Africa Centre of Excellence on Technology Enhanced Learning
National Open University of Nigeria

A Secure Data-Centric Model for Digital Learning in Higher Institutions

by

Ismaila Idris Sinan
ACE21150003

Thesis Submitted to African Centre of Excellence on Technology Enhanced Learning National Open University of Nigeria for the Award of Doctorate in Cyber Security

Supervisor:
Dr Vivian Nwaocha

Co-Supervisors:
Prof. Jules Dégila
&
Prof. Adebukola Onashoga

Date of Submission:
22/12/2023

METADATA

Title: A Secure Data-Centric Model for Digital Learning in Higher Institutions

Student Name: Ismaila Idris Sinan

Supervisor: Dr Vivian Nwaocha

Co-Supervisors: Prof. Jules Dégila & Prof Adebukola Onashoga

Department: Africa Centre of Excellence on Technology Enhanced Learning

Qualification: Doctorate in Cyber Security

Institution: National Open University of Nigeria

Keywords: Data-centric model, security model, comparative analysis, proof of concept

Document Date: 22/12/2023

Sponsor: Digital Science and Technology Network

DECLARATION

I, Ismaila Idris Sinan ACE21150003, affirm that I conducted this thesis independently, and it has not been submitted for the fulfilment of any academic prerequisites or awards by any other means.

Ismaila Idris Sinan ________________ 21/12/2023

Student Name Signature Date

Signature of the Supervisor

I, Vivian Nwaocha, herewith declare that I accept this thesis for my supervision.

Signature________________________________ Data: 21/12/2023

Signature of Co-Supervisors:

I, Jules Dégila, herewith declare that I accept this thesis for my supervision

Signature_______________. Data: 17/01/2024

I, Adebukola Onashoga, herewith declare that I accept this thesis for my supervision.

Signature__________________. Data: 18/01/2024

DEDICATION

This work is dedicated to the cherished memory of my Late Dad, Ash-Sheikh Ismaila Idris Ibn Zakariya, whose wisdom, and support have been an enduring source of inspiration. His influence continues to shape my journey. Additionally, I dedicate this work to my beloved Mother, whose love and encouragement have been my pillars of strength. May the soul of my dear Dad rest in peace, and may this work reflect the values he instilled in me.

ACKNOWLEDGEMENTS

I express my profound gratitude to my esteemed supervisors, Dr Vivian Nwaocha, Prof. Jules Dégila, and Prof. Adebukola Onashoga, for their exceptional guidance, unwavering support, and valuable insights throughout the entire research journey. The depth of their expertise and commitment significantly contributed to the success of this work.

A special acknowledgement is extended to Prof. Valerie Viet Triem Tong, my internship supervisor, and Prof. Grace Jokthan, Centre Director, for their mentorship and encouragement, which played a pivotal role in shaping the research direction.

I am grateful to my family members, including my brothers and sisters, whose encouragement and understanding provided a strong foundation. I extend my appreciation to my colleagues at work, friends, especially my wife Hajara Abubakar Maihali, and children Suwaiba Sinan, Rufaida Sinan, and Nusaiba Sinan, for their unwavering support.

I would like to acknowledge the financial support from Digital Science and Technology Network (DSTN) Confidentiality, Integrity, Disponibilité, Repartition (CIDRE) team, Africa Centre of Excellence on Technology Enhanced Learning (ACETEL), Institut de Recherche pour le Développement (IRD), and the invaluable contributions of Prof. Goussou Camara, Sopho, Helen, and my proposed wife Amina Aliyu Abdullahi. Their support and encouragement have been instrumental in the completion of this research.

Furthermore, I appreciate the collaborative efforts of the research community, including the DSTN, CIDRE team, ACETEL, IRD, for their financial support, which greatly facilitated the successful execution of this research.

In essence, the collective support, mentorship, and encouragement from these individuals and institutions have been crucial to this research's completion, and I am sincerely grateful.

LIST OF PUBLICATIONS FROM THE THESIS

1. I. I. Sinan, J. Degila, V. Nwaocha and S. A. Onashoga, "Data Architectures' Evolution and Protection," *2022 International Conference on Electrical, Computer and Energy Technologies (ICECET)*, 2022, pp. 1-6, doi: 10.1109/ICECET55527.2022.9872597. https://ieeexplore.ieee.org/document/9872597
2. I. I. Sinan, V. Nwoacha, J. Degila and S. A. Onashoga, "A Comparison of Data-Driven and Data-Centric Architectures using E-Learning Solutions," 2022 International Conference Advancement in Data Science, E-learning and Information Systems (ICADEIS), Bandung, Indonesia, 2022, pp. 1-6, doi: 10.1109/ICADEIS56544.2022.10037358.https://ieeexplore.ieee.org/document/10037358
3. I. I. Sinan, V. Nwaocha, J. Degila and S. A. Onashoga, E-Learning Digitalization' Evolution and Transformation. Proceedings of the 33rd ECOWAS iSTEAMS Emerging Technologies, Scientific, Business, Social Innovations & Cyber Space Ecosystem Multidisciplinary Conference. University of Ghana/Academic City University College, Ghana. 29th Sept – 1st Oct, 2022. Pp 119-129. www.isteams.net/ecowasetech2022. dx.doi.org/10.22624/AIMS-/ECOWASETECH2022P23.https://www.isteams.net/_files/ugd/185b0a_41b1497099c2407aad194f6782a92eda.pdf
4. I. I. Sinan, V. Nwaocha, J. Degila and S. A. Onashoga, Data-architectures and the prevalent cyberattacks Encountered by West African Institutions in the COVID—19 Era, journal of Infrastructure, Policy and Development. **Accepted for publication**.
5. I. I. Sinan, V. Nwaocha, Valerie Viet Triem Tong, J. Degila and S. A. Onashoga, Enhancing Security and Privacy in Educational Environments: A Secure Grade Distribution Scheme with Moodle Integration, Journal of Infrastructure, Policy, and Development. **Accepted for publication.**
6. I.I. Sinan, Valérie Viet Triem Tong, V. Nwoacha, and J. Degila, "A Comprehensive Scheme for Secure Grade Distribution in Educational Settings: Integration of Cryptographic Techniques and User-Centric Security in Moodle," Under Review" European Interdisciplinary Cybersecurity Conference

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

Digital learning universities face unique challenges in managing, securing, and extracting information from the vast amount of data generated in educational settings. This research focused on identifying the most suitable data architecture for these Universities in West Africa. The methodology involves a comparative analysis of 109 e-learning solution use cases, classified based on their data architectures. A total of 983 user reviews from the e-learning industry further inform the analysis, the study proceeds to develop a data-centric model specifically designed to meet the distinct needs of digital learning universities. The proposed data-centric model integrates essential components, such as data sources, a data hub, efficient data streaming through Apache Kafka, and data governance and security using Apache Ranger and Apache Atlas. The security framework embedded within the model employs Role-Based Access Control (RBAC), encryption through AES, and countermeasures to address identified threats. The research's innovation extends to a Secure Grade Distribution Scheme, practically implemented within the Moodle learning management system. This scheme leverages advanced features, including Diffie-Hellman key exchange, Hardware security module (HSM) and Message Integrity Code (MIC) verification, showcasing its adaptability and effectiveness in enhancing security within educational environments. The integrated proof of concept provides a practical demonstration of both the Secure Grade Distribution Scheme and the proposed Data-Centric Model within a controlled lab environment. This comprehensive approach ensures the validation of the research findings and their potential impact on the secure and efficient management of data in digital learning universities.

Keywords: Data-centric model, security model, comparative analysis, proof of concept

CHAPTER ONE: INTRODUCTION

1.1 Background of the Study

The integration of digital learning in higher education has become an imperative response to the changing educational landscape globally. The advent of the COVID-19 pandemic has acted as a catalyst, compelling educational institutions to swiftly adopt and adapt to digital methodologies to ensure the continuity of learning (Aulakh et al., 2023). This shift has, in turn, heightened the significance of robust data architectures to support the seamless functioning of digital learning platforms.

The challenges posed by the pandemic have been particularly pronounced in West African Universities, where the intersection of limited resources and a growing demand for digital education has amplified the complexities of the data management (Djeki et al., 2023). The vulnerabilities of existing data architectures have been exposed as universities grapple with the sudden surge in cyber-attacks targeting academic databases and communication channels (Sinan, Nwoacha, et al., 2022). These challenges necessitate a thorough investigation into the prevailing data architectures and their resilience to cyber threats during the pandemic.

Understanding the unique context of West African Universities is crucial in this study. These institutions often face resource constraints, both in terms of technological infrastructure and financial capabilities (Aborode et al., 2020; Bervell & Umar, 2017). This context adds layers of complexity to the task of ensuring data security in a digital learning environment. The dynamic nature of cyber threats and the evolving landscape of digital technologies further compound the challenges faced by these universities (Sun et al., 2023).

Scholarly discussions on the impact of digital learning in higher education have acknowledged the need for a nuanced understanding of the contextual factors influencing its implementation (Setiawan et al., 2023). The intersection of digital learning, data architecture, and cybersecurity in the specific context of West African Universities remains underexplored. This research aims to fill this gap by providing insights that can inform both local and global strategies for enhancing data security in the rapidly evolving domain of higher education.

1.2 Statement of the Problem

The intersection of digital learning, data architectures, and cybersecurity in West African Universities presents a multifaceted challenge that demands comprehensive investigation. As higher education institutions rapidly transition to digital platforms, the vulnerabilities within existing data architectures become increasingly apparent, exacerbated by the unforeseen disruptions brought about by the COVID-19 pandemic. The surge in cyber-attacks targeting academic databases and communication channels has exposed critical shortcomings in the resilience of these architectures (Sinan, Degila, et al., 2022a; Sinan, Nwoacha, et al., 2022). This raises pressing concerns about the integrity, confidentiality, and availability of academic data essential for the smooth functioning of digital learning environments.

The challenges faced by West African Universities in this context are compounded by resource constraints, both in terms of technological infrastructure and financial capabilities (Aborode et al., 2020; Bervell & Umar, 2017). These constraints not only limit the ability of institutions to invest in sophisticated cybersecurity measures but also underscore the need for context-specific solutions that consider the unique socio-economic and technological landscape of the region. The absence of a dedicated exploration into the existing data architectures, cybersecurity challenges faced during the pandemic, and the countermeasures implemented by these universities leave a critical gap in understanding the holistic picture of data security in the digital learning domain.

Furthermore, the dynamic nature of cyber threats and the evolving landscape of digital technologies add an additional layer of complexity to the problem (Sun et al., 2023). There is a palpable urgency to address these challenges comprehensively, ensuring that any proposed solutions are not only effective in mitigating current threats but also adaptable to the evolving nature of cybersecurity risks in the higher education sector. Thus, the overarching problem to be addressed is the inadequacy of current data architectures in West African Universities to withstand cyber threats, particularly in the context of the rapid shift towards digital learning platforms during and beyond the COVID-19 pandemic.

Considering these, this research seeks to investigate the following key aspects: the prevailing data architectures in West African Universities, the types of cyber-attacks faced during the pandemic, the countermeasures implemented by these universities, and ultimately, the design of a robust security model and a Secure Grade Distribution Scheme tailored to the specific needs of digital learning environments in the region.

1.3 Aim of the study

This study aims to provide insights into the prevailing data architectures, the types of cyber-attacks faced, and the countermeasures implemented, ultimately contributing to the development of a secure and resilient framework for digital learning.

1.4 Specific objectives

1. Comprehensive Survey of Existing Data Architectures, Cybersecurity Challenges, and Countermeasures in West African Universities during the COVID-19 Pandemic
    a. Conduct an extensive survey to identify prevailing data architectures within West African Universities.
    b. Analyse the types of cyber-attacks faced by these architectures during the COVID-19 pandemic.
    c. Identify and assess the countermeasures implemented by these universities to mitigate cybersecurity challenges.
2. Comparative Analysis for Optimal Digital Learning University Architecture Using E-learning Use Cases
    a. Conduct a rigorous comparative analysis of identified data architectures, focusing on e-learning use cases to determine the most suitable one for meeting the evolving demands of a digital learning university.
3. Design and Proposal of a Security Model for Digital Learning Universities
    a. Develop a robust security model tailored to address the specific threats identified in digital learning environments within universities.
    b. Propose enhancements to existing security measures to fortify data protection.

