Information Systems and e-Business Management (2021) 19:313–362
https://doi.org/10.1007/s10257-020-00500-5

ORIGINAL ARTICLE

Enterprise architecture management as a solution for addressing general data protection regulation requirements in a big data context: a systematic mapping study

Georgios Georgiadis[1] · **Geert Poels**[1]

Received: 19 August 2019 / Revised: 10 November 2020 / Accepted: 8 December 2020 /
Published online: 18 January 2021

Abstract

Context Big Data Analytics is a rapidly emerging IT practice whose applications offer benefits for a wide variety of business areas across an organisation. Given the wide scope of applications, the many types of processing involved, including those for purposes not yet foreseen, and the inherent privacy concerns resulting from collecting and storing personal data, the newly introduced General Data Protection Regulation (GDPR) poses specific challenges for safeguarding the security and protection of big data. These challenges are not limited to the IT function but extend across the entire organisation. This raises the question whether Enterprise Architecture Management (EAM), as an approach for ensuring the coherence, strategic alignment and focus on value creation of all organisational resources, offers guidance for addressing those challenges in a holistic manner, and thus provides a fruitful ground for developing an approach for complying to GDPR requirements in a Big Data context.

Objective This study surveys the state-of-the-art in research on security, privacy, and protection of big data. The focus is on investigating which specific issues and challenges have been identified and whether these have been linked to GDPR requirements. Further, it examines whether previous research has investigated the potential of EAM in addressing those challenges and what the main findings of those studies are.

Method We used Systematic Mapping Review (SMR), which is a methodology for literature review aimed at surveying the state-of-the-art in a research field as it is documented in the scientific literature. Further, we used Template Analysis, which is a thematic analysis technique, for coding the texts of the selected papers, classifying the research studies, and interpreting the different themes addressed in the literature.

Supplementary Information The online version of this article (https://doi.org/10.1007/s10257-020-00500-5) contains supplementary material, which is available to authorized users.

Extended author information available on the last page of the article

Results Our study indicates that only few researchers have explored the use of EAM practices in relation to data security and protection in a Big Data context. We further identified seven trends within the areas under consideration that could be subjects for further research.
Conclusions Our study does not invalidate the potential of EAM to help addressing GDPR requirements in a Big Data context. However, how EAM practices may contribute to risk management and data governance in environments where big data are being processed, is still a huge research gap, which we intend to address in our future research.

Keywords Big data · Data protection · Data protection directive · Enterprise architecture management · General data protection regulation · Governance · Information security · Privacy · Systematic literature mapping

1 Introduction

Big Data Analytics (BDA) is a rapidly emerging area for both business and IT professionals involving technology, people and organisations (Alharthi et al. 2017) and generally refers to the use of analytical applications that process a wide variety of voluminous, rapidly changing, non-transactional datasets that contain data of varying degrees of structuredness and quality. Although the benefits of BDA applications are far-ranging and affect different business fields (from medicine to social networking) and organisational functions,, their reach also introduces great exposure to security and privacy risks (Lee 2017) Given the types of processing involved and the ways data is used, which may have nothing to do with the purpose set during their collection (Rubinstein 2012), concerns regarding security and privacy matters have vastly increased, putting organisations under immense pressure to protect their information assets and themselves from hefty legal repercussions of probable breaches (Salleh and Janczewski 2016).

In Europe, the General Data Protection Regulation (GDPR) (EU 2016) is the new law backing data privacy and data protection. Its leading role is to ensure the fair processing of personal data by both the public and private sectors. To do so, it imposes stronger rules on data protection, enabling individuals (data subjects (DSs) in GDPR terms) to have more control over their data. Instead of dealing with the different levels of national data protection laws, where something that is allowed in one EU member state (MS) could be unlawful in another, the GDPR becomes the common data privacy law across all EU MSs. The repealed law was based on EU Data Protection Directive (95/46/EC or DPD) (EP 1995). DPD outlined the desired outcomes to be achieved by MSs and did not induce legal obligation to them. Its interpretation, therefore, varied from one MS to another.

Considering the great potential of BDA in digital marketing and the profound effects of GDPR (Ibraimova 2017), especially when referring to processing with analytical algorithms and techniques, one wonders if there is an effective way to address all these effects adequately and methodically, while avoiding any

trial-and-error approaches resulting in an increase in risk and unnecessary waste of valuable time and resources. Current research confirms that security and privacy of big data have received relatively less focus despite both being strategically vital (Akoka et al. 2017; EDPS 2016; Liu et al. 2017).

As a result, companies or organisations with BDA applications must take quick steps to fulfil all the requirements imposed by the GDPR and deal with their complexity and numerous facets at stake. For all exposed data, entities must be able to point out the justification or legal purpose, regardless the area in the organisation where the big data or insights obtained from analysing these big data, are being used. For these reasons, we argue that organisations have no choice other than to consider incorporating the new law in their different structures and models. In other words, an approach is needed to cover the entire organisational architecture (Wagter et al. 2005). For example, (Sauer and Willcocks 2003) discussed the lessons learnt from Oracle and Macquarie Bank to highlight the strong link between technology and organisation. They concluded that organisations without a structured overview of their organisational architecture have a hard time developing business visions that are informed by the potentiality of technology and establishing technology platforms that 'reflect future needs under conditions of great uncertainty' (Sauer and Willcocks 2003). Also, the concept of *data protection by design as a default* (Rubinstein and Good 2013), which is key in addressing GDPR requirements, entails in practice an organisation-wide and architectural component whereby data protection is not restricted to information technology (IT) and is embedded in all business processes and their entire development cycle.

The essence and benefits of managing the organisational architecture have been subject to several past studies (Ethiraj and Levinthal 2004; Mendelson 2003; Nadler et al. 1992; Sauer and Willcocks 2003, Sauer and Willcocks 2004). The notion behind architecture is the idea of a holistic approach, and a widely prescribed method to pursue such an approach is through Enterprise Architecture Management (EAM) (Ahlemann et al. 2012; Kappelman et al. 2008; Kehrer et al. 2016a; Kotusev et al. 2015; Lagerström et al. 2011; Radeke 2010, 2011; Wißotzki et al. 2013). EAM practices enable organisations to foster digital trust by understanding the strategic benefits from the use of their data and better view their capabilities, applications, and systems, including how they are all aligned with the overall business strategy and the payoff of IT investment (Anthony Byrd et al. 2006).

The aim of this paper, therefore, is twofold. The first aim is to survey existing research to gain an overview of the state-of-the-art on the three interweaved concerns of security, privacy, and protection of big data. As this is a broad field to pursue, we focus our research on matters that can be associated with the two European regulatory data protection frameworks: the repealed Data Protection Directive (DPD) and the recently introduced GDPR. This means that research proposing merely technological solutions for securing big data but without explicit consideration for addressing regulatory requirements or challenges imposed by such regulation, falls outside the scope of our survey. Second, we will critically examine a carefully selected set of studies to analyse their results further and generate possible scenarios on how to tackle arising business needs regarding these two frameworks. We look particularly for studies that investigate how EAM can offer guidance and

solutions for addressing GDPR requirements in organisations that make use of BDA applications. With this in mind, we aim to assess whether EAM practices could offer a solution for addressing the challenges imposed by GDPR requirements in a Big Data context.

With these objectives in mind, we apply the Systematic Mapping Review (SMR) methodology for literature-based survey research (Petersen et al. 2008) to the literature strand in Big Data research that addresses issues of data security, privacy, and protection. Our survey of the state-of-the-art in the research on these topics is guided by the following six mapping research questions (MRQs):

MRQ1 How many studies on these topics were published over the years?
MRQ2 What are the most frequently used publication outlets?
MRQ3 What are the main issues and challenges with respect to data security, protection and privacy in big data?
MRQ4 What does big data research show regarding the rising needs for DPD or GDPR?
MRQ5 Considering EAM, is there any approach or method for dealing with big data, and what is its perceived role regarding big data security, protection, and privacy aspects?
MRQ6 Can we identify any topics within these areas that could be subjects for further research?

MRQs 1 and 2 are intended to obtain an overview of the research area and publication space. MRQs 3 and 4 relate to our first objective and are intended to clarify how research has approached the idiosyncrasies of data security, protection and privacy in a Big Data context (MRQ3) and which specific challenges to these issues posed by the European regulatory requirements have been identified (MRQ4). Based on this understanding, MRQ5 looks specifically into possible solutions for these challenges that researchers have proposed starting from EAM practices, which relates to our second objective. While MRQ 5 is instrumental for our further research, MRQ6 covers both objectives and intents to contribute to the field by outlining avenues for future research on the nexus of GDPR and Big Data, based on the understanding of the state-of-the-art that we obtain through our SMR study.

The remainder of this paper is organised as follows. In Sect. 2, we present the background needed for our research. Section 3 describes in more detail the methodology we applied to carry out the literature survey. Section 4 presents the findings while answering each of the above listed MRQs. Section 5 discusses the main findings in terms of research gaps and future research. Finally, Sect. 6 presents a conclusion and outlook.

2 Background

In this section, we provide additional background information. We summarise the purpose of storing and analysing big data and exemplify recent issues that have arisen with the use of big data. Next, we explain why data security, privacy, and

protection have gradually received considerable attention and discuss the expectations regarding challenges posed by the GDPR. Finally, we introduce the reader to EAM and the roles it can play in contemporary organisations, including ensuring compliance with data protection and privacy requirements.

2.1 Big data

According to (Manyika et al. 2011), big data has a great potential to create tremendous value for the global economy. (Shirer 2016; Shirer and Goepfert 2019), who supported this view, predicted that the expected worldwide revenues for big data and analytics will be $274.3 billion by 2022. In Europe, the market size for the same technologies is projected to reach $105.82 billion by 2027 (Borasi et al. 2020). Despite this potential, there are also specific security and privacy risks associated to BDA which typically involves processing personal data. First the legitimacy of the analytical applications that process big data is something notoriously difficult to show as 'necessary for the performance of a contract' under the GDPR (Gandomi and Haider 2015). Second, if a security breach occurs, it may affect the data of a much larger number of people.

Two characteristic examples of a breach of a large scale with widespread publicity in the press are Equifax and Cambridge Analytica. Equifax Inc. is the oldest of the three largest American credit agencies (along with Experian and TransUnion). The company was confronted with a serious cybersecurity incident that affected the sensitive personal information of about 143 million consumers, including European residents (BBC 2017; Bracy 2017). Cambridge Analytica's case is more recent and involved the illicit harvesting of at least 50 million users' personal data posted on Facebook. Information published by the Privacy Rights Clearinghouse (PrivacyRightsClearinghouse 2017) shows that, in just the US, close to 8000 data breaches have been made public since 2005. The tendency of the number and effects of data breach incidents to grow for different industry sectors is alarmingly great, as shown in a year-end review report (ITRC 2019) and depicted in Fig. 1.[1] Another study prepared by (Ponemon Institute 2018) has indicated that the average total cost of a data breach amounts to $3.86 M, and $148 is the average cost of each stolen record. Despite the various breaching incidents, there is a growing number of companies interested in investing in new technologies whose core business model relies on the processing of big data (Guggenheim 2016; IDC 2019). Considering that in a Big Data context, virtually any data is identifiable and in conjunction with other data can reveal sensitive information (McMahon et al. 2020; Mourby et al. 2018), and combining that with the huge data storage in several locations, we may expect a wide array of complex and inherently costly data breach incidents to occur in the years to come, unless specific attention is paid to big data security, protection and privacy.

[1] https://www.statista.com/statistics/273550/data-breaches-recorded-in-the-united-states-by-number-of-breaches-and-records-exposed/

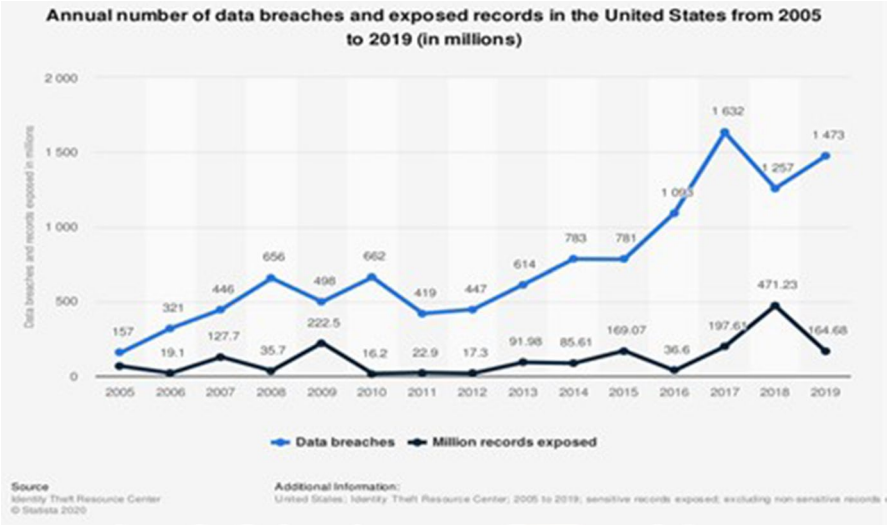

Fig. 1 Number (in millions) of the data breach incidents in the United States from 2005 to 2019

2.2 The GDPR

Understanding the effects of a data breach on the company's shares (Moyer 2017) leaves no doubt that data security and privacy are becoming vital factors for the competitiveness and survival of a business. Consumers and employees alike highly value privacy protection (Spiekermann 2012; Willemsen and Bhajanka 2017).

As two distinct terms, data security and data privacy have both been the subject of highly debated discussions among researchers and lawmakers. A good reason for that is their aim. The prime aim of *data security* is risk reduction. To do so, specific information about a resource is usually requested. *Data privacy*, on the other hand, comes into effect when such information involves the data of an individual, group, or institution (Westin 1968). Privacy is an umbrella concept that contains many meanings and dimensions (Solove 2002). Its purpose, among other things, is the protection of information, home, body, life, secrecy, and communications. Moreover, it is recognised as a fundamental human right in many countries, allowing citizens to be free from intrusions by the state (UN 1948).

Along with data privacy, there is the parallel notion of *data protection*. The two terms are often used interchangeably and are within the same set of rights protected by EU law. The right to privacy (or respect for private and family life) can be found in Article 7 of the EU Charter of Fundamental Rights (ECHR), whereas the right to the protection of personal data is in Article 8 of the charter (EC 2000) and Article 16 of the Treaty on the Functioning of the European Union (EU 2012). Data protection primarily deals with data generated from the increasing use of information and communication technology (ICT) and is necessary but insufficient for privacy. In Europe, it plays a leading role by ensuring the fair processing of personal data by both the public and private sectors (Kuneva 2009).

In its recent reform, namely the General Data Protection Regulation, also referred to as the 'Regulation' or more commonly as GDPR, considerable improvements and changes to the existing DPD were introduced, whose main flaw was ambiguous jurisdiction. It is evident that the GDPR rules are not minor revisions of those laid down by the DPD. In general, the new rules address concerns and weaknesses that have been reported by field expert organisations and individual researchers since the early 2000s (Birnhack 2008; Robinson et al. 2009; Wong 2011). Being aware of the arising complexity of data protection, while under increasing requirements for ensuring compliance, many organisations are apprehensive, as they view the GDPR as more of a threat than an opportunity.

Notwithstanding the approaches being pursued and given that compliance to the GDPR is red-flagged as one of the top organisational priorities, surveys conducted from advisory companies have shown that many organisations will probably not manage to succeed before 2019 (Crowd Research Partners 2018; Gartner 2017; ISACA 2018). Few years later, two research studies conducted by Mikkelsen et al. (2019) and Capgemini (2019) in 2019 have shown that the same organisational challenges in terms of achieving and maintaining compliance as well as improving IT landscape remain still unresolved and many companies still strive to implement solutions that are durable and long-lasting. For example, one area that looks quite problematic and that could lead to additional delays is the area of unstructured data such as the data stored in email servers. It is exactly this type of data that is exemplary for BDA applications.

Further, emerging trends in social media, including the proliferation of increasingly capable mobile devices, all of which constantly feed the big data ecosystem with vastly large datasets, add another dimension to data privacy threats. Research findings (Smith et al. 2012) in public social media point towards a posteriori approach, as they argue that pre-emptive measures cannot stop one from making potentially damaging content public. Such incidents, that often occur unintentionally, do not solely concern private use. In this digital age, companies and large non-profit organisations make extensive use of social media platforms for trading and communicating and engaging with their stakeholders as well as the public (Culnan et al. 2010; Lovejoy and Saxton 2012). For instance, these platforms are used for capturing and then analysing the emotional state of the electorate in the course of large political campaigns (Parnet et al. 2016).

When considering how the value of data protection and privacy is perceived by individuals in Europe, our attention again shifts to pertinent large-scale surveys conducted by the European Commission in the period between 2003 and 2015.[2] The key takeaway from their analyses is the noticeably low level of trust in online businesses and citizens' widespread concerns about the consequences of breaches, leading to the conclusion that companies should be provided with clear standards and rules to which they need to adhere. Viewing that from the perspective of BDA applications processing personal data, it also indicates that many organisations must truly redefine their strategies during the collection and processing of data.

[2] https://ec.europa.eu/commfrontoffice/publicopinion/index.cfm/Survey/index#p=1&search=data protection.

2.3 EAM and its role in contemporary organisations

Contemporary organisations operate in changing and fiercely competitive environments. In order to survive, they must adapt by redesigning their structures and processes and by leveraging their information systems (Ahlemann et al. 2012). EAM is a strategy-driven management discipline that aims to support the development and evolution of the organisation's infrastructure that is needed to achieve strategic objectives (Kehrer et al. 2016a; Lagerström et al. 2011; Simon et al. 2014). The focus is on the different components that comprise this infrastructure, their interrelationships, and their relationships to the environment, hence the notion of organisational or *enterprise architecture*. EAM practices have as main instrument a structured overview of the organisational architecture, specific aspects of it or perspectives (or viewpoints) on it. The structured overview is built up from enterprise architecture models. Based on such models, EAM ensures that the organisational resources that compose the organisation's infrastructure are integrated, aligned and directed towards the common goal of value creation in accordance to strategic objective.

EAM can have a positive effect on organisations by promoting holistic thinking, leading to the development of a consistent business strategy, an improved IT/IS landscape,and a proper business-IT strategy alignment (Löhe and Legner 2014). Apart from strategic IT/IS planning, EAM also facilitates and improves corporate and business planning processes and ensures business continuity, compliance, and risk management (Winter and Schelp 2008).

In the context of ensuring compliance with GDPR rules, EAM can prove highly beneficial for organisations as it allows to have a comprehensive overview of what is required, across the different units, departments and functions of an organization, through legal (e.g. Data Protection Impact Assessment, retention period and Record of processing activities), business (e.g. data types and flows) and technical (e.g. information security and IT applications) means and to deal methodically with all limitations (Huth et al. 2020). This is particularly promising for GDPR compliance in organisations with BDA applications, as big data are frequently obtained from different sources, collected for different reasons, and processed for different purposes, and the insights gained from their processing are applied across an organisation' structure (e.g., for informing marketing campaigns, for new product/service development, for improved customer relationships, for strategic decision-making, for auditing, etc.). Hence, a holistic approach like EAM that does not reduce an organisation to its parts, but emphasizes the linkages between those parts, might be instrumental for a comprehensive coverage of big data security, protection and privacy issues, in response to GDPR requirements.

3 Research methodology

We first present the Systematic Mapping Review (SMR) methodology for literature-based survey research, which is related to the better known Systematic Literature Review (SLR) methodology (Kitchenham 2004; Kitchenham et al. 2009; Petersen

et al. 2008). Next, the SMR process is briefly presented. The rest of the section describes how we conducted this process.

3.1 Systematic mapping review

Systematic Mapping Review also called Systematic Literature Mapping, shares concepts and techniques with SLR, as they were originally developed in the field of medicine (Kitchenham et al. 2011b). Their substantial difference, however, lies in their goals (Petersen et al. 2015). The SMR has a high level of granularity because it provides a bird's-eye view of the current state of a research area. It allows a general overview of the type of research that has been performed and usually results in devising a scheme for classifying the discovered articles. Due to its broad coverage, the terms used during the literature search are likely to return a relatively large number of articles. It is, however, a time efficient and effective method of getting ready for follow-up research, as it enables researchers to identify the quantity and type of research conducted, while studying the results achieved within it (Kitchenham et al. 2011a). SLR, on the other hand, is of evidence-driven nature (Kitchenham et al. 2004) and intends to provide a meta-review of empirical research findings. Its primary aim is thus providing a scientifically rigorous synthesis of primary studies. Similar to the SMR, SLR enables identifying gaps in the current research and gives an explicit account of how the literature was systematically searched and relevant studies were selected. Generally, SLRs have a finer-grained scope and therefore require a different search process strategy and quality evaluation approach (Petersen et al. 2015).

As we wish to known the specific issues regarding security, privacy and protection of big data that have been addressed in the scientific literature, and enquire whether GDPR poses specific challenges regarding these issues, and further find out whether researchers have considered EAM as a possible approach to tackle these challenges, an SMR study is adequate. Should we discover a considerable number of studies that have already applied EAM practices for ensuring compliance with GDPR requirements in a Big Data context, then an SLR can be performed to synthesize the research findings.

3.2 The systematic mapping review process

We perform the five SMR process steps suggested by (Petersen et al. 2008):

1. Defining the mapping research questions (MRQs);
2. Conducting a search for primary studies in the chosen area of study;
3. Screening of pertinent papers based on pre-defined inclusion/exclusion criteria;
4. Paper classification, and
5. Extraction and aggregation of data for producing the desired overview.

We opted to adhere to all steps not only because we wanted to be consistent with the SMR process but also to ensure high quality regarding the completeness and rigour of our literature survey. The MRQs (i.e., step 1) were presented in the introduction of the paper, where they were related to the objectives of our survey. The methodology of steps 2–4 is explained in the next sub-sections, whereas the results of step 5 are presented in Sect. 4.

3.3 Conducting the search

The identification of the primary studies was done using search strings on digital scientific databases. Our MRQs served as the means for discovering the search keywords or terms. For the search strings to be created, the chosen keywords or their synonyms were combined with Boolean (AND, OR) operators. All these yielded the following three search strings:

1. *'big data' AND (security OR privacy OR 'data privacy' OR 'data protection' OR dpd OR 'eu dpd' OR 'Data Protection Directive' OR 'gdpr' OR 'general data protection regulation' OR 'data protection regulation')*;
2. *'big data' AND (security OR privacy OR 'data privacy' OR 'data protection' OR dpd OR 'eu dpd' OR 'Data Protection Directive' OR 'gdpr' OR 'general data protection regulation' OR 'data protection regulation') AND ('enterprise architecture management' OR EAM OR 'ea management' OR ea OR 'enterprise architecture')*; and
3. *'big data' AND ('enterprise architecture management' OR eam OR 'ea management' OR ea OR 'enterprise architecture').*

The first search string is based on the concepts of MRQ3 and MRQ4 and relates to our objective of identifying studies that address security, privacy and protection of Big Data or discuss the EU data protection regulations in relation to big data. The second search string includes the first search string but combines it in a conjunction with terms related to EAM. With this search string we specifically look for studies related to our second objective. As the number of studies that we can find using this string might be low, we added a third search string that looks for research on Big Data and EAM, but without explicitly requiring these studies to address issues of data security, privacy and protection. We believe that such papers might still inform us on GDPR related concerns, like risk management and data governance. Together, search strings two and three aim at delivering information that allows us to answer MRQ5. The more general MRQs 1,2 and 6, don't contain specific concepts to search for, but are typical questions directing a SMR with the overall aim of creating a structured overview of the research in the field of interest and identifying research gaps.

As common for SMRs, the search strings are applied to the title, abstract, and keyword fields. For the sake of replicability, Table 4 in the appendix, documents the strings as written in the query language supported by each database. Due to the interdisciplinary nature of our research topic, we considered both specific subject

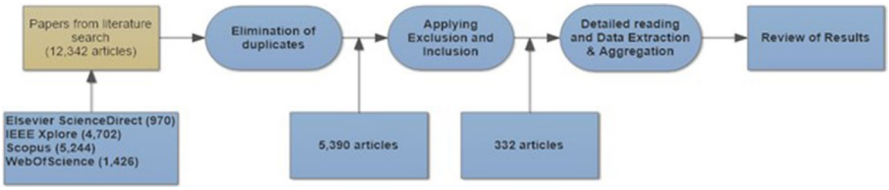

Fig. 2 Flow diagram of the literature mapping review

and more generic databases. Given that we endeavoured to study an area of research rather than practice, we decided to restrict our search to databases that contain predominantly or (almost) exclusivley contributions in peer-reviewed journals and conference proceedings. We finally settled down on a choice of four databases that we expected to contain the relevant material for our SMR: Scopus, IEEEXplore, ScienceDirect (Elsevier), and Web of Science. In cases where a database allowed specifying a domain area, we selected computer science, engineering, business/management, and law. As it was not easy to find all relevant articles via automatic searching and also to avoid the risk of disregarding papers that do not contain the exact term 'Big Data', we carried out, after paper screening (see Sect. 3.4) a parallel manual search using the reference list or authors and their works cited in the reference section of the selected papers, and we then applied the backward and forward snowballing technique described by Jalali and Wohlin (2012). We have not imposed any start or end dates to the publication range.

The search in all four databases was conducted in August 2020. In total, we retrieved 12,342 articles, all of which were imported into the EndNote X9 reference management application. This enabled us to detect and eliminate 6952 duplicates, leaving 5390 articles for the screening process Fig. 2.

3.4 Screening for pertinent papers

The paper screening was done against a set of well-defined inclusion and exclusion criteria. The inclusion criteria were directly related to the intent of our MRQs 3, 4 and 5, thus, only papers that were deemed relevant to answering these questions were included. For instance, an article returned by one of the search strings but with a central theme that has nothing to do with big data, was considered irrelevant and thus not considered for inclusion. Our exclusion criteria are described and motivated in Table 1.

The screening consisted of two consecutive steps. Initially, as common for SMRs, we reviewed papers against the inclusion and exclusion criteria based on title, keywords, and abstract. Considering the massive number of articles returned from the search, those without an abstract had to be excluded. For the few that did not have an abstract but whose title and keywords appeared quite relevant to our MRQs, we examined the introduction and conclusion sections of the papers. Next, a similar

Table 1 Exclusion criteria

Exclusion criterion	Reason
Source language other than English	Research contributions in peer-reviewed outlets are predominantly published in English
Contributions focusing solely on techniques, technologies, or tools	We are interested in general solutions or managerial approaches to address the issues of interest rather than the technical aspects of specific solution components. For example, papers discussing exclusively technological data security solutions for big data ecosystems (e.g. Apache Hadoop) without addressing the more general issues related to big data security or GDPR requirements or EAM practices, were excluded
Grey literature including books, book chapters, articles in press, notes, etc.	To ensure that the reviewed works are peer-reviewed, which is required for research contributions to be considered as part of the state-of-the-art in a research field
Domain areas other than computer science, engineering, business, management, and law	To exclude other perspectives than the ones mentioned (e.g., psychology, political sciences) to focus our survey eventually on studies that could potentially relate to EAM

type of review based on reading the paper's introduction and conclusions was performed for the articles containing abstracts.

As shown in Fig. 2, 332 unique articles survived the screening based on the inclusions and exclusion criteria. Following the approach of Petersen et al. (2008), we categorised these 332 articles into three broadly defined facets. The facets are based upon combinations of the major topics that are considered by our MRQs. Hence, we initially categorised the contributions according to the following facets:

1. Privacy and security with big data,
2. Big data and EAM and,
3. EAM and data privacy with a focus on the two European regulatory frameworks in a Big Data context.

3.5 Constructing the classification scheme

The construction of our classification scheme follows the methodological approach presented in the paper by Petersen et al. 2008). This approach is based on extracting keywords from the papers, for which we used MAXQDA 2020,[3] which is a CAQDAS[4] tool supporting qualitative and mixed-methods research. We fed the MAXQDA's database by the metadata and, where it was feasible,[5] the pdf file of each article, which were stored in EndNote. The use of a qualitative data analysis tool enabled us

[3] https://www.maxqda.com.

[4] Computer-aided qualitative data analysis software.

[5] This is due to access limitations imposed by some publishers.

to extract terms from the different articles and easily group them in thematic categories or themes. The resulting structure forms the desired classification scheme.

To do all this, we applied a flexible and highly iterative research technique, namely Template Analysis,[6] which is a generic form of thematic analysis offering a structured approach to hierarchical data coding. Initially, we started with an a priori code structure consisting of the three above defined facets. We then assigned codes to segment the text of the examined articles. While progressing, we identified additional codes and subcodes corresponding to reoccurring themes. Lastly, we organised the discovered themes into meaningful clusters to produce the most general higher order themes which provided the classification dimensions of our schema.

We considered Template Analysis as more appropriate than Grounded Theory (Charmaz 2006) as our current was not to develop a theory of how EAM could help addressing GDPR challenges in a Big Data context, but rather assess whether researchers have explored this avenue. Template Analysis enabled us to use an easily adaptable, 'lightweight' approach with many similarities to the Grounded Theory coding processes (King et al. 2018), whereby the research topics relevant to our study, including the connections between them, could be discovered and studied.

4 Mapping results

In this section, we present the classification scheme and show the summary mapping results for each MRQ.

4.1 Classification scheme

The application of Template Analysis revealed many dimensions for classifying the contributions. Whereas the three initial facets were based on the MRQs, the following classification scheme emerged from the selected articles.

Considering that EAM is the central theme of our research, we adopted the suggestion that, within its overarching context, both security and privacy as key elements in the big data lifecycle (Cagle 2015; Chen and Yan 2016) should put forward an enterprise-wide view (Bernard and Ho 2010). In this regard, EAM could play an important role in designing, implementing, and operating the necessary controls. Thus, the classification scheme makes use of key concepts that contribute to adhering to such a view. It consists of the following five dimensions: (1) data security and privacy with a focus on the two European data protection regulatory frameworks, (2) the triad of data governance, data management, and privacy risks, (3) emerging trends with regard to the data security and protection of big data, (4) the fundamental principles behind the processing of personal data in a big data context, and (5) research type. The last dimension was added specifically to assess the maturity of the state-of-the-art, as is common in SMRs. Note that EAM is absent in these dimensions as only few studies were found (see sub-Sect. 4.2).

[6] https://research.hud.ac.uk/research-subjects/human-health/template-analysis/what-is-template-analysis/.

