows one to reap the benefits deriving from the technological integration efforts implemented by companies and by society in general, as it is able not only to contribute to renewed digital-based processes fully but to appreciate the true final value of the collected data and to use it as a productive factor and strategic resource for future planning.

In this sense, the correct innovation management and digitization efforts may allow the creation of a real data economy based on the ability of companies to manage the growing amount of digital information. Within this scenario, data is a fundamental resource for offering economic actors access to a vast range of products and services in the public and private sectors, ultimately guaranteeing significant returns in economies of scale.

Our findings, related to the emerging topics, unveiled interesting theoretical and managerial implications and some suggestions for policymakers. From a theoretical perspective, our study systematizes knowledge in the domain of Data Economy. To the best of the authors' knowledge, the current work is the first effort to organise the corpus of knowledge around this topic through an appropriate classification of emerging trends in combination with the three main enabling technologies of the Data Economy. More importantly, our study advances knowledge in the domains related to emerging topics. For instance, findings may contribute to the field of study related to the relevance of Data Security in innovation management strategies, requiring an appropriate balance between the exploitation of the knowledge deriving from the data themselves and the legitimate interests attributable to those to whom those data are attributable (Reinhardt, 2022). Furthermore, we add knowledge on the domain related to the technology enablers of Data Economy by identifying the main three activators of Data Economy-based actions, namely: Big Data, IoT, and Cloud. They correspond to multiple facets of the same phenomenon, able to promote digital transformation processes (Sestino *et al.*, 2020). Moreover, as for Business implications, our findings clarify the field of the study, highlighting the unforeseen opportunities deriving from a fully-digitalized world. In such a scenario, marketers and managers may be able to exploit the value that can be extracted from every single data to enhance their business and propose increasingly customized value propositions (Fosso Wamba *et al.*, 2019).

Additionally, as for Social Implications, our findings confirm the relevance of a Data Economy-driven approach in digitalization strategy to exploit technological opportunities for building a better society (Chantamas, 2021). New technologies represent the turning point of the current society, which breaks the link with the previous one and becomes an information society. In such a society, the exploitation of data and the incentive to use new technologies can be exploited in favour of end users and for the construction of a better world, and not just unethically, for pure business purposes. By considering the Political Framework topics, our findings not only may contribute to theory by adding knowledge on how properly manage and incentivize the adoption of those technologies building the Data Economy world but may also suggest interesting implications for policymakers. Indeed, despite the recognized relevance of the subject in the Data Economy, the maturity in Europe is lower than that of the USA and China (e.g., as revealed by Liu, 2021), which are more advanced in terms of digital business models exploiting data, information, and technologies. More specifically, in Europe, at a systemic level (companies, institutions, consumers), reaching data sovereignty should be considered a necessary and strategic ambition. In fact, the full exploitation of

the Data Economy depends on digital technology providers who are not Europeans (Kaloudis, 2022). Additionally, the investments guaranteed at the European level and by the Member States should aim at the diffusion of broadband connectivity which amplifies the Data Economy impact (allowing the creation of numerous digital businesses), guaranteeing faster connections, greater data generation and an unprecedented wealth of information (European Commission, 2022). Thus, in light of the hyper-competitive international scenario, Europe must achieve genuine strategic autonomy in the technological field.

Regarding Legal Enablers, the opportunity derived by the Data Economy to promote fully-digitalized ecosystems should be fully exploited both in an attempt to reach a better society by capturing the hidden value of data for social and business purposes and to boost the technological innovation adoption in strategic sectors. However, the Data Economy does not hide the limits still rooted in end users' minds regarding barriers to adopting new technologies and sharing personal data. Although some studies (e.g. Gutierrez et al., 2019) explain how in a hyper-digitalized society, end users are aware of having to give up part of their data to live to have full access to value-adding services, technophobia is still rooted. Marketers, managers, and policymakers' efforts should thus be aimed at mitigating this concern in consumers and, through appropriate communication and awareness campaigns, enhance the added value deriving from the information economy. Finally, as for Data Marketplace, the Data Economy requires appropriate data exchange platforms to create a synergistic ecosystem of innovation and information whose vast amounts of data can be processed to extract the greatest possible value. Despite some timid current efforts, such as the European project called Gaia X aimed to support the development of a federation of data infrastructure and service providers for Europe to ensure European digital sovereignty (Braud et al., 2021), theoretical and managerial advances are requested in the attempt to build ad hoc virtual spaces, which can pool the innovation efforts promoted by businesses, research institutions, and government agencies for the purpose to create a single synergistic ecosystem of innovation.

In spite of the promising findings emerging from our research, this paper is not exempt from limitations. Firstly, coherently with our methodological approach, we only considered research articles written in English that appeared in Scopus' most critical managerial, marketing, and engineering top journals. Consequently, some findings and insights related to non-English contributions and other types of publications may not have been considered. Additionally, since our method for choosing the topic model and characterizing subjects is mainly based on human judgment, it is possible that we overlooked novel or less obvious conceptual elements of the Data Economy concept.