4. Development of a Secure Grade Distribution Scheme
    a. Create a Secure Grade Distribution Scheme specifically designed for digital learning universities, ensuring the integrity and confidentiality of students' grades.
5. Proof of Concept for Data Architecture and Grade Distribution Scheme
    a. Provide a practical demonstration of the proposed data architecture and grade distribution scheme to validate their effectiveness in a real-world digital learning setting.

1.5 Scope of the Study

This research will focus specifically on West African Universities, recognizing the unique contextual factors that influence digital learning, data architectures, and cybersecurity in this region. The scope encompasses a comprehensive exploration of existing data architectures, cybersecurity challenges faced during the pandemic, and the development of targeted solutions to enhance the overall security posture of digital learning environments.

1.6 Significance of the Study

The significance of this study lies in its potential to address critical gaps in understanding and fortifying data security within the context of West African Universities. As these institutions grapple with the challenges of digital learning adoption exacerbated by the COVID-19 pandemic, insights gained from this research can inform strategic decisions at various levels. By unravelling the intricacies of existing data architectures, cybersecurity challenges, and countermeasures implemented by universities, this study contributes valuable knowledge that can be leveraged by educational policymakers, administrators, and technologists. Furthermore, the findings hold broader implications for higher education globally, offering a nuanced understanding of the intersection between digital learning, data architecture, and cybersecurity.

1.7 Definition of Terms

1. Data-Centric Architecture: Refers to an approach in system design where data is the primary focus, and applications are built around the data. It emphasizes efficient data management, accessibility, and usability.
2. E-Learning: The use of electronic technologies to facilitate learning and education. It involves the use of computers, digital resources, and the internet to deliver educational content and support interactive learning experiences.
3. Hybrid Storage: A storage solution that combines different types of storage technologies, such as Network-Attached Storage (NAS), Storage Area Network (SAN), Cloud Storage, Local Storage, and Tape Storage, to optimize performance and cost efficiency.
4. Data Governance: The overall management of the availability, usability, integrity, and security of data used in an organization. It involves defining data-related policies, standards, and processes to ensure effective data management.
5. Security Model: A structured framework that defines and enforces security policies, measures, and controls to protect data, systems, and resources from unauthorized access, attacks, and potential threats.
6. Data as a Service (DaaS): A service-oriented approach that provides on-demand access to data, allowing users to access and utilize data without the need for extensive local storage and management.
7. TensorFlow: An open-source machine learning library developed by Google that facilitates the development and deployment of machine learning models. It is widely used for tasks such as data analysis, classification, and predictive modelling.
8. KConnect: A connector software utilized in data-centric architectures to establish seamless communication and integration between different data sources and systems.
9. Apache Ranger: An open-source tool that provides centralized security administration for Hadoop-based data systems. It enables fine-grained access control and security policies.

10. Apache Atlas: An open-source tool for metadata management and data governance in Hadoop-based ecosystems. It allows organizations to create, share, and manage metadata and data lineage information.
11. Kibana & Elastic Search: Tools used for creating user-centric applications. Kibana provides visualization capabilities, while Elastic Search offers efficient data retrieval, enhancing the user experience.
12. Kafka: A distributed streaming platform used for building real-time data pipelines and streaming applications. It ensures reliable and scalable data streaming between different components of a system.
13. Proof of Concept: A demonstration or experiment that validates the feasibility and functionality of a proposed model or system, typically involving simulated scenarios and real-world testing.
14. Digital Learning Universities: Institutions of higher education that leverage digital technologies, e-learning platforms, and online resources to enhance and deliver educational content to students.

1.8 Organization of the Thesis

This thesis is structured to facilitate a coherent exploration of the research objectives. Chapter 2 provides an extensive literature review, offering a foundation for understanding the complexities of digital learning, data architecture, and cybersecurity in various educational settings. Chapter 3 details the research methodology, outlining the design, participants, and data analysis procedures. Subsequent chapters delve into survey findings, comparative analysis, security model design, Secure Grade Distribution Scheme development, technological recommendations, and a proof of concept. The final chapters offer a comprehensive discussion of the implications, contributions, limitations, and future recommendations arising from this research.

This thesis is structured to facilitate a coherent exploration of the research objectives. Chapter 2 provides an extensive literature review, offering a foundation for understanding the complexities of digital learning, data architecture, and cybersecurity in various educational settings. Chapter 3 details the research methodology, outlining the design, participants, and data analysis procedures. Subsequent chapters delve into survey findings, comparative analysis, security model design, Secure Grade Distribution Scheme development, technological recommendations, and a proof of concept. The final chapters offer a comprehensive discussion of the implications, contributions, limitations, and future recommendations arising from this research.

CHAPTER TWO: LITERATURE REVIEW

2.1 Preamble

The literature review in this chapter unfolds as a comprehensive exploration of the foundational concepts essential to understanding the intricate interplay of digital learning, data architectures, security models, and grade distribution schemes within higher education. By delving into the evolving landscape of these domains, this chapter seeks to unearth critical insights, identify existing gaps, and establish a theoretical framework that aligns with the research objectives.

Education, particularly in higher institutions, has experienced a profound shift with the integration of digital technologies. To comprehend this transformative journey, the literature review embarks on an exploration of digital learning in higher education. The evolution, trends, challenges, and opportunities within this dynamic realm lay the groundwork for understanding the broader context in which data architectures, security models, and grade distribution schemes operate.

Moving seamlessly into the realm of data architectures, the review scrutinizes the existing models adopted by universities. This includes a nuanced examination of centralized and decentralized architectures, distributed databases, and data warehouses. Simultaneously, a spotlight is cast on the cybersecurity challenges inherent in these architectures, unravelling the intricacies of safeguarding digital assets against an evolving landscape of threats.

The subsequent sections navigate through security models in digital learning environments, surveying previous approaches and discerning their limitations. It is within this critical evaluation that the chapter aims to identify areas ripe for innovation and enhancement, contributing to the development of a robust security model. The exploration concludes with an in-depth analysis of grade distribution schemes, evaluating current practices and probing into the security concerns that underscore the dissemination of academic assessments in digital settings.

This chapter not only reviews relevant literature but also synthesizes and analyses it, providing a foundation for the subsequent research methodology and data analysis. As the narrative unfolds, the reader is invited to traverse the terrain of digital transformation in higher education, where the amalgamation of technology, security, and academic assessment converges, setting the stage for the novel contributions and insights that this research endeavours to unveil.

2.2 Theoretical Framework

2.2.1 Digital Learning in Higher Education

The landscape of higher education has undergone a profound transformation with the integration of digital learning methodologies. This section explores the evolution, trends, challenges, and opportunities that define the intricate relationship between technology and education in the higher academic sphere.

2.2.1.1 Evolution and Trends

The evolution of digital learning in higher education represents a transformative journey from its embryonic stages to the sophisticated models witnessed today. Early experiments with computer-assisted instruction set the foundation for the development of more advanced Learning Management Systems (LMS) and collaborative online platforms (Bervell & Umar, 2017). The initial focus on digitizing content gradually shifted towards more interactive and learner-centric approaches, emphasizing personalized educational experiences.

Historically, the evolution has been marked by a move from static, one-size-fits-all educational content to dynamic, adaptive learning technologies. Artificial Intelligence (AI) has played a crucial role in this evolution, offering the capability to analyze vast amounts of student data. AI-driven tools, such as personalized learning platforms, can tailor educational content to individual learning styles and pace (Bezovski & Poorani, 2016). This evolution represents a paradigm shift towards a more individualized and responsive educational environment.

Digital learning's roots can be traced back to early experiments with programmed instruction, with B.F. Skinner's teaching machine being a noteworthy example (Skinner, 1957). However, it was the advent of the internet and the subsequent development of online learning platforms that propelled digital learning into mainstream education. The introduction of Learning Management Systems (LMS), such as Blackboard and Moodle, marked a significant leap, providing a centralized platform for course management and content delivery (Allen & Seaman, 2017). This phase laid the groundwork for the subsequent evolution by bringing education into the digital realm.

As the 21st century unfolded, digital learning witnessed a shift towards more dynamic and interactive models. The concept of Massive Open Online Courses (MOOCs) emerged, offering scalable and accessible educational content to a global audience (Daniel, 2012) MOOCs exemplify the democratization of education, breaking down geographical barriers and providing learners with the flexibility to engage with course materials at their own pace. This period saw a transition from traditional, instructor-centred models to more learner-centric approaches, reflecting a growing understanding of the diverse needs and preferences of students in a digital age.

2.2.1.2 Trends in Digital Learning

The trends shaping digital learning in higher education are dynamic, responding to the evolving needs of learners and advancements in technology. Mobile learning, facilitated by the ubiquity of smartphones and tablets, has become a prevalent trend, providing learners with the flexibility to access educational content anytime and anywhere (Ferguson et al., 2019) This trend aligns with the societal shift towards a mobile-centric lifestyle, making education more accessible and convenient.

Moreover, the integration of immersive technologies has become a defining trend in digital learning. Virtual Reality (VR) and Augmented Reality (AR) are reshaping traditional learning environments. VR immerses learners in computer-generated scenarios, facilitating experiential learning in fields like science and medicine (Maurice et al., 2014). AR overlays digital content onto the real world, creating interactive learning experiences. These technologies contribute to a more engaging and interactive educational experience.

Simultaneously, social learning platforms have gained prominence as a trend in digital education. Platforms like Edmodo and Schoology provide spaces for collaborative learning, enabling students to interact, share resources, and engage in discussions beyond the confines of the physical classroom (Dabbagh & Kitsantas, 2012). Social learning leverages the power of online communities, fostering a sense of belonging and facilitating peer-to-peer learning.

The recent global response to the COVID-19 pandemic has accelerated existing trends. The widespread adoption of remote and online learning has become a dominant trend, emphasizing the importance of digital learning in ensuring continuity during times of disruption (Hodges et al., 2020). This trend has showcased the adaptability and resilience of digital learning models in the face of unprecedented challenges.

2.2.1.3 Challenges and Opportunities

Digital learning, while offering transformative possibilities, is not without its challenges. One significant obstacle is the existence of a digital divide, representing disparities in access to technology and the internet among different demographic groups (Warschauer & Matuchniak, 2010). Bridging this gap is crucial for achieving inclusive education, as learners without adequate access may be excluded from the benefits of digital learning. Initiatives to address the digital divide must be comprehensive, considering infrastructural, economic, and educational aspects.

Ensuring the quality of online education is another pressing challenge. The shift to virtual classrooms necessitates careful considerations regarding the effectiveness of digital learning experiences (Selwyn, 2016). Developing robust strategies for designing and delivering high-quality content, assessments, and interactive elements is essential. Moreover, issues related to digital literacy and the ability of learners to navigate online platforms can impact the overall quality of the learning experience (Mackey & Jacobson, 2011).

Teacher training stands out as a critical challenge in the digital learning landscape. Educators need to be equipped with the skills and knowledge to effectively leverage digital tools and technologies for teaching and learning. This involves not only technical proficiency but also pedagogical strategies that integrate digital resources seamlessly into the curriculum (Ertmer et al., 2012). Continuous professional development becomes imperative to ensure educators remain adept in the rapidly evolving digital landscape.

The rapid pace of technological advancements compounds the challenges in digital learning. Continuous adaptation is required not only from educators but also from institutions and policymakers to keep abreast of the latest innovations and best practices (Bates, 2019). Ensuring that educational institutions have the capacity and flexibility to integrate emerging technologies responsibly is a multifaceted challenge that requires strategic planning and collaboration.

Issues related to student engagement in virtual classrooms also merit attention. Maintaining a sense of connection and interaction in digital environments poses unique challenges. Strategies to enhance student engagement need to be explored, encompassing both synchronous and asynchronous elements of digital learning (Crompton, 2013). Balancing flexibility with structured engagement becomes crucial to foster a sense of community among learners.