(1) *Data security and privacy*

The immeasurable mass of data available today will continue to grow exponentially as time goes by. Seeking to address data security and privacy together and not apart is inevitable in the Big Data era (Kshetri 2014; Thuraisingham 2015).

Both concepts deal with different types of data 'protection' and should be regarded as complementary. Data security through confidentiality, integrity, availability, authentication, authorisation, and accounting lays the foundation for implementing data protection. It is in fact one of the key enablers of data protection. Put simply, security is essential for protecting data assets.

Data protection is more about protecting the rights of individuals over the use of their personal data (i.e., data privacy) than it is about securing that data. The DPD and GDPR particularly maintain and expand the focus on data security and protection by underlying the need for processing data lawfully and securely. However, even if data is compliant with any of the regulatory frameworks, that does not necessarily mean data is secure as well.

(2) *Governance and management of data and risks*

Apart from volume, big data contains a complex mix of structured and unstructured fast changing data that cannot be managed by improvident and conventional methods. A fiduciary duty of C-level officers is therefore to ensure that their organisation has taken appropriate steps to comply with the relevant law. Data governance can assist with such objectives by offering processes for proper planning, oversight, and compliance control over the use and management of all data-related resources (DAMA 2017). Apart from reinforcing proper handling of confidential information (Thompson et al. 2015), effective data governance yields numerous benefits to an organisation, and its absence can lead to a massive overhaul to meet the various standards, such as those raised by the GDPR (Beckett 2017).

Data management, on the other hand, includes a wide range of activities from getting strategic value to the way data as assets are operationalised. It involves the activities surrounding the entire data lifecycle. In this regard, data governance is seen as a part of the overall practice of data management. By placing it at the centre of its framework (DAMA 2017) has argued that data governance indicates to data management what should be done to support the organisation (see Fig. 3). In other words, their co-existence is vital for an effective and legally compliant use of the organisation's big data assets (Kemp 2014). The research study by (Aiken and Gorman 2013) revealed that most organisations suffer from poor data management implementation with more than 90% of CIOs not being 'data-knowledgeable' (Aiken and Gorman 2013). Conversely, organisations that have achieved data management success and shaped an organisation-wide data lifecycle management culture were the ones to exploit strategic data advantage. In the context of data protection, personal data should properly be managed from the initial data acquisition to data archiving and deletion. Thereby, data must be managed as an asset rather than a liability, and an effective way to accomplish that is by adopting a proactive attitude. Moreover, the proliferation of digital technologies and the fact that businesses have started employing scientific meth-

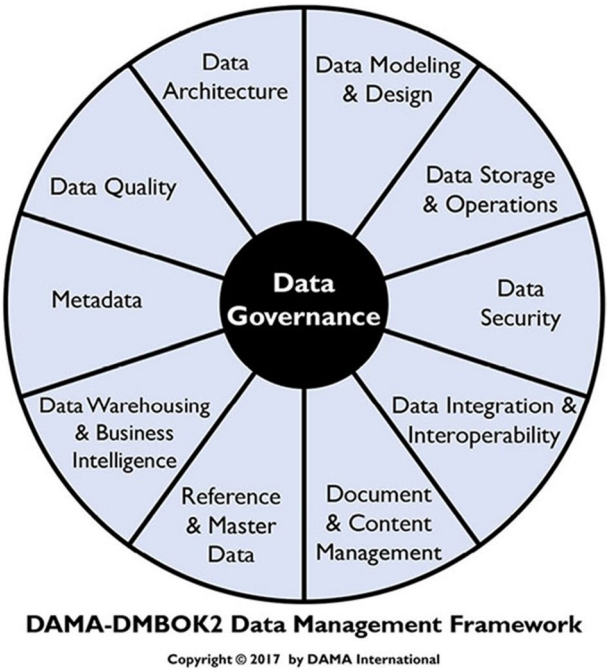

Fig. 3 DAMA-DMBOK2 Data Management Framework (DAMA 2017)

ods to benefit from the opportunities presented with collection and data mining for a large variety of data sources, notably data provenance and lineage, have increased the importance of data management and simultaneously turned into a difficult problem to address with big data (Bertot and Choi 2013; Demchenko et al. 2014; Kaisler et al. 2013).

In our analysis, we have considered data management and governance separately and looked for their references in the literature under review. We avoided, however, doing the same for information governance and data governance because these are very often used interchangeably in the literature (Olaitan et al. 2016). Despite that part of the data management includes managing risks associated with data, we considered risk management separately as well. The combination of governance, risk management, and compliance can help organisations employ an integrated and unified approach protecting them from the ill effects of (organisational) silos, reducing risk, and thus improving overall control effectiveness (ISACA 2013; Mansfield-Devine 2017). If data is not secure, organisations can end up in a non-compliant stage. A published report by the EU Parliament rapporteur Ana Gomez (Gomez 2017) stressed the fact that opportunities and prospects that big data brings to citizens, including in the private and public sectors as well as the academic community, can only be guaranteed when compliance with the EU data protection is ensured. Data compliance alone does not necessarily mean that data are also secure. To enforce proper compliance with current EU data protection law, both security and risk management

are necessary (EU 2016; NIST 2013). In this context, the GDPR has adopted a rather uncompromising stance on risks through Privacy Impact Assessments (DPIA) mandating a risk-based approach of finding and mitigating compliance risks and not simply carrying out a 'box-ticking exercise' (Gellert 2018). Within the same context, the element of risk has a 'dual nature' (Fritsch and Abie 2008), considering that a breach can affect organisations and users—the persons whom the data are about—alike. In conclusion, the successful combination of data governance and management of data and risks will not only help organisations achieve legally compliant data use but will also bring competitive advantages to them (Kemp 2014).

(3) *Emerging trends*

The early published papers (before 2016) in our set point towards the increasing use of technology trends led by the introduction of cloud computing and social media services for storing, exchanging, and analysing large quantities of high velocity, complex, and variable data.

Cloud computing encompasses various platforms where stored data are diffused across a potentially large number of sites, which in turn can be located in jurisdictions with less strict data laws (Khajeh-Hosseini et al. 2010; Pearson and Benameur 2010). Social media networks (SMN) spearheaded by Facebook, YouTube, and Instagram can be regarded a subset of cloud computing, given that their data are stored in the cloud (Campisi et al. 2009). One of their notable features that has led to many privacy-related disputes is the automatic sharing of their data with third parties. This phenomenon has given rise to what is known as 'data in wild' (Campisi et al. 2009).

Users access their data in the SMN using an assortment of portable and stationary devices over the Internet, which is commonly referred to as Internet of Things (IoT) (Minerva et al. 2015). Despite offering considerable benefits, such as ease and ubiquitous data reach, the risks that they entail are also considerable. Adding the blurry line between professional and personal use to this, emerging trends of this kind can pose serious threats regarding the security and privacy of data.

A trend that has received a lot of attention over the last 2 years is the impact of Big Data on ethical and social values. Big Data can generate through inference new knowledge and perspectives. Although this can create new opportunities, there are raising concerns about the influence of those technologies on society and more specifically for issues related to data abuse, discrimination, fairness, privacy and security (EESC 2017). Some reasons that increase the importance of this trend are the existing mechanisms relying on privacy by policy, such as (informed and explicit) consent and privacy notices or 'notice-and-consent' models (Mantelero 2014), fall short of providing sufficient transparency and control for individuals to know how their data have been collected and used. In addition, the increasing risk of statistical discrimination (Favaretto et al. 2019), a by-product of BDA, may cause non-material damage to DSs in accordance with Recital 75 of the GDPR.

The integration of ethical insights into the day-to-day work and technological advancements centrally to Big Fata processing is something that has also been

strongly stressed and promoted by the EU's independent data protection authority, EDPS (EDPS 2018).

(4) *Principles related to the processing of personal data*

The GDPR lays down a set of six privacy principles to guide how organisations manage personal data. These principles show up as recurring themes in the set of selected articles.

(a) Lawfulness, fairness, and transparency

In clear terms, data processing must be legal, fair, and transparent in relation to the DSs. A known way to demonstrate how well this principle is embraced is when asking the DSs to give their consent for processing their data for one or several purposes.

(b) Purpose limitation

Data must be collected for specific, explicit, and legitimate purposes. In simple words, it is not allowed for entities to do more with data than stated. Among the few exceptions are when data is processed for archiving, scientific, and historical research purposes.

(c) Data minimisation

A common old practice to collect more data than that is needed is no longer lawful. It is therefore mandatory to collect the bare minimum of what is required for the purpose for which the data is processed.

(d) Accuracy

From a data privacy and protection perspective, data must be kept up to date and complete on the planned and performed data processing. Inaccuracies must be rectified without delay, meaning the overall data quality of personal data must be ensured.

(e) Storage limitation

Data must be kept no longer that is necessary and only for the purposes for which the personal data are processed. As with the purpose limitation, the GDPR describes data processing purposes that do not apply to this principle.

(f) Integrity and confidentiality

Data must be kept safe while under processing. In other words, appropriate security measures must be put in place to ensure secure access to personal data and protection against unauthorised or unlawful processing, including accidental loss or destruction.

(5) *Research technique*

The research technique uses the classification scheme proposed by (Wieringa et al. 2006). It basically describes the research approach followed by the authors by mapping works to the following six classes:

- Validation Research

  The paper reports on the validation of innovative techniques that have no real-world implementation (e.g., through experiments).
- Evaluation Research

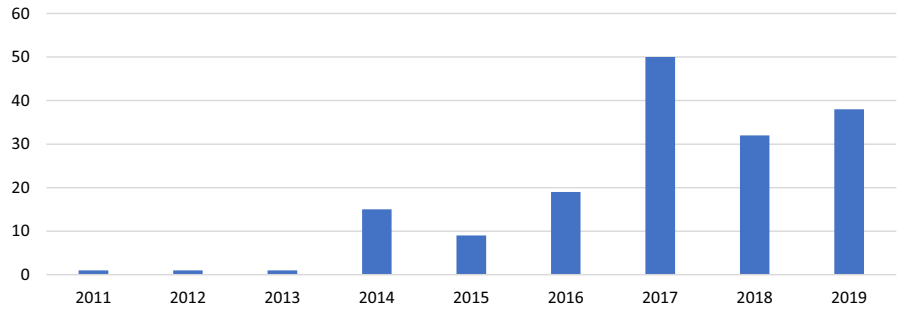

Fig. 4 Distribution of articles according to the publication year

Contrary to validation research, the paper assesses existing, practical implementations.

- Solution Proposal
  The paper presents improvements to existing techniques but without providing a tangible evidence (i.e., neither validation nor evaluation).
- Philosophical[7] Paper
  The paper provides an ideal and abstract view of how things might be examined.
- Opinion Paper
  The paper airs the author's opinion but without producing any proof from related work.
- Experience Paper
  Instead of opinion, the paper is based on the experience of the author and the lessons learnt from having been involved in similar projects.

4.2 Answering the mapping review questions

The answers to our MRQs were obtained from aggregating and summarizing the reviewed papers, as aided by our classification scheme (from MRQ3 onwards).

4.2.1 MRQ1: How many studies on these topics were published over the years?

Figure 4 shows that out of the 332 selected papers, 99% of them were published from 2013 onwards. Although our search queries have not imposed any date boundaries, almost no papers were found for the years prior to 2011. The graph indicates that the papers of interest to our survey had two sudden increases in 2014 and 2017, whereas their number on average has remained steady ever since. In this graph, we do not consider year 2020 as it takes some time for articles to appear in the databases.

The reason that virtually no papers were discovered for the years prior to 2011 can be explained by the remark by Özköse et al. (2015) along with the findings by

[7] It is also referred to as 'conceptual' (Jalali and Wohlin 2012).

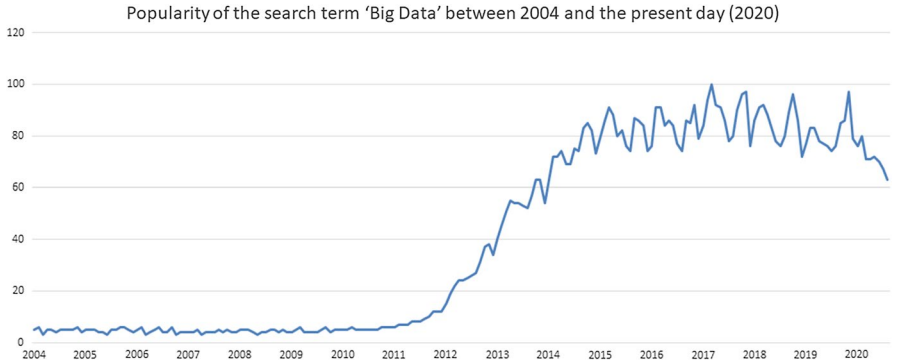

Fig. 5 Popularity of the search term 'Big Data' between 2004 and 2020 in Google Trends Analytics (https://trends.google.com/trends/explore?q=Big%20Data&date=all#TIMESERIES)

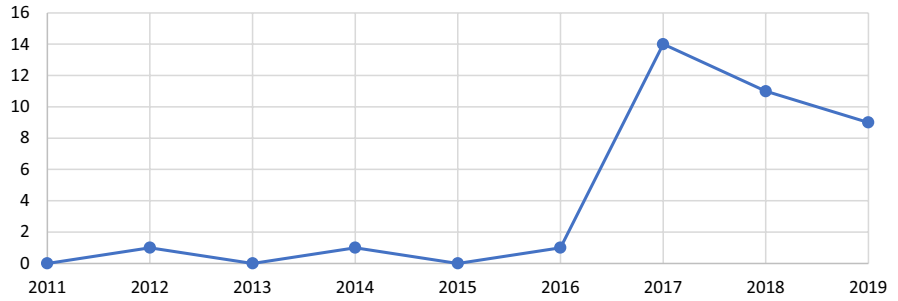

Fig. 6 Distribution of articles addressing DPD or GDPR matters according to the publication year

Gandomi and Haider (2015) that the academic research on Big Data has only been noticeably active since 2012. This is reflected in the information depicted in the Google Trend Analysis graph in Fig. 5, which shows the overall popularity of online searches where the term 'Big Data' was used in the period between 2004 and middle 2020.

What appears fairly interesting is the steep increase in 2017 of articles discussing matters directly related to the DPD or GDPR, as shown in Fig. 6 the tendency of producing papers about the GDPR has reduced slightly over the next 2 years but we believe will be growing, as future research will increasingly be focusing on discussing concerns and implications for businesses as well as the lives of individuals. A recent example is the outbreak of COVID-19 pandemic, which has created many new opportunities in using Big Data for combatting the disease as well as serious risks for applying intrusive measures that violate GDPR rules and fundamental rights to privacy given that the tracking of individuals tends to become a newly established norm (Deloitte 2020).

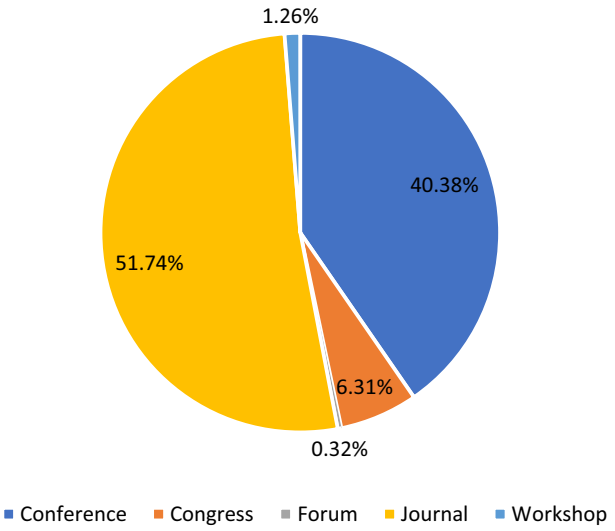

Fig. 7 Proportion of selected studies by venue type

4.2.2 MRQ2: What are the most frequently used publication outlets?

Table 5 in the Appendix, depicts the list of the publication venues and types of the selected papers. The 332 articles are distributed over 182 different publication venues. The leading venues are shown in the top rows of the table. *Computer Law & Security Review* (*CLSR*),[8] a journal that addresses IT law and computer security is where most journal publications were found. The absolute number of papers in that journal is, however, still low (21) compared to the total number of articles. Next to *CLSR,* follows the IEEE's Congress on Big Data that covers emerging theme-topics around Big Data and the ACM International Conference Proceeding Series (ICPS), which allows the easy publishing of proceedings from many conferences. Looking at the long list of conferences where IEEE is present, it is evident that it has made a significant contribution towards the Big Data domain, especially for advanced technology matters, such as analytics engines for processing and techniques for data protection.

It is, however, evident that the research surveyed is published in a wide variety of venues. Regarding the proportion of venue types, journals and conferences are the two main publication types, which account for 51.74% (164 out of 332 studies) and 40.06% (128 out of 332 studies) of the articles, respectively (see Fig. 7). Congress papers constitute 6.31% or 20 out of 332 of the total selected studies with the remaining 1.58% (5 out of 332 studies) in forums and workshops.

[8] https://www.journals.elsevier.com/computer-law-and-security-review.

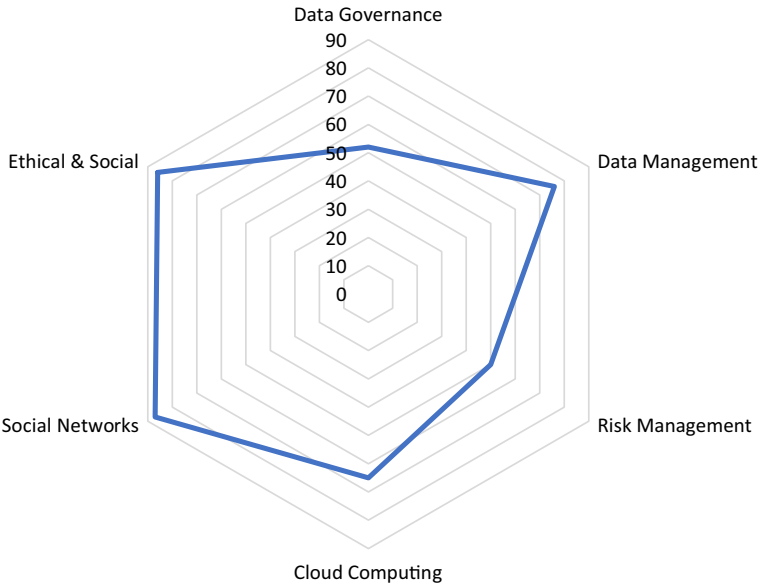

Fig. 8 Number of articles for topics in the classification scheme where data security and privacy on Big Data are addressed

4.2.3 MRQ3: What are the main issues and challenges with respect to data security, protection and privacy in Big Data?

Figure 8 was constructed based on the classification of papers according to the first three framework dimensions. The fact that data security and privacy have received the most attention is because we were favourably inclined towards papers that have them as the central theme. Only the third search string could deliver other papers, according to the second initial facet (i.e., EAM and Big Data). Figure 8 depicts how papers addressing data security and privacy in a Big Data context relate to data management, data governance and risk management (second dimension) and the major emerging trends discovered (third dimension).

While exploring for topics addressed by the subset of papers that address security and privacy of Big Data (i.e., the first facet in our categorisation of papers), we note in Table 2 that data governance and risk management do not exceed 26% in the set of papers[9] we analysed. These findings reveal that a relatively greater focus was given on exploiting data management in conjunction with ethical and societal issues and along with social networks and cloud computing at the expense of addressing issues and challenges that arise when data governance and risk management are not successfully established. Research in pertinent domains and examples from top consultancy companies have shown that data governance is one of the most effective

[9] Some papers address more than a single topic.

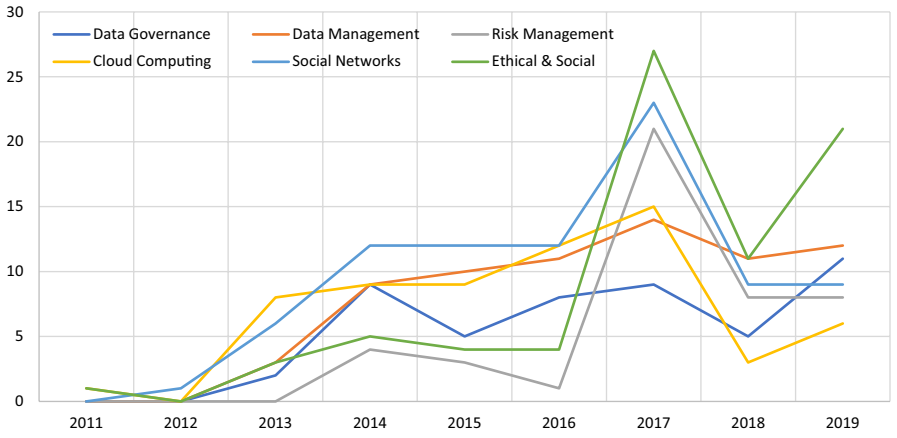

Fig. 9 Topics of interest throughout the publication years

means for exerting control, especially for areas related to availability, integrity, confidentiality, and data quality (Newman and Logan 2006; Thompson et al. 2015).

As discussed above, the management of data risks followed by the development and deployment of countermeasures to mitigate those risks is an element of long-term prediction. It also offers reassurance whereby an organisation manifests that it takes the processing of its data assets seriously. Privacy protection and compliance requirements add another dimension of risk to data processing activities. Data risks are not always privacy-related but can be entailed from the way data is processed from ethical and social perspectives (EESC 2017; Mantelero 2017). On the other hand, lessons learnt from recent cyberattacks, like those targeting governments and international organisations working on the COVID-19 pandemic response, and the overall data handling of sensitive information during the same period, have put the management of risk in question (CSC 2020; Deloitte 2020). Hence, GDPR puts extra emphasis upon placing the appropriate security measures for mitigating risk at an acceptable level.

Another important observation is made from the information depicted in Fig. 9. It shows us that authors' interest in data governance, data management and risk management have slightly increased over the years.[10] On the other hand, the interest in cloud computing and social networks has progressively declined since 2014. Perhaps the research on these two topics has become more mature and thus saturated. It could also be that they are more attractive to papers discussing technical matters such as the application of analytics algorithms, storage frameworks and programming techniques. What we find impressive, however, is the rapid increase of papers discussing the impact of Big Data on ethical and societal values. As BDA is gradually used by public bodies (e.g. political parties and research groups), large

[10] We avoided to include the figures from 2020 because we believe that the number of papers discovered for this year are not yet representative.

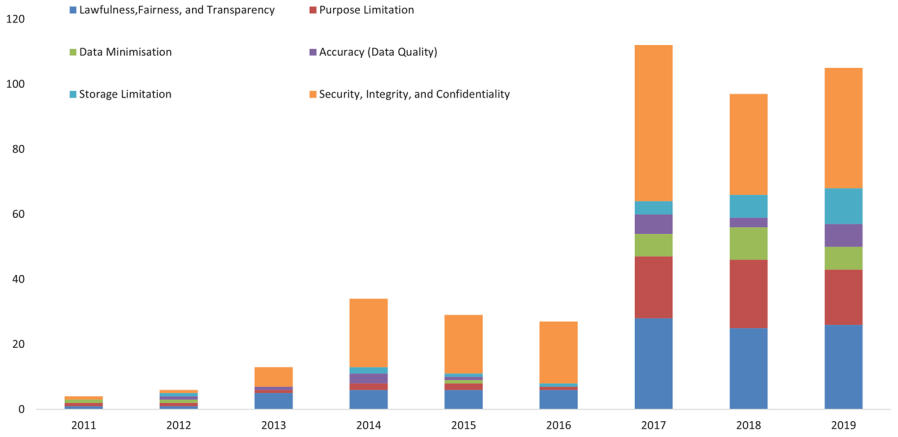

Fig. 10 Literature coverage of the six data protection principles

enterprises active in the area of healthcare, and governments, the ethical handing of personal data has raised many serious concerns for identifying the appropriate legal and organisational remedies all of which need to be carefully considered (Bertino et al. 2019; EC 2018; Jurkiewicz 2018; Saltz and Dewar 2019). The GDPR alone, for instance, does not adequately address that concern (Hijmans and Raab 2018; McMahon et al. 2020). Finally, the peak occurring in 2017 could be explained from the fact that GDPR was just introduced and had attracted the attention of several researchers.

4.2.4 MRQ4: What does Big Data research show regarding the rising needs for DPD or GDPR?

As the number of papers where the two European data protection regulatory frameworks are involved, has increased especially over the last 3 years, we decided to carry out a complementary qualitative content analysis through which an attempt was made to interpret how well authors of the selected papers have embodied the six data protection principles in their analysis (i.e. the fourth dimension of our classification framework). This resulted in the creation of Figs. 10 and 11 that show a progressive change to attention to all principles and not just some of them. Most studies still focus on the security, integrity, and confidentiality principles. A reason for that is because these are the most well perceived and researched principles, as they involve interdisciplinary techniques, including a manifold of advanced technologies by means of tools known as privacy and transparency enhancing tools or PETs (Gürses 2010) and TETs (Murmann and Fischer-Hübner 2017). These also happen to constitute the mass in the wide range of papers found in this study.

However, authors of recently published papers in years 2018 and 2019 have specifically addressed two of the most challenging GDPR concerns of storage limitation and purpose limitation, which in their turn can have profound effect in the precision and reliability of BDA.

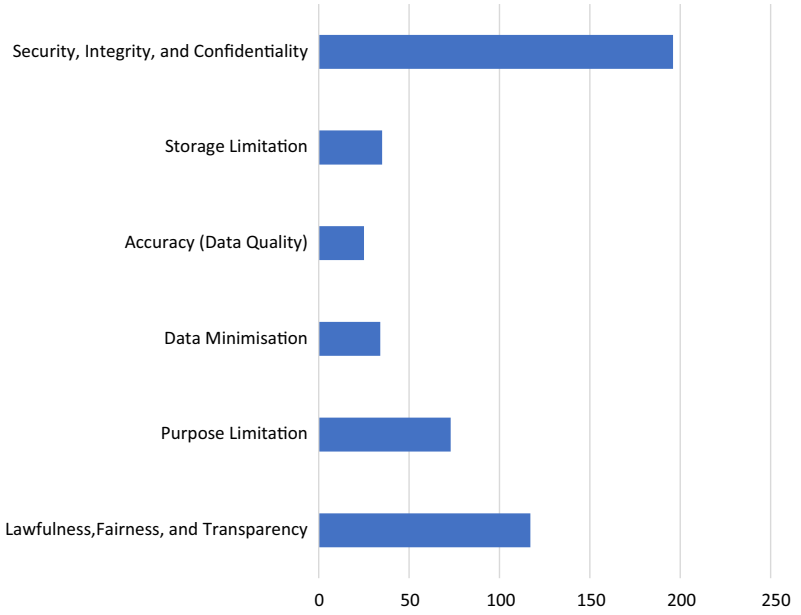

Fig. 11 Coverage all data protection principles for all years from 2011 to 2019

4.2.5 *MRQ5:* Considering EAM, is there any approach or method for dealing with Big Data, and what is its perceived role regarding Big Data security and privacy aspects?

Table 3 lists the sixteen papers discovered during the literature review which propose EAM as a solution for dealing with a variety of Big Data and GDPR-related issues (i.e., the papers in the second and third facets of the initial categorisation).

Their relatively small number confirms the observed gap of research on the nexus of EAM and Big Data or GDPR. The papers address either the EAM—Big Data relationship or the EAM–GDPR relationship, but only one paper (i.e., (Chao 2018) simultaneously addresses all three areas). The fact that most of those papers are published from 2018 onwards indicates that the research interests in applying EAM for big data protection and privacy compliance in environments with BDA applications, is quite recent.

Our insights obtained from these papers in relation to MRQ5 are summarized as follows: In an increasingly complex and changing business environment comprising highly integrated systems and information technologies with multifaceted interdependencies, conventional management approaches lack strategic vision (Ahlemann et al. 2012). Without a clearly defined strategic vision, negative consequences on performance and risks can ensue. Understanding the value that Big Data and BDA bring to the business along with the compliance risks imposed, multiple factors and perspectives must be considered (Vanauer et al. 2015). It is only then possible to create or take advantage of opportunities that lie ahead and respond to new challenges like the urgent need for full compliance with the GDPR requirements.

The research conducted by Chao (2018; Gong and Janssen 2017, 2020; Lněnička et al. 2017) supports the idea that EAM is an effective means to address these needs, as it enables organisations to easily adapt and manage, to a certain degree, the complexity of their corporate information systems, increase the business value of BDA besides developing the appropriate internal capabilities. In this regard (Kehrer et al. 2016a, b) pointed out that many organisations endeavour to integrate Big Data technologies into existing IT and business architectures. On that basis, they argued that key challenges that arise in Big Data management, such as data privacy, data quality, governance, intellectual property, and data security, are implicitly related to EAM. They stressed that approaches and methods provided by the Big Data literature are making implicit use of EAM's underlying principles. In this respect, they identified techniques that apply similarly within EAM. Contrary to expectation, they were not able to single out existing EAM approaches or methods that can be readily used to address all Big Data challenges. One reason for that could be the appearance of technologies, such as the BDA that introduce new types of architectural challenges. This also implies that EAM cannot be used straight out of the box as a fix-all solution. It requires proper institutionalisation and integration into existing processes (Kearny et al. 2016; Vanhoorelbeke et al. 2020).

As regards the compliance with the GDPR (Huth et al. 2020) argues that EAM allows to identify and address several regulatory concerns. Although their work is limited to organisations located in German-speaking areas, it provides valuable insights where EAM could be useful to ensure the protection of data subject rights and support organisation's efforts for the digitalisation of business processes. For example, they show in Huth et al. (2019) how EAM models could be used for creating the Record of Processing Activities (RPA), which one of the strict legal obligations, as per Art.30 of the GDPR. The RPA includes significant information about data processing, data categories including the purpose of their processing and must be available to the authorities upon request (EU 2016). This obligation poses major challenges for companies and organisations that work with personal data as its accuracy and comprehensiveness are closely associated with the level of embracement of the data protection principles. Building on the same idea (Blanco-Lainé et al. 2020; Burmeister et al. 2019, 2020) argue that EAM can play a key role in the GDPR implementation by addressing the privacy and security related concerns of various classes of stakeholders and by doing so to gain a better understanding of how an effective information governance could be implemented in the organisation. Referring to the size of the organisation, (Rozehnal and Novak 2018) argue that EAM can bring equal benefits to SMEs, where resources for the support of GDPR compliance processes are inheritably limited. Moreover, EAM can help with the development of a data strategy and can define and adopt the data-centric principles to guide the data strategy goals and objective criteria needed for assessing data leveraging, which is crucial as far as the proper management of Big Data is concerned.

4.2.6 MRQ6: Can we identify any topics within these areas that could be subjects for further research?

Throughout this study, we identified the following concerns and open questions, which merit further investigation:

(1) Adapting and applying Big Data technologies to existing application areas and the IT landscape, including the prediction of their effects, remains a daunting challenge (Gong and Janssen 2017). For example, through analytical techniques and tools, Big Data has increased the scope of personally identifiable information or information that can be used alone or combined with other information to identify an individual (Crawford and Schultz 2014; Sampson 2014; Tene and Polonetsky 2011; Thuraisingham 2015). As a result, organisations are not yet capable of utilising Big Data to its fullest potential.
(2) The influence of Big Data transformation processes has exacerbated the problem of addressing needs in a holistic manner. The so-called 'management by magazine' is a phenomenon that emerged from this problem whereby decisions are taken without a prior adequate analysis of direct and indirect risks. As a consequence, consumers have little faith in companies for safeguarding their data (Chibba and Cavoukian 2015). It is therefore imperative that when Big Data value is assessed, it should be done in parallel with concerns related to data security and privacy (Suresh 2014).
(3) Certain domains like health care, and technologies like cloud computing, may be affected more than others by the adoption of Big Data. It would, however, be a wrong assumption that the same set of data security and privacy protection measures could be uniformly applicable. Hence, prior to any consideration, it is necessary to analyse them from different organisational perspectives (Holm and Ploug 2017; Kunz et al. 2014; Li and Guo 2014; Patil and Seshadri 2014; Tse et al. 2018).
(4) The increasing strategic role of Big Data in competitive environments requires the use of a thorough legal model to understand the legal rights and responsibilities (Kemp 2014; Tene and Polonetsky 2012) and thus cope with legal uncertainties that surround its management (Spiekermann et al. 2015). The existence of such a model can drive data governance and management and act as a key enabler of actively and effectively managing the privacy risks, a big part of which are posed by re-identification from the processing of large quantities of correlative non-personal data and inadequate security (Joo et al. 2017; Yu 2016).
(5) The impact of Big Data and BDA on ethical and societal values including pertinent concerns adds another dimension of complexity to data protection in terms of the GDPR (Cuquet and Fensel 2018). Notions like the personal consent on which regulatory frameworks have traditionally been based, do not offer adequate remedies for dealing with data subject rights and data protection principles. Hence, there are raising concerns about the influence of Big Data technologies on the society and more specifically for issues related to discrimination and fairness of data processing (Amalina et al. 2020; EESC 2017).

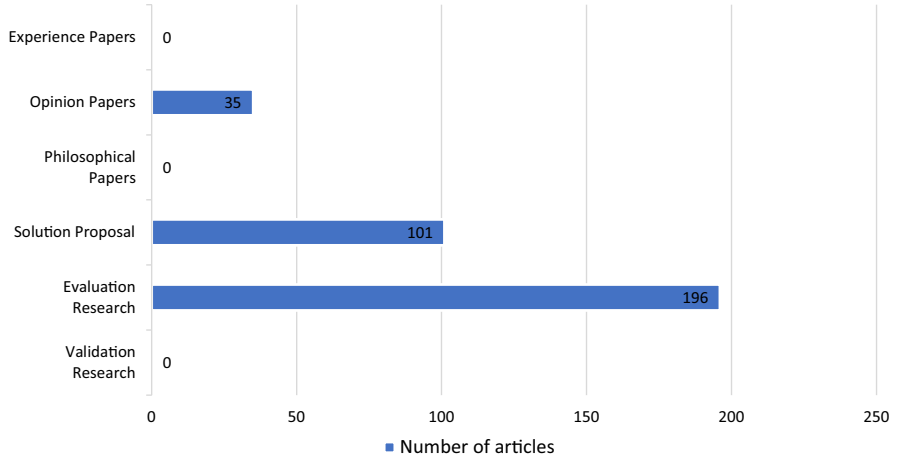

Fig. 12 Number of articles classified as each research type

Table 2 Number of articles along with their percentages for topics in the classification scheme where data security and privacy on Big Data are addressed

Topics	# Papers	%
Social Networks	87	43.28%
Ethical & Social	86	42.79%
Data Management	76	37.81%
Cloud Computing	65	32.34%
Data Governance	52	25.87%
Risk Management	50	24.88%

(6) Allowing the analytics to produce useful insight while preserving anonymity and individual privacy is highly challenging and constitutes one of the mounting concerns in Big Data (Basso et al. 2016; Gahi et al. 2016; Jensen 2013). Individuals can become widely vulnerable to security and privacy violations given that huge amounts of their behavioural data are being collected from online footprints, sensors, and sensory systems, notably without a pre-established goal (van der Sloot 2015), and are stored in dispersed databases thereby presenting an attractive target for attackers. Finally, data are also harnessed by analytics to acquire and infer inside knowledge (McKInsey 2016). Many of these processes often occur without the knowledge of individuals (Tan and Pivot 2015).

(7) Having analysed the selected papers against the research technique type classification that was first introduced by (Wieringa et al. 2006), another concern emerged. Although their classification scheme was originally made for evaluating requirement engineering research papers, (Wohlin et al. 2013) generalised it while stressing that the choice for the type is bit subjective. In their view, it could apply to different fields, as it can convey useful information about the research approach thus far. Figure 12 shows that a clear majority of authors preferred

Table 3 Research papers discussing the use of EAM for dealing with Big Data and GDPR requirements

Reference	Year	Research methodology and research topic
Vanauer et al. (2015)	2015	Proposal of a EAM-based ideation, assessment, and implementation methodology for introducing Big Data in organisations
Kehrer et al. (2016a, b)	2016	Systematic Literature Review (SLR) to identify the conceptual categories of EAM requirements that can support the adoption of Big Data by the organisations
Kearny et al. (2016)	2016	Systematic Literature Review (SLR) for the integration of Big Data and BDA into the TOGAF Architecture Development Method
Lněnička et al. (2017)	2017	Design Science Research aiming at a solution for identifying BDA requirements from a strategic management point of view
Gong and Janssen (2017)	2017	Action Research through the study of a large administrative organisation to identify how EAM can support the effort required for the adoption of Big Data
Rozehnal and Novak (2018)	2018	Case Study for the possible usage of EA artefacts to implement the GDPR in a small or medium-sized enterprise (SME)
Aldea et al. (2018)	2018	Systematic Literature Review (SLR) to study the relationship between EAM and Big Data (Analytics)
Chao (2018)	2018	Literature Study to explore the role of EAM in building a sustainable, regulation-guided holistic ecosystem that involves Big Data technologies
Burmeister et al. (2019)	2019	Systematic Literature Review (SLR) to study the approaches for ensuring organisation's compliance with the GDPR requirements
Huth et al. (2019)	2019	Literature Study combined with the interviews of domain experts and focus groups to evaluate the use of existing EAM models for improving organisation's compliance with the GDPR Record of Processing Activities requirement
Huth et al. (2020)	2020	Literature Review combined with Qualitative Interviews to explore the possible synergies between EAM and Data Protection Management in the context of the GDPR implementation
Blanco-Lainé et al. (2020)	2020	Literature Review to study the way to use EAM for ensuring GDPR compliance
Burmeister et al. (2020)	2020	Systematic Literature Review (SLR) to study the interplay of EAM and information governance using the GDRP implementation as an example
Vanhoorelbeke et al. (2020)	2020	Literature Study combined with domain expert interviews to understand the current state of the art along with the challenges concerning IOT and Enterprise Architecture Framework (EAF)
Gong and Janssen (2020)	2020	Literature Study with thet aim to study the EA roles and the capabilities required for adopting and implementing BDA in large public organisations

'solution proposal' and 'evaluation research' types. The former, corresponding to 30% of the papers, is used to describe and discuss a research problem but without conducting any validation or evaluation, whereas the latter, corresponding to 59% of the papers, investigates the outcome of techniques, methods, and tools in a practical context. The last 11% were classified as 'opinion papers'. Papers of this type express the opinion of the authors who did not rely on any research methodology. The fact that zero papers were found in the other categories is because the validation type refers to the testing of solutions based on novel techniques that often takes place in a lab setting. In the context of our research, there are proposals of solutions, but independent validation is apparently lacking. Moreover, this could be a sign of a research domain that is not yet mature, but gaining interest, which is confirmed by the increasing number of evaluation and solution related publications. The experience papers refer to author's personal experiences and lessons learnt. We presume that this type of papers will appear in the future when the research involving GDPR, EAM in connection with Big Data and BDA become more mature and finds its way to practice. Finally, philosophical papers are less pertinent, but their number could increase once research moves towards questioning underlying ontological, epistemological and methodological foundations of a maturing research field.

5 Discussion

In this section, we interpret the results. We discuss the main findings in terms of research gaps and future research.

5.1 Extent of studies in the domain of data security and privacy of big data

During the analysis of the mapping results, we noticed that the papers where the DPD or GDPR are discussed had an increase since 2016 with a sharp peak on 2017. It is apparent that the profound effects on the GDPR along with all its uncertainties and confusion have introduced a new dynamic. In this regard, we expect that the central themes of future papers would be on examining its effects, while involving discussion over possible solutions and validating pertinent techniques, methods, and tools. Given that many of the legal terms used in the GDPR are hard to grasp, especially by actors with no law background, it could result in another stream of research attempts through which the different terms used are approached from a conceptual and empirical standpoint. Adding to that, are all considerations about ethical and societal values involved, directly or indirectly, in the processing of Big Data and BDA, which according to our mapping results have drawn a lot of attention over the last 2 years (i.e. 2018 and 2019).

5.2 Coverage of key issues and challenges

Our analysis has also revealed that both data governance and risk management have not received the appropriate level of attention despite their order of magnitude in addressing compliance matters with legitimate personal data processing. As already mentioned, the GDPR puts extra emphasis on the sound management of data risks. By having explicitly expressed the privacy-by-design approach (Rubinstein and Good 2013) as one of its legal requirements, it makes PIAs or DPIAs (in GDPR's parlance) mandatory in certain circumstances. For example. a DPIA is required in situation where data processing is likely to result in high risk to individuals. Failing to conduct a DPIA where mandated constitutes a serious breach. In addition, it is suggested to expand the scope of DPIA to overcome limitations on existing risk assessment models to cover the impact from social and ethical values in Big Data or BDA project contexts. Our future research paper that employs SLR shall therefore aim at studying data governance alongside the adequate management of data protection risks including a representative set of pertinent (DPIA and PIA) methodologies.

5.3 Considering EAM to employ an adaptive and integrated management approach

We noticed that EAM opens up an array of opportunities to engineering the appropriate Data Protection by Design-inspired technical and organisational measures by employing an adaptive and integrated management approach in Big Data contexts. The literature discovered for the purpose of this paper acknowledges the business value of EAM and its means for achieving such a goal. This includes the general adoption and integration of new technologies (Gong and Janssen 2017), the governing and undergoing of complex transformations on all layers from an as-is to a to-be state (Kehrer et al. 2016a; Vanauer et al. 2015), and solving technical challenges of Big Data management projects in mainstream sectors like banking, insurance, air transport industry, healthcare, and the public sector (Kemp 2014). Bearing in mind that Big Data problems are not solely an IT exercise but rather a strategic function, future research should therefore be enriched with papers discussing practical cases with the aim to assess the role of EAM in holistically addressing the adoption of DBA or Big Data while taking into consideration ethical and societal concerns and in the scope of conducting risk management operations with the use of a more improved version of the GDPR DPIA.

5.4 Addressing different trends

The seven trends we have identified offer additional opportunities for research. Depending on the problem at hand, the ideas that stem from each trend could be combined or used alone.

6 Conclusion and outlook

For most organisations, data are regarded as an indispensable asset. The explosive growth of social media platforms and digital information devices have catapulted the data volume, resulting in an emerging area involving technology, people and organisations and known as Big Data. The purpose of this paper was to conduct an SMR to study EAM's general perception in managing the effect of the GDPR compliance requirements in a Big Data context. We searched four digital libraries through which we discovered and then analysed 332 pertinent articles. Our study has shown that Big Data has radically expanded the range of personally identifying an individual, leading to more risk and complexity. In sensitive sectors like health care, this has significantly increased security, privacy, and data protection concerns. Daunting challenges in the context of understanding and handling Big Data issues linked to security and data privacy concerns are somehow linked to the lack of holistic view. As a result, there are different approaches for Big Data adoption that introduced new frameworks, methods, and technologies but without considering problems in the long run and from a wider perspective.

We have identified seven trends that merit researchers' attention regarding the effect of Big Data technology and the role of the different organisational levels. The vast body of existing literature focuses mainly on solutions for technical problems like the optimisation of existing BDA algorithms or implementation of Big Data frameworks (e.g. Apache Hadoop, NoSQL, etc.) and much less on pressing organisational matters, such as risk management and data governance, which are barely addressed in the context of Big Data and BDA.

Finally, despite the existence of articles acknowledging the business value of EAM in a Big Data context and the positive effects of employing it during the implementation of GDPR to address several concerns or risks such as the management of the creating of RPA ensuring organisation's compliance with data subjects' rights, there are still needs for additional research to address other data protection concerns like ensuring better protection of data subjects' rights besides the ethical processing of their personal data.

Appendix

See Tables 4 and 5.

Table 4 Table with the query strings of the literature searching

Database	Query #1	Query #2	Query #3
Elsevier ScienceDirect[a]	Title-Abstr-Key ("big data" AND (security OR privacy OR "data privacy" OR "data protection" OR dpd OR "eu dpd" OR "Data Protection Directive" OR "gdpr" OR "general data protection regulation" OR "data protection regulation"))	Title-Abstr-Key ("big data" AND (security OR privacy OR "data privacy" OR "data protection" OR dpd OR "eu dpd" OR "Data Protection Directive" OR "gdpr" OR "general data protection regulation" OR "data protection regulation") OR ("enterprise architecture management" OR EAM OR "ea management" OR ea OR "enterprise architecture"))	Title-Abstr-Key ("big data" AND ("enterprise architecture management" OR EAM OR "ea management" OR ea OR "enterprise architecture"))
IEEE Xplore[b]	"Document Title": "big data" AND (security OR privacy OR "data privacy" OR "data protection" OR dpd OR "eu dpd" OR "Data Protection Directive" OR gdpr OR "general data protection regulation" OR "data protection regulation") "Abstract":"big data" AND (security OR privacy OR "data privacy" OR "data protection" OR dpd OR "eu dpd" OR "Data Protection Directive" OR gdpr OR "general data protection regulation" OR "data protection regulation") "Author Keywords":.QT.big data.QT. AND (security OR privacy OR .QT.data privacy. QT. OR .QT.data protection.QT. OR dpd OR .QT.eu dpd.QT. OR .QT.Data Protection Directive.QT. OR gdpr OR .QT.general data protection regulation.QT. OR .QT.data protection regulation.QT.)		"Document Title": "big data" AND ("enterprise architecture management" OR EAM OR "ea management" OR ea OR "enterprise architecture")

Table 4 (continued)

Database	Query #1	Query #2	Query #3
Scopus[c]	TITLE-ABS-KEY ("big data" AND (security OR privacy OR "data privacy" OR "data protection" OR dpd OR "eu dpd" OR "Data Protection Directive" OR "gdpr" OR "general data protection regulation" OR "data protection regulation")) AND (LIMIT-TO (SRCTYPE , "p ") OR LIMIT-TO (SRCTYPE , " j ")) AND (EXCLUDE (SUBJAREA , "MEDI ") OR EXCLUDE (SUBJAREA , " PHYS ") OR EXCLUDE (SUBJAREA , " MATE ") OR EXCLUDE (SUBJAREA , " BIOC ")) AND (LIMIT-TO (LANGUAGE , "English ")) AND (EXCLUDE (DOCTYPE , "cr") OR EXCLUDE (DOCTYPE , "ip") OR EXCLUDE (DOCTYPE , "re") OR EXCLUDE (DOCTYPE , "ed") OR EXCLUDE (DOCTYPE , "no") OR EXCLUDE (DOCTYPE , "ch") OR EXCLUDE (DOCTYPE , "sh")) AND (EXCLUDE (SUBJAREA , "DECI") OR EXCLUDE (SUBJAREA , "SOCI") OR EXCLUDE (SUBJAREA , "MATH") OR EXCLUDE (SUBJAREA , "ENER") OR EXCLUDE (SUBJAREA , "ENVI") OR EXCLUDE (SUBJAREA , "MULT") OR EXCLUDE (SUBJAREA , "ARTS")) AND (EXCLUDE (LANGUAGE , "German") OR EXCLUDE (LANGUAGE , "French") OR EXCLUDE (LANGUAGE , "Spanish"))		TITLE-ABS-KEY ("big data" AND ("enterprise architecture management" OR eam OR "ea management" OR ea OR "enterprise architecture")) AND (EXCLUDE (DOCTYPE , "ch") OR EXCLUDE (DOCTYPE , "re") OR EXCLUDE (DOCTYPE , "bk")) AND (EXCLUDE (SUBJAREA , "MATH") OR EXCLUDE (SUBJAREA , "DECI") OR EXCLUDE (SUBJAREA , "SOCI") OR EXCLUDE (SUBJAREA , "MATE") OR EXCLUDE (SUBJAREA , "PHYS") OR EXCLUDE (SUBJAREA , "CHEM") OR EXCLUDE (SUBJAREA , "ENER")) AND (EXCLUDE (LANGUAGE , "German"))

Table 4 (continued)

Database	Query #1	Query #2	Query #3
Web of Science[d]	((TI=("big data" AND (security OR privacy OR "data privacy" OR "data protection" OR dpd OR "eu dpd" OR "Data Protection Directive" OR "gdpr" OR "general data protection regulation" OR "data protection regulation")))) AND LANGUAGE: (English) AND DOCUMENT TYPES: (Article) Indexes=SCI-EXPANDED, SSCI, A&HCI, CPCI-S, CPCI-SSH, ESCI Timespan=All years		(TS=("big data" AND ("enterprise architecture management" OR eam OR "ea management" OR ea OR "enterprise architecture"))) AND LANGUAGE: (English) AND DOCUMENT TYPES: (Article) Timespan: All years. Indexes: SCI-EXPANDED, SSCI, A&HCI, CPCI-S, CPCI-SSH, ESCI

[a] https://auth.elsevier.com/ShibAuth/institutionLogin?entityID=urn%3Amace%3Akuleuven.be%3Akulassoc%3Akuleuven.be&appReturnURL=http%3A%2F%2Fwww.sciencedirect.com

[b] http://ieeexplore.ieee.org/Xplore/home.jsp

[c] https://www.scopus.com/search/

[d] https://apps.webofknowledge.com/

Table 5 Number of selected studies over the publication venues

Publication venue	Venue type	Number of papers
Computer Law and Security Review (CLSR)	Journal	25
International Congress on Big Data (IEEE)	Congress	19
ACM International Conference Proceeding Series	Conference	15
International Conference on Big Data (IEEE)	Conference	15
IEEE Access	Journal	10
Journal of Big Data	Journal	9
Procedia Computer Science	Conference	9
Ethics and Information Technology	Journal	8
Information & Communications Technology Law	Journal	8
International Conference on Computing, Communication and Automation	Conference	6
International Journal of Engineering and Advanced Technology	Journal	4
Business Horizons	Journal	4
Business Process Management Journal	Journal	3
China Communications	Journal	2
Conference on Computer Communications Workshops	Conference	2
Conference on IT in Business, Industry and Government	Conference	2
Future Internet	Journal	2
Health and Technology	Journal	2
International Advance Computing Conference	Journal	2
International Conference for Internet Technology and Secured Transactions	Conference	2
International Conference on Collaborative Computing	Conference	2
International Conference on I-SMAC	Conference	2
International Enterprise Distributed Object Computing Workshop	Workshop	2
International Journal of Information Management	Journal	2

Table 5 (continued)

Publication venue	Venue type	Number of papers
Big Data and Society	Journal	2
International Journal of Recent Technology and Engineering	Journal	2
Journal of Bioethical Inquiry	Journal	2
Journal of Business Ethics	Journal	2
Lecture Notes in Informatics (LNI), Proceedings - Series of the Gesellschaft fur Informatik (GI)	Conference	2
Technology in Society	Journal	2
The ORBIT Journal	Journal	2
ACM Conference on Data and Application Security and Privacy	Conference	2
ACM Symposium on Access Control Models and Technologies	Symposium	2
Advanced Information Management, Communicates, Electronic and Automation Control Conference	Conference	2
Africa Week Conference	Conference	2
American journal of law & medicine	Journal	2
Americas Conference on Information Systems	Conference	2
Annals of the New York Academy of Sciences	Conference	2
Annual International Conference on Digital Government Research	Journal	2
Annual Privacy Forum	Conference	2
Bioethics	Journal	1
CEUR Workshop Proceedings	Conference	1
Cleveland State Law Review	Journal	1
Communications of the Association for Information Systems	Journal	1
Competition and Change	Journal	1
Computer Networks	Journal	1
Computers	Journal	1

Table 5 (continued)

Publication venue	Venue type	Number of papers
Computers & Security	Journal	1
Concurrency and Computation-Practice & Experience	Journal	1
Conference on Big Data Security on Cloud	Conference	1
Current Opinion in Systems Biology	Journal	1
Digital Enterprise Computing	Journal	1
Economy and Society	Journal	1
Electronic Markets	Journal	1
Energies	Journal	1
Epj Data Science	Journal	1
ERA Forum	Journal	1
European Conference on Information Systems	Conference	1
European Intelligence and Security Informatics Conference	Conference	1
European Journal of Health Law	Journal	1
European Journal of Human Genetics	Journal	1
Frontiers in Neuroinformatics	Journal	1
Future Generation Computer Systems	Journal	1
Hawaii International Conference on System Sciences	Forum	1
Health Policy and Technology	Journal	1
HeinOnline	Journal	1
IEEE Security and Privacy	Journal	1
IEEE Technology and Society Magazine	Journal	1
IET Conference Publications	Conference	1
Ifla Journal-International Federation of Library Associations	Journal	1

Table 5 (continued)

Publication venue	Venue type	Number of papers
Information (Switzerland)	Journal	1
Intelligent Systems and Computer Vision	Journal	1
International Conference "Quality Management, Transport and Information Security, Information Technologies	Conference	1
International Conference for Convergence in Technology	Conference	1
International Conference on Advanced Cloud and Big Data	Conference	1
International Conference on Application of Information and Communication Technologies (AICT)	Conference	1
International Conference on Automated Software Engineering	Conference	1
International Conference on Availability, Reliability and Security	Conference	1
International Conference on Big Data Analysis	Conference	1
International Conference on Big Data and Smart Computing	Conference	1
International Conference on Big Data Science and Computing	Conference	1
International Conference on Cloud and Green Computing	Conference	1
International Conference on Cloud Computing in Emerging Markets	Conference	1
International Conference on Collaboration Technologies and Systems	Conference	1
International Conference on Communication Technology	Conference	1
International Conference on Computational Intelligence and Communication Networks	Conference	1
International Conference on Computational Intelligence and Security	Conference	1
International Conference on Computational Science and Engineering	Conference	1
International Conference on Computer Engineering and Systems	Conference	1
International Conference on Computing for Sustainable Global Development	Conference	1
International Conference on Computing Sciences	Conference	1
International Conference on Computing, Communication & Automation	Conference	1
International Conference on Computing, Communication and Networking Technologies	Conference	1

Table 5 (continued)

Publication venue	Venue type	Number of papers
International Conference on Cyber Warfare and Security	Conference	1
International Conference on Data Management, Analytics and Innovation	Conference	1
International Conference on Dependable, Autonomic and Secure Computing	Conference	1
International Conference on Eco-friendly Computing and Communication Systems	Conference	1
International conference on Electronics, Communication and Aerospace Technology	Conference	1
International Conference on Energy, Communication, Data Analytics and Soft Computing	Conference	1
International Conference on Engineering and Technology	Conference	1
International Conference on e-Science	Conference	1
International Conference on Geoinformatics	Conference	1
International Conference on Global Security, Safety and Sustainability	Conference	1
International Conference on ICT for Smart Society	Conference	1
International Conference on Identification, Information and Knowledge in the Internet of Things	Conference	1
International Conference on Image Information Processing	Conference	1
International Conference on Information Communication and Embedded Systems	Conference	1
International Conference on Information Management and Processing	Conference	1
International Conference on Information Reuse and Integration	Conference	1
International Conference on Information Science and Applications	Conference	1
International Conference on Information Technology and Electronic Commerce	Conference	1
International Conference on Innovations in Information, Embedded and Communication Systems	Conference	1
International Conference On Internet of Things: Smart Innovation and Usages	Conference	1
International Conference on Inventive Communication and Computational Technologies	Conference	1
International Conference on IT Convergence and Security	Conference	1
International Conference on Machine Learning, Big Data, Cloud and Parallel Computing	Conference	1

Table 5 (continued)

Publication venue	Venue type	Number of papers
International Conference on Recent Trends in Advance Computing	Conference	1
International Conference on Reliability, Infocom Technologies and Optimization	Conference	1
International Conference on Research and Innovation in Information Systems	Conference	1
International Conference on Research in Computational Intelligence and Communication Networks	Conference	1
International Conference on Robots & Intelligent System	Conference	1
International Conference on Smart City/SocialCom/SustainCom together with DataCom 2015 and SC2 2015	Conference	1
International Conference on Smart Grid and Electrical Automation	Conference	1
International Conference on Software, Knowledge, Information Management and Applications	Conference	1
International Conference on Strategic Management and its Support by Information Systems	Conference	1
International Conference on System Sciences	Conference	1
International Conference on Systems, Man, and Cybernetics	Conference	1
International Conference On Trust, Security And Privacy In Computing And Communications	Conference	1
International Conference on Utility and Cloud Computing	Conference	1
International Conference on Virtual Reality and Intelligent Systems	Conference	1
International Congress on Advanced Applied Informatics	Congress	1
International Data Privacy Law	Journal	1
International Database Engineering & Applications Symposium	Symposium	1
International Enterprise Distributed Object Computing Conference	Conference	1
International Journal of Engineering, Transactions A: Basics	Journal	1
International Journal of Information Technology and Decision Making	Journal	1
International Journal of Innovative Technology and Exploring Engineering	Journal	1
International Journal of Interactive Mobile Technologies	Journal	1
International Journal of Law and Information Technology	Journal	1

Table 5 (continued)

Publication venue	Venue type	Number of papers
International Journal of Law and Psychiatry	Journal	1
International Journal of Managing Projects in Business	Journal	1
International Journal of Operations and Production Management	Journal	1
International Journal of Software Engineering and its Applications	Journal	1
International Symposium on Biometrics and Security Technologies	Symposium	1
International Symposium on Communications and Information Technologies	Symposium	1
International Symposium on Multimedia	Symposium	1
International Symposium on Networks, Computers and Communications	Symposium	1
International Workshop on BIG Data Software Engineering	Workshop	1
International Workshop on Privacy and Secuirty of Big Data	Workshop	1
ITU Kaleidoscope: Trust in the Information Society	Journal	1
Journal of Computers (Taiwan)	Journal	1
Journal of Data and Information Quality	Journal	1
Journal of Sport and Social Issues	Journal	1
Latin-American Symposium on Dependable Computing	Journal	1
Medical law review	Journal	1
Meditari Accountancy Research	Journal	1
Monash University Law Review	Journal	1
Multimedia Tools and Applications	Journal	1
NELLCO	Journal	1
Northwestern Journal of Technology and Intellectual Property	Journal	1
Politics Philosophy & Economics	Journal	1
Proceedings of Future Technologies Conference	Conference	1

Table 5 (continued)

Publication venue	Venue type	Number of papers
Ps-Political Science & Politics	Journal	1
Public Integrity	Journal	1
Quality - Access to Success	Journal	1
Religions	Journal	1
Risk Analysis	Journal	1
Science and Public Policy	Journal	1
Seton Hall Law Review	Journal	1
SIGCAS Comput. Soc.	Journal	1
Social Science Computer Review	Journal	1
Social Sciences in China	Journal	1
Sociological Research Online	Journal	1
Stanford Law Review Online	Journal	1
Sustainability	Journal	1
Symposium on Computers and Communication	Journal	1
Technological Forecasting and Social Change	Journal	1
Technology Innovation Management Review	Journal	1
Telecommunications Policy	Journal	1
The Lancet Oncology	Journal	1
Theology	Journal	1
Washington Law Review	Journal	1
Yale Law Journal	Journal	1

References

Ahlemann F, Stettiner E, Messerschmidt M, Legner C (2012) Strategic enterprise architecture management. Springer, Berlin. https://doi.org/10.1007/978-3-642-24223-6

Aiken P, Gorman M (2013) The case for the chief data officer: recasting the C-suite to leverage your most valuable asset. Elsevier, Burlington. https://doi.org/10.1016/C2012-0-02692-0

Akoka J, Comyn-Wattiau I, Laoufi N (2017) Research on Big Data—systematic mapping study. Comput Stand Inter 54:105–115. https://doi.org/10.1016/j.csi.2017.01.004

Aldea A, Iacob M-E, Wombacher A, Marlon H, Franck T (2018) Enterprise architecture 4.0—a vision, an approach and software tool support. In: 2018 IEEE 22nd international enterprise distributed object computing conference (EDOC), 2018. pp 1–10. https://doi.org/10.1109/EDOC.2018.00011

Alharthi A, Krotov V, Bowman M (2017) Addressing barriers to big data. Bus Horiz 60:285–292. https://doi.org/10.1016/j.bushor.2017.01.002

Amalina F, Targio Hashem IA, Azizul ZH, Fong AT, Firdaus A, Imran M, Anuar NB (2020) Blending big data analytics: review on challenges and a recent study. IEEE Access 8:3629–3645. https://doi.org/10.1109/access.2019.2923270

Anthony Byrd T, Lewis BR, Bryan RW (2006) The leveraging influence of strategic alignment on IT investment: an empirical examination. Inf Manage 43:308–321. https://doi.org/10.1016/j.im.2005.07.002

Basso T, Matsunaga R, Moraes R, Antunes N (2016) Challenges on anonymity, privacy, and big data. In: 2016 Seventh Latin-American symposium on dependable computing (LADC), 19–21 October. IEEE, pp 164–171. https://doi.org/10.1109/ladc.2016.34

BBC (2017) Equifax says almost 400,000 Britons hit in data breach. https://www.bbc.com/news/technology-41286638. Accessed 15 September 2017

Beckett P (2017) GDPR compliance: your tech department's next big opportunity, vol 2017. Elsevier, Amsterdam. https://doi.org/10.1016/S1361-3723(17)30041-6

Bernard S, Ho SM (2010) Enterprise architecture as context and method for designing and implementing information security and data privacy controls in Government Agencies. In: Saha P (ed) Advances in government enterprise rchitecture. IGI Global, pp 340–370. https://doi.org/10.4018/978-1-60566-068-4.ch015

Bertino E, Kundu A, Sura Z (2019) Data transparency with blockchain and AI ethics. ACM J Data Inf Qual 11. https://doi.org/10.1145/3312750

Bertot JC, Choi H (2013) Big Data and e-Government: issues, policies, and recommendations. In: 14th Annual international conference on digital government research, pp 1–10. https://doi.org/10.1145/2479724.2479730

Birnhack MD (2008) The EU data protection directive: an engine of a global regime. Comput Law Security Rev 24:508–520. https://doi.org/10.1016/j.clsr.2008.09.001

Blanco-Lainé G, Sottet JS, Dupuy-Chessa S (2020) Using an enterprise architecture model for GDPR compliance principles. In: 12th IFIP conference on practice of entreprise modeling (POEM), 2020. Springer International Publishing, Berlin, pp 199–214. https://doi.org/10.1007/978-3-030-35151-9_13

Borasi P, Khan S, Kumar V (2020) Europe Big Data and business analytics market size & share. https://www.alliedmarketresearch.com/europe-big-data-and-business-analytics-market-A06533. Accessed 05 June 2020

Bracy J (2017) The Equifax breach, response, and fallout. https://iapp.org/news/a/equifax-data-breach-affects-143-million-consumers/.

Burmeister F, Drews P, Schirmer I (2019) A privacy-driven enterprise architecture meta-model for supporting compliance with the general data protection regulation. In: 52nd Hawaii international conference on system sciences, pp 6052–6061. https://doi.org/10.24251/hicss.2019.729

Burmeister F, Huth D, Drews P, Schirmer I, Matthes F (2020) Enhancing information governance with enterprise architecture management: design principles derived from benefits and barriers in the GDPR implementation. In: 53rd Hawaii international conference on system sciences, pp 5593–5602. https://doi.org/10.24251/hicss.2020.688

Cagle K (2015) Understanding the Big Data life-cycle. https://www.linkedin.com/pulse/four-keys-big-data-life-cycle-kurt-cagle/. Accessed 10 February 2016

Campisi P, Maiorana E, Neri A (2009) Privacy protection in social media networks a dream that can come true? In: 2009 16th International conference on digital signal processing. IEEE, pp 1–5. https://doi.org/10.1109/icdsp.2009.5201125

Capgemini (2019) Championing data protection and privacy: a source of competitive advantage in the digital century. https://www.capgemini.com/be-en/research-reports/championing-data-protection-and-privacy/

Chao DF (2018) Sustainability for the Holistic ecosystem: regulation guide designing for the prevalent technology development of China emerging communications: Big Data, IoT, 5G etc. In: 2018 IEEE International symposium on innovation and entrepreneurship (TEMS-ISIE), 2018. IEEE, pp 1–8. https://doi.org/10.1109/TEMS-ISIE.2018.8478529

Charmaz K (2006) Constructing grounded theory. A practical guide through qualitative analysis. SAGE Publications Ltd, CA

Chen H, Yan Z (2016) Security and privacy in big data lifetime: a review. In: 2016 International Workshops, TrustData, TSP, NOPE, DependSys, BigDataSPT, and WCSSC (SpaCCS). Springer International Publishing, Berlin, pp 3–15. https://doi.org/10.1007/978-3-319-49145-5_1

Chibba M, Cavoukian A (2015) Privacy, consumer trust and big data: privacy by design and the 3 C'S. In: 2015 ITU Kaleidoscope: Trust in the Information Society (K-2015), 2015. pp 1–5. https://doi.org/10.1109/Kaleidoscope.2015.7383624

Crawford K, Schultz J (2014) Big data and due process: toward a framework to redress predictive privacy harms. BCL Rev 55:93

Crowd Research Partners (2018) GDPR compliance report. https://crowdresearchpartners.com/portfolio/gdpr-compliance-report/

CSC (2020) Cyberspace Solarium Commission White Paper #1: Cybersecurity Lessons From The Pandemic. Cyberspace Solarium Commission, https://www.solarium.gov/public-communications/pandemic-white-paper

Culnan MJ, McHugh PJ, Zubillaga JI (2010) How large U.S. companies can use twitter and other social media to gain business value. MIS Quart Executive 9:243–259

Cuquet M, Fensel A (2018) The societal impact of big data: a research roadmap for Europe. Technol Soc 54:74–86. https://doi.org/10.1016/j.techsoc.2018.03.005

DAMA (2017) DAMA-DMBOK: Data Management Body of Knowledge, 2nd edn. Dama International, WA

Deloitte (2020) Privacy and Data Protection in the age of COVID-19. https://www2.deloitte.com/be/en/pages/risk/articles/privacy-and-data-protection-in-the-age-of-covid-19.html

Demchenko Y, Laat Cd, Membrey P (2014) Defining architecture components of the Big Data Ecosystem. In: 2014 International conference on collaboration technologies and systems (CTS), 2014, pp 104–112. https://doi.org/10.1109/CTS.2014.6867550

EC (2000) Charter of Fundamental Rights of the European Union vol C. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:12012P/TXT

EC (2018) Ethics and data protection. https://ec.europa.eu/info/sites/info/files/5._h2020_ethics_and_data_protection_0.pdf

EDPS (2016) EDPS Opinion on coherent enforcement of fundamental rights in the age of big data. https://edps.europa.eu/data-protection/our-work/publications/opinions/big-data_en

EDPS (2018) Ethics: European Data Protection Supervisor. https://edps.europa.eu/data-protection/our-work/ethics_en. Accessed 10 March 2018

EESC (2017) The Ethics of Big Data: Balancing Economic Benefits and Ethical Questions of Big Data in the EU Policy Context. https://www.eesc.europa.eu/en/our-work/publications-other-work/publications/ethics-big-data

EP (1995) Directive 95/46/EC. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A31995L0046

Ethiraj SK, Levinthal D (2004) Bounded rationality and the search for organizational architecture: an evolutionary perspective on the design of organizations and their evolvability. Admin Sci Quart 49:404–437

EU (2016) (EU) 2016/679 GDPR. https://eur-lex.europa.eu/eli/reg/2016/679/oj

Favaretto M, De Clercq E, Elger BS (2019) Big Data and discrimination: perils, promises and solutions. A systematic review. J Big Data 6. https://doi.org/10.1186/s40537-019-0177-4

Fritsch L, Abie H (2008) Towards a research road map for the management of privacy risks in information systems. In: SICHERHEIT 2008 Gesellschaft für Informatik eV (GI).

Gahi Y, Guennoun M, Mouftah HT (2016) Big data nalytics: Security and privacy challenges. In: 2016 IEEE symposium on computers and communication (ISCC), Messina, 2016. IEEE, pp 952–957. https://doi.org/10.1109/ISCC.2016.7543859

Gandomi A, Haider M (2015) Beyond the hype: Big data concepts, methods, and analytics. Int J Inf Manage 35:137–144. https://doi.org/10.1016/j.ijinfomgt.2014.10.007

Gartner (2017) Gartner says organizations are unprepared for the 2018 European Data Protection Regulation. https://www.gartner.com/newsroom/id/3701117. Accessed 10 May 2017

Gellert R (2018) Understanding the notion of risk in the general data protection regulation. Comput Law Security Rev 34:279–288. https://doi.org/10.1016/j.clsr.2017.12.003

Gomez A (2017) Fundamental rights implications of big data: privacy, data protection, non-discrimination, security and law-enforcement. Committee on Civil Liberties, Justice and Home Affairs, https://www.europarl.europa.eu/doceo/document/A-8-2017-0044_EN.html

Gong Y, Janssen M (2017) Enterprise architectures for supporting the adoption of big data. In: Hinnant CC, Ojo A (eds) 2017–18th Annual international conference on digital government research, NY, 2017. ACM, New York, pp 505–510. https://doi.org/10.1145/3085228.3085275

Gong Y, Janssen M (2020) Roles and capabilities of enterprise architecture in big data analytics technology adoption and implementation. J Theor Appl Electronic Commerce Res 16:37–51. https://doi.org/10.4067/S0718-18762021000100104

Guggenheim (2016) Technological Innovation Portfolio , Series 11. https://www.guggenheiminvestments.com/uit/trust/atec011

Gürses S (2010) PETs and their users: a critical review of the potentials and limitations of the privacy as confidentiality paradigm. Identity Inf Soc 3:539–563. https://doi.org/10.1007/s12394-010-0073-8

Hijmans H, Raab CD (2018) Ethical dimensions of GDPR: commentary on the general data protection regulation. Edward Elgar Publishing, UK

Holm S, Ploug T (2017) Big Data and health research—the governance challenges in a mixed data. Econ Bioethical Inquiry 14:515–525. https://doi.org/10.1007/s11673-017-9810-0

Huth D, Tanakol A, Matthes F (2019) Using enterprise architecture models for creating the record of processing activities (Art. 30 GDPR). In: 2019 IEEE 23rd International enterprise distributed object computing conference (EDOC), France, 2019. IEEE, pp 98–104. https://doi.org/10.1109/EDOC.2019.00021

Huth D, Burmeister F, Matthes F, Schirmer I (2020) Empirical results on the collaboration between enterprise architecture and data protection management during the implementation of the GDPR. In: 2020 Hawaii international conference on system sciences (HICSS), 2020. https://doi.org/10.24251/HICSS.2020.715

Ibraimova A (2017) How does the GDPR apply to Big Data? http://www.iptechblog.com/2017/03/how-does-the-gdpr-apply-to-big-data/. Accessed 20 April 2018

IDC (2019) Worldwide Public Cloud Services Spending Forecast to Reach $160 Billion This Year, According to IDC. https://www.businesswire.com/news/home/20190228005137/en/Worldwide-Public-Cloud-Services-Spending-Forecast-Reach. 28 February 2019

ISACA (2013) COBIT 5: Enabling Information. www.isaca.org

ISACA (2018) GDPR: The end of the beginning. www.isaca.org

ITRC (2019) 2018 Annual data breach year-end review. https://www.idtheftcenter.org/2018-data-breaches/

Jalali S, Wohlin C (2012) Systematic literature studies: database searches vs . backward snowballing 2012 ACM-IEEE international symposium on empirical software engineering and measurement (ESEM), pp 29–38. https://doi.org/10.1145/2372251.2372257

Jensen M (2013) Challenges of privacy protection in Big Data analytics. In: 2013 IEEE International congress on Big Data, Santa Clara, CA, 2013. IEEE, pp 235–238. https://doi.org/10.1109/BigData.Congress.2013.39

Joo MH, Yoon SP, Kim SY, Kwon HY (2017) Research on distribution of responsibility for de-identification policy of personal information. In: 2017 18th Annual international conference on digital government research, 2017. pp 74–83. https://doi.org/10.1145/3085228.3085243

Jurkiewicz CL (2018) Big Data, big concerns: ethics in the digital age. Public Integrity 20:S46–S59. https://doi.org/10.1080/10999922.2018.1448218

Kaisler S, Armour F, Espinosa JA, Money W (2013) Big Data: issues and challenges moving forward. In: 2013 46th Hawaii international conference on system sciences, Wailea, Maui, HI, 2013. IEEE, pp 995–1004. https://doi.org/10.1109/hicss.2013.645

Kappelman L, McGinnis T, Pettite A, Sidorova A, McGinnis T (2008) Enterprise architecture: charting the territory for Academic Research. In: 2008 Americas Conference on Information Systems (AMCIS). Boca Raton, pp 96–110. https://doi.org/10.1201/9781439811146

Kearny C, Gerber A, Van Der Merwe A (2016) Data-driven enterprise architecture and the TOGAF ADM phases. In: 2016 IEEE International conference on systems, man, and cybernetics (SMC), 2016. IEEE, pp 4603–4608. https://doi.org/10.1109/SMC.2016.7844957

Kehrer S, Jugel D, Categorizing ZA (2016a) Requirements for Enterprise Architecture Management in Big Data Literature. In: (2016) IEEE 20th International Enterprise Distributed Object Computing Workshop (EDOCW), Vienna. IEEE. https://doi.org/10.1109/EDOCW.2016.7584352

Kehrer S, Jugel D, Zimmermann A (2016b) A systematic literature review of big data literature for EA evolution. 2016 Digital Enterprise Computing (DEC). Oblingen, Germany. LNI. Gesellschaft für Informatik e.V., pp 209–220

Kemp R (2014) Legal aspects of managing Big Data. Comput Law Security Rev 30:482–491. https://doi.org/10.1016/j.clsr.2014.07.006

Khajeh-Hosseini A, Sommerville I, Sriram I (2010) Research challenges for enterprise cloud. Comput J Internet Serv Appl. https://doi.org/10.1007/s13174-010-0007-6

King N, Brooks J, Tabari S (2018) Template analysis in business and management research. In: Qualitative methodologies in organization studies, vol II: Methods and possibilities. Palgrave Macmillan, Cham, pp 179–206. https://doi.org/10.1007/978-3-319-65442-3_8

Kitchenham B (2004) Procedures for performing systematic reviews. Keele University, UK 33:1–26

Kitchenham B, Budgen D, Brereton OP (2011) The value of mapping studies – a participant-observer case study. Inf Softw Technol 53:638–651. https://doi.org/10.1016/j.infsof.2010.12.011

Kitchenham B, Budgen D, Pearl Brereton O (2011) Using mapping studies as the basis for further research—a participant-observer case study. Inf Softw Technol 53:638–651. https://doi.org/10.1016/j.infsof.2010.12.011

Kitchenham B, Dyba T, Jorgensen M (2004) Evidence-based software engineering. In: 26th International conference on software engineering, Edinburgh, UK, 2004. pp 273–281. https://doi.org/10.1109/ICSE.2004.1317449

Kitchenham B, Pearl Brereton O, Budgen D, Turner M, Bailey J, Linkman S (2009) Systematic literature reviews in software engineering—a systematic literature review. Inf Softw Technol 51:7–15. https://doi.org/10.1016/j.infsof.2008.09.009

Kotusev S, Singh M, Storey I (2015) Consolidating enterprise architecture management research. In: 2015 48th Hawaii international conference on system sciences, Kauai, HI, 2015. IEEE, pp 4069–4078. https://doi.org/10.1109/HICSS.2015.489

Kshetri N (2014) Big data's impact on privacy, security and consumer welfare. Telecommun Policy 38:1134–1145. https://doi.org/10.1016/j.telpol.2014.10.002

Kuneva M (2009) Keynote Speech—Roundtable on online data collection, targeting and profiling. https://ec.europa.eu/commission/presscorner/detail/en/SPEECH_09_156. Accessed 11 May 2017

Kunz T, Selzer A, Waldmann U (2014) On the measurement of data protection compliance of cloud services. In: Gesellschaft fur Informatik (GI), 2014. GI-Jahrestagung, pp 289–296

Lagerström R, Sommestad T, Buschle M, Ekstedt M (2011) Enterprise architecture management's impact on information technology success. In: 2011 44th Hawaii International Conference On System Sciences, Kauai, HI, 2011. IEEE, pp 1–10. https://doi.org/10.1109/HICSS.2011.187

Lee I (2017) Big data: dimensions, evolution, impacts, and challenges. Bus Horizons 60:293–303. https://doi.org/10.1016/j.bushor.2017.01.004

Li P, Guo S (2014) Load balancing for privacy-preserving access to big data in cloud. In: 2014 IEEE Conference on computer communications workshops (INFOCOM WKSHPS), Toronto, ON, 2014. IEEE, pp 524–528. https://doi.org/10.1109/INFCOMW.2014.6849286

Liu Q, Srinivasan A, Hu J, Wang G (2017) Preface: security and privacy in big data clouds. Future Generation Comput Syst 72:206–207. https://doi.org/10.1016/j.future.2017.03.033

Lněnička M et al. (2017) Components of Big Data analytics for strategic management of enterprise architecture. In: 2017 12th International conference on strategic management and its support by information systems, 2017, pp 398–406

Löhe J, Legner C (2014) Overcoming implementation challenges in enterprise architecture management: a design theory for architecture-driven IT Management (ADRIMA). Inf Syst E-Bus Manage 12:101–137. https://doi.org/10.1007/s10257-012-0211-y

Lovejoy K, Saxton GD (2012) Information, community, and action: how nonprofit organizations use social media. J Comput-Mediated Commun 17:337–353. https://doi.org/10.1111/j.1083-6101.2012.01576.x

Mansfield-Devine S (2017) Data governance: going beyond compliance vol 2017. Elsevier Ltd. Amsterdam. https://doi.org/10.1016/S1361-3723(17)30052-0

Mantelero A (2014) The future of consumer data protection in the E.U. Re-thinking the "notice and consent" paradigm in the new era of predictive analytics. Comput Law Security Rev 30:643–660. https://doi.org/10.1016/j.clsr.2014.09.004

Mantelero A (2017) Regulating big data. The guidelines of the Council of Europe in the context of the European data protection framework. Comput Law Security Rev 33:584–602. https://doi.org/10.1016/j.clsr.2017.05.011

Manyika J, Chui M, Brown B, Bughin J, Dobbs R, Roxburgh C, Hung Byers A (2011) Big data: the next frontier for innovation , competition , and productivity. https://www.mckinsey.com/business-functions/mckinsey-digital/our-insights/big-data-the-next-frontier-for-innovation

McKInsey (2016) How companies are using big data and analytics. https://www.mckinsey.com/business-functions/mckinsey-analytics/our-insights/how-companies-are-using-big-data-and-analytics

McMahon A, Buyx A, Prainsack B (2020) Big Data governance needs more collective responsibility: the role of harm mitigation in the governance of data use in medicine and beyond. Med Law Rev 28:155–182. https://doi.org/10.1093/medlaw/fwz016

Mendelson H (2003) Organizational architecture and success in the information technology. Ind Manage Sci 46:513–529. https://doi.org/10.1287/mnsc.46.4.513.12060

Mikkelsen D, Soller H, Strandell-jansson M, Wahlers M (2019) GDPR compliance since May 2018: a continuing challenge. https://www.mckinsey.com/business-functions/risk/our-insights/gdpr-compliance-after-may-2018-a-continuing-challenge

Minerva R, Biru A, Rotondi D (2015) Towards a definition of the Internet of Things (IoT). https://iot.ieee.org/images/files/pdf/IEEE_IoT_Towards_Definition_Internet_of_Things_Revision1_27MAY15.pdf

Mourby M et al (2018) Are 'pseudonymised' data always personal data? Implications of the GDPR for administrative data research in the UK. Comput Law Security Rev 34:222–233. https://doi.org/10.1016/j.clsr.2018.01.002

Moyer L (2017) Equifax shares sink as Massachusetts prepares lawsuit over breach. https://www.cnbc.com/2017/09/13/equifax-shares-sink-9-percent-as-massachusetts-prepares-lawsuit-over-breach.html. Accessed 20 August 2018

Murmann P, Fischer-Hübner S (2017) Tools for achieving usable ex post transparency: a survey. IEEE. Access 5:22965–22991. https://doi.org/10.1109/ACCESS.2017.2765539

Nadler DA, Gerstein MS, Shaw RB (1992) Organizational architecture: designs for changing organizations. Jossey-Bass Inc Pub, San Francisco

Newman D, Logan D (2006) Governance is an essential building block for enterprise information management. https://www.gartner.com/en/documents/492444/governance-is-an-essential-building-block-for-enterprise-

NIST (2013) Security and privacy controls for Federal information systems and organizations, vol 800. https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-53r4.pdf. https://doi.org/10.6028/NIST.SP.800-53r4

Olaitan O, Herselman M, Wayi N (2016) Taxonomy of literature to justify data governance as a prerequisite for information governance. In: 28th Annual conference of the Southern African Institute of Management Scientists (SAIMS), Pretoria, South Africa, 2016, pp 586–605

Özköse H, Arı ES, Gencer C (2015) Yesterday, today and tomorrow of Big Data. Procedia Soc Behav Sci 195:1042–1050. https://doi.org/10.1016/j.sbspro.2015.06.147

Parnet O, Benoit K, Nulty P, Theocharis Y, Popa SA (2016) Social media and political communication in the 2014 elections to the European Parliament. Electoral Stud 44:429–444. https://doi.org/10.1016/j.electstud.2016.04.014

Patil HK, Seshadri R, (2014) Big Data security and privacy issues in healthcare. In: (2014) IEEE International congress on Big Data, Anchorage, AK, 2014. IEEE. https://doi.org/10.1109/BigData.Congress.2014.112

Pearson S, Benameur A (2010) Privacy, security and trust issues arising from cloud computing. In: 2010 IEEE second international conference on cloud computing technology and science, Indianapolis, IN, 2010. IEEE, pp 693–702. https://doi.org/10.1109/CloudCom.2010.66

Petersen K, Feldt R, Mujtaba S, Mattson M (2008) Systematic mapping studies in software engineering. In: 2008 12th International conference on evaluation and assessment in software engineering (EASE), Italy, 2008. BCS Learning & Development Ltd, pp 68–77. https://doi.org/10.1142/S0218194007003112

Petersen K, Vakkalanka S, Kuzniarz L (2015) Guidelines for conducting systematic mapping studies in software engineering: an update. Inf Softw Technol 64:1–18. https://doi.org/10.1016/j.infsof.2015.03.007

Ponemon Institute LC (2018) 2018 Cost of data breach study, global overview. https://www.illusivenetworks.com/

PrivacyRightsClearinghouse (2017) Data breaches. https://www.privacyrights.org/data-breaches

Radeke F (2010) Awaiting explanation in the field of enterprise architecture management. In, (2010) Americas conference on information systems (AMCIS). Lima, Peru 2010:442

Radeke F (2011) Toward understanding enterprise architecture management's role in strategic change: antecedents, processes, outcomes. In: 10th International conference on Wirtschaftsinformatik, WI, 2011, p 62

Robinson N, Graux H, Botterman M, Valeri L (2009) Review of the European data protection directive. https://ico.org.uk/media/about-the-ico/documents/1042349/review-of-eu-dp-directive.pdf

Rozehnal P, Novak V (2018) The core of enterprise architecture as a management tool: GDPR implementation case study 2018 26th interdisciplinary information management talks, pp 359–366

Rubinstein I (2012) Big data: the end of privacy or a new beginning? Int Data Privacy Law. https://doi.org/10.2139/ssrn.2157659

Rubinstein I, Good N (2013) Privacy by design: a counterfactual analysis of Google and Facebook privacy incidents. Berkeley Technol Law J 28:1333–1413. https://doi.org/10.2139/ssrn.2128146

Salleh KA, Janczewski L (2016) Technological, organizational and environmental security and privacy issues of Big Data: a literature review. Proc Comput Sci 100:19–28. https://doi.org/10.1016/j.procs.2016.09.119

Saltz JS, Dewar N (2019) Data science ethical considerations: a systematic literature review and proposed project framework. Ethics Inf Technol 21:197–208. https://doi.org/10.1007/s10676-019-09502-5

Sampson F (2014) 7 Rights of Individuation: the need for greater protection of individual rights in Big Data. Paper presented at the 2014 IEEE/ACM 7th International Conference on Utility and Cloud Computing, London

Sauer C, Willcocks L (2003) Establishing the business of the future: the role of organizational architecture and information technologies. Euro Manage J 21:497–508. https://doi.org/10.1016/S0263-2373(03)00078-1

Sauer C, Willcocks L (2004) Strategic alignment revisited: connecting organizational architecture and IT infrastructure. Paper presented at the 2004 37th Hawaii International Conference on System Sciences, Big Island, HI

Shirer M (2016) Double-digit growth forecast for the Worldwide big data and business analytics market through 2020 led by banking and manufacturing investments, according to IDC. https://www.businesswire.com/news/home/20161003005030/en/Double-Digit-Growth-Forecast-for-the-Worldwide-Big-Data-and-Business-Analytics-Market-Through-2020-Led-by-Banking-and-Manufacturing-Investments-According-to-IDC.

Shirer M, Goepfert J (2019) IDC Forecasts revenues for Big Data and business analytics solutions will reach $189.1 Billion This Year with Double-Digit Annual Growth Through 2022. https://www.businesswire.com/news/home/20190404005662/en/IDC-Forecasts-Revenues-for-Big-Data-and-Business-Analytics-Solutions-Will-Reach-189.1-Billion-This-Year-with-Double-Digit-Annual-Growth-Through-2022#:~:text=By%202022%2C%20IDC%20expects%20worldwide,Data%20and%20Analytics%20Spending%20Guide. Accessed 05 April 2019

Simon D, Fischbach K, Schoder D (2014) Enterprise architecture management and its role in corporate strategic management. Inf Syst E-Bus Manage 12:5–42. https://doi.org/10.1007/s10257-013-0213-4

Smith M, Szongott C, Henne B, Von Voigt G (2012) Big data privacy issues in public social media. In: 6th IEEE International conference on digital ecosystems and technologies (DEST), Campione d'Italia, 2012. IEEE. https://doi.org/10.1109/DEST.2012.6227909

Solove DJ (2002) Conceptualizing privacy. California Law Rev 90:1087–1155. https://doi.org/10.2307/3481326

Spiekermann S (2012) The challenges of privacy by design, vol 55. ACM, New York. https://doi.org/10.1145/2209249.2209263

Spiekermann S, Acquisti A, Böhme R, Hui KL (2015) The challenges of personal data markets and privacy. Electron Markets 25:161–167. https://doi.org/10.1007/s12525-015-0191-0

Suresh J (2014) Bird's eye view on big data management. In: 2014 Conference on IT in business, industry and government (CSIBIG), Indore, 2014. IEEE, pp 1–5. https://doi.org/10.1109/CSIBIG.2014.7056930

Tan Q, Pivot F (2015) Big Data privacy: changing perception of privacy. In: 2015 IEEE International conference on smart city/SocialCom/SustainCom (SmartCity), Chengdu, 2015. IEEE, pp 860–865. https://doi.org/10.1109/SmartCity.2015.176

Tene O, Polonetsky J (2011) Privacy in the age of big data: a time for big decisions. Stanford Law Rev Online 64:63–69

Tene O, Polonetsky J (2012) Big Data for all: privacy and user control in the age of analytics. Northwestern J Technol Intellectual Property, 11

Thompson N, Ravindran R, Nicosia S (2015) Government data does not mean data governance: lessons learned from a public sector application audit. Govern Inf Quart 32:316–322. https://doi.org/10.1016/j.giq.2015.05.001

Thuraisingham B (2015) Big Data Security and Privacy. Paper presented at the 2015 5th ACM Conference on Data and Application Security and Privacy (CODASPY), San Antonio, Texas, USA,

Tse D, Chow C, Ly T, Tong C, Tam K (2018) The challenges of Big Data governance in healthcare. In: 2018 17th IEEE International conference on trust, security and privacy in computing and communications/12th IEEE international conference on big data science and engineering (TrustCom/BigDataSE), New York, NY, 2018. IEEE, pp 1632–1636. https://doi.org/10.1109/TrustCom/BigDataSE.2018.00240

UN (1948) Universal declaration of human rights. https://www.un.org/en/universal-declaration-human-rights/

van der Sloot B (2015) How to assess privacy violations in the age of Big Data? Analysing the three different tests developed by the ECtHR and adding for a fourth one. Inf Commun Technol Law 24:74–103. https://doi.org/10.1080/13600834.2015.1009714

Vanauer M, Böhle C, Hellingrath B (2015) Guiding the introduction of big data in organizations: a methodology with business- and data-driven ideation and enterprise architecture management-based implementation. In: 2015 48th Hawaii international conference on system sciences, Kauai, HI, 2015. IEEE, pp 908–917. https://doi.org/10.1109/HICSS.2015.113

Vanhoorelbeke F, Snoeck M, Serral E (2020) Identifying the challenges and requirements of enterprise architecture frameworks for IoT systems. Research Challenges in Information Science, pp 576–581. https://doi.org/10.1007/978-3-030-50316-1_41

Wagter R, Van den Berg M, Luijpers J, Van Steenbergen M (2005) Dynamic enterprise architecture how to make it work. Cambridge University Press, Cambridge. https://doi.org/10.1017/CBO9781107415324.004

Westin AF (1968) Privacy and freedom. Washington Lee Law Rev 56:911–914. https://doi.org/10.2307/3479272

Wieringa R, Maiden N, Mead N, Rolland C (2006) Requirements engineering paper classification and evaluation criteria: a proposal and a discussion. Requirements Eng 11:102–107. https://doi.org/10.1007/s00766-005-0021-6

Willemsen B, Bhajanka P (2017) The Four Do's and Don'ts of implementing your privacy program. https://www.onetrust.com/blog/gartner-report-four-dos-donts-implementing-privacy-program/. Accessed 26 August 2018

Winter R, Schelp J (2008) Enterprise architecture governance: the need for a business-to-IT Approach, pp 548–552. https://doi.org/10.1145/1363686.1363820

Wißotzki M, Koç H, Weichert T, Sandkuhl K (2013) Development of an enterprise architecture management capability catalog. In: Kobyliński A, Sobczak A (eds) Perspectives in business informatics research (BIR). Springer, Berlin, pp 112–126. https://doi.org/10.1007/978-3-642-40823-6_10

Wohlin C, Runeson P, da Mota Silveira Neto PA, Engström E, do Carmo Machado I, de Almeida ES, (2013) On the reliability of mapping studies in software engineering. J Syst Softw 86:2594–2610. https://doi.org/10.1016/j.jss.2013.04.076

Wong R (2011) Data protection: the future of privacy. Comput Law Security Rev 27:53–57. https://doi.org/10.1016/j.clsr.2010.11.004

Yu S (2016) Big privacy: challenges and opportunities of privacy study in the age of Big Data. IEEE Access 4:2751–2763. https://doi.org/10.1109/ACCESS.2016.2577036

Affiliations

Georgios Georgiadis[1] · **Geert Poels**[1]

✉ Georgios Georgiadis
Georgios.Georgiadis@UGent.be

Geert Poels
Geert.Poels@UGent.be

[1] Faculty of Economics and Business Administration, Ghent University, Tweekerkenstraat 2, 9000 Gent, Belgium