In conclusion, digitization is rapidly impacting many aspects of people's lives and companies worldwide, which must adapt to societal changes and review the functioning of their businesses. At the same time, digital technologies extend well beyond the concepts of digitization and automation. Leveraging a Data Economy-oriented approach can have the advantage of accelerating technological innovation by enabling the total exploitation of the main output of the technological crafts, i.e. the data. This approach can result in a synergistic innovation ecosystem, giving full dignity to the digitization processes initiated by businesses, governments, and societies.

Acknowledgements

The authors would like to thank Carlo Mantini for his valuable contribution to the development of this study.

References

Abbas, A.E., Agahari, W., van de Ven, M., Zuiderwijk, A. and de Reuver, M. (2021), “Business data sharing through data marketplaces: a systematic literature review”, Journal of Theoretical and Applied Electronic Commerce Research. doi: 10.3390/jtaer16070180.

Abdelrazek, A., Eid, Y., Gawish, E., Medhat, W. and Hassan, A. (2023), "Topic modeling algorithms and applications: a survey", Information Systems, Vol. 112, 102131, doi: 10.1016/j.is.2022.102131.

Airoldi, E.M., Blei, D., Fienberg, S. and Xing, E. (2008), "Mixed membership stochastic blockmodels", in Koller, D., Schuurmans, D., Bengio, Y. and Bottou, L. (Eds), Advances in Neural Information Processing Systems, Vol. 21, Curran Associates.

Airoldi, E.M., Blei, D.M., Erosheva, E.A. and Fienberg, S.E. (2015), Handbook of Mixed Membership Models and Their Applications, CRC Press, Oxford.

Akter, S. and Wamba, S.F. (2016), "Big data analytics in E-commerce: a systematic review and agenda for future research", Electronic Markets, Vol. 26, pp. 173-194.

Alouffi, B., Hasnain, M., Alharbi, A., Alosaimi, W., Alyami, H. and Ayaz, M. (2021), "A systematic literature review on cloud computing security: threats and mitigation strategies", IEEE Access, Vol. 9, pp. 57792-57807.

Al-Sarawi, S., Anbar, M., Abdullah, R. and Al Hawari, A.B. (2020), "Internet of things market analysis forecasts, 2020-2030", 2020 Fourth World Conference on smart trends in systems, security and sustainability (WorldS4), IEEE, pp. 449-453.

Alastalo, M. and Helén, I. (2021), "A code for care and control: the PIN as an operator of interoperability in the Nordic welfare state", History of the Human Sciences, Vol. 35 No. 1, pp. 242-265, doi: 10.1177/09526951211017731.

Aleem, A. and Ryan Sprott, C. (2013), "Let me in the cloud: analysis of the benefit and risk assessment of cloud platform", Journal of Financial Crime, Vol. 20 No. 1, pp. 6-24, doi: 10.1108/13590791311287337.

Allam, Z. and Dhunny, Z.A. (2019), "On big data, artificial intelligence and smart cities", Cities, Vol. 89, pp. 80-91, doi: 10.1016/j.cities.2019.01.032.

Almgerbi, M., De Mauro, A., Kahlawi, A. and Poggioni, V. (2022), "A systematic review of data analytics job requirements and online-courses", Journal of Computer Information Systems, Vol. 62 No. 2, pp. 422-434, doi: 10.1080/08874417.2021.1971579.

Alzakholi, O., haji, L., Shukur, H., Zebari, R., Abas, S. and Sadeeq, M. (2020), "Comparison among cloud technologies and cloud performance", Journal of Applied Science and Technology Trends, Vol. 1 No. 2 SE, pp. 40-47, doi: 10.38094/jastt1219.

Antonopoulos, N. and Gillam, L. (2010). Cloud computing (Vol. 51, No. 7). London: Springer.

Ardito, L., Cerchione, R., Mazzola, E. and Raguseo, E. (2022), "Industry 4.0 transition: a systematic literature review combining the absorptive capacity theory and the data–information–knowledge hierarchy", Journal of Knowledge Management, Vol. 26 No. 9, pp. 2222-2254, doi: 10.1108/JKM-04-2021-0325.

Armbrust, M., Fox, A., Griffith, R., Joseph, A.D., Katz, R., Konwinski, A. and Zaharia, M. (2010), "A view of cloud computing", Communications of the ACM, Vol. 53 No. 4, pp. 50-58.

Attaran, M. (2017), "The internet of things: limitless opportunities for business and society", Journal of Strategic Innovation and Sustainability, Vol. 12 No. 1, pp. 11-17.

Ben-Daya, M., Hassini, E. and Bahroun, Z. (2019), "Internet of things and supply chain management: a literature review", International Journal of Production Research, Vol. 57 Nos 15-16, pp. 4719-4742, doi: 10.1080/00207543.2017.1402140.

Benthall, S. and Goldenfein, J. (2021), "Artificial intelligence and the purpose of social systems", Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society, New York, Association for Computing Machinery, pp. 3-12, doi: 10.1145/3461702.3462526.

Berman, S.J., Kesterson-Townes, L., Marshall, A. and Srivathsa, R. (2012), "How cloud computing enables process and business model innovation", Strategy and Leadership, Vol. 40 No. 4, pp. 27-35, doi: 10.1108/10878571211242920.

