

On the Data Security of Information Systems: Comparison of Approaches and Challenges

Farah Abdmeziem[1(✉)], Saida Boukhedouma[1], and Mourad Chabane Oussalah[2]

[1] LSI Laboratory, University of Science and Technology Houari Boumediene, Algiers, Algeria
{fabdmeziem,sboukhedouma}@usthb.dz
[2] LS2N Laboratory, University of Nantes, Nantes, France
Mourad.oussalah@univ-nantes.fr

Abstract. Information Systems (IS) are in constant evolution to cope with new technologies like cloud computing and Big data. Nevertheless, these new technologies bring several security issues, some caused by the numerous access points which make companies prone to an increasing number of attacks. The architecture of an information system can be structured in three main layers: data layer, application layer, and technology layer. Security must be reinforced in each of the previous layers to prevent internal and external attacks. In this paper, we focus on the security of the data layer, thus we survey the research work dealing with security related to this layer. For that, we propose a data meta-model with a high level of abstraction that shows the relationship between data, security issues, and security mechanisms covering the three aspects of security (confidentiality, integrity, availability) and considering different types of data and novel technologies (big data and cloud computing). We use the main concepts of this meta-model to classify security mechanisms according to security goals, to compare the different approaches proposed in the literature, and to identify open challenges in the data security field. Our data meta-model can be easily extended and serves as a reference to compare other approaches in the literature.

Keywords: Information system architecture · Security · Data layer · Security mechanisms · Security issues · Meta-model

1 Introduction

Information System (IS) is an essential asset of companies. Its main objectives consist in collecting, storing, and sharing data between the stakeholders of the companies. Nowadays, information systems are becoming more and more collaborative and dynamic to meet business needs. New technologies like Cloud

© Springer Nature Switzerland AG 2021
O. Gervasi et al. (Eds.): ICCSA 2021, LNCS 12951, pp. 240–255, 2021.
https://doi.org/10.1007/978-3-030-86970-0_18

Computing, Internet of Things (IoT), Big data, and Blockchain foster the collaborative work of companies where data and business processes are exchanged and shared between stakeholders of the same company or other companies.

Moreover, information System Architecture (ISA) attracts the interest of a lot of researchers and several works have been proposed. The most known frameworks for ISA design are: *the ZACHMAN framework* [44], *the TOGAF*[1], and *the DODAF*[2] frameworks. Nevertheless, and despite the researchers' efforts, ISA lacks a reference model. Overall, ISA can be divided into three levels [41]: the **data level** that supports business, the **application level** where the applications needed for data management and business support are defined and the **technological level** that represents the main technologies used for IS implementation. It is worth noting that these three layers are found, among others, in the ISA frameworks mentioned above. Furthermore, security is one of the most important non-functional requirements of information systems and in particular collaborative information systems. Specifically, it is about securing the data and the processes handling this data. Each layer of the ISA can be prone to security attacks, such as the data layer where we can count several known attacks including *SQL injection* [5] and *Man in the Middle* [21]. These attacks are not without consequences on the application layer too as these two layers are strongly linked and this can induce considerable losses for companies. It is becoming then crucial to secure the companies' information systems from the design phase onwards. In this paper we dig into the data layer of ISA by proposing a meta-model that gives our perception of IS' data security. The meta-model highlights the main concepts attached to data, security issues, and security mechanisms. The proposed meta-model is used to classify security mechanisms according to security goals; and also to compare the research work surveyed in this paper, according to a set of identified criteria. Notice that our meta-model can be easily used/extended to consider other approaches not mentioned in this survey. The remainder of this paper is structured as follows: Sect. 2 introduces our global ISA and security meta-model. Section 3 presents a generic data meta-model followed by the most known data security issues and security mechanisms. In the same Section, we present a classification of security mechanisms relative to security goals and finally the data security meta-model. In Sect. 4, we use our data security meta-model to compare research works discussed in Sect. 3 and identify open challenges in data security. Finally, Sect. 5 concludes the paper and gives our future work directions.

2 Information System and Security

Starting from the global Information System Architecture (ISA), we present in Fig. 1 a meta-model that shows the link between information system layers and security. The meta-model depicts the three main layers of an ISA that consist of

[1] The Open Group Architecture Framework. https://www.opengroup.org/togaf.

[2] Federal Enterprise Architecture Framework. https://dodcio.defense.gov/Library/DoD-Architecture-Framework/.

(i) the data layer containing data used/produced by business applications, (ii) the application layer where the applications for business support are defined, (iii) the technological layer that represents the main technologies used in application's implementation, data storage/access and the infrastructures that provide an environment for IS deployment (for instance, building services using web services technology or Cloud Computing). Information systems can carry some weaknesses and security vulnerabilities which make them prone to attacks and security issues like unauthorized accesses and data loss. These security issues compromise the security of the three layers of the IS. Security is defined through three main concepts called the CIA (Confidentiality, Integrity, Availability):

- Confidentiality: is described as the assurance that the IS assets are kept secret;
- Integrity: refers to the inability to alter or destroy IS assets by accident or malfeasance;
- Availability is the ability to access the IS assets whenever it is needed.

Security mechanisms supported by tools or frameworks are proposed to mitigate IS security issues and hence ensure security. However, these solutions have an impact on the IS non-functional properties; this will be detailed in Sect. 3.5.

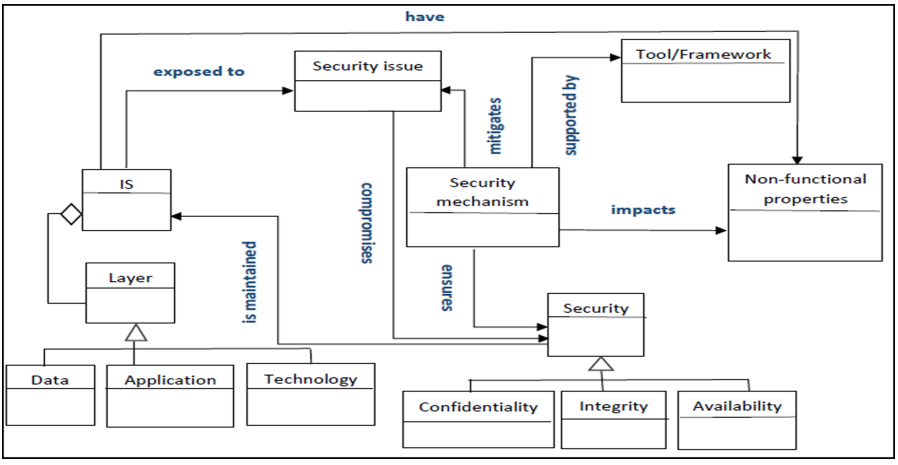


Fig. 1. IS and security

3 Data Security

Data is at the heart of companies and represents their key asset. Indeed, among other benefits, it helps companies to make better decisions and improve their processes. However, data is also the main target of cyber-attacks that cause considerable damages to companies. In this section, we begin by presenting our data meta-model followed by a brief description of the security issues that the

data layer could be prone to. Then, we give an overview of the proposed security mechanisms in the literature to ensure data security and their impact on the IS non-functional properties. Finally, we present our data security meta-model.

3.1 Data Meta-model

Figure 2 depicts the main concepts associated with the data layer. Data can be of different types mainly structured, semi-structured or unstructured. It is stored, processed and, accessed via databases that can be SQL or NoSQL databases. These databases can be located on premise (on the organisation's site) or in a cloud environment. When data with its different types is with high volume and velocity, we talk about Big data. Data with its different types can be in two states: *at rest* when it is stored in the databases or *in motion* when it is travelling across a network.

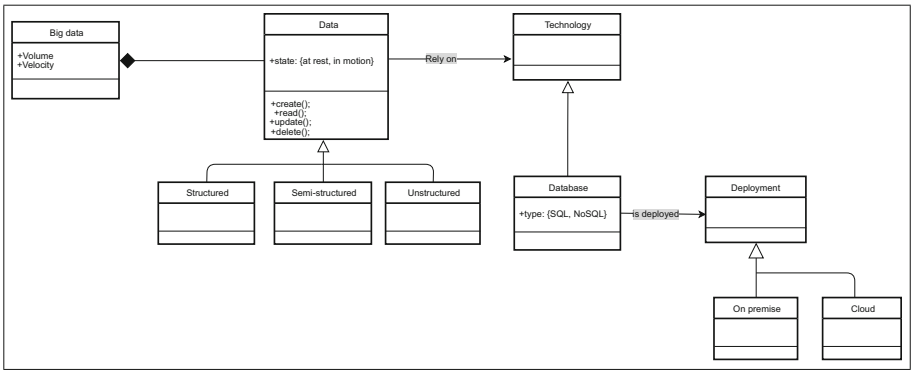


Fig. 2. Data meta-model

3.2 Data Security Issues

Data is exposed to security issues like unauthorised access, tempering of sensitive data, leakage, and data breach. Some of the security issues are carried with the used technologies like cloud computing. Those security issues compromise the confidentiality, integrity, and availability of data and come with an immense cost to companies. Thus it is becoming urgent to ensure data security by proposing a set of data security mechanisms.

3.3 Data Security Mechanisms: Overview of the Literature

Many researchers have worked on the issue of data security; in this Section, we provide an overview of the proposed solutions. We choose to review the well-known models on data security and the works that deals with the cloud computing, blockchain and big data technologies.

Access Control and Encryption Mechanisms: Access control and encryption are well-known and used techniques to ensure data confidentiality and integrity. Encryption [34], on one hand, involves transforming data to make it unreadable to those who don't have the key to decrypt it. Access control, on the other hand, is the decision to permit or deny a subject access to system objects [12]. Objects in our work scope refer to data.

Access Control (AC). Is achieved through models that provide a framework and a set of conditions on which objects, subjects, operations, and rules can be combined to generate and enforce an access control decision. AC can be static or dynamic i.e. whether the access control mechanism uses contextual information or is adjustable and adaptable to environmental changes. The most known access control models are Role-Based Access Control (RBAC) [8] and Attribute-Based Access Control (ABAC). Overall, RBAC employs pre-defined roles that carry a specific set of privileges associated with them and to which subjects are assigned. ABAC controls access to objects by evaluating rules against the attributes of entities (subject and object), operations, and the environment relevant to a request [13]. However, these mechanisms are not well suited for highly dynamic contexts like cloud computing or to deal with big data characteristics (the 3V). Hence, new access control mechanisms have been proposed. In [45] authors suggest Content-Based Access Control (CBAC) model, in which access control decisions are made using data content. More precisely CBAC makes access control decisions based on the content similarity between user credentials and data content dynamically. In [14] the authors propose a methodology with six phases for deploying reliable and efficient access control policies. The proposed solution aims to maintain coherence and compliance between the high-level definition of an access control policy and its low-level definition. To do so, the authors use a formal verification of the specified access control policy before proceeding to its implementation and a reverse engineering method to extract the enforced policy. On another note, [16] propose a generic and dynamic architecture for an access control meta-model that is composed of 6 levels. The proposed solution enables to derive different access control models and to combine them to create hybrid models. Last but not least, *the Bell-LaPadula (BLP) model*, a state machine-based multilevel security [3]. This model is based on subjects (users) and labelled objects from the most sensitive to the least sensitive. A user is restricted from reading up an object with higher label and writing down in object with lower label, it is the No Read Up, No Write Down principle.

Encryption Mechanism. Is the process of encoding information in a way that only authorized parties can access it. This mechanism is used for example in [4], where the authors propose to split the data into several parts and store each part on different cloud providers. Furthermore, the authors point out the fact that big data owners need to consider the cost of encryption in term of money and time. Hence, to deal with this issue they suggest to encrypt only storage path of big data. Finally, to improve availability, the proposed scheme stores more copies of the data parts and their index information on the cloud provider

storage. To achieve a secure data transmission between users and cloud storage services, the authors in [18] propose an approach based on the combination of asymmetric (RSA) and symmetric (AES) encryption techniques. The symmetric technique was used to send files from user to the cloud service storage and the asymmetric technique was used to retrieve files from the cloud service storage. The idea behind this approach greatly strengthen the security of file transmission in cloud environment. However, it should be noticed that this technique poses a problem concerning the number of the generated keys for each file stored in the cloud, a number that will keep increasing in context of big data. In the same vein, Sandeep [36] propose an approach that combines three security parameters: data classification, index and encryption, MAC (Message Authentication Code) and authentication techniques to secure data in motion and at rest in a cloud environment. The solution consists of, first to divide data into three sections according to a sensitivity rating value, afterwards, data owner defines keywords to data and their corresponding indexes. Next, data, keywords and indexes are encrypted using the 128-bit SSL encryption algorithm and sent to the cloud. An authentication mechanism is also defined to access data stored in the cloud. The combination of security techniques as proposed in this approach can give successful results in regard to data security, however in case of massive amount of data, index management can become a cumbersome task and will cause performance problems. *Attribute-Based Encryption* (ABE) introduced by Cloud security Alliance (CSA) is another technique that consists on encrypting data as part of an access policy, so that only the user who has permission to access the data can decrypt it [31]. However, according to [43], policy updating poses heavy computation burden on data owner because this latter has to decrypt the data then re-encrypt it again according to the new access policy before sending it to the cloud. Hence the authors propose a dynamic policy updating.

Data Anonymization Mechanism: Data anonymization and data encryption both serve the purpose of securing data that is in motion or stored and can, sometimes, be considered as the same technique. However, as mentioned in [34], encryption can be a useful tool for anonymization, but it is neither necessary nor sufficient for anonymization, and encrypted data is not necessarily anonymized. Data anonymization is the process in which changes are made to data that will be used or in a way that will prevent the identification of key information [15]. The authors of [24] discussed and compared five anonymization techniques: suppression, generalization, swapping, masking, and distortion. The pros and cons of each technique have been discussed and the suppression technique that is "removing an entire part of data by changing the value to one value that doesn't have meaning [24]" was found as the most efficient one [24]. The challenge with anonymization is, first, to find the most suitable technique to apply to sensitive data as this latter depends on many other criteria [24] like the type of data and the appropriate anonymization level. Second, it is important to know when to stop anonymizing data to balance between

privacy and data utility [6]. K-anonymity is a key concept of anonymization; it was introduced in [40]. A dataset containing sensitive information is transformed using K-anonymity in a way that guarantees protection against identity disclosure. However, to be efficient, this technique requires some conditions and is vulnerable to attacks like the background knowledge attack and the homogeneity attack. Furthermore, authors in [34] conduct a Proof Of Concept (POC) to show that anonymization is a viable technique to secure data stored in the public cloud. The POC principle is storing anonymized logs in public cloud-based Software as a Service (SaaS) application and then analysing logs to gather security and performance data. Virtual Machines (VMs) send the anonymized data and the de-anonymization takes place in a secure enclave. Successful results have been obtained on the fact that anonymization enhances public cloud security. [11] propose an efficient data anonymization scheme where anonymization and deanonymization are performed by a secure enclave rather than at the client's premises so that computational power is saved. The main drawback of the proposed solution is that it depends on the security of the enclave entity. In [26], Prakash et al. propose a personalized anonymization approach which consists of a top-down greedy algorithm that splits the data into partitions according to a chosen value. Then, a second algorithm is proposed to check if the partitioning violates the privacy requirement. Experimental results have shown that the proposed personalized approach gives better efficiency than other anonymization techniques like k-anonymity.

Mirroring, Checksumming, and Digital Signature: Mirroring, Checksumming, and Digital signature are well-known techniques to ensure data integrity. Mirroring [35] principle is to create two or more copies of data and the integrity checks can be made by comparing the copies. Checksumming is another technique based on checksums (also called hashes) computed using hash functions including MD5 and SHA1. Message Authentication Code (MAC) [20] is an application of the checksumming technique which involves generating cryptographic checksum on data to ensure that data has not been modified while in transit. Digital signature [17] is a mechanism that employs encryption, hashing, and digital signature algorithms to guarantee data integrity. In [25] the authors propose data integrity as a service (DIaaS) to address the integrity issue in cloud storage services. The proposed solution is based on the checksumming technique. Sandeep in [36] proposes a general framework to secure data in the cloud. One of the framework's phases is to generate a Message Authentication Code (MAC) to ensure data integrity. On another note, researchers investigate blockchain technology to ensure data integrity. Indeed, blockchain is a public ledger that works like a log by keeping a record of all transactions in chronological order, secured by an appropriate consensus mechanism and providing an immutable record. Its exceptional characteristics include immutability, irreversibility, decentralization, persistence and anonymity [28]. The integrity and persistence of the blocks in the blockchain are ensured using the hash principle and cryptographic signature as well as a consensual protocol. These characteristics make the blockchain

technology an appropriate tool to guarantee data integrity. In [9] and [46] authors tackle the data integrity issue and use the blockchain characteristics to propose a blockchain-based database that ensures data integrity.

Data Backup and Recovery: Regarding the third concept of security: availability, backup plans and recovery mechanisms are proposed. Data backups and recovery mechanisms refer to all the activities and policies implemented to recover resources after an outage event [38] (in the scope of this work we are interested only in data recovery). This outage event can be due to a malign human intervention or a catastrophic event [1]. One of the most known techniques used to backup data are *Full backup, incremental backup and data de-duplication. Full backup* consists of making a copy of all data to a storage device. Such backup has to be made while a database is off-line and unavailable to its users, and would take more time and costs as data become larger [29]. To overcome the disadvantages of *Full backup*, the *Incremental backup* have been proposed and which consist of making copies of only files which have changed or are new since a prior backup [29]. With regard to the last technique: *Data de-duplication* [37,42] is used to find duplicate data and store a sole copy, then uses a pointer to replace the rest of duplicate data [39]. De-duplication concept reduces storage needs, costs and the amount of energy and processing power used to run the storage system [10]. Rahumed et al. in [30] propose *FadeVersion*, a secure cloud backup system based on version control technique and provide the property of assured deletion. The version control technique is based on the concept of de-duplication and the assured deletion property is guaranteed employing two-layer encryption. In [27] the authors present a simple recovery service from disaster based on multiple cloud service providers. Backing up data in different sites in the cloud gives maximum availability and greater reliability at the expense of cost and resource consumption. Sambrani et al. propose in [32] a seed block algorithm that aims to provide a secure backup and recovery process in a cloud environment. Experiments was conducted to show the ability of the proposed solution in maintaining the data without any lost during the recovering process.

3.4 What Security Mechanism for What Security Goal?

Based on the security mechanisms surveyed in Sect. 3.3, we propose Table 1 that classifies the security mechanisms according to security concepts (CIA) they tackle and data type (structured and big data; we consider that semi-structured and unstructured types are covered in big data). This work allowed us to come up with the following conclusions: (i) Confidentiality is the most addressed concept by the proposed security mechanisms despite the fact that data integrity and availability are equally important. (ii) We were not able to find a security mechanism that satisfies the three concepts of security all at once. (iii) Moreover, Big data security mechanisms lack standard models unlike the security mechanism for structured data like RBAC, CBAC... (4i) Finally, the same security mechanisms could be applied for structured data and big data (like

Table 1. Security mechanisms for security goals

Data type/security goal	Confidentiality	Integrity	Availability
Structured data	- RBAC, - CBAC, -BLP, - ABE, - RSA, - AES, -Data classification, - Indexes, - K-anonymity, - Generalization, - Suppression, - Swapping, - Masking, - Distortion	- MAC, - DIaaS - Mirroring, - Digital signature, - Blockchain-based databases	- Full backup, - Incremental backup, de- duplication, - FadeVersion, - Backing up data in different sites in the cloud.
Big data	- CBAC, - Dynamic policy updating for ABE, - A top-down greedy algorithm	- MAC, - Mirroring, - Digital signature, - DIaaS	- Incremental backup, de- duplication, - Backing up data in different sites in the cloud

the mirroring technique to ensure data integrity) however, a security mechanism that suits structured data could impact the IS non-functional properties in case of big data (see Sect. 3.5).

3.5 Impact of Security Mechanisms on Non-functional IS Properties

One of the important criteria to consider when implementing security mechanisms is the impact of these solutions on IS non-functional properties. In our work, we consider four IS non-functional properties, namely: the management of the security solution, data transmission time, monetary cost of the solution, and the computational burden caused by the solution. Indeed, some of the proposed security solutions can be hard to manage/maintain and have negative impacts on the information system's performances causing heavy computation burden or extending the response time for data access requests, especially in the case of big data. Moreover, it is also important to evaluate the security mechanism with regard to the cost price.

3.6 Data Security Meta-model

Based on the security approaches surveyed in Sect. 3.3 and the work done in Sect. 3.4 and Sect. 3.5 we propose in Fig. 3 below, our data security meta-model which is made of three main components: Data that was detailed via our data meta-model in Sect. 2, security with its three main goals (CIA) and finally the proposed security mechanisms and their impact on the IS non-functional properties. The main objectives of our meta-model are as follow:

- Identify the security techniques used to ensure security goals/needs (CIA);
- Serves as a referential to compare data security mechanisms according to data type, security goal, and the impact of the solution on the IS non-functional properties.
- Easy to extend with other existing security mechanisms and non-functional properties;
- Use the meta-model to link between data type, security goal, and security mechanisms used to achieve this goal. This will help to find research directions not yet covered in the literature. In addition to these points, our meta-model is not bound to a single technology as is the case of [7,19,23] where the proposed meta-model is closely linked to a given technology.

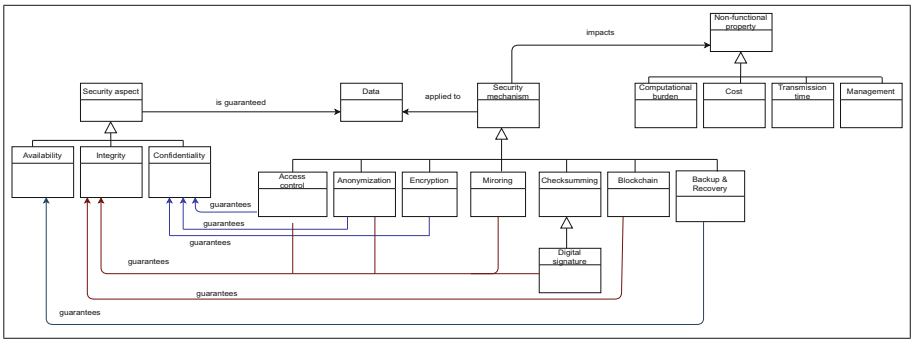


Fig. 3. Data security meta-model

4 Comparison of Approaches and Open Questions

In this section, we use our data security meta-model to compare the approaches discussed in Sect. 3.3. For this purpose, we distinguish the three security goals (CIA) and present three tables: Table 2, Table 3, and Table 4 that tackle respectively data confidentiality, data integrity, and data availability. We criticise the approaches in the drawback column in each table according to the non-functional properties presented in the meta-model. Lastly, regarding the tool/framework column we are interested in whether the paper in question gives implementation details about the proposed approach or proposes the solution only as a conceptual framework. The tables' cells with the "-" mark mean that we were unable to find any information in the paper to provide a conclusion or appreciation about the element.

Table 2. Security mechanisms for data confidentiality

Approach	Data type	Security mechanism	Technology	Tool/Framework	Drawback
[8]	Structured at rest	Static access control	On premise	Framework	- Cause "role explosion" and become hard to manage and maintain
[13]	Structured at rest	Dynamic access control	On premise	Framework	- Efforts to define attributes - Cause "attributes explosion" and become hard to manage
[45]	Big data at rest	Dynamic access control	On premise	Tool	- The experiments show that CBAC generates overhead
[14]	Structured, semi-structured and unstructured data at rest	Dynamic access control	On premise	Tool	–
[16]	Structured at rest	Dynamic access control	On premise	Framework	–
CSA	Structured in motion	Static access control and encryption.	On premise		- Policy updating poses heavy computational burden
[43]	Big data at rest and in motion	Dynamic access control and encryption.	Cloud	Tool	–
[34]	Structured at rest	Anonymization	Cloud	Tool	- Could pose computational burden in the case of big data
[11]	Structured at rest	Anonymization		Tool	–
[4]	Big data	Encryption	Cloud	Tool	–
[18]	Structured in motion	Encryption	Cloud	Tool	- Could become hard to manage the generating encryption keys; - Poses computational burden in case of big data
[36]	Structured at rest and in motion	Encryption	Cloud	Framework	- Data transmission time could increase; - Computational burden due to the index management in case of large amount of data
[40]	Structured in motion	Anonymization	On premise and cloud	Framework	–
[26]	Structured and big data at rest	Anonymization	On premise	Tool	–
[33]	Big data	Access control	Cloud	Framework	–
[3]	Structured and big data at rest	Access control	On premise and cloud	Framework	- A subject can write up but cannot read what he just wrote

Table 3. Security mechanisms for data "integrity"

Approach	Data type	Security mechanism	Tool / Framework	Technology	Drawback
[35]	Structured and big data at rest	Mirroring	Tool	On premise and cloud	- More copies equal higher costs
[20]	Structured and big data at rest	MAC	Tool	On premise and cloud	- Costly in time in case of big data; - Consumes more bandwidth in the case of big data [46]
[17]	Structured and big data at rest	Digital signature	Tool	On premise and cloud	- Costly in money
[25]	Structured and big data at rest	DIaaS	Tool	Cloud	–
[9, 46]	Structured at rest	Blockchain	Tool	-	- Blockchain can't deal with big data; - Performance issues due to the consensus protocol

Table 4. Security mechanisms for data "availability"

Approach	Data type	Security mechanism	Technology	Tool / Framework	Drawback
[37, 42]	Structured, Semi-structured, Unstructured and big data	De-duplication	On premise and cloud	Tool	- For data that must be encrypted, the de-duplication cannot play its optimization role [39]
[30]	Structured at rest	Backup	Cloud	Tool	- Adds performance overhead in terms of monetary cost and upload time
[27]	Structured, Semi-structured and unstructured at rest	Backup	Cloud	Tool	- As data amount increases the monetary cost increases accordingly; - The recovery time is not taken into account in the experimentation
[32]	Structured, Semi-structured and unstructured at rest	Backup	Cloud	Tool	–
[36]	Structured at rest and in motion	Combined mechanisms	Cloud	Framework	- Data transmission time could increase; - Computational burden could be noticed due to the index management in case of a large amount of data

From our comparison tables we have come to the following conclusions:

- **Security goals:** The proposed approaches don't consider all three security goals. Hybrid approaches that combine two or more security mechanisms could be investigated to ensure data confidentiality, integrity, and availability;
- **Encryption mechanism:** Approaches proposing encryption as a security mechanism to protect data don't give much attention to *key management* i.e. the process of protecting cryptographic keys, as this latter could be the target of security attacks. Besides, the proposed algorithms (symmetric

and asymmetric) to encrypt structured data can be used to encrypt unstructured and big data, however, in this case, these solutions induce higher computational burden and cost;
- **Data type:** Data type plays a crucial role in choosing the right security mechanism. As noticed in the comparison tables, a data security approach could give satisfying results when securing structured data but presents computational burden and performance problems when dealing with big data;
- **Big data security:** The existing security solutions for big data are still not mature compared to security solutions for structured data. Big data security lacks standard tools and models. Moreover, patterns could be proposed to help the automatic detection of sensitive information in unstructured data and big data as this becoming a challenge due to the diversity of content in these types of data;
- **Dynamic data:** One of the main characteristics of unstructured and big data is the constant changing of data and therefore, the security mechanisms and frameworks should be as dynamic as possible to cope with these changes.

5 Conclusions and Future Work Directions

Security is one of the most important properties of information systems. Specifically, it is about securing the data and the processes handling this data. In this paper, we began by proposing a general meta-model describing the three main layers of an information system linked with security aspects. Then, we surveyed research works dealing with the security of the data layer and we classified the proposed security approaches according to the data type and security goals. Finally, we put forward a data security meta-model made of three main aspects: data, security issues, security mechanisms, and the impacted non-functional properties (cost, management, computational burden, and data transmission time). We used our meta-model as a reference to compare different approaches in the literature and to identify the shortcomings of each of them. A number of criteria should be taken into consideration when proposing solutions to secure data, namely its type, its state, the technology used, and the non-functional properties impacted by the solution. It is about finding a balance between security mechanisms and the system's non-functional properties. Thus, the purpose of our work is to present a conceptual framework allowing to compare the existing approaches in the literature, and we plan as future work to extend our meta-model with dynamic security mechanisms as proposed in [2,22]. Moreover, we intend to present our work to professionals, in order to validate our meta-model on reel case scenarios, and evaluate the trade-off of each security solution. We are currently working on this meta-model and its application to the e-health domain in a Cloud context. Otherwise, we are about to propose another meta-model that covers the security of the application layer. Our goal is to build a generic meta-model to be used as a reference to secure the three layers of an information system.

References

1. Alshammari, M.M., Alwan, A.A., Nordin, A., Al-Shaikhli, I.F.: Disaster recovery in single-cloud and multi-cloud environments: Issues and challenges. In: 2017 4th IEEE International Conference on Engineering Technologies and Applied Sciences (ICETAS), pp. 1–7. IEEE (2017)
2. Brindha, K., Jeyanthi, N.: A novel approach to secure cloud data storage and dynamic data auditing in a cloud. In: Proceedings of the Second International Conference on Internet of things, Data and Cloud Computing, pp. 1–5 (2017)
3. Cankaya, E.C.: Bell-LaPadula Confidentiality Model, pp. 71–74. Springer, Boston (2011). https://doi.org/10.1007/978-1-4419-5906-5_773
4. Cheng, H., Rong, C., Hwang, K., Wang, W., Li, Y.: Secure big data storage and sharing scheme for cloud tenants. China Commun. **12**(6), 106–115 (2015)
5. Clarke-Salt, J.: SQL Injection Attacks and Defense. Elsevier, Amsterdam (2009)
6. El Emam, K., Rodgers, S., Malin, B.: Anonymising and sharing individual patient data. BMJ **350**, 1–6 (2015)
7. Erraissi, A., Belangour, A.: A big data security layer meta-model proposition. Adv. Sci. Technol. Eng. Syst. J. **4**(5), 409–418 (2019)
8. Ferraiolo, D., Kuhn, D.R., Chandramouli, R.: Role-Based Access Control. Artech House, Boston (2003)
9. Gaetani, E., Aniello, L., Baldoni, R., Lombardi, F., Margheri, A., Sassone, V.: Blockchain-based database to ensure data integrity in cloud computing environments (2017)
10. Geery, D.: Reducing the storage burden via data deduplication. Computer **41**(12), 15–17 (2008)
11. George, R.S., Sabitha, S.: Data anonymization and integrity checking in cloud computing. In: 2013 Fourth International Conference on Computing, Communications and Networking Technologies (ICCCNT), pp. 1–5. IEEE (2013)
12. Hu, V.C., et al.: Guide to attribute based access control (abac) definition and considerations (draft). NIST Spec. Publ. **800**(162), 1–54 (2013)
13. Hu, V.C., Kuhn, D.R., Ferraiolo, D.F., Voas, J.: Attribute-based access control. Computer **48**(2), 85–88 (2015)
14. Jaïdi, F., Labbene-Ayachi, F., Bouhoula, A.: Advanced techniques for deploying reliable and efficient access control: application to e-healthcare. J. Med. Syst. **40**(12), 262 (2016)
15. Karle, T., Vora, D.: Privacy preservation in big data using anonymization techniques. In: 2017 International Conference on Data Management, Analytics and Innovation (ICDMAI), pp. 340–343. IEEE (2017)
16. Kashmar, N., Adda, M., Atieh, M., Ibrahim, H.: Deriving access control models based on generic and dynamic metamodel architecture: industrial use case. Procedia Comput. Sci. **177**, 162–169 (2020)
17. Kaur, R., Kaur, A.: Digital signature. In: 2012 International Conference on Computing Sciences, pp. 295–301. IEEE (2012)
18. Khanezaei, N., Hanapi, Z.M.: A framework based on RSA and AES encryption algorithms for cloud computing services. In: 2014 IEEE Conference on Systems, Process and Control (ICSPC 2014), pp. 58–62. IEEE (2014)
19. Kritikos, K., Massonet, P.: An integrated meta-model for cloud application security modelling. Procedia Comput. Sci. **97**, 84–93 (2016)
20. Liu, D., et al.: Chapter 3 - an introduction to cryptography. In: Next Generation SSH2 Implementation, pp. 41–64. Syngress, Burlington (2009)

21. Mallik, A.: Man-in-the-middle-attack: Understanding in simple words. Cyberspace: Jurnal Pendidikan Teknologi Informasi **2**(2), 109–134 (2019)
22. Mathur, N., Bansode, R.: Aes based text encryption using 12 rounds with dynamic key selection. Procedia Comput. Sci. **79**, 1036–1043 (2016)
23. Menzel, M., Meinel, C.: A security meta-model for service-oriented architectures. In: 2009 IEEE International Conference on Services Computing, pp. 251–259. IEEE (2009)
24. Murthy, S., Bakar, A.A., Rahim, F.A., Ramli, R.: A comparative study of data anonymization techniques. In: 2019 IEEE 5th Internatioinal Conference on Big Data Security on Cloud (BigDataSecurity), IEEE Internatioinal Conference on High Performance and Smart Computing, (HPSC) and IEEE Internatioinal Conference on Intelligent Data and Security (IDS), pp. 306–309. IEEE (2019)
25. Nepal, S., Chen, S., Yao, J., Thilakanathan, D.: Diaas: Data integrity as a service in the cloud. In: 2011 IEEE 4th International Conference on Cloud Computing, pp. 308–315. IEEE (2011)
26. Prakash, M., Singaravel, G.: An approach for prevention of privacy breach and information leakage in sensitive data mining. Comput. Electr. Eng. **45**, 134–140 (2015)
27. Prathyakshini, M., Ankitha, K.: Data storage and retrieval using multiple cloud interfaces. Int. J. Adv. Res. Comput. Commun. Eng. **5**(4), 1–4 (2016)
28. Puthal, D., Malik, N., Mohanty, S.P., Kougianos, E., Das, G.: Everything you wanted to know about the blockchain: its promise, components, processes, and problems. IEEE Consum. Electron. Mag. **7**(4), 6–14 (2018)
29. Qian, C., Huang, Y., Zhao, X., Nakagawa, T.: Optimal backup interval for a database system with full and periodic incremental backup. JCP **5**(4), 557–564 (2010)
30. Rahumed, A., Chen, H.C., Tang, Y., Lee, P.P., Lui, J.C.: A secure cloud backup system with assured deletion and version control. In: 2011 40th International Conference on Parallel Processing Workshops, pp. 160–167. IEEE (2011)
31. Sahafizadeh, E., Nematbakhsh, M.A.: A survey on security issues in big data and nosql. Adv. Comput. Sci. Int. J. **4**(4), 68–72 (2015)
32. Sambrani, Y., Rajashekarappa: Efficient data backup mechanism for cloud computing. Int. J. Adv. Res. Comput. Commun. Eng. **5**(7), 1–4 (2016)
33. Sarkar, B.K.: Big data for secure healthcare system: a conceptual design. Complex Intell. Syst. **3**(2), 133–151 (2017)
34. Sedayao, J., Enterprise Architect, I.I.: Enhancing cloud security using data anonymization. White Paper, Intel Coporation (2012)
35. Sivathanu, G., Wright, C.P., Zadok, E.: Ensuring data integrity in storage: techniques and applications. In: Proceedings of the 2005 ACM Workshop on Storage Security and Survivability, pp. 26–36 (2005)
36. Sood, S.K.: A combined approach to ensure data security in cloud computing. J. Netw. Comput. Appl. **35**(6), 1831–1838 (2012)
37. Storer, M.W., Greenan, K., Long, D.D., Miller, E.L.: Secure data deduplication. In: Proceedings of the 4th ACM International Workshop on Storage Security and Survivability, pp. 1–10 (2008)
38. Suguna, S., Suhasini, A.: Overview of data backup and disaster recovery in cloud. In: International Conference on Information Communication and Embedded Systems (ICICES2014), pp. 1–7. IEEE (2014)
39. Sun, G.Z., Dong, Y., Chen, D.W., Wei, J.: Data backup and recovery based on data de-duplication. In: 2010 International Conference on Artificial Intelligence and Computational Intelligence, vol. 2, pp. 379–382. IEEE (2010)

40. Sweeney, L.: k-anonymity: a model for protecting privacy. Int. J. Uncertainty Fuzziness Knowl. Based Syst. **10**(05), 557–570 (2002)
41. Vasconcelos, A., da Silva, M.M., Fernandes, A., Tribolet, J.: An information system architectural framework for enterprise application integration. In: . Proceedings of the 37th Annual Hawaii International Conference on System Sciences 2004, p. 9 IEEE (2004)
42. Xia, W., et al.: A comprehensive study of the past, present, and future of data deduplication. Proc. IEEE **104**(9), 1681–1710 (2016)
43. Yang, K., Jia, X., Ren, K.: Secure and verifiable policy update outsourcing for big data access control in the cloud. IEEE Trans. Parallel Distrib. Syst. **26**(12), 3461–3470 (2014)
44. Zachman, J.A.: A framework for information systems architecture. IBM Syst. J. **26**(3), 276–292 (1987)
45. Zeng, W., Yang, Y., Luo, B.: Access control for big data using data content. In: 2013 IEEE International Conference on Big Data, pp. 45–47. IEEE (2013)
46. Zikratov, I., Kuzmin, A., Akimenko, V., Niculichev, V., Yalansky, L.: Ensuring data integrity using blockchain technology. In: 2017 20th Conference of Open Innovations Association (FRUCT), pp. 534–539. IEEE (2017)