Moreover, the diversity of digital learning tools and platforms introduces challenges related to standardization and interoperability. Integrating various technologies seamlessly into the learning environment can be complex and may require consistent standards to ensure a cohesive experience for both educators and learners (Bates, 2019). Achieving interoperability can enhance the efficiency and effectiveness of digital learning ecosystems.

2.2.1.4 Opportunities in Digital Learning:

Despite these challenges, digital learning presents unprecedented opportunities for the higher education sector. The recent global response to the COVID-19 pandemic showcased the resilience and adaptability of digital learning models (Hodges et al., 2020). Opportunities include expanded access to education, as learners can participate in courses and programs from anywhere in the world. The digital landscape also fosters collaboration and engagement through various online platforms, allowing for interactive and dynamic learning experiences.

Innovation in pedagogical approaches is a significant opportunity presented by digital learning. The flexibility of digital tools enables educators to explore new ways of delivering content, fostering critical thinking, and promoting active learning (Bates, 2019).

Virtual laboratories, simulations, and gamified elements provide avenues for experiential learning, enhancing the depth and breadth of educational experiences(Maurice et al., 2014). This opens possibilities for educators to tailor instruction to individual learning styles, catering to diverse student needs.

Additionally, digital learning facilitates the creation of inclusive learning environments, accommodating diverse learning styles and preferences. Customizable learning paths, adaptive assessments, and personalized feedback contribute to an individualized learning experience (Means & Neisler, 2021). This inclusivity extends beyond geographical boundaries, offering educational opportunities to learners who might face constraints in traditional settings.

The widespread use of analytics in digital learning platforms presents an opportunity to gather valuable insights into student performance and engagement. Learning analytics can inform educators about effective teaching strategies, areas where students may need additional support, and the overall effectiveness of the learning materials (Siemens & Long, 2011). Harnessing the power of data-driven insights allows for continuous improvement in educational practices.

Moreover, digital learning offers the potential for lifelong learning and continuous skill development. Online courses, micro-credentials, and digital badges provide avenues for learners to acquire new skills and knowledge throughout their lives, fostering a culture of continuous learning (Hodges et al., 2020). This aligns with the evolving needs of the workforce, where adaptability and upskilling are increasingly crucial.

2.3 Data Architectures in Higher Education

The orchestration of data architectures assumes a pivotal role in shaping the digital learning landscape within higher education institutions. This section intricately explores the diverse structures and frameworks that underpin the storage, management, and utilization of data in the complex realm of educational settings.

2.3.1 Existing Models

Within the realm of higher education, a rich tapestry of data architectures has unfolded, each offering a distinctive approach to addressing the intricate needs of digital learning environments. The centralized data architecture stands as a prominent model, characterized by a singular repository serving as the focal point for all educational data (Boh Podgornik et al., 2016). Praised for its simplicity in management and uniform data access, this model contributes to fostering a cohesive learning environment. However, potential challenges in scalability and adaptability may emerge, especially when confronted with the dynamic landscape of evolving digital learning technologies.

In contrast, federated data architectures introduce a decentralized paradigm, distributing data across multiple repositories maintained by distinct departments or units within a university (Guo & Zeng, 2020). This approach fosters autonomy in data governance but may encounter challenges related to data consistency and interoperability between disparate systems.

The advent of cloud-based data architectures has ushered in a new era, leveraging cloud computing services for the storage and processing of educational data (Al-Malah et al., 2021). Renowned for scalability, flexibility, and accessibility, cloud-based solutions empower institutions to swiftly adapt to changing demands. However, careful consideration is essential regarding data security, privacy, and the potential implications of relying on external service providers.

Graph-based data architectures have gained prominence for their adeptness in representing intricate relationships within educational datasets (Nakagawa et al., 2019). Graph databases excel in capturing connections between various data points, offering a nuanced understanding of student interactions, course dependencies, and institutional dynamics. This model aligns seamlessly with the emphasis on personalized and interconnected educational experiences in digital learning environments.

Hybrid data architectures have emerged as a pragmatic approach, combining elements of centralized, federated, and cloud-based models (Guo & Zeng, 2020). This versatile approach allows institutions to tailor their data architecture to specific needs, striking a harmonious balance between standardization and customization.

2.3.2 Cybersecurity Challenges

In the intricate landscape of data architectures within higher education, cybersecurity emerges as a paramount concern, shaping the resilience and integrity of digital learning environments. This section delves into the multifaceted cybersecurity challenges faced by existing data architectures, emphasizing the importance of safeguarding educational data.

Cybersecurity challenges within higher education data architectures are diverse and dynamic, demanding vigilant attention to mitigate potential risks. One significant challenge arises from the increasing sophistication of cyber threats targeting educational institutions (Joksimović et al., 2019). Malicious actors often exploit vulnerabilities within data architectures to gain unauthorized access, compromise sensitive information, or disrupt essential educational services. The rapid evolution of these threats necessitates continuous adaptation and proactive measures to ensure the security of educational data.

Data privacy concerns constitute a critical facet of cybersecurity challenges in higher education data architectures (Carvalho Ota et al., 2020). As educational institutions amass vast amounts of student and faculty data, ensuring compliance with data protection regulations becomes paramount. Unauthorized access, data breaches, or inadvertent disclosures can result in severe consequences, not only compromising individuals' privacy but also undermining institutional trust and reputation.

The interconnectedness of data architectures in digital learning environments amplifies the challenge of securing sensitive information. The sharing and exchange of data between various components of the educational ecosystem create potential vulnerabilities that malicious actors may exploit (Djeki et al., 2023). This interconnectedness necessitates a holistic approach to cybersecurity, addressing vulnerabilities at the system, network, and application levels.

Moreover, the increasing reliance on cloud-based data architectures introduces new cybersecurity challenges. While cloud solutions offer scalability and flexibility, they also expose educational institutions to risks associated with third-party service providers (Al-Malah et al., 2021). Data breaches, service disruptions, or inadequate security measures implemented by cloud providers can have profound implications on the confidentiality and availability of educational data.

Insider threats pose a significant cybersecurity challenge, highlighting the importance of robust internal controls and user awareness (Bhatia & Maitra, 2018). Individuals within educational institutions, intentionally or unintentionally, may compromise data security. Mitigating insider threats requires a combination of technical controls, employee training, and effective monitoring to detect and respond to suspicious activities.

The integration of emerging technologies, such as Internet of Things (IoT) devices and Artificial Intelligence (AI), into educational data architectures introduces additional cybersecurity challenges (Li et al., 2019). Ensuring the security of these technologies and their seamless integration with existing data architectures is crucial to prevent potential exploitation by cyber adversaries.

2.4 Security Models in Digital Learning Environments

The secure operation of digital learning environments relies on resilient security models that protect sensitive data and uphold the integrity of educational processes. This section delves into the evolution of security models within the realm of digital learning, with a specific focus on exploring vulnerabilities addressed by previous approaches to fortify the resilience of these environments.

2.4.1 Previous Approaches

Security models in digital learning environments have undergone transformations in response to dynamic cyber threats and the evolving landscape of educational technologies.

Previous approaches predominantly emphasized perimeter-based security, concentrating on fortifying external boundaries to thwart unauthorized access (Bhatia & Maitra, 2018). Utilizing firewalls, intrusion detection systems, and secure network configurations, this approach sought to create a secure perimeter around digital learning systems. However, it grappled with limitations, particularly in addressing internal threats and countering sophisticated cyber-attacks that could circumvent traditional perimeter defences.

Authentication and access control mechanisms constituted critical components of earlier security models in digital learning (Aissaoui & Azizi, 2017). Implementing user authentication through passwords, multi-factor authentication, and role-based access control aimed to ensure that only authorized individuals could access sensitive educational data and resources. Despite their effectiveness to a certain extent, these mechanisms faced challenges related to password vulnerabilities, user compliance, and the dynamic nature of user roles within educational institutions.

Cryptography played a fundamental role in previous security models, focusing on encrypting data to safeguard its confidentiality during both transmission and storage (Alassery, 2021). Widely used protocols such as Secure Sockets Layer (SSL) and Transport Layer Security (TLS) were employed to establish secure communication channels. However, challenges surfaced in maintaining cryptographic protocols up-to-date, securely managing cryptographic keys, and addressing vulnerabilities associated with encryption algorithms.

The shift towards a more holistic and adaptive security posture led to the development of risk-based security models in digital learning environments (Mihailescu et al., 2020). Rather than relying solely on predetermined security measures, risk-based models assess contextual factors, user behavior, and emerging threats to dynamically adjust security controls. This approach acknowledges the dynamic nature of digital learning environments and aims to strike a balance between security, usability, and adaptability.

Furthermore, the Zero Trust Security Model gained prominence as a response to the limitations of perimeter-based approaches (Arabi, 2021). In the Zero Trust model, trust is never assumed, and verification is required from anyone attempting to access resources, even within the internal network. This approach minimizes the potential impact of insider threats, operating on the assumption that the internal network is as untrusted as external networks.

2.4.2 Limitations and Gaps

While previous security approaches in digital learning environments have made strides in fortifying the integrity of educational processes, they are not without their limitations and identifiable gaps. This section scrutinizes the vulnerabilities and shortcomings inherent in these models, shedding light on areas that demand further attention and innovation.

One significant limitation lies in the reliance on perimeter-based security models, which, while providing a degree of protection, struggle to address internal threats effectively (Bhatia & Maitra, 2018). The traditional emphasis on securing external boundaries often leaves digital learning environments susceptible to insider threats and sophisticated attacks that navigate through the established defences. The permeability of these perimeters poses a persistent challenge in safeguarding against threats originating within the educational ecosystem.

Authentication and access control mechanisms, while essential, grapple with vulnerabilities tied to human behaviour and compliance (Aissaoui & Azizi, 2017). Password-based authentication remains susceptible to issues such as weak password practices, password reuse, and the challenge of enforcing robust password policies across diverse user groups. Additionally, as user roles evolve within educational institutions, maintaining an accurate and dynamic representation of access privileges becomes an ongoing challenge, leading to potential security gaps.

Cryptography, a stalwart in data protection, encounters limitations in managing the complexity of cryptographic protocols and keys (Alassery, 2021). Keeping cryptographic protocols up to date with emerging standards and mitigating vulnerabilities associated with encryption algorithms necessitate continuous attention. The secure management of cryptographic keys, crucial for maintaining the confidentiality of data, demands robust practices to prevent unauthorized access and potential compromise.

Risk-based security models, while providing adaptability, introduce complexities in assessing and responding to dynamic contextual factors (Mihailescu et al., 2020). The effectiveness of these models hinges on accurate risk assessments, which can be challenging given the evolving nature of cyber threats and the intricate interplay of factors influencing security postures. Striking the right balance between security measures, usability, and adaptability remains a delicate challenge in the implementation of risk-based models.

The Zero Trust Security Model, although a paradigm shift, presents challenges in practical implementation and cultural adaptation within educational institutions (Arabi, 2021).. Overcoming ingrained trust assumptions and seamlessly integrating the Zero Trust approach into existing digital learning environments requires comprehensive planning and organizational readiness.

Moreover, the increasing integration of emerging technologies, such as the Internet of Things (IoT) and Artificial Intelligence (AI), poses new challenges in terms of security (Lu & Da Xu, 2018). Ensuring the security of these technologies and their harmonious coexistence with existing security models demand continuous vigilance and adaptation.

2.4.3 Review of Related Work

In the exploration of security models within digital learning environments, an exhaustive review of related literature has been conducted, encompassing the below notable works that provide diverse insights into the development, challenges, and advancements in this critical domain.

Security models and frameworks are important in both e-learning and university systems to ensure the protection of sensitive information and prevent cyber threats. In the context of e-learning, the use of security and cyber security countermeasures has been found to have a significant impact on students' frequent use and participation in the system (Kale et al., 2023). Additionally, student feedback and communication about their e-learning experience can help address security concerns and increase participation(Al-Sherideh et al., 2023). In the context of university automation systems, information security frameworks have been proposed to ensure the overall safety of the system(Hasan et al., 2022). These frameworks aim to protect the information of different confidentiality levels and ensure sustainable information security (Ramanauskaitė et al., 2021). Furthermore, a security concept model has been developed for distance learning, consisting of security assurance, users, and organizational processes, with technical measures provided at the system administrator level.

In conclusion, these works collectively contribute a nuanced understanding of security models in digital learning environments, each offering valuable insights, methodologies, and recommendations. While presenting significant advancements, these studies also acknowledge inherent limitations, underscoring the evolving nature of research in this dynamic field (Table 1).

Table 1 Summary of Related Work 1

Reference	Objectives	Method	Contribution	Limitation
(Kale et al., 2023)	The provided paper proposes an information security framework for a university automation system	Blog crawling and Traditional document search	• Identification of shared resources in university automation system. • Proposal of an information security framework for overall system safety.	The research did not provide secure data storage.
(Al-Sherideh et al., 2023)	The paper focuses on developing e-learning security model, the impact of security measures on students' academic achievements.	Collecting information on e-learning security measures and students' perceptions. Designing a questionnaire and selecting the right sample of respondents.	Assessing the impact of security measures on students' academic achievements. Development of a security model to detect cyberattacks.	The model only detect attack does not prevent it
(Hasan et al., 2022)	The provided paper proposes an information security framework for a university automation system.	Identification of shared resources in university automation system Proposal of information security framework for overall system safety	Identification of shared resources in university automation system. Proposal of an information security framework for overall system safety.	The model did not provide data security for the university

(Jusas et al., 2022)	The paper provides a framework for implementing security systems in e-learning environment	Explaining and establishing frameworks for implementing security systems and Modelling threats posed by a malicious hacker	Identification of shared resources in university automation system. Proposal of an information security framework for overall system safety.	Lack of proper mitigation technique
(Ramanauskaitė et al., 2021)	The paper proposes a security level estimation model for educational organizations	The article developed a security level estimation model for educational organizations.	Proposal of a security level estimation model for educational organizations. Validation of the proposed model through use case analysis and expert evaluation.	Lack of security level modelling for educational organizations.
(Modesti, 2020)	The provided paper proposes an information security framework for a university automation system.	Integration of formal methods for security research into teaching practice and adoption of a conceptual model aligned with high-level representation of cryptographic and communication primitives	Identification of shared resources in university automation system. Proposal of an information security framework for overall system safety.	No efficient method of preventing identified cyberattacks

2.5 Grade Distribution Schemes

Grade distribution schemes play a pivotal role in assessing and communicating student performance within educational institutions. This section examines the current practices employed in grade distribution schemes, shedding light on the methodologies and frameworks used to evaluate and disseminate academic achievements.

2.5.1 Current Practices

In contemporary educational landscapes, various institutions employ diverse grade distribution schemes to represent student accomplishments fairly and transparently. The prevalent approach involves the utilization of letter grades, ranging from A to F, each corresponding to a specific level of achievement. This traditional system allows for a straightforward classification of performance, with 'A' typically denoting excellent performance and 'F' indicating failure.

Percentage-based grading is another widely adopted practice wherein students receive a numerical score reflective of their performance relative to the maximum achievable score. This system provides a granular representation of achievement, allowing for subtle distinctions in performance. However, challenges may arise in cases of standardized testing, where a fixed scale may not align with the difficulty level of different assessments.

Some institutions embrace a pass/fail system, simplifying the evaluation process by categorizing students as either passing or failing without assigning specific grades. This approach aims to reduce academic stress and foster a focus on learning rather than grades. However, it may lack the granularity necessary for detailed academic assessments.

In recent years, competency-based grading has gained traction, emphasizing the mastery of specific skills and knowledge rather than traditional letter or percentage grades. This approach is particularly prevalent in competency-driven programs and is designed to provide a more nuanced understanding of a student's capabilities.

Additionally, narrative evaluations offer a qualitative alternative, providing detailed written assessments of a student's performance. This personalized approach allows instructors to offer tailored feedback, emphasizing strengths and areas for improvement. While narrative evaluations can offer a comprehensive view of a student's progress, they may lack the standardization and comparability of more quantitative systems.

The shift towards digital learning and assessment tools has prompted the exploration of automated grading systems. These systems utilize algorithms to assess assignments, quizzes, and exams, providing quick feedback to students. While efficient, concerns persist regarding the potential limitations in capturing the nuanced aspects of student performance that may be evident in qualitative assessments.

2.5.2 Security Concerns

As educational institutions continue to embrace digital platforms for grade distribution, a set of security concerns emerges, demanding vigilant attention to safeguard the integrity and confidentiality of academic assessments. This section scrutinizes the potential security challenges associated with digital grade distribution schemes, emphasizing the need for robust measures to mitigate risks.

One primary security concern revolves around data breaches and unauthorized access to grade databases. The digital storage and transmission of sensitive student information necessitate stringent measures to prevent malicious actors from gaining unauthorized access. Encryption protocols and secure authentication mechanisms become imperative to safeguard against unauthorized intrusion and data tampering.

The potential manipulation of grades poses another significant security threat. Malicious actors may attempt to alter grades either for personal gain or to create disruptions within the educational system. Ensuring the integrity of the grade distribution system requires implementing measures such as digital signatures and audit trails to detect and prevent unauthorized grade changes.

Phishing attacks targeting students or faculty members represent a tangible threat to the confidentiality of grade information. Cybercriminals may employ deceptive tactics to trick individuals into divulging login credentials, enabling unauthorized access to the grade distribution platform. Educational institutions must prioritize cybersecurity awareness programs to mitigate the risks associated with social engineering attacks.

The reliance on third-party grading platforms introduces concerns related to the security practices of external service providers. Educational institutions must rigorously assess and monitor the security measures implemented by these platforms to ensure compliance with data protection regulations and prevent potential vulnerabilities that may compromise student data.

The potential for distributed denial-of-service (DDoS) attacks on grade distribution systems poses a disruptive threat. An orchestrated DDoS attack could overwhelm the system, rendering it temporarily inaccessible and causing disruptions during critical grading periods. Implementing robust network infrastructure and DDoS mitigation strategies becomes crucial to maintain system availability and resilience.

Moreover, the vulnerability of automated grading systems to algorithmic biases and errors raises ethical and security concerns. Biases in algorithms may disproportionately impact certain student groups, leading to unfair assessments. Ensuring transparency in algorithmic decision-making and regularly auditing automated grading systems can address these ethical and security considerations.

2.5.3 Review of Related Work

In this (Plyer et al., 2022), the authors created a unique method for grading chemistry examinations in Moodle. Their plugin can properly grade chemistry tests, and the mark is safely stored in Moodle. Other authors (Pérez et al., 2017) suggested a method for detecting any modification of Moodle grades and alerting the users in charge to maintain the grades' security. The article focuses on SQL injection, a code injection attack that targets data-driven systems that introduce malicious SQL statements into a field for execution. The suggested solution may detect student grade changes and inform the instructor. It was created using PHP. However, the research was limited to detecting SQL injection and did not include prevention methods. (Abdelsalam et al., 2023) In their study aimed to enhance the security of Moodle's grade distribution system. They proposed a new encryption scheme to protect grade data during transmission. The research introduces cryptographic techniques to safeguard sensitive information. (Cyoy, 2022) focuses on implementing two-factor authentication in Moodle to ensure secure access to grade-related information. The study explores methods to add an extra layer of protection to prevent unauthorized access to student grades. (Korać et al., 2022b) investigated the vulnerabilities of Moodle's gradebook and proposed strategies to strengthen its security. The study delves into potential threats and provides recommendations to address weaknesses in the Moodle platform's grade management system. (Elmaghrabi & Eljack, 2019) provided a comprehensive review of existing security measures in Moodle's grade distribution is presented in this research. The authors analyze the strengths and weaknesses of current methods and suggest improvements to enhance overall system security (Table 2).

Table 2. Summary of Related Work 2

Reference	Objective	Contribution	Limitations
(Pérez et al., 2017)	The objective of the research is to prevent changes in student grades in the Moodle platform	The study suggested a solution that will detect any change in a student's status and inform the instructor of it.	The research has limitations in detecting SQL injection and did not include prevention methods. The research only provides means of detecting changes in grades, not preventing them.
(Plyer et al., 2022)	Providing a new grading method for chemistry exams and safe grade storage inside the Moodle platform is the primary objective of the work.	The study developed and installed a Moodle plugin for grading chemistry examinations.	The research did not develop any security technique for preventing data breaches in grades in the Moodle platform.
(Abdelsalam et al., 2023)	Enhancing the security of Moodle's grade distribution system using a new encryption scheme	Introduction of cryptographic techniques to safeguard sensitive information	The research did not provide a secured way of sharing grades with staff and students
(Cyoy, 2022)	Implementing two-factor authentication in Moodle to secure access to grade-related information	Exploration of methods to add an extra layer of protection	The research did not provide encryption for student's grades.
(Korać et al., 2022b)	Investigating vulnerabilities in Moodle's gradebook and proposing strategies for improvement	In-depth analysis of potential threats and recommendations	Limited information on the practical implementation of suggested strategies
(Elmaghrabi & Eljack, 2019)	Reviewing existing security measures in Moodle's grade distribution	Analysis of strengths and weaknesses, suggestions for improvements	Lack of empirical testing for proposed enhancements

CHAPTER THREE: METHODOLOGY

3.1 Preamble

The research design serves as the architectural framework that guides the systematic inquiry into the multifaceted aspects of a secure data-centric model for digital learning in higher education, particularly within the context of West African universities. This section delineates the blueprint and methodology employed to rigorously explore the existing data architectures, cybersecurity challenges, and grade distribution schemes prevalent in the evolving landscape of digital education.

A judicious selection of a mixed-methods research design is deemed imperative for its capacity to amalgamate the strengths of both qualitative and quantitative methodologies. This combination facilitates a comprehensive and nuanced investigation into the intricacies of data security and digital learning. The qualitative dimension unfolds through in-depth interviews, engaging key stakeholders to extract rich narratives and perspectives. Concurrently, the quantitative facet leverages surveys and statistical analyses to quantify prevailing trends, assess the efficacy of security measures, and gauge satisfaction levels with current grade distribution systems.

The integrity and robustness of the research endeavour hinge on the meticulous design and calibration of data collection instruments. Interview guides, tailored to the intricacies of data architectures and digital learning, are poised to elicit profound insights. The structured questionnaire for surveys draws on validated measures to ensure the reliability and validity of the gathered quantitative data.

A purposive sampling strategy has been strategically devised to assemble a diverse cohort of participants, encompassing educators, IT administrators, and students from West African universities. This approach seeks to capture a spectrum of perspectives reflective of the regional nuances and challenges inherent in the subject matter.

As ethical considerations stand as a cornerstone of responsible research, this section underscores the commitment to upholding ethical standards. Rigorous adherence to participant confidentiality, informed consent, and responsible data handling are paramount.

The research protocol will be subjected to scrutiny and approval by the pertinent institutional review board, affirming the ethical rigor of the study.

3.2 Problem Formulation

The formulation of the research problem serves as the compass directing the inquiry into the secure data-centric model for digital learning in higher education, particularly within the dynamic context of West African universities. This section meticulously defines and articulates the challenges and gaps that motivate the research, providing a clear trajectory for the investigation.

3.2.1 Data Architectures

In the intricate landscape of higher education, the efficacy and security of data architectures stand as pivotal determinants of the digital learning experience. This subsection delves into the nuanced realm of data architectures within West African universities, aiming to meticulously identify vulnerabilities that may compromise the integrity, accessibility, and confidentiality of academic information.

3.2.2 Systemic Analysis of Existing Data Architectures

The foundation of any digital learning environment lies in its data architecture. Through a comprehensive survey, this research endeavours to conduct a systemic analysis of the prevailing data architectures across West African universities. This includes an exploration of the infrastructure, databases, and data storage mechanisms employed, with a keen focus on understanding their design principles and implementation intricacies.

I. Examination of Cybersecurity Incidents

A critical dimension of identifying vulnerabilities is an in-depth examination of past cybersecurity incidents. By scrutinizing the historical landscape, this research aims to catalogue and analyse instances of cyber-attacks faced by West African universities. Understanding the modus operandi of these incidents is imperative for uncovering potential weak points within data architectures and formulating targeted strategies for fortification.

II. Evaluation of Countermeasures Implemented

In tandem with identifying vulnerabilities, an assessment of the countermeasures implemented by universities becomes paramount. This research will investigate the proactive measures taken by institutions to mitigate and respond to cybersecurity challenges. The evaluation extends beyond technological solutions to encompass policies, training programs, and organizational protocols designed to bolster the resilience of data architectures.

III. Regional and Institutional Variances

Recognizing the diversity across West African universities, this research will be attentive to regional and institutional variances in data architectures. Understanding the unique challenges faced by different institutions ensures that interventions and recommendations are contextually relevant. Factors such as infrastructure limitations, resource availability, and regional threat landscapes will be considered in this nuanced analysis.

IV. Integration of Lessons Learned from the COVID-19 Pandemic

The paradigm shift in educational practices catalyzed by the COVID-19 pandemic has underscored the significance of robust data architectures. Lessons learned from this transformative period will be integrated into the identification of vulnerabilities, considering the specific challenges and adaptations made by West African universities during this global crisis.

3.2.3 Cybersecurity Challenges during the Pandemic

The unprecedented shift to remote and digital learning catalyzed by the COVID-19 pandemic has brought forth a myriad of cybersecurity challenges within the realm of higher education. This subsection delves into the multifaceted landscape of cybersecurity challenges faced by West African universities during the pandemic, aiming to unravel the intricacies of digital vulnerabilities and potential threats to academic data.

I. Surge in Phishing and Social Engineering Attacks

With the surge in digital communication channels, the pandemic ushered in an alarming increase in phishing and social engineering attacks. Malicious actors exploited the uncertainties and urgency surrounding the pandemic, targeting students, faculty, and administrators. This research will dissect the methodologies employed in these attacks, shedding light on the vulnerabilities exposed within the digital communication infrastructure of universities.

II. Scalability Issues and Technological Gaps

The abrupt transition to digital learning exposed scalability issues and technological gaps in existing cybersecurity infrastructures. West African universities faced challenges in scaling up their security measures to accommodate the sudden influx of online activities. This research seeks to identify the specific technological gaps and scalability bottlenecks that impeded effective cybersecurity responses during the pandemic.

III. Data Privacy Concerns in Remote Learning Environments

Remote learning, while crucial for continuity, introduced concerns regarding data privacy. The transition to online platforms for lectures, examinations, and collaborative projects raised questions about the protection of sensitive student information. This research will investigate the extent of data privacy concerns, examining the adequacy of existing measures and proposing strategies to enhance the safeguarding of student data.

IV. Adapting to Evolving Cyber Threats

The dynamic nature of cyber threats demands continual adaptation from educational institutions. This research aims to analyze how West African universities adapted to evolving cyber threats during the pandemic. It will explore the agility of existing cybersecurity frameworks, the integration of threat intelligence, and the responsiveness of incident response mechanisms.

V. Impact of Increased Network Traffic

The surge in online activities during the pandemic led to a substantial increase in network traffic within educational institutions. This subsection will explore the impact of heightened network usage on cybersecurity, assessing the resilience of network infrastructures, potential bottlenecks, and strategies employed to maintain network security while accommodating increased demand.

VI. Collaboration and Communication Security

As collaboration tools became integral to remote learning, ensuring the security of communication channels and collaborative platforms became paramount.

This research will scrutinize the cybersecurity challenges associated with the adoption of virtual communication tools, emphasizing the need for secure channels for academic discourse and collaboration.

3.2.4 Grade Distribution Scheme Vulnerabilities

The distribution of grades is a cornerstone of academic evaluation, and the transition to digital learning platforms has brought forth a unique set of vulnerabilities within grade distribution schemes. This subsection meticulously examines the vulnerabilities inherent in the systems responsible for disseminating student grades in West African universities, aiming to fortify the confidentiality, accuracy, and overall integrity of the grading process.

I. Integrity and Authenticity of Digital Grade Repositories

The digitalization of grade repositories introduces challenges related to the integrity and authenticity of academic records. This research will scrutinize the vulnerabilities associated with the storage and management of digital grades, including the risk of unauthorized access, tampering, or manipulation. Strategies for ensuring the trustworthiness of these repositories will be explored.

II. Potential Exploitation of Online Examination Systems

With the surge in online examinations, concerns arise regarding the potential exploitation of these systems. This research delves into the vulnerabilities associated with digital examination platforms, including the risk of cheating, impersonation, or manipulation of examination results. Strategies for enhancing the security of online examination systems will be considered.

III. Privacy Concerns in Digital Grade Transmission

The transmission of grades in digital formats raises privacy concerns, especially regarding the secure and confidential communication of academic results. This subsection explores vulnerabilities in the transmission process, including the risk of interception or unauthorized access. Recommendations for ensuring encrypted and secure grade transmission will be proposed.

IV. Accessibility and Inclusivity Challenges

While digital grade distribution offers convenience, it may inadvertently introduce accessibility challenges. This research examines vulnerabilities related to the inclusivity of digital grade distribution systems, considering factors such as internet access disparities, technological barriers, and the potential exclusion of certain student groups. Strategies for fostering inclusivity will be addressed.

V. Technological Infrastructure Limitations

The effectiveness of digital grade distribution is contingent on the technological infrastructure supporting it. This research investigates vulnerabilities arising from technological limitations, such as server downtimes, bandwidth constraints, or compatibility issues. Recommendations for bolstering technological resilience in grade distribution schemes will be explored.

VI. Risk of Algorithmic Biases in Automated Grading

The adoption of automated grading systems introduces the risk of algorithmic biases. This subsection explores vulnerabilities related to the fairness and impartiality of automated grading algorithms, considering potential disparities in grading outcomes based on demographic or contextual factors. Strategies for mitigating algorithmic biases will be scrutinized.

3.2.5 Impact on Digital Learning Experience

The dynamic integration of digital technologies into higher education has redefined the learning experience, yet this transformation is not without challenges. This subsection delves into the multifaceted impacts on the digital learning experience within the context of West African universities. By examining both positive and negative implications, the research seeks to provide a holistic understanding of the consequences of the digital shift on students, educators, and the academic ecosystem.

I. Enhanced Accessibility and Flexibility

One of the positive impacts of the digital learning experience is the enhanced accessibility and flexibility it offers. Students can engage with educational materials and participate in classes from virtually anywhere. This subsection will explore how these advantages have positively influenced the learning experience, enabling more inclusive and flexible educational practices.

II. Challenges in Technological Adaptation

Conversely, the rapid shift to digital learning has brought about challenges in technological adaptation. This research will scrutinize how students and educators navigate the learning curve associated with digital tools, examining potential barriers and disparities in technological proficiency that may affect the overall learning experience.

III. Interactive Learning Opportunities

Digital learning platforms often facilitate interactive learning opportunities through forums, collaborative projects, and virtual discussions. This subsection aims to highlight the positive impact of these interactive elements on student engagement, knowledge retention, and the overall quality of the learning experience.

IV. Social Isolation and Reduced Engagement

On the flip side, the digital learning experience has been associated with social isolation and reduced engagement for some students. This research will investigate the impact of virtual learning environments on social interactions, community-building, and the sense of belonging within the academic community.

V. Adaptation of Pedagogical Approaches

The adoption of digital tools has prompted a reconsideration of pedagogical approaches. This subsection will delve into how educators have adapted their teaching methods to the digital landscape, exploring innovations in online teaching, assessment strategies, and the integration of multimedia resources.

VI. Digital Fatigue and Cognitive Overload

Digital learning, if not well-managed, can contribute to digital fatigue and cognitive overload. This research aims to understand the negative impact of prolonged screen time, constant connectivity, and information overload on students and educators, exploring strategies to mitigate these challenges.

VII. Opportunities for Lifelong Learning

The digital learning experience opens avenues for lifelong learning and continuous skill development. This subsection will explore how digital platforms have facilitated ongoing education, professional development, and the acquisition of new skills beyond traditional academic settings.

VIII. Impact on Academic Performance

The research will assess the impact of the digital learning experience on academic performance. This includes an analysis of the correlation between digital engagement, grades, and overall student success, providing insights into the effectiveness of digital learning methodologies.

3.3 Proposed Solution

Addressing the identified vulnerabilities and challenges within the digital learning landscape requires a strategic and robust approach. This section outlines a proposed solution designed to fortify data architectures, enhance cybersecurity measures, and ensure the integrity of grade distribution schemes within West African universities.

3.3.1 Comprehensive Security Model

In response to the identified vulnerabilities within West African universities' digital learning landscapes, a Comprehensive Security Model is proposed. This model aims to establish a robust and adaptive security infrastructure, safeguarding data architectures, and fortifying against cyber threats (Figure 1). The components of this model include:

Encryption Protocols: To ensure the confidentiality and integrity of data, the implementation of advanced encryption protocols is paramount. This involves encrypting data both in transit and at rest, utilizing industry-standard algorithms. The adoption of encryption mechanisms will secure sensitive information from unauthorized access or tampering.

Multi-Factor Authentication (MFA): Enhancing user authentication is critical to thwarting unauthorized access attempts. MFA, incorporating factors such as passwords, biometrics, or security tokens, adds an additional layer of protection. This reduces the risk of compromised user credentials and strengthens overall system security.

Role-Based Access Control (RBAC): offers a robust framework that brings about numerous advantages to security management within systems. Its structured approach simplifies administration by consolidating access control under predefined roles, mitigating the complexities of assigning individual permissions. This not only streamlines the process but also significantly reduces the likelihood of human error and unauthorized access. RBAC enhances security by precisely aligning permissions with job functions, fostering a principle of least privilege where users only gain access necessary for their roles, minimizing potential vulnerabilities. Its scalability and adaptability empower organizations to efficiently manage access rights, seamlessly accommodating changes in personnel or roles by adjusting role assignments. RBAC stands as a cornerstone of access control, promoting a balance between stringent security measures and operational efficacy.

Regular Security Audits: A proactive approach is taken through regular and thorough security audits. These audits assess the effectiveness of existing security measures, identify potential vulnerabilities, and ensure compliance with cybersecurity best practices. Continuous monitoring and assessment contribute to a dynamic and resilient security posture.

Incident Response Plan:

Preparation for cybersecurity incidents is addressed by formulating a well-defined incident response plan. This plan outlines the steps to be taken in the event of a security breach, ensuring a swift and coordinated response to mitigate potential damages. Regular drills and updates refine the incident response strategy.

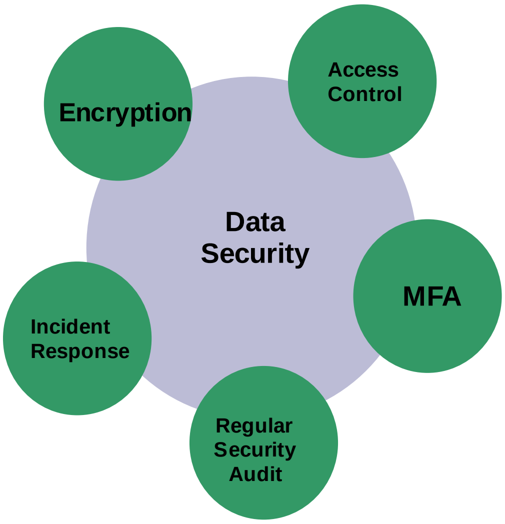


Figure 1: Comprehensive Security Model

3.3.2 Adaptive Data Architecture

In response to the diverse landscape of West African universities, an Adaptive Data Architecture is proposed to fortify the foundations of digital learning environments. This framework is designed to ensure scalability, resilience, and efficiency in managing academic data (Figure 2). The key components of the Adaptive Data Architecture include:

Scalable Infrastructure: Recognizing the fluctuating demands of digital learning, the architecture incorporates a scalable infrastructure. This entails the ability to dynamically adjust resources to accommodate variations in user activity, ensuring optimal performance during peak usage periods and efficient resource utilization during low-demand periods.

Cloud Integration: Strategic integration of cloud technologies forms a pivotal element of the Adaptive Data Architecture. Leveraging cloud services facilitates enhanced storage capacity, accessibility, and seamless data backup capabilities. This integration provides flexibility, scalability, and cost-effectiveness in managing academic data.

Redundancy Measures: To mitigate the impact of system failures or cyber-attacks, the architecture incorporates redundancy measures. Redundancy ensures that critical data is duplicated and distributed across multiple servers or locations, reducing the risk of data loss and enhancing overall system reliability.

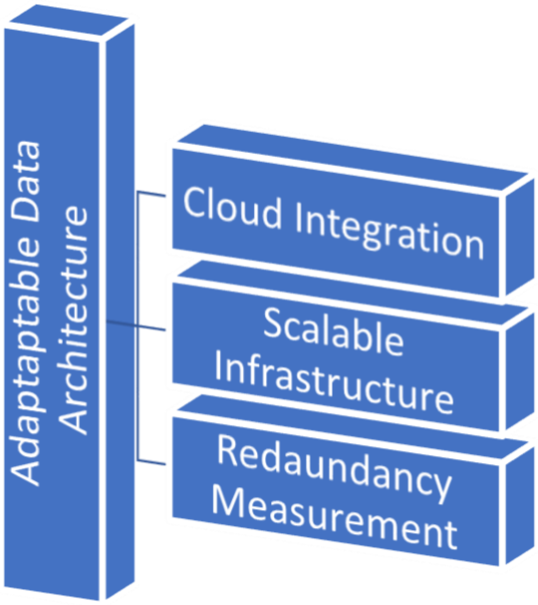


Figure 2: Adaptable Data Architecture

3.3.3 Secure Grade Distribution Scheme

In the pursuit of fortifying the grade distribution process within West African universities, a Secure Grade Distribution Scheme is proposed. This scheme integrates advanced cryptographic techniques, including AES (Advanced Encryption Standard), HSM (Hardware Security Module), Diffie-Hellman key exchange, and MIC (Message Integrity Check) verification. The amalgamation of these elements aims to ensure the confidentiality, integrity, and secure transmission of academic grades.

AES Encryption: The utilization of AES encryption stands as a foundational element in securing grade distribution. AES, a symmetric encryption algorithm, ensures the confidentiality of transmitted grades. Each grade is encrypted using a unique key, providing a robust defence against unauthorized access or tampering.

Hardware Security Module (HSM): To elevate the security posture, a Hardware Security Module is integrated into the scheme. HSM serves as a secure enclave for storing cryptographic keys, preventing unauthorized access. By utilizing HSM, the scheme enhances key management practices, safeguarding encryption keys from potential vulnerabilities.

Diffie-Hellman Key Exchange: The Diffie-Hellman key exchange protocol is incorporated to establish secure communication channels. This ensures that the encryption keys are exchanged securely between the sender and receiver without the risk of interception. Diffie-Hellman enhances the confidentiality of the grade distribution process.

Message Integrity Check (MIC) Verification: MIC verification is employed to ensure the integrity of transmitted grades. This involves attaching a cryptographic hash value to each grade, allowing the recipient to verify the authenticity of the received data. MIC verification acts as a crucial defence against any unauthorized alterations during transmission.

3.3.3.1 Benefits of the Secure Grade Distribution Scheme:

1. Confidentiality: AES encryption ensures that grades are transmitted confidentially and securely.

2. Key Protection: HSM safeguards cryptographic keys, reducing the risk of key compromise and unauthorized access.
3. Secure Communication Channels: Diffie-Hellman key exchange establishes secure channels for key transmission, enhancing overall communication security.
4. Data Integrity: MIC verification provides a robust mechanism for ensuring the integrity of transmitted grades, minimizing the risk of tampering.

3.3.3.2 Implementation Considerations:

1. Key Management: A robust key management strategy ensures the secure generation, distribution, and storage of encryption keys.
2. User Authentication: Implementing strong user authentication mechanisms ensures that only authorized individuals have access to grade distribution processes.

3.4 Tools used in the implementation.

In the implementation phase of this research, various tools were employed to facilitate the development and assessment of the proposed data-centric architecture and security model. The following tools played a crucial role in ensuring the effectiveness and reliability of the implemented solutions:

1. KConnect: KConnect served as the primary connector software in the data-centric architecture. Its robust capabilities and compatibility made it an ideal choice for seamlessly integrating diverse data sources and ensuring efficient communication between different components of the architecture.
2. Apache Ranger and Apache Atlas: These tools were instrumental in enforcing data governance and security measures within the architecture. Apache Ranger provided fine-grained access control and policy enforcement, while Apache Atlas facilitated metadata management and lineage tracking, enhancing overall data governance.
3. Hybrid Storage Solution: The implementation leveraged a hybrid storage approach, combining various storage solutions such as Network-Attached Storage (NAS), Storage Area Network (SAN), Cloud Storage, Local Storage, and Tape Storage. This

combination was essential for optimizing performance, ensuring cost efficiency, and enhancing data availability in the digital learning environment.

5. TensorFlow: TensorFlow, a powerful open-source machine learning library, was utilized for data analytics within the architecture. Its versatility and scalability allowed for the development of advanced analytics models, contributing to informed decision-making based on the analyzed educational data.
6. Kibana & Elastic Search: These tools were employed in the creation of user-centric applications. Kibana, with its visualization capabilities, and Elastic Search, providing efficient data retrieval, collectively enhanced the user experience by delivering intuitive and responsive applications.
7. Kafka: Kafka played a pivotal role in the implementation of the messaging system, ensuring seamless communication and data transfer between different components of the architecture. Its distributed and fault-tolerant nature contributed to the reliability of data streaming processes.
8. Data as a Service (DaaS): The concept of Data as a Service was integrated into the architecture, allowing for on-demand access to data. This facilitated efficient data sharing and collaboration among users, contributing to a more dynamic and interactive learning environment.

These tools collectively formed a cohesive and integrated technological ecosystem, enabling the development and deployment of a robust data-centric architecture tailored to the unique requirements of digital learning universities in West Africa. Their functionalities spanned from ensuring data security and governance to optimizing data storage, analytics, and user-centric applications, contributing to the overall success of the implemented model.

3.5 Approach and Technique(s) for the Proposed Solution

The successful implementation of the proposed Comprehensive Security Model, Adaptive Data Architecture, and Secure Grade Distribution Scheme within West African universities necessitates a systematic approach and the application of specific techniques. The chosen approach is structured and methodical, incorporating a series of well-defined steps and techniques tailored to each component of the overall solution.

3.5.1 Comprehensive Security Model

3.5.1.1 Approach:

1. Risk Assessment: Conduct a thorough risk assessment to identify potential threats and vulnerabilities specific to the digital learning environment of West African universities. Collaborate with cybersecurity experts and stakeholders to comprehensively analyze the risk landscape.
2. Baseline Security Measures: Establish baseline security measures by defining access controls, implementing network segmentation, and ensuring regular security updates across the digital infrastructure. Develop a foundational security framework to address common vulnerabilities.
3. Implementation of SIEM: Integrate a Security Information and Event Management (SIEM) system to centralize and analyze security event data in real-time. Configure SIEM to collect and correlate logs from various sources, allowing for proactive threat detection and response.
4. Endpoint Protection Deployment: Deploy advanced endpoint protection solutions on all devices within the digital learning environment. These solutions should encompass antivirus, anti-malware, and intrusion detection capabilities to safeguard against a spectrum of cybersecurity threats.

3.5.1.2 Techniques:

Penetration Testing: Conduct regular penetration testing exercises to simulate cyber-attacks and identify potential weaknesses in the security infrastructure. This technique provides insights into system vulnerabilities and ensures proactive security enhancements.

Continuous Monitoring: Implement continuous monitoring using SIEM tools to detect and respond to security incidents promptly. Real-time analysis of security events enables rapid incident response, reducing the impact of potential cyber threats.

Security Awareness Training: Provide security awareness training for faculty, staff, and students to promote a culture of cybersecurity. Educate users on recognizing phishing attempts, following secure practices, and reporting security incidents.

3.5.1.3 Benefits of the Approach and Techniques:

- Holistic Security Coverage: The approach ensures a holistic security coverage by addressing risks comprehensively and implementing baseline measures across the digital learning environment.
- Proactive Threat Mitigation: Techniques such as penetration testing and continuous monitoring contribute to proactive threat mitigation, allowing for the identification and remediation of security issues before they escalate.
- User Empowerment: Security awareness training empowers users to actively contribute to the security posture, fostering a collaborative approach to cybersecurity within the academic community.

3.5.2 Adaptive Data Architecture

3.5.2.1 Approach:

1. Data Classification: Initiate a comprehensive data classification process to categorize information based on sensitivity. This process will inform the implementation of security measures commensurate with the importance and confidentiality of the data.
2. Cloud Integration Strategy: Develop a strategic plan for seamless integration with reputable cloud service providers, such as Amazon Web Services (AWS) or Microsoft Azure. This strategy ensures scalable storage, efficient resource management, and reliable data backup capabilities.
3. Containerization Implementation: Implement containerization platforms like Docker or Kubernetes to enhance the deployment and scalability of applications within the adaptive data architecture. Containerization facilitates efficient resource utilization and accelerates the development lifecycle.
4. Load Balancing Configuration: Configure load balancing solutions, utilizing tools like HAProxy or Nginx, to optimize the distribution of network traffic. This ensures that the digital learning environment maintains optimal performance even during varying workloads.

3.5.2.2 Techniques:

- Data Encryption: Implement robust encryption mechanisms for sensitive data stored in the cloud. Encryption safeguards data confidentiality during transmission and storage, aligning with regulatory requirements and security best practices.
- Container Orchestration: Leverage container orchestration tools such as Kubernetes to automate the deployment, scaling, and management of containerized applications. This technique streamlines operations and enhances the adaptability of the adaptive data architecture.
- Regular Security Audits: Conduct regular security audits to assess the effectiveness of implemented security measures. These audits help identify potential vulnerabilities and ensure ongoing compliance with security standards.

3.5.2.3 Benefits of the Approach and Techniques:

- Scalability and Efficiency: The approach ensures the scalability of data storage and efficient resource management, allowing the digital learning environment to adapt to varying workloads.
- Resource Utilization: Containerization and load balancing techniques optimize resource utilization, enhancing the overall performance of applications and services.
- Adaptability to Changes: Techniques such as container orchestration enhance the adaptability of the architecture to changes in user demand and technological advancements.

3.5.3 Secure Grade Distribution Scheme

3.5.3.1 Approach:

1. Key Management Plan: Develop a robust key management plan to govern the generation, distribution, and secure storage of encryption keys within the Secure Grade Distribution Scheme. This plan ensures that cryptographic keys remain confidential and are appropriately managed throughout their lifecycle.
2. Integration of Cryptographic Libraries: Integrate advanced cryptographic libraries supporting the AES encryption algorithm, Hardware Security Modules (HSMs), and Diffie-Hellman key

exchange. These libraries form the foundational elements of the scheme, providing the necessary cryptographic functions for secure grade distribution.

3. Implementation of MIC Verification: Incorporate hashing libraries for Message Integrity Check (MIC) verification into the grade distribution process. This technique ensures the integrity of transmitted grades by attaching cryptographic hash values, allowing recipients to verify the authenticity of the received data.

3.5.3.2 Techniques:

- Diffie-Hellman Key Exchange: Implement the Diffie-Hellman key exchange technique to securely exchange encryption keys between the sender and receiver. This ensures a secure and confidential key distribution process within the grade distribution scheme.
- Hardware Security Module Usage: Utilize Hardware Security Modules (HSMs) for secure key storage and cryptographic operations. HSMs enhance the security of cryptographic keys by providing a dedicated and tamper-resistant hardware environment.
- AES Encryption: Implement AES encryption libraries to encrypt grades using symmetric key cryptography. This technique ensures the confidentiality of transmitted grades, preventing unauthorized access.

3.5.3.3 Benefits of the Approach and Techniques:

- Confidentiality and Integrity: The approach ensures the confidentiality and integrity of transmitted grades through the integration of cryptographic techniques such as AES encryption and MIC verification.
- Secure Key Exchange: Techniques like Diffie-Hellman key exchange facilitate a secure and confidential process for exchanging encryption keys, enhancing overall communication security.
- Key Protection and Management: The use of HSMs provides a secure enclave for key protection and management, minimizing the risk of key compromise and unauthorized access.

3.6 Research Design

The research design for this study draws inspiration from the framework proposed by Georgiadou et al. (2021), encompassing a systematic approach to ensure methodological rigour. The distinct phases, namely survey methodology, case studies for comparative analysis, and subsequent data analyses, are strategically aligned to address the research questions effectively.

3.6.1 Participants and Sampling

In alignment with the peculiarities of the survey, targeted participants included technical staff directly involved in managing learning management systems, university websites, and portals, as well as directors of IT units and academic planning. A deliberate sampling strategy aimed to secure a robust dataset, targeting a minimum of 1,000 responses from at least 90 institutions, constituting 70% of the 128 West African universities registered with the Association of African Universities (AAU, 2022).

3.7 Data Collection

3.7.1 Survey Methodology

The survey methodology served as the primary data collection technique, driven by a comprehensive set of 20 survey questions (SQs) designed in both English and French languages (Appendix A). Each question was meticulously crafted to address specific research questions (RQs), covering diverse aspects such as data architectures, cyber threats, countermeasure techniques, and the integration of data in decision-making.

3.7.1.1 Validity Testing

Before wide dissemination, rigorous validity testing involved a diverse group comprising survey specialists, experienced researchers, certified security and technology officers, and non-technical staff. This phase employed respondent debriefing, cognitive interviewing, think-aloud, and verbal probing techniques to refine the survey instrument. Feedback from this phase informed the development of the final survey version (see Appendix A).

3.7.1.2 Dissemination and Analysis

The survey, initiated on February 26th, 2022, was disseminated to West African universities, both public and private, through email and WhatsApp channels. The three-month circulation period from March 1st to May 31st, 2022, allowed for comprehensive data collection. Specific eligibility criteria limited participation to technical staff, emphasizing the importance of their roles in technological infrastructure.

A total of 1,164 responses were received from 93 universities, representing approximately 72% of West African universities registered with the AAU. To ensure data integrity, duplicate responses were avoided, and 109 responses indicating the absence of workshops or training during the COVID-19 pandemic were excluded. The remaining 1,055 responses formed the basis for the study.

Utilizing a four-point Likert scale for nuanced responses, where Extensively = 4, Moderately = 3, A little = 2, and Not at all = 1, facilitated a nuanced understanding of participant perspectives. The collected data underwent analysis using SPSS, ensuring a robust and systematic exploration of the research questions.

3.7.2 Methodology for Comparative Analysis

The comparative analysis methodology commenced with the identification of pertinent keywords, namely (i) e-learning or "online learning" or "digital learning," (ii) solutions, and (iii) "use case." These keywords were employed to formulate a comprehensive search string: (e-learning or "online learning" or "digital learning") and solutions and "use case." The Google search conducted on November 23, 2021, produced 38 entries, all of which were meticulously downloaded into Zotero, a reference management application.

Subsequently, on November 25, 2021, a meticulous screening process was initiated to ensure alignment with our research objectives. Out of the initial pool, 18 papers were excluded as they were deemed irrelevant to the research. Simultaneously, 20 publications were identified as eligible, contributing to a total of 109 use cases, each associated with a specific e-learning solution and meticulously mapped. This comprehensive analysis phase concluded on November 28, 2021.

Building upon this foundation, a thorough investigation into the data architecture of each mapped e-learning solution transpired on December 15, 2021. This investigation involved scrutinizing the websites and scrutinizing published white papers. Among the 20 identified e-learning solutions, 14 were found to employ data-driven architecture, while the remaining six utilized data-centric architecture. This phase of investigation was successfully concluded on December 25, 2021.

Advancing the research, a total of 983 user reviews were collected from the e-learning industry, with 696 emanating from identified data-driven e-learning solutions and 287 from data-centric counterparts. To analyse this voluminous dataset, a conceptual framework was developed, aligning with current e-learning requirements obtained from [5], [23]–[25] (Figure 3). The analysis transpired from December 26, 2021, to January 02, 2022.

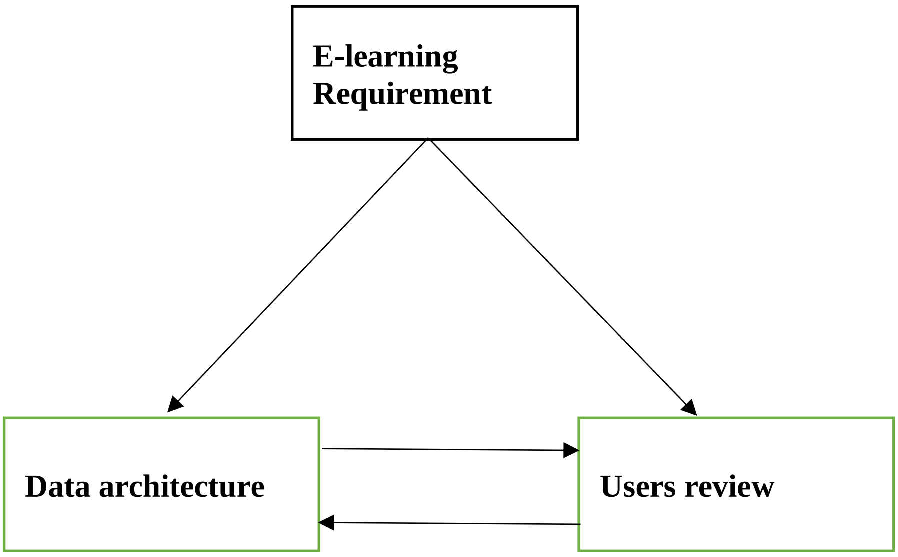


Figure 3. Proposed conceptual framework.

CHAPTER FOUR: RESULTS

4.1 Preamble

The fourth chapter of this thesis embarks on the practical implementation of the proposed frameworks, Comprehensive Security Model, Adaptive Data Architecture, and Secure Grade Distribution Scheme, within the dynamic context of West African universities' digital learning settings. This chapter serves as the bridge between theoretical concepts and real-world application, elucidating the strategic deployment strategies, the evaluation criteria, and the anticipated challenges inherent in implementing these innovative solutions.

As we delve into the practical realm, it is imperative to acknowledge the symbiotic relationship between theory and application. The envisaged security enhancements, data architecture adaptability, and secure grade distribution scheme are about to undergo a transformative journey from conceptualization to operationalization. This chapter provides a comprehensive account of the hands-on aspects of translating theoretical constructs into tangible solutions.

The successful implementation of these frameworks requires a meticulous approach, considering the unique nuances of each university's digital learning infrastructure. From configuring security measures to optimizing data architecture and securing grade distribution, this chapter navigates through the details of translating theoretical excellence into practical realities.

By detailing the step-by-step procedures, the selection and configuration of tools, and the nuances of integrating these solutions into existing university systems, this chapter aims to serve as a practical guide for information technology professionals, academic administrators, and other stakeholders involved in the implementation process. It is a testament to the commitment to fortifying the digital learning landscape and ensuring the integrity, security, and adaptability of academic processes.

In essence, this chapter signifies the convergence of vision and action, theory and practice, as we embark on the journey of transforming West African universities' digital learning environments into fortified bastions of academic excellence, resilience, and security.

4.2 Survey Findings

4.2.1 Overview of Data Architectures in West African Universities

Different researchers have established definitions for data architectures (DA), but the definition by (Zheng et al., 2010) is the one that is most frequently used. It describes DA as a collection of models, policies, guidelines, and standards that regulate the collected data types and how they are organized, integrated, stored, and utilized in data systems and organizations. According to (Ascend, 2020; Carol, 2021; Kampakis, 2018; Sinan, Degila, et al., 2022a), between 400 BC and 2022, DA progressed through four major stage:

1. Traditional architecture
2. Data-informed architecture
3. Data-driven architecture
4. Data-centric architecture

Our study will concentrate on data-informed, data-driven, and data-centric architectures because these are the only ones now in use by businesses and institutions (Sinan, Degila, et al., 2022a); the following are definitions taken from the literature:

- Data-informed Architecture: Data is collected from various sources, including flash drives, computers' internal and external hard drives, and so on. The data is analyzed using a spreadsheet, and the results are used as inputs in the decision-making (Ascend, 2020).
- Data-driven Architecture: In this approach, algorithms are utilized to generate decisions based on the data gathered from several data silos, including the cloud, data lakes, and other sources (Alfonso, 2018). (Kampakis, 2018) defines it as a DA in which storage devices or silos are scattered across several places and algorithms are used to preserve, analyze, and derive decisions from the analysis result. It is defined as a distributed storage architecture employing technology to gather and analyze data to make better business decisions (Kampakis, 2018).
- Data-centric Architecture: (Alfonso, 2018) refers to a system in which data is the primary and permanent asset, whereas applications come and go. In (Vista, 2021) and (Dave, 2020), organizations and institutions create a single data model that is shared by all of the organization’s information systems, data science is used as the bedrock for decision-making, and all data are linked and connected using a graph database to eliminate data silos and redundancy.

4.2.2 Demography

The first part of the survey is for demography including six survey questions, this aids in getting the descriptive data of the Universities and participants' behaviors towards securing their university data. Figure 1 presents the breakdown of the universities surveyed according to their countries with Nigeria being the highest with forty-nine (49) institutions, Ghana 23, Gambia 3, Senegal 3, Sierra leone 3, Burkina Faso 2, Cote d'voire 2, Niger 2, Togo 2. Benin, Liberia, Mauritania and Mali with the fewest amount of one (1) each, making a total of ninety-three Universities (Figure 4).

Additionally, seventy-seven (77) are public Universities and sixteen (16) are private, these institutions employed different modes of delivery (MOD), 44.4 % of the responses came from universities using face-to-face, 45.4% from e-learning institutions and 10.1% from blended MOD institutions, and all universities regularly create vast amounts of data as a result of the plethora of online activity. Of the respondents, 10.5% claimed their institutions only complete applications (A) online, compared to 56.6% who completed application and registration (AR), 6.6% who completed applications, registrations, and examinations (ARE), 8.9% who agreed that their institutions are always online for applications, registrations, and lectures (ARL), and 24.4% who agreed on applications, registrations, lectures, and examinations (ARLE) (Table 1).

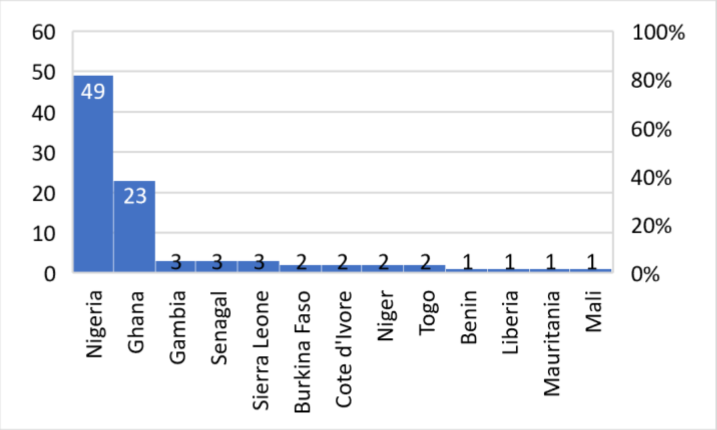


Figure 4 Number of universities according to countries

This demonstrates the significant reliance on online resources for the efficient operation of WAU. In terms of DA, 4.1% of the participants believed they employed data-centric architecture, 24.1% data-driven architecture and 71.8% data-informed architecture (Table 3)

Table 3 Demography of Universities

Online Activities			
Frequency	**Frequency**	**Percent**	**Cumulative percent**
Application	111	10.5	10.5
Application and Registration	591	56.0	66.5
Application, Registration and Examination	6	.6	67.1
Application, Registration and Lectures	94	8.9	76.0
Application, Registration, Lectures and Examination	253	24.0	100.0
Total	1055	100.0	
Mode of Delivery			
Blended learning	117	10.1	10.1
E-learning	518	44.5	54.6
Face-to-face learning	529	45.4	100.0
Total	1164	100.0	
Data Architecture			
Data-Centric Architecture	43	4.1	4.1
Data Driven Architecture	254	24.1	28.1
Data Informed Architecture	758	71.8	100.0
Total	1055	100.0	

The survey received huge responses from both males and females, 76% are males and 24% are females, this is of particular importance, as the ratio of females to males is the ideal ratio for productive work in a cybersecurity environment (Fatokun et al., 2019). Additionally,18.6% of participants were under 25 years, followed by 39.6% between 26 and 35 years, 27.7% from 35 to 45 years, 10% from 46 to 55 years, and 4.7% from participants over 56years of age, Figure 5 shows a histogram of age having mean of 2.43, this is vital it entails that, universities staff have the ideal age to learn new emerging cybersecurity techniques.

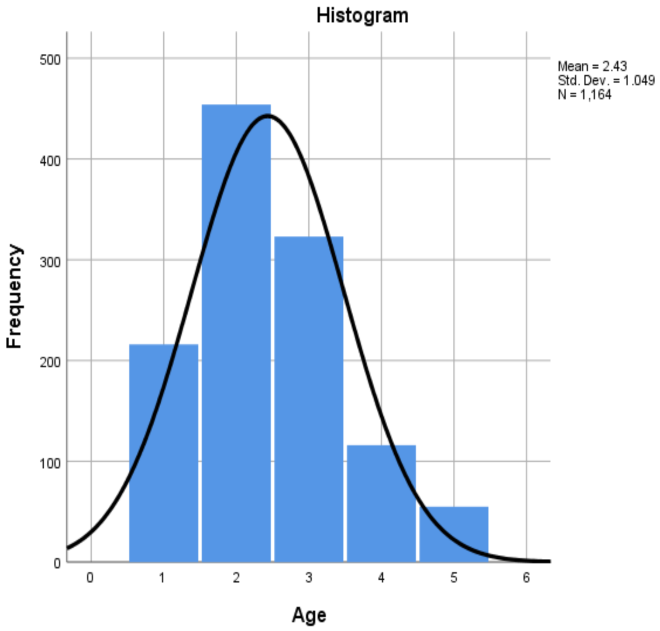


Figure 5. Histogram of age analysis

Furthermore, among the participants, 19.8% have a diploma, 46.1% have a bachelor's degree, 20.6 % have completed their master's, and 13.3% have a PhD (Table 4)

4.2.3 Data Application and Usability

In this survey, participants were given several questions on the use of data and analysis results in making decisions using a four-point Likert scale. The responders were initially questioned on the types of data they gather for analysis prior to making decisions, the tools they used to execute the analysis, and the type of decisions they made.

Table 4 Demography of participants

Frequency	**Frequency**	**Percent**	**Cumulative percent**
Gender			
Female	279	24.0	24.0
Male	885	76.0	100.0
Total	1164	100.0	
Age			
26 – 35 years	454	39.0	39.0
36 – 45 years	323	27.7	66.8
46 – 55 years	116	10.0	76.7
56 and above	55	4.7	81.4
Below 25 years	216	18.6	100.0
Total	1164	100.0	
Qualification			
Bachelor	537	46.1	46.1
Diploma	231	19.8	66.0
Masters	240	20.6	86.6
PGD	1	.1	86.7
PhD	155	13.3	100.0
Total	1164	100.0	

Figure 6 presents the details of the type of data used for analysis for WAU, 35.1% believed their institutions don't use data at all for decision-making they rely on their gut feeling and experience, 36.1% claimed they use data about what happened in the recent past (e.g last year or last quarter), 21.1% agreed that their Universities use past and recent data including some longer-term trends analysis and 7.6% said they use past, present and forward-looking data.

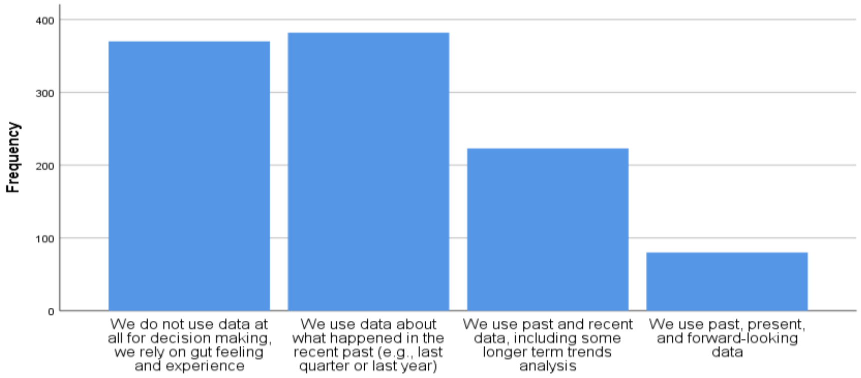


Figure 6 Breakdown of the types of data employed by WAU

Table 5 shows descriptive statistics of the tools employed for analysis by WAU, spreadsheets (e.g charts, counts, tables) have a mean of 3.09, website analytics (e.g google analytics) with 2.22, database (e.g CRM analytics, reports) with 2.65 and specialised tools (e.g SAS, R, Stata, Python, SPSS, GIS mapping) has 2.57.

Table 5 Descriptive statistics of tools used for analysis

Tools	N	Minimum	Maximum	Mean	Std. Deviation
Spreadsheet	1055	1	4	3.09	.951
Website Analytics	1055	1	4	2.22	1.208
Database	1055	1	4	2.65	.933
SpecializeTools	1055	1	4	2.57	.974
Valid N (listwise)	1055				

The data analysis result is used by WAU to make decisions on different categories, in terms of academic development decisions it has a mean of 2.74, employment 2.70, environmental impacts 2.70, other societal impacts 2.68, research opportunities 2.70 and student satisfaction has 3.11 (Table 6)

Table 6. Descriptive statistics on the use of data analysis results

Data Usage	N	Mean	Std. Deviation
Academic Development	1055	2.74	.913
Employment	1055	2.70	.971
Environmental Impacts	1055	2.70	.964
Other Societal Impacts	1055	2.68	.956
Research Opportunities	1055	2.70	.975
Students Satisfaction	1055	3.11	1.022
Valid N (listwise)	1055		

4.2.4 Cyberattacks and Countermeasures.

During the Covid-19 pandemic, WAU were severely targeted by cyberattacks. 87.4% of the responders indicated that they were victims of cyberattacks, and 12.6% were not. Of the victims who are knowledgeable enough about the security vulnerabilities at their institutions, 621 agreed their institutions were attacked by SQL injection, 752 by a denial-of-service attack, 565 by ransomware, 451 by a virus, 214 by a worm, 335 by a phishing attack, and 1 participant reported not knowing about any cyberattacks (Figure 7).

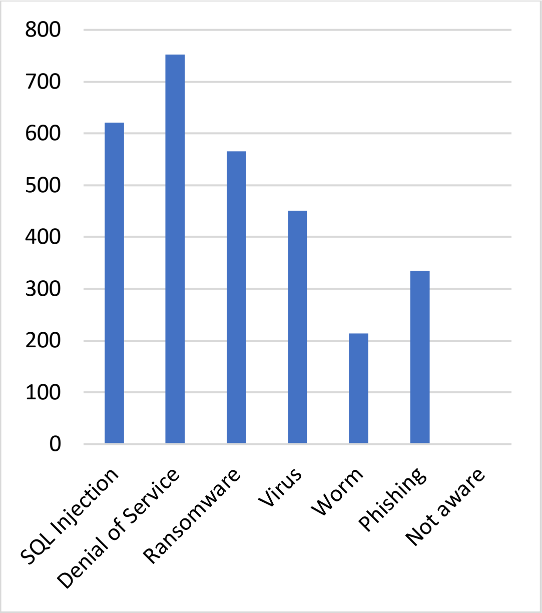


Figure 7. Summary of cyberattacks faced by WAU

Moreover, the participants receive cybersecurity training, but only 8.1% complete it after 3 months, 10.2% do so after 6 months, 40.3% do so after 12 months, and 41.1% have never attended any cybersecurity training (Table 7)

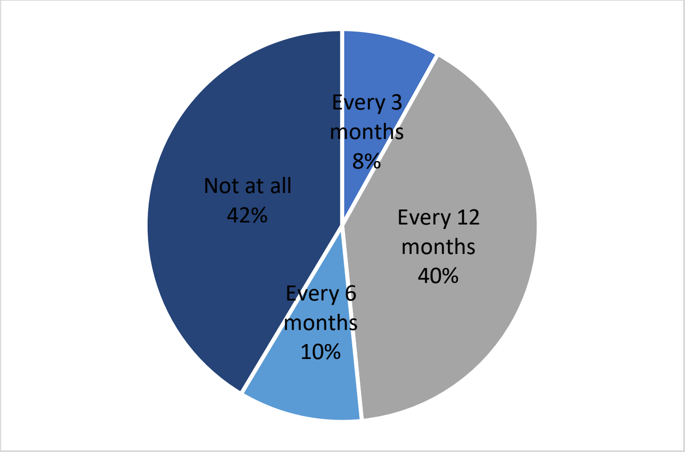


Figure 8. Summary of staff cybersecurity training

Moreover, a variety of countermeasures are used by WAU, Table 6 breaks down these techniques. These institutions used a variety of techniques to ensure secure cyberspace for learning, and many participants (52.9%) claimed their institutions only used firewalls and antivirus software for security, while 0.1% thought their institutions used firewalls, intrusion detection systems, and intrusion prevention system.

Table 7 Countermeasures

Countermeasures Technique	Frequency	Per cent	Cumulative Percent
Anti-virus	28	2.7	2.7
Anti-virus, Intrusion detection system	173	16.4	19.1
Anti-virus, Intrusion detection system, Intrusion prevention system	12	1.1	20.2
Anti-virus, Intrusion prevention system	4	.4	20.6
Firewall	22	2.1	22.7
Firewall, Anti-virus	558	52.9	75.5
Firewall, Anti-virus, Intrusion detection system	176	16.7	92.2
Firewall, Anti-virus, Intrusion detection system, Intrusion prevention system	38	3.6	95.8
Firewall, Anti-virus, Intrusion prevention system	9	.9	96.7
Firewall, Intrusion detection system	13	1.2	97.9
Firewall, Intrusion detection system, Intrusion prevention system	1	.1	98.0
Intrusion detection system	9	.9	98.9
Intrusion detection system, Intrusion prevention system	8	.8	99.6
Intrusion prevention system	2	.2	99.8
Not known	2	.2	100.0
Total	1055	100.0	

Responders were asked about the level of satisfaction they had with their institution’s data protection techniques; Table 8 shows that it has a mean 2.24.

Table 8 Descriptive Statistics on Satisfaction

Items	N	Mean	Std. Deviation
Satisfaction	1055	2.24	.795
Valid N (listwise)	1055		

4.2.5 Discussion

This study created a survey and distributed it to WAU to determine the security vulnerability of their data architectures, techniques for preventing cyberattacks, and the effect of data analysis on decision-making. In this section, first and foremost, we will discuss demographic analysis, data analysis and usability, and cyberattacks and countermeasures.

Looking at figure 2's age analysis, it has a mean of 2.43 (std. Dev 1.049) which indicates that the majority age of the participants is 25-35 years, and 46.1% have bachelor's degrees, which is the perfect age and educational background for the staff to learn new skills for fending off cyberattacks. Furthermore, analysis demonstrates that the gender ratio is favorable for staff to co-exist for effective work in a cybersecurity environment (Fatokun et al., 2019). In addition, WAU has quickly made the switch to digital learning; 45.1% of the institutions surveