Proc. of the International Conference on Electrical, Computer and Energy Technologies (ICECET 2022)
20-22 June 2022, Prague-Czech Republic

Data Architectures' Evolution and Protection

Ismaila Idris Sinan
Africa Centre of Excellence on Technology Enhanced Learning, National Open University of Nigeria,
Abuja, Nigeria
isinan@noun.edu.ng

Jules Degila
Africa Centre of Mathematical Sciences, Computer Science and Applications, University of Abomey
Porto-Novo, Benin
jules.degila@imsp-uac.org

Vivian Nwaocha
Africa Centre of Excellence on Technology Enhanced Learning, National Open University of Nigeria,
Abuja, Nigeria
onwaocha@noun.edu.ng

Saidat Adebukola Onashoga
Department of Computer Science, Federal University of Agriculture,
Abeokuta, Nigeria
onashogasa@funaab.edu.ng

2022 International Conference on Electrical, Computer and Energy Technologies (ICECET) | 978-1-6654-7087-2/22/$31.00 ©2022 IEEE | DOI: 10.1109/ICECET55527.2022.9872597

***Abstract*— Nowadays, data architecture allows us to store, manage, analyze, and visualize a limitless quantity of data. It has gone through several stages to get to this point, overcoming various hurdles and challenges. This study highlights these stages of evolution, starting with traditional architecture and progressing to data-informed architecture, data-driven architecture, and finally, data-centric architecture. It also devises a conceptual framework to identify the requirements for transitioning from one stage to other benefits of each architecture. The study's results include a discussion of security vulnerabilities based on confidentiality, integrity, and availability (CIA Triad) and recommendations for future research.**

Keywords*— *Security, Challenges, Traditional architecture, Data-informed architecture, Data-driven architecture, Data-centric architecture

I. Introduction

Many researchers have defined data architectures (DA), [1] describes it as a set of models, policies, rules, and standards that govern which data is collected and how it is stored, formatted, integrated, and used in data systems and organizations. In other terms, [2] defines it as the higher-level perspective of how an enterprise handles its data, categorizes, integrates, and stores it. [3] describes it as a framework of models, policies, regulations, and standards that an organization uses to manage data and its movement across the company. According to [4]–[6], there are four types of data architecture.

1. Traditional architecture (TA)
2. Data-informed architecture (DIA)
3. Data-driven architecture (DDA)
4. Data-centric architecture (DCA)

DA is a storage device that has evolved from 400BC to 2022, and it contains some intelligent features such as [6]:

- ability to store an unlimited volume of data with the advent of cloud storage and data lake.
- categorizes data by setting rules and policies peculiar to the organization
- provides decisions from the data.

Moreover, due to the massive increase in data generation, DA has evolved from one stage to the next. It was predicted that by the end of 2022, the world will have produced 94 zettabytes of data, up from 74 zettabytes in 2021[7], one of the reasons for the constant evolution of data architectures is to meet the needs of this rapid growth. Furthermore, researchers have been developing policies, models, data structures, and security mechanisms to improve data architectures [8]. Findings show that as the complexity of DA grows, so do security threats and cyber-attacks [9], [10] and that these attacks increased dramatically during the Covid-19 pandemic [7]. Within the epidemic period, it was estimated that 75% of businesses, institutions, and healthcare organizations were exposed to cyber-attacks, resulting in $6 trillion in losses to hackers [11], [12]. To mitigate these attacks and upgrade DA to handle the rapid growth of data, institutions, companies, and organizations are dueling in research in academic journals, conference proceedings, reports, and books. However, DA research is limited and mainly on data storage devices [13]–[15]; little or no attention has been paid to architecture and security issues. In light of the foregoing, this research is initiated with the following objectives:

1. To investigate and identify data architectures
2. To explore the implementation challenges of the identified architectures
3. To highlight each architecture's security vulnerabilities
4. Device a conceptual framework for determining the requirements for transitioning from one architecture to another
5. Use the conceptual framework to identify the requirements for moving from one architecture to another and each step's benefits.

In essence, this study aims to answer the following research questions:

1. What are the various data architectures in the literature?
2. What are their benefits, challenges, and security vulnerabilities?
3. Which architecture has the highest advantage and fewer security vulnerabilities?

The following is how this article is structured: Section I is the Introduction part, Section II is a literature review, Section III is an introduction to the study methods, Section IV is a

978-1-6654-7087-2/22/$31.00 ©2022 IEEE

presentation of the research findings, and Section V is a discussion of the contribution. Section VI is the limitation, and Section VII is the paper's final section and serves as the conclusion and future research direction.

II. LITERATURE REVIEW

A small number of researchers have examined the evolution of data architectures. [13] identified six distinct phases in data management, its challenges, and benefits between 400 BC and 1995, [14] investigated the evolution of the hard disk drive (HDD) from 1950 to 2000, [15] studied storage systems from 1950 to 2006 to see how they've changed and to assess the benefits of cloud storage from 1950 to 2008 [16] looked at the evolution of storage system challenges. Table I presents the related work.

TABLE I: RELATED WORK

Reference	Objectives	Research Questions	Range
[13]	This study looks at how data management systems have evolved through time.	What are the phases of data management? What are the challenges of data management?	400BC-1995
[14]	This article looks at how storage systems have changed over the last five decades to satisfy changing consumer demands.	What is the evolution of HDD? What are the requirements of IT? What are the components of data storage?	1950-2000
[15]	This study looks at how storage systems evolved from early mechanical systems to magnetic, optical, and cloud storage.	What are storage systems? What courses storage system evolution? What are the benefits of cloud storage?	1950- 2006
[16]	The study highlights recent advancements in storage systems as well as future challenges and offers recommendations for solutions.	What is the evolution of storage systems? What are the challenges associated with storage systems?	1950-2008

III. METHODOLOGY

For this study, the research technique was adopted from [16]. It starts with the formulation of search strings, followed by sorting, screening, data analysis, data extraction, and developing a conceptual framework.

Security AND (vulnerability OR issues OR challenge) AND "traditional architecture" AND data-driven AND data-centric were used as a search string on Google, Google Scholar, Research Gate, Scopus, IEEE Xplore, and Science Direct, relevant publications relating to the search string, including articles, reviews, chapters, and conference papers, were discovered for further analysis, which yielded 11 results; after that, the search string was adjusted by omitting the words security AND (vulnerability OR issue OR challenge) and 98 items were realized. The search was conducted between November 11th and 12th, 2021, and a total of 109 articles were found. These articles' references and abstracts were downloaded into Endnote, a reference management software application. On November 13, 2021, an abstract screening was conducted to verify whether each article's abstract matched our research objective and to identify duplicate publications. 22 articles were found to be duplicates, 58 were determined to be inappropriate for this investigation based on abstract evaluation, and 29 publications were judged to be eligible and included in the study. On November 14, 2021, this analysis was completed

Conceptual Framework

The development of the conceptual framework was based on previous research on data-storage devices. It was argued by [14], [15] that each architecture's benefits are determined by the requirements needed to transit from one architecture to another, Starting with the TA, moving on to DIA, DDA, and finally DCA. This framework was designed to derive benefits based on the criteria provided in the literature (Fig 1).

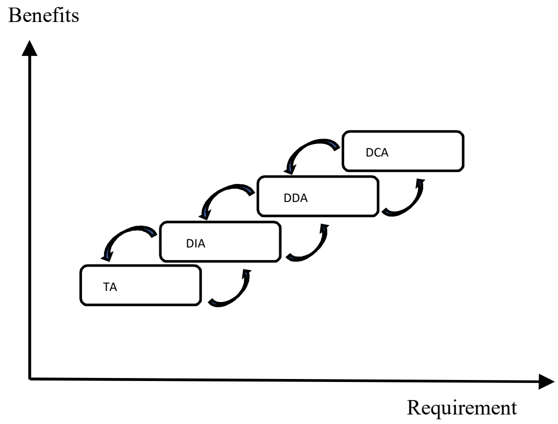

Fig. 1. Conceptual framework

Confidentiality, Integrity, and Availability (CIA Triad)

The CIA triad is a concept that emphasizes the need to maintain a balance between data confidentiality, integrity, and availability [17]–[19]. For a system to be secured, it needs to have confidentiality, integrity, and availability (Fig 2); where:

1) Confidentiality: Its purpose is to restrict data access to just authorized users [17], [19]. Confidentiality is defined as "the degree to which a product or system ensures that data is available only to those authorized to have access," *[20]*. Unauthorized disclosure of data is referred to as a breach of confidentiality.

Fig. 2. CIA Triad [21]

2) Integrity: It guarantees that the data is valid and that only authorized users have access to it [17]. Integrity is defined as "the degree to which a system, product, or component prevents unauthorized access to, or change of, computer programs or data," according to [19]. The unauthorized modification or erasure of data is referred to as a loss of integrity.

3) Availability: Data should be available to authorized users whenever they need it, just as it is critical to keep unauthorized users out of an organization's data [19]

IV. RESULTS

The results produced from the studied literature are now presented and discussed. First and foremost, we will discuss the evolution of data architectures, their challenges, benefits, security issues, and the requirement needed for transitioning from one architecture to another.

A. *Evolution*

DA has evolved into four distinct stages, each with its own set of features [22]–[24]:

- Traditional architecture
- Data-informed architecture
- Data-driven architecture
- Data-centric architecture

Traditional architecture, which is the oldest, began existence with the invention of the punch card in 1800 [13], [15], and can store a small quantity of data. Magnetic tapes supplanted them in 1950 because they had a higher storage capacity; one magnetic tape roll could contain as much data as 10,000 punch cards. It was an instant hit and remained the most prevalent method of storing computer data until the early-1980s. As a result, users have no trouble structuring and making analyses from the data as they deal with small amounts of data.

The invention of the first hard drive by IBM in 1956 [15] kick-started the data-informed architecture era. It was revolutionary because it could hold up to 4.4MB (5 million characters) of data, which was massive. The hard drive is still a product that is being improved upon [13], [24]. In 1969, the floppy disk was introduced to give backup choices[15]. It was a read-only 8-inch disk that could hold 80kB of data. In 1973, a comparable floppy disk with the same capacity could hold 256kB of data and was rewritable [15], [24]. Since then, the trend has remained consistent: smaller floppy disks with more data storage capacity. In the late 1990s, a 3-inch disk could easily store 250 MB of data. Floppy disks were hailed as a game-changing medium for moving data from one computer to another. They couldn't hold as much data as hard drives, but they were far less expensive and versatile [15], [22]. Thus, they became prevalent. Of course, this tendency had an impact on the backup industry. Later, devices with increased storage capacity, such as CD-Recordable (CD-R) and CD-Rewritable (CD-RW) discs with storage capacities of up to 700 megabytes, and flash drives with sufficient storage capacity. Data warehouses were developed [13], [15], [22], [25], a dashboard or excel is used to analyze the data, and the results are used as part of inputs in decision making [22]. In this era, users have begun to appreciate the benefits of making decisions based on data, and the demand for more data storage devices has increased drastically.

The creation of cloud storage was thought to have kicked off the data-driven architecture period in 1960 [14], [15]. However, it wasn't wholly implemented until mid-1990, when CompuServe began to provide its clients with tiny amounts of disk space to store some of their files[26]. Automation tools were created due to data-driven architectures to make data analysis easier. The introduction of silos, cloud computing, big data, and data lakes ushered in a transformation in which just storing data is no longer sufficient; data integration is required to analyze and get insight from the data [23]–[25], [27]. Companies and organizations who use this architecture see it as a decision-making strategy that incorporates successful data storage, modeling, availability, and artificial intelligence (AI) and human knowledge workers [28], resulting in a positive transformation in data management.

Data-centric architecture is a system in which data is the primary and permanent asset, whereas applications come and go [19]. It first appeared in the early 20s. Users create a single data model that is shared across all information systems. Data science is at the heart of decision-making. All data is linked and connected using a graph database, eliminating data redundancy and silos.

Data ownership, data integrity, data traceability, data access permission, data insights, and data interoperability benefit from a data-centric architecture. Data-centricity makes security, integration, portability, and analysis considerably more manageable, and faster insights are supplied along the entire data value chain [13], [16], [17]. Fig. 3 depicts data-centric and data-driven approaches in decision making

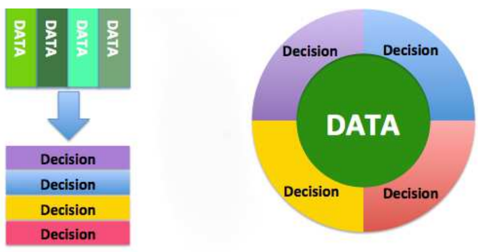

Fig. 3. Data-driven vs data-centric [29]

B. *Challenges*

The findings suggest that there were challenges during the traditional architectural period. The traditional architecture, like the first, was unable to store a large amount of data due to a lack of appropriate devices [13]–[15], [25], [30], [31]. It was also argued that it lacked a backup option [13]–[15], [25], [31] and no capabilities for analyzing data and identifying patterns [13], [15], [31], and it was prolonged and required a lot of processing time (Table II)

Data-informed architecture, on the other hand, has problems storing enough data [15], [23], [31], analyzing data [23]–[25], automating data [14], [22], [30], [31], and taking too much data as input [22], [23], [25].

Data redundancy [27], [35] , delayed real-time analysis [27], [35], the establishment of silos [4], [27], [35], [36], significant maintenance costs [27] , and no data ownership [4], [27], [37], [38] are among issues that data-driven architecture faces.

TABLE II: EVOLUTION OF DATA ARCHITECTURES

Evolution	Challenge(s)	Benefit(s)	Security Issue(s)
Traditional data architecture	No trends identification[15] [13], [31] No enough memory storage [15] [13], [14], [25], [30], [31] No backup option[13]–[15], [25], [31]	Low maintenance cost [32] Easy to maintained [22] Saves little amount of data [13], [14], [25], [30], [31], [33]	Availability [13], [15], [32]
Data-informed architecture	No automation [14], [22], [30], [31] Complexity of too much input [22], [23], [25] Difficulty explaining decision rationale [22], [23], [25] No enough memory storage[15], [23], [31]	Complement data with other decision inputs [22] [25] Trends Identification [23] Low maintenance cost [15], [31] Saves more data [15], [25], [30], [31] Backup option [13]–[15], [25], [31], [33]	Confidentiality [32], [34] Integrity [15], [34] Availability [13]–[15], [22]
Data-driven architecture	Data redundancy [27], [35] Delayed real time analysis [27], [35] Creation of data silos [4], [27], [35], [36] High maintenance cost [27] No data ownership [4], [27], [37], [38]	Automation of decision [22], [23], [35] Reduce decision cycle time [22], [23] Capacity for frequent high-volume decision [4], [22], [27] Eliminate partisan influences [22], [23] Saves unlimited amount of data [31], [36], [36]	Confidentiality [15], [32], [34], [39]–[46] Integrity [15], [32], [39]–[46]
Data-centric architecture	High cost of implementation [47]	Reduce data redundancy [37], [48] Eliminate data silos [4], [27], [37], [48] Data ownership [4], [27], [37] The unified governance of data and processes [27], [37] Improves data quality and reliability [48] Eliminate complex data transformation [38], [48] Makes data more accessible to key stakeholders [38], [48]	Confidentiality [49]

The most modern and sophisticated design, data-centric architecture, has the fewest concerns, with just high implementation costs as a hurdle[47]

Security Concerns

Using the CIA triad, it was observed that availability security difficulties were most prevalent at the traditional and data-informed architectural stages [13], [15], [22], [32], [47]. Availability difficulties were at their lowest during the time of data-driven architecture and data-centric architecture (Fig. IV)

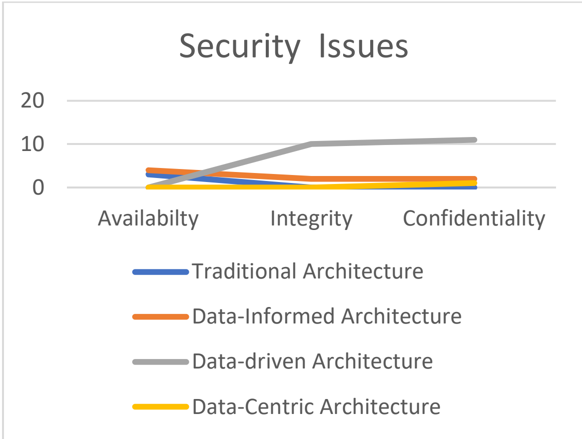

Fig. IV. Security issues

Traditional architectures and data-centric architectures have fewer data integrity challenges, whereas data-driven architectures have the most, followed by data-informed architectures [15], [32], [39]–[46]

Traditional architecture and data-informed architectures, according to [15], [32], [39]–[45], [49], [50], have less confidential concerns, but data-driven designs have the most issues, followed by data-centric architecture..

Transition requirements and benefits

In this paper, we use the proposed conceptual framework to highlight the requirements for transitioning from one architecture to another and the benefits of each architecture. To transition from traditional architecture to data-informed architecture, the advent of floppy disks, hard drives, flash drives, and data warehouses was required as various tools for data analysis using graphs. Users were able to profit from this by storing more data, identifying trends, and creating data backups.

To transition from data-informed to data-driven architecture, it took the invention of silos, cloud storage, big-data, and data-lake and the advent of tools for analyzing and automating the stored data, such as Hadoop and rapid miner. Users benefited from this by preserving adequate data and automating data to gain insight and make decisions based on these criteria (Table III).

To transition from a data-driven to a data-centric architecture required the creation of a graph database, eliminating silos, and introducing various tools for analyzing and automating the stored data, such as Egnyt and ermetic. Users could profit from this by owning their data, less data redundancy, and improved data quality and reliability.

TABLE III: TRANSITION'S REQUIREMENT

Transition state(s)	Requirement(s)	Data architecture	Benefit(s)
Traditional architecture to data-informed architecture	Floppy disk Hard drives Flash drives Data warehouse	Data-informed architecture	Saves more data Trend's identification Backups
Data-informed architecture to data-driven architecture	Data Silos Cloud storage Big data Data lake Tools, e.g., Hadoop, rapid miner, etc.	Data-driven architecture	Saves enough data Data automation
Data-driven architecture to data-centric architecture	Removal of data silos Graph database Tools, e.g., Egnyt, ermetic etc.	Data-centric architecture	Data ownership Remove data redundancy Improve data quality and reliability Makes data more accessible to key stakeholders

V. CONTRIBUTION

This study investigates the evolution of data architecture and its challenges, benefits, and security flaws. Almost all firms, organizations, and institutions with a significant role in implementing, maintaining, and upgrading data architectures could benefit from this. Moreover, by building on the work of [15], this study bridged the gap between the evolution of data storage devices and the evolution of data architectures. Our study has identified four distinct stages of data architecture transition between 1800 and 2022, and DA users will be able to identify their stage and the requirement needed to move to the next stage. Furthermore, based on scholarly contributions, our findings reveal that data-centric architecture is the most advanced level of DA, with high benefits, fewer difficulties, and a high implementation cost. As a result, it would be fascinating to investigate the possibilities and requirements for deploying data-centric architecture. Additionally, identifying data-architectures security risks, benefits, and challenges allows researchers better to understand the most promising aspects of each architecture.

VI. LIMITATION

Although the literature review was completed utilizing six significant databases, the search was not exhaustive regarding the article's limitations. Furthermore, the research did not look into the firms and organizations that employ these architectures. Therefore, the size of the databases should be increased in future research and the firms that use the architectures.

VII. CONCLUSION AND DIRECTION FOR FUTURE RESEARCH

This study provides a comprehensive review of papers relating to the evolution of data storage devices and the classification of four data architectures: traditional architecture, data-informed architecture, data-driven architecture, and data-centric architecture. We proposed and implemented a conceptual framework that uses requirements to realize the benefits of each transition stage from one architecture to the next. In addition, we look at the implementation challenges of each architecture and, most crucially, the security concerns that each architecture poses. We believe that, based on the findings of this study, more research is needed into ways to mitigate the confidentiality and security challenges connected with data-centric architecture.

ACKNOWLEDGMENT

The authors would like to express their gratitude to the IRD and AFD for their moral and financial support

REFERENCES

[1] X. Zheng, Q. Li, and L. Kong, "A Data Storage Architecture Supporting Multi-level Customization for SaaS," in *2010 Seventh Web Information Systems and Applications Conference*, Aug. 2010, pp. 106–109. DOI: 10.1109/WISA.2010.18.

[2] W. H. Inmon, D. Linstedt, and M. Levins, "Chapter 1.1 - An Introduction to Data Architecture," in *Data Architecture (Second Edition)*, W. H. Inmon, D. Linstedt, and M. Levins, Eds. Academic Press, 2019, pp. 1–5. DOI: 10.1016/B978-0-12-816916-2.00001-2.

[3] Simplilearn, "What Is Data Architecture? Overview and Best Practices," *Simplilearn.com*, June 15, 2021. https://www.simplilearn.com/what-is-data-architecture-article (accessed January 01, 2022).

[4] D. Carol, "The Difference Between Data-centric and Data-driven," *Applied Software*, February 04, 2021. https://www.asti.com/the-difference-between-data-centric-and-data-driven/ (accessed November 11, 2021).

[5] Ascend, "Data-Informed, Data-Driven, and Data-Centric: What's the Difference?," *Ascend Venture Capital*. https://www.ascendstl.com/press/2020/4/28/data-driven-and-data-centric-whats-the-difference (accessed January 01, 2022).

[6] D. S. Kampakis, "What are the differences between data-driven, data-informed and data-centric?" *The Data Scientist*, October 22, 2018. https://thedatascientist.com/data-driven-data-informed-data-centric/ (accessed January 01, 2022).

[7] CloverDX, "How much data will the world produce in 2021?" 2021. https://www.cloverdx.com/blog/how-much-data-will-the-world-produce-in-2021 (accessed November 12, 2021).

[8] J. M. Hatfield, "Social engineering in cybersecurity: The evolution of a concept," *Comput. Secure.*, vol. 73, pp. 102–113, Mar. 2018, DOI: 10.1016/j.cose.2017.10.008.

[9] A. S. Arora, L. Raja, and B. Bahl, "Data-Centric Security Approach: A Way to Achieve Security & Privacy in Cloud Computing," Social Science Research Network, Rochester, NY, SSRN Scholarly Paper ID 3168615, Apr. 2018. DOI: 10.2139/ssrn.3168615.

[10] J. Ahmed and Q. Tushar, "Covid-19 Pandemic: A New Era Of Cyber Security Threat And Holistic Approach To Overcome," in *2020 IEEE Asia-Pacific Conference on Computer Science and Data Engineering (CSDE)*, Dec. 2020, pp. 1–5. DOI: 10.1109/CSDE50874.2020.9411533.

[11] B. Dixon, "COVID-19 shows the urgent need to address the cyber poverty gap," *World Economic Forum*, 2020. https://www.weforum.org/agenda/2020/03/covid-19-pandemic-shows-the-urgency-for-addressing-the-cyber-poverty-gap/ (accessed April 20, 2022).

[12] J. Chigada and R. Madzinga, "Cyberattacks and threats during COVID-19: A systematic literature review," *South Afr. J. Inf. Manag.*, vol. 23, no. 1, pp. 1–11, 2021, DOI: 10.4102/sajim.v23i1.1277.

[13] J. Gray, "Evolution of data management," *Computer*, vol. 29, no. 10, pp. 38–46, Oct. 1996, DOI: 10.1109/2.539719.

[14] R. J. T. Morris and B. J. Truskowski, "The evolution of storage systems," *IBM Syst. J.*, vol. 42, no. 2, pp. 205–217, 2003, DOI: 10.1147/sj.422.0205.

[15] B. Trikha, "A Journey from floppy disk to cloud storage," vol. 02, no. 04, p. 4, 2010.

[16] S. Fosso Wamba, A. Anand, and L. Carter, “A literature review of RFID-enabled healthcare applications and issues,” *Int. J. Inf. Manag.*, vol. 33, no. 5, pp. 875–891, Oct. 2013, doi: 10.1016/j.ijinfomgt.2013.07.005.
[17] H. Ghadeer, “Cybersecurity Issues in Internet of Things and Countermeasures,” in *2018 IEEE International Conference on Industrial Internet (ICII)*, Oct. 2018, pp. 195–201. DOI: 10.1109/ICII.2018.00037.
[18] P. Eugen and D. Petruț, “Exploring the New Era of Cybersecurity Governance,” *Analele Univ. Ovidius Constanta*, vol. XVIII, pp. 358–363, Jan. 2019.
[19] L. O. Nweke, “Using the CIA and aaa models to explain cybersecurity activities,” *PM World J.*, vol. 6, pp. 1–2, 2017.
[20] C. Chigozie-Okwum, E. Peter, and J. Odii, “Adoption of Learning Management Systems in Nigerian Tertiary Institutions: Issues and Challenges,” *Int. J. Comput. Appl.*, vol. 181, pp. 5–10, Nov. 2018, DOI: 10.5120/ijca2018918173.
[21] Diesec, “Data Security – DIESEC™,” 2021. https://diesec.com/data-security/ (accessed April 22, 2022).
[22] W. Dave, “Data-Informed vs. Data-Driven: Comparing the Approaches,” 2021. https://www.eckerson.com/articles/from-analysis-to-action-data-driven-vs-data-informed (accessed November 11, 2021).
[23] T. C. Team, “Data-Driven Vs Data-Informed: Which is More Strategic?,” *Trapica*, January 19, 2021. https://medium.com/trapica/data-driven-vs-data-informed-which-is-more-strategic-53efe194ebb6 (accessed November 11, 2021).
[24] M. Dave, “The Data-Centric Revolution: Data-Centric vs. Data-Driven,” *TDAN.com*, 2020. https://tdan.com/the-data-centric-revolution-data-centric-vs-data-driven/20288 (accessed August 16, 2021).
[25] B. Bhushan, “Historical evolution of magnetic data storage devices and related conferences,” *Microsyst. Technol.*, vol. 24, no. 11, pp. 4423–4436, Nov. 2018, DOI: 10.1007/s00542-018-4133-6.
[26] Jake Frankenfield, “Cloud Storage,” *Investopedia*, 2021. https://www.investopedia.com/terms/c/cloud-storage.asp (accessed November 11, 2021).
[27] EBS Integrator, “BASICS: Data-Centric vs Data-Driven: the whys’ and whats’,” *EBS Integrator*, August 23, 2020. https://ebs-integrator.com/blog/basics-data-centric-vs-data-driven-the-whys-and-whats/ (accessed November 11, 2021).
[28] K. Andrews, S. Steinau, and M. Reichert, “Enabling runtime flexibility in data-centric and data-driven process execution engines,” *Inf. Syst.*, vol. 101, p. 101447, Nov. 2021, DOI: 10.1016/j.is.2019.101447.
[29] Vista, “Data-centric Architecture — A Different Way of Thinking | Vista Projects,” 2021. https://www.vistaprojects.com/blog/data-centric-architecture/ (accessed August 16, 2021).
[30] B. Pirenne, “Data Storage Technology: Coping by the Evolution,” *Symp. - Int. Astron. Union*, vol. 161, pp. 339–346, ed 1994, DOI: 10.1017/S0074180900047616.
[31] H. K. Ramapriyan, “Evolution of Archival Storage (from tape to memory),” *Retrieved April*, vol. 13, p. 2019, 2015.
[32] M. S. Chithik Raja, A. Munir, and M. Rabbani, “Big Data Analytics Security Issues in Data-Driven Information System,” vol. 3297, pp. 2320–9801, Oct. 2014.
[33] D.-Y. Lin and I. Neamtiu, “Collateral evolution of applications and databases,” in *Proceedings of the joint international and annual ERCIM workshops on Principles of software evolution (IWPSE) and software evolution (Evol) workshops*, 2009, pp. 31–40.
[34] Z. Chunying and G. Weiqing, “Risk management based on data warehouse of securities companies,” in *2009 4th International Conference on Computer Science Education*, Jul. 2009, pp. 819–822. DOI: 10.1109/ICCSE.2009.5228164.
[35] S. Luke and D. Matt, “Becoming a Data-Driven Payer Organization: 6 Challenges,” 2018. https://www.healthcatalyst.com/insights/becoming-data-driven-payer-organization-6-challenges/ (accessed November 11, 2021).
[36] A. Siddiqa, A. Karim, and A. Gani, “Big data storage technologies: a survey,” *Front. Inf. Technol. Electron. Eng.*, vol. 18, no. 8, pp. 1040–1070, Aug. 2017, doi: 10.1631/FITEE.1500441.
[37] F. Alfonso, “Data-driven versus data-centric,” *Stratio*, February 13, 2018. https://blog.stratio.com/datadriven-versus-datacentric/ (accessed November 11, 2021).
[38] D. Alvarez-Coello, D. Wilms, A. Bekan, and J. Marx Gómez, “Towards a Data-Centric Architecture in the Automotive Industry,” *Procedia Comput. Sci.*, vol. 181, pp. 658–663, 2021, DOI: 10.1016/j.procs.2021.01.215.
[39] R. Danda and D. Ronald, “Cybersecurity in Big Data Era: From Securing Big Data to Data-Driven Security,” 2019. https://ieeexplore.ieee.org/abstract/document/8673585/ (accessed November 12, 2021).
[40] M. Derfouf, A. Mimouni, and M. Eleuldj, “Vulnerabilities and storage security in cloud computing,” in *2015 International Conference on Cloud Technologies and Applications (CloudTech)*, Jun. 2015, pp. 1–5. DOI: 10.1109/CloudTech.2015.7337002.
[41] S. Dilmaghani, M. R. Brust, G. Danoy, N. Cassagnes, J. Pesaro, and P. Bouvry, “Privacy and Security of Big Data in AI Systems: A Research and Standards Perspective,” in *2019 IEEE International Conference on Big Data (Big Data)*, Dec. 2019, pp. 5737–5743. DOI: 10.1109/BigData47090.2019.9006283.
[42] Q. Liu, P. Li, W. Zhao, W. Cai, S. Yu, and V. C. M. Leung, “A Survey on Security Threats and Defensive Techniques of Machine Learning: A Data-Driven View,” *IEEE Access*, vol. 6, pp. 12103–12117, 2018, DOI: 10.1109/ACCESS.2018.2805680.
[43] M. Paryasto, A. Alamsyah, B. Rahardjo, and Kuspriyanto, “Big-data security management issues,” in *2014 2nd International Conference on Information and Communication Technology (ICoICT)*, Bandung, Indonesia, May 2014, pp. 59–63. DOI: 10.1109/ICoICT.2014.6914040.
[44] A. Shameli-Sendi, “An efficient security data-driven approach for implementing risk assessment,” *J. Inf. Secure. Appl.*, vol. 54, p. 102593, Oct. 2020, DOI: 10.1016/j.jisa.2020.102593.
[45] B. Thuraisingham *et al.*, “A Data-Driven Approach for the Science of Cyber Security: Challenges and Directions,” in *2016 IEEE 17th International Conference on Information Reuse and Integration (IRI)*, Pittsburgh, PA, USA, Jul. 2016, pp. 1–10. DOI: 10.1109/IRI.2016.10.
[46] S. Venkatraman and R. Venkatraman, “Big data security challenges and strategies,” *AIMS Math.*, vol. 4, no. 3, pp. 860–879, 2019.
[47] William morse, “Improving Your Institution with Data-Centric Design,” *The EvoLLLution*, March 08, 2021. https://evolllution.com/managing-institution/operations_efficiency/improving-your-institution-with-data-centric-design/ (accessed November 11, 2021).
[48] S. Robert, “Why and how to adopt a data-centric architecture,” *SearchConvergedInfrastructure*, 2021. https://searchconvergedinfrastructure.techtarget.com/tip/Why-and-how-to-adopt-a-data-centric-architecture (accessed August 16, 2021).
[49] A. E. Dadzie and S. Kant, “A Quick Survey of Security and Privacy Issues in Cloud and a Proposed Data-Centric Security Model for Data Security,” in *Evolutionary Computing and Mobile Sustainable Networks*, Singapore, 2021, pp. 431–443. DOI: 10.1007/978-981-15-5258-8_41.
[50] M.-L. Clausen, “Justifying military intervention: Yemen as a failed state,” *Third World Q.*, vol. 40, no. 3, pp. 488–502, Mar. 2019, DOI: 10.1080/01436597.2019.1573141.