Berte, D.R. (2018), "Defining the IoT", Proceedings of the International Conference on Business Excellence, Vol. 12 No. 1, pp. 118-128.

Biswas, S., Akhter, T., Kaiser, M.S. and Mamun, S.A. (2014), "Cloud based healthcare application architecture and electronic medical record mining: an integrated approach to improve healthcare system", 2014 17th international conference on computer and information technology (ICCIT), IEEE, pp. 286-291.

Blei, D.M. (2012), "Probabilistic topic models", Commun. ACM, Vol. 55 No. 4, pp. 77-84, doi: 10.1145/2133806.2133826.

Borgman, C.L. (2007), Scholarship in the Digital Age: Information, Infrastructure, and the Internet, The MIT Press, doi: 10.7551/mitpress/7434.001.0001.

Borgman, C.L. (2015), Big Data, Little Data, No Data: Scholarship in the Networked World, The MIT Press, New York.

Braud, A., Fromentoux, G., Radier, B. and Le Grand, O. (2021), "The road to European digital sovereignty with Gaia-X and IDSA", IEEE Network, Vol. 35 No. 2, pp. 4-5.

Bui, T.H. and Nguyen, Van P. (2023), "The impact of artificial intelligence and digital economy on Vietnam's legal system", International Journal for the Semiotics of Law – Revue Internationale de Sémiotique Juridique, Vol. 36 No. 2, pp. 969-989, doi: 10.1007/s11196-022-09927-0.

Cai, L. and Zhu, Y. (2015), "The challenges of data quality and data quality assessment in the big data era", Data Science Journal, Vol. 14 No. 2, p. 2.

Chang, J., Gerrish, S., Wang, C., Boyd-graber, J. and Blei, D. (2009), "Reading tea leaves: how humans interpret topic models", in Bengio, Y., Schuurmans, D., Lafferty, J., Williams, C. and Culotta, A. (Eds), Advances in Neural Information Processing Systems, Vol. 22, Curran Associates.

Chantamas, M. (2021), "Book review: marketing 5.0 technology for humanity", ABAC Journal, Vol. 41 No. 2, pp. 271-275.

Chen, F., Deng, P., Wan, J., Zhang, D., Vasilakos, A.V. and Rong, X. (2015a), "Data mining for the internet of things: literature review and challenges", International Journal of Distributed Sensor Networks, Vol. 11 No. 8, 431047, doi: 10.1155/2015/431047.

Chen, F., Deng, P., Wan, J., Zhang, D., Vasilakos, A.V. and Rong, X. (2015b), "Data mining for the internet of things: literature review and challenges", International Journal of Distributed Sensor Networks, Vol. 11 No. 8, 431047.

Chong, D. and Shi, H. (2015), "Big data analytics: a literature review", Journal of Management Analytics, Vol. 2 No. 3, pp. 175-201.

Chuang, I.-H., Weng, T.-C., Tsai, J.-S., Horng, M.-F. and Kuo, Y.-H. (2018), "A reliable IoT data economic system based on edge computing", 2018 IEEE 29th Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC), pp. 1-5, doi: 10.1109/PIMRC.2018.8580742.

Cong, L.W., Xie, D. and Zhang, L. (2021), "Knowledge accumulation, privacy, and growth in a data economy", Management Science, Vol. 67 No. 10, pp. 6480-6492, doi: 10.1287/mnsc.2021.3986.

Constantinescu, E.M. (2019), "New marketing solutions for IoT market", Knowledge Horizons. Economics, Vol. 11 No. 2, pp. 26-31.

Cumbley, R. and Church, P. (2013), "Is "big data" creepy?", Computer Law and Security Review, Vol. 29 No. 5, pp. 601-609.

Cuquet, M. and Fensel, A. (2018), "The societal impact of big data: a research roadmap for Europe", Technology in Society, Vol. 54, pp. 74-86, doi: 10.1016/j.techsoc.2018.03.005.

Dahlman, E., Parkvall, S. and Skold, J. (2016), 4G, LTE-Advanced Pro and the Road to 5G, Academic Press, New York.

DaSilva, C.M., Trkman, P., Desouza, K. and Lindič, J. (2013), "Disruptive technologies: a business model perspective on cloud computing", Technology Analysis and Strategic Management, Vol. 25 No. 10, pp. 1161-1173, doi: 10.1080/09537325.2013.843661.

Davenport, T.H. and Patil, D.J. (2012), "Data scientist", Harvard Business Review, Vol. 90 No. 5, pp. 70-76.

De Hert, P., Papakonstantinou, V., Malgieri, G., Beslay, L. and Sanchez, I. (2018), "The right to data portability in the GDPR: towards user-centric interoperability of digital services", Computer Law and Security Review, Vol. 34 No. 2, pp. 193-203, doi: 10.1016/j.clsr.2017.10.003.

De Mauro, A., Greco, M. and Grimaldi, M. (2015), "What is big data? A consensual definition and a review of key research topics", AIP Conference Proceedings, Vol. 1644 No. 1, pp. 97-104.

De Mauro, A., Greco, M. and Grimaldi, M. (2016), "A formal definition of Big Data based on its essential features", Library Review.

De Mauro, A., Greco, M., Grimaldi, M. and Ritala, P. (2018), "Human resources for Big Data professions: a systematic classification of job roles and required skill sets", Information Processing and Management, Vol. 54 No. 5, pp. 807-817, doi: 10.1016/j.ipm.2017.05.004.

De Mauro, A., Greco, M. and Grimaldi, M. (2019), "Understanding big data through a systematic literature review: the ITMI model", International Journal of Information Technology and Decision Making, Vol. 18 No. 4, pp. 1433-1461.

De Mauro, A., Sestino, A. and Bacconi, A. (2022), "Machine learning and artificial intelligence use in marketing: a general taxonomy", Italian Journal of Marketing, pp. 1-19.

Del Giudice, M. (2016), "Discovering the Internet of Things (IoT) within the business process management: a literature review on technological revitalization", Business Process Management Journal, Vol. 22 No. 2, pp. 263-270.

Del Giudice, M., Scuotto, V., Papa, A., Tarba, S.Y., Bresciani, S. and Warkentin, M. (2021), "A self-tuning model for smart manufacturing SMEs: effects on digital innovation", Journal of Product Innovation Management, Vol. 38 No. 1, pp. 68-89.

Delen, D. and Crossland, M.D. (2008), "Seeding the survey and analysis of research literature with text mining", Expert Systems with Applications, Vol. 34 No. 3, pp. 1707-1720, doi: 10.1016/j.eswa.2007.01.035.

Dijkman, R.M., Sprenkels, B., Peeters, T. and Janssen, A. (2015), "Business models for the internet of things", International Journal of Information Management, Vol. 35 No. 6, pp. 672-678.

Dillon, T., Wu, C. and Chang, E. (2010), "Cloud computing: issues and challenges", 2010 24th IEEE international conference on advanced information networking and applications, IEEE, pp. 27-33.

Duan, Y., Edwards, J.S. and Dwivedi, Y.K. (2019), "Artificial intelligence for decision making in the era of Big Data – evolution, challenges and research agenda", International Journal of Information Management, Vol. 48, pp. 63-71, doi: 10.1016/j.ijinfomgt.2019.01.021.

Durao, F., Carvalho, J.F.S., Fonseka, A. and Garcia, V.C. (2014), "A systematic review on cloud computing", The Journal of Supercomputing, Vol. 68, pp. 1321-1346.

Elia, G., Polimeno, G., Solazzo, G. and Passiante, G. (2020), "A multi-dimension framework for value creation through Big Data", Industrial Marketing Management, Vol. 90, pp. 617-632, doi: 10.1016/j.indmarman.2020.03.015.

Erevelles, S., Fukawa, N. and Swayne, L. (2016), "Big Data consumer analytics and the transformation of marketing", Journal of Business Research, Vol. 69 No. 2, pp. 897-904.

Europa.eu (2022), "Results of the new European data market study 2021-2023", available at: https://digital-strategy.ec.europa.eu/en/library/results-new-european-data-market-study-2021-2023

European Commission (2020), "European Data Strategy: making the EU a role model for society empowered by data", available at: https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/europe-fit-digital-age/european-data-strategy_en

European Commission (2022a), "A European strategy for data", available at: https://digital-strategy.ec.europa.eu/en/policies/strategy-data

European Commission (2022b), "Digital economy and society index DESI", available at: https://digital-strategy.ec.europa.eu/en/library/digital-economy-and-society-index-desi-2022

European Data Act (2022), "European data Act", available at: https://digital-strategy.ec.europa.eu/en/news/data-act-commission-proposes-measures-fair-and-innovative-data-economy

Eviette, M. and Simpson, A. (2021), "Towards models for privacy preservation in the face of metadata exploitation BT – privacy and identity management", in Friedewald, M., Schiffner, S. and Krenn, S. (Eds), Springer International Publishing, Cham, pp. 247-264.

Fainmesser, I.P., Galeotti, A. and Momot, R. (2022), "Digital privacy", Management Science. doi: 10.1287/mnsc.2022.4513.

Fleisch, E., Weinberger, M. and Wortmann, F. (2015), "Business models and the internet of things (extended abstract) BT", in Podnar Žarko., I., Pripužić, K. and Serrano, M. (Eds), Interoperability and Open-Source Solutions for the Internet of Things, Springer International Publishing, Cham, pp. 6-10.

Fosso Wamba, S., Akter, S., Trinchera, L. and De Bourmont, M. (2019), "Turning information quality into firm performance in the big data economy", Management Decision, Vol. 57 No. 8, pp. 1756-1783, doi: 10.1108/MD-04-2018-0394.

Fung, B.C.M., Wang, K. and Ester, M. (n.d.), "Hierarchical document clustering using frequent itemsets", Proceedings of the 2003 SIAM International Conference on Data Mining (SDM), pp. 59-70, doi: 10.1137/1.9781611972733.6.

Giessmann, A. and Legner, C. (2016), "Designing business models for cloud platforms", Information Systems Journal, Vol. 26 No. 5, pp. 551-579.

Gokhale, P., Bhat, O. and Bhat, S. (2018), "Introduction to IOT", International Advanced Research Journal in Science, Engineering and Technology, Vol. 5 No. 1, pp. 41-44.

Griffiths, T.L. and Steyvers, M. (2004), "Finding scientific topics", Proceedings of the National Academy of Sciences, National Acad Sciences, Vol. 101, pp. 5228-5235.

Grimaldi, M. and Cricelli, L. (2009), "Intangible asset contribution to company performance: the 'hierarchical assessment index'", Vine, Vol. 39 No. 1, pp. 40-54.

Günther, W.A., Rezazade Mehrizi, M.H., Huysman, M. and Feldberg, F. (2017), "Debating big data: a literature review on realizing value from big data", The Journal of Strategic Information Systems, Vol. 26 No. 3, pp. 191-209, doi: 10.1016/j.jsis.2017.07.003.

Gutierrez, A., O'Leary, S., Rana, N.P., Dwivedi, Y.K. and Calle, T. (2019), "Using privacy calculus theory to explore entrepreneurial directions in mobile location-based advertising: identifying intrusiveness as the critical risk factor", Computers in Human Behavior, Vol. 95, pp. 295-306, doi: 10.1016/j.chb.2018.09.015.

Halim, D.K. and Hutagalung, S. (2022), "Towards data sharing economy on internet of things: a semantic for telemetry data", Journal of Big Data, Vol. 9 No. 1, p. 1, doi: 10.1186/s40537-021-00549-0.

Harikrishna, B. and Kiran, S. (2020), "Cloud computing technology for digital economy", i-manager's Journal on Cloud Computing, Vol. 7 No. 1, pp. 12-21.

Haroon, A., Shah, M.A., Asim, Y., Naeem, W., Kamran, M. and Javaid, Q. (2016), "Constraints in the IoT: the world in 2020 and beyond", International Journal of Advanced Computer Science and Applications, Vol. 7 No. 11, pp. 112-132.

Hassija, V., Chamola, V., Saxena, V., Jain, D., Goyal, P. and Sikdar, B. (2019), "A survey on IoT security: application areas, security threats, and solution architectures", IEEE Access, Vol. 7, pp. 82721-82743.

Ho, S.M., Ocasio-Velázquez, M. and Booth, C. (2017), "Trust or consequences? Causal effects of perceived risk and subjective norms on cloud technology adoption", Computers and Security, Vol. 70, pp. 581-595, doi: 10.1016/j.cose.2017.08.004.

IDC 4EU (2022), "The European data market study", available at: https://european-big-data-value-forum.eu/wp-content/uploads/2022/09/EDM-2021-23_EBDVF2022_v1.0.pdf

Jeon, C., Cho, D. and Chu, H.-Y. (2022), "How much will you pay to use open data?: evidence from the Seoul metropolitan government", Public Performance and Management Review, Vol. 45 No. 2, pp. 308-328, doi: 10.1080/15309576.2021.1989313.

Johannis, W., Bietenbeck, A., Malchau, G. and Streichert, T. (2020), "Point-of-care testing (POCT) and IT security concepts", Vol. 44 No. 2, pp. 107-111, doi: 10.1515/labmed-2019-0199.

Kamilaris, A., Kartakoullis, A. and Prenafeta-Boldú, F.X. (2017), "A review on the practice of big data analysis in agriculture", Computers and Electronics in Agriculture, Vol. 143, pp. 23-37.

Kodali, R.K., Swamy, G. and Lakshmi, B. (2015), "An implementation of IoT for healthcare", in 2015 IEEE Recent Advances in Intelligent Computational Systems (RAICS), IEEE, pp. 411-416.

Koskinen, J., Knaapi-Junnila, S. and Rantanen, M.M. (2019), "What if we had fair, people-centred data economy ecosystems?", in IEEE SmartWorld, Ubiquitous Intelligence and Computing, Advanced and Trusted Computing, Scalable Computing and Communications, Cloud and Big Data Computing, Internet of People and Smart City Innovation, Leicester, pp. 329-334, doi: 10.1109/SmartWorld-UIC-ATC-SCALCOM-IOP-SCI.2019.00100.

Kostakis, P. and Kargas, A. (2021), "Big-data management: a driver for digital transformation?", Information, Vol. 12 No. 10, p. 411.

Krotov, V. (2017), "The Internet of Things and new business opportunities", Business Horizons, Vol. 60 No. 6, pp. 831-841, doi: 10.1016/j.bushor.2017.07.009.

Lageson, S.E., Webster, E. and Sandoval, J.R. (2021), "Digitizing and disclosing personal data: the proliferation of state criminal records on the internet", Law and Social Inquiry, Vol. 46 No. 3, pp. 635-665, doi: 10.1017/lsi.2020.37.

Lai, S.S. and Flensburg, S. (2021), "Invasive species of the app ecosystem: exploring the political economy of mobile communication", International Journal of Communication, Vol. 15, p. 18.

Lammi, M. and Pantzar, M. (2019), "The data economy: how technological change has altered the role of the citizen-consumer", Technology in Society, Vol. 59, 101157, doi: 10.1016/j.techsoc.2019.101157.

Liddell, K., Simon, D.A. and Lucassen, A. (2021), "Patient data ownership: who owns your health?", Journal of Law and the Biosciences, Vol. 8 No. 2, lsab023, doi: 10.1093/jlb/lsab023.

Liu, L. (2021), "The rise of data politics: digital China and the world", Studies in Comparative International Development, Vol. 56, pp. 45-67.

Löhr, H., Sadeghi, A.R. and Winandy, M. (2010), "Securing the e-health cloud", Proceedings of the 1st acm international health informatics symposium, pp. 220-229.

Luan, H., Geczy, P., Lai, H., Gobert, J., Yang, S.J., Ogata, H. and Tsai, C.C. (2020), "Challenges and future directions of big data and artificial intelligence in education", Frontiers in Psychology, Vol. 11, pp. 580820-580825.

Maher, M. and Khan, I. (2022), "From sharing to selling: challenges and opportunities of establishing a digital health data marketplace using blockchain technologies", Blockchain in Healthcare Today, Vol. 5, SE-Narrative/Systematic Review/Meta-Analysis, doi: 10.30953/bhty.v5.184.

Malik, R. and Rybkowska, K. (2023), "Business processes powered by big data: current issues and new research directions", in Anna, V., Troisi, O. and Grimaldi, M. (Eds), Big Data and Decision-Making: Applications and Uses in the Public and Private Sector, Emerald Publishing, pp. 145-161, doi: 10.1108/978-1-80382-551-920231010.

Marelli, L., Lievevrouw, E. and Van Hoyweghen, I. (2020), "Fit for purpose? The GDPR and the governance of European digital health", Policy Studies, Vol. 41 No. 5, pp. 447-467, doi: 10.1080/01442872.2020.1724929.

Marjani, M., Nasaruddin, F., Gani, A., Karim, A., Hashem, I.A.T., Siddiqa, A. and Yaqoob, I. (2017), "Big IoT data analytics: architecture, opportunities, and open research challenges", IEEE Access, Vol. 5, pp. 5247-5261.

Marjanovic, O. (2022), "A novel mechanism for business analytics value creation: improvement of knowledge-intensive business processes", Journal of Knowledge Management, Vol. 26 No. 1, pp. 17-44, doi: 10.1108/JKM-09-2020-0669.

Mashiko, T. (2020), "Big data, IoT, AI, and smart cities", in Telecommunications Policies of Japan, pp. 181-200.

Milligan, G.W. and Cooper, M.C. (1985), "An examination of procedures for determining the number of clusters in a data set", Psychometrika, Vol. 50, pp. 159-179.

Misra, N.N., Dixit, Y., Al-Mallahi, A., Bhullar, M.S., Upadhyay, R. and Martynenko, A. (2020), "IoT, big data, and artificial intelligence in agriculture and food industry", IEEE Internet of Things Journal, Vol. 9 No. 9, pp. 6305-6324.

Morgan, J. (2014), "Introduction to IOT", International Advanced Research Journal in Science, Engineering and Technology, Vol. 5 No. 1, pp. 41-44.

Moro, S., Cortez, P. and Rita, P. (2014), "A data-driven approach to predict the success of bank telemarketing", Decision Support Systems, Vol. 62, pp. 22-31, doi: 10.1016/j.dss.2014.03.001.

Murtagh, M.J., Demir, I., Jenkings, K.N., Wallace, S.E., Murtagh, B., Boniol, M., Bota, M., et al. (2012), "Securing the data economy: translating privacy and enacting security in the development of DataSHIELD", Public Health Genomics, Vol. 15 No. 5, pp. 243-253, doi: 10.1159/000336673.

North, K. and Kumta, G. (2020), Knowledge Management: Value Creation through Organizational Learning, Springer.

Oliver, N., Lepri, B., Sterly, H., Lambiotte, R., Deletaille, S., De Nadai, M. and Vinck, P. (2020), "Mobile phone data for informing public health actions across the COVID-19 pandemic life cycle", Science Advances, Vol. 6 No. 23, pp. eabc0764-eabc0769.

Pennino, D., Pizzonia, M., Vitaletti, A. and Zecchini, M. (2022), "Blockchain as IoT economy enabler: a review of architectural aspects", Journal of Sensor and Actuator Networks. doi: 10.3390/jsan11020020.

Peter, O., Pradhan, A. and Mbohwa, C. (2023), "Industrial internet of things (IIoT): opportunities, challenges, and requirements in manufacturing businesses in emerging economies", Procedia Computer Science, Vol. 217, pp. 856-865, doi: 10.1016/j.procs.2022.12.282.

Ploennigs, J., Schumann, A. and Lécué, F. (2014), "Adapting semantic sensor networks for smart building diagnosis", in Mika, P., Tudorache, T., Bernstein, A., Welty, C., Knoblock, C., Vrandečić, D., Groth, P., Noy, N., Janowicz, K. and Goble, C. (Eds), ISWC 2014, LNCS, Springer, Cham, pp. 308-323.

Ponweiser, M. (2012), Latent Dirichlet allocation in R, Wu Vienna University of Economics and Business, Vienna.

Quinn, M., Strauss, E. and Kristandl, G. (2014), "The effects of cloud technology on management accounting and business decision-making", Financial Management, Vol. 10 No. 6, pp. 1-12.

Reinhardt, J. (2022), "Realizing the fundamental right to data protection in a digitized society BT – personality and data protection rights on the internet: Brazilian and German approaches", in Albers, M. and Sarlet, I.W. (Eds), Springer International Publishing, Cham, pp. 55-68, doi: 10.1007/978-3-030-90331-2_4.

Reyna, A., Martín, C., Chen, J., Soler, E. and Díaz, M. (2018), "On blockchain and its integration with IoT. Challenges and opportunities", Future Generation Computer Systems, Vol. 88, pp. 173-190.

Russom, P. (2013), "Managing big data", TDWI Best Practices Report, TDWI Research, pp. 1-40.

Sadowski, J. (2019), "When data is capital: datafication, accumulation, and extraction", Big Data and Society, Vol. 6 No. 1, pp. 549-553.

Santoro, G., Fiano, F., Bertoldi, B. and Ciampi, F. (2019), "Big data for business management in the retail industry", Management Decision, Vol. 57 No. 8, pp. 1980-1992.

Santoro, G., Thrassou, A., Bresciani, S. and Giudice, M.D. (2021), "Do knowledge management and dynamic capabilities affect ambidextrous entrepreneurial intensity and firms' performance?", IEEE Transactions on Engineering Management, Vol. 68 No. 2, pp. 378-386, doi: 10.1109/TEM.2019.2907874.

Schneiders, A., Fell, M.J. and Nolden, C. (2022), "Peer-to-peer electricity trading and the sharing economy: social, markets and regulatory perspectives", Energy Sources, Part B: Economics, Planning, and Policy, Vol. 17 No. 1, 2050849, doi: 10.1080/15567249.2022.2050849.

Sestino, A. and De Mauro, A. (2022a), "Leveraging artificial intelligence in business: implications, applications and methods", Technology Analysis and Strategic Management, Vol. 34 No. 1, pp. 16-29, doi: 10.1080/09537325.2021.1883583.

Sestino, A. and De Mauro, A. (2022b), "Leveraging artificial intelligence in business: implications, applications and methods", Technology Analysis and Strategic Management, Vol. 34 No. 1, pp. 16-29.

Sestino, A., Prete, M.I., Piper, L. and Guido, G. (2020), "Internet of Things and Big Data as enablers for business digitalization strategies", Technovation, Vol. 98, 102173, doi: 10.1016/j.technovation.2020.102173.

Sestino, A., Simonetti, A. and Gastaldi, L. (2021), "COVID-19 and digital acceleration: an investigation about citizens' perception of digitalized public services in Italy", 4th INICART International Conference 'Digital Transition and Green Sustainable Economy', pp. 137-146.

Sestino, A. and De Mauro, A. (2022), "Leveraging artificial intelligence in business: implications, applications and methods", Technology Analysis and Strategic Management, Vol. 34 No. 1, pp. 16-29.

Sestino, A., Giraldi, L., Cedrola, E., Zamani, S.Z. and Guido, G. (2022), "The business opportunity of blockchain value creation among the internet of value, 'global business review'", in press, doi: 10.1177/09721509221115012.

Shaikh, F.B. and Haider, S. (2011), "Security threats in cloud computing", 2011 International conference for Internet technology and secured transactions, IEEE, pp. 214-219.

Sharma, P., Lawrenz, S. and Rausch, A. (2020), "Towards trustworthy and independent data marketplaces", Proceedings of the 2020 The 2nd International Conference on Blockchain Technology, Association for Computing Machinery, New York, NY, pp. 39-45, doi: 10.1145/3390566.3391687.

Shenoy, J. and Pingle, Y. (2016), "IOT in agriculture", 2016 3rd International Conference on Computing for Sustainable Global Development (INDIACom), IEEE, pp. 1456-1458.

Singh, A. and Chatterjee, K. (2017), "Cloud security issues and challenges: a survey", Journal of Network and Computer Applications, Vol. 79, pp. 88-115.

Sivarajah, U., Kamal, M.M., Irani, Z. and Weerakkody, V. (2017), "Critical analysis of Big Data challenges and analytical methods", Journal of Business Research, Vol. 70, pp. 263-286.

Strømmen-Bakhtiar, A. (2019), "Digital economy, business models, and cloud computing", in Global Virtual Enterprises in Cloud Computing Environments, IGI Global, pp. 19-44.

Sunikka, A. and Bragge, J. (2012), "Applying text-mining to personalization and customization research literature – who, what and where?", Expert Systems with Applications, Vol. 39 No. 11, pp. 10049-10058, doi: 10.1016/j.eswa.2012.02.042.

Surya, L. (2015), "An exploratory study of AI and Big Data, and it's future in the United States", International Journal of Creative Research Thoughts (IJCRT), Vol. 5, pp. 2320-2882.

Suzumura, T., Sugiki, A., Takizawa, H., Imakura, A., Nakamura, H., Taura, K., Kudoh, T., et al. (2022), "Mdx: a cloud platform for supporting data science and cross-disciplinary research collaborations", 2022 IEEE Intl Conf on Dependable, Autonomic and Secure Computing, Intl Conf on Pervasive Intelligence and Computing, Intl Conf on Cloud and Big Data Computing, Intl Conf on Cyber Science and Technology Congress, DASC/PiCom/CBDCom/CyberSciTech, pp. 1-7, doi: 10.1109/DASC/PiCom/CBDCom/Cy55231.2022.9927975.

Taddy, M. (2012), "On estimation and selection for topic models", in Artificial Intelligence and Statistics, PMLR, pp. 1184-1193.

Tosheva, E. (2020), "COVID-19 crisis as an accelerator of digital transformation and digital economy", International Scientific Conference "Towards a Better Future: Human Rights, Organized Crime and Digital Society", pp. 179-187.

Trabucchi, D. and Buganza, T. (2019), "Data-driven innovation: switching the perspective on Big Data", European Journal of Innovation Management, Vol. 22 No. 1, pp. 23-40.

UNCTAD Digital Economy Report (2021), "UNCTAD digital economy report", available at: https://unctad.org/publication/digital-economy-report-2021

Vertesi, J. and Dourish, P. (2011), "The value of data: considering the context of production in data economies", Proceedings of the ACM 2011 Conference on Computer Supported Cooperative Work, New York, NY, Association for Computing Machinery, pp. 533-542, doi: 10.1145/1958824.1958906.

Vo, D.-T. and Ock, C.-Y. (2015), "Learning to classify short text from scientific documents using topic models with various types of knowledge", Expert Systems with Applications, Vol. 42 No. 3, pp. 1684-1698, doi: 10.1016/j.eswa.2014.09.031.

Volpert, S., Eichhammer, P., Held, F., Huffert, T., Reiser, H.P. and Domaschka, J. (2023), "The view on systems monitoring and its requirements from future cloud-to-thing applications and infrastructures", Future Generation Computer Systems, Vol. 141, pp. 243-257, doi: 10.1016/j.future.2022.11.024.

Wired (2021), "The future of creative economy is connection and data", available at: https://www.wired.com/sponsored/story/the-future-of-the-creative-economy-is-connection-and-data/

Xu, X. (2012), "From cloud computing to cloud manufacturing", Robotics and Computer-Integrated Manufacturing, Vol. 28 No. 1, pp. 75-86.

Yannopoulos, P. (2011), "Impact of the Internet on marketing strategy formulation", International Journal of Business and Social Science, Vol. 2 No. 18, pp. 110-121.

Zhao, Y. and Zhou, Y. (2022), "Measurement method and application of a deep learning digital economy scale based on a big data cloud platform", Journal of Organizational and End User Computing (JOEUC), Vol. 34 No. 3, pp. 1-17.

Zillner, S., Gomez, J.A., García Robles, A., Hahn, T., Le Bars, L., Petkovic, M. and Curry, E. (2021), "Data economy 2.0: from big data value to AI value and a European data space BT – the elements of big data value: foundations of the research and innovation ecosystem", in Curry, E., Metzger, A., Zillner, S., Pazzaglia, J.-C. and García Robles, A. (Eds), Springer International Publishing, Cham, pp. 379-399, doi: 10.1007/978-3-030-68176-0_16.

About the Authors

Andrea Sestino is an adjunct professor of Competitive Strategy at the LUISS Guido Carli University, in Rome, Research Fellow in Management and Marketing at the University of Rome Tre, Research Assistant in Management and Innovation in Healthcare Services, and Technology Innovation Governance, at the Catholic University Sacro Cuore, in Rome. Andrea holds a Ph.D. in Business Management and Marketing at the University of Bari Aldo Moro, and a M.Sc. in Business Management at Sapienza University, Rome, Italy. He has been also Expert Collaborator for the Cabinet of the Italian Minister of Economic Development, in fields related to Artificial Intelligence, Internet-of-Things applications and business digitalization, and he is currently also R&D Specialist in the field of applied industrial research for business digital transition in several private and public research projects. He has published two books, and several articles in international peer-reviewed journals, such as Technovation, Technology Analysis and Strategic Management, British Food J., Technology in Society, Global Business Review, Int. J. of Healthcare and Pharmaceutical Marketing,

Italian J. of Marketing, Italian J. of Management, SN Business and Economics, Psychological Reports, Int. J. of Marketing Studies, and so on.

Adham Kahlawi is an adjunct professor at the University of Florence, Italy. He received his PhD in applied statistics from the University of Florence. His current research interests include statistical information systems, decision support systems, data quality, machine learning, text mining and big data analytics. He serves as a Data Scientist at Hubpay Limited.

Andrea De Mauro has more than 15 years of international experience managing Data Analytics and Data Science organizations. He is currently Head of Data and Analytics at Vodafone Italy. Previously, he served as Director of Business Analytics at Procter and Gamble Europe. He holds a PhD in Management Engineering from Rome Tor Vergata University, a Master of Science in Electrical and Computer Engineering from the University of Illinois at Chicago, a Master's degree in ICT Engineering from Polytechnic of Turin, and a diploma in Innovation from Alta Scuola Politecnica in Milan. He is an adjunct professor of Marketing Analytics and Applied Machine Learning at the Universities of Florence and Bari, Italy. He is the author of popular science books on data and several research articles in international peer-reviewed journals. In his research, he investigates the essential components of Big Data as a phenomenon and the impact of AI and Data Analytics on companies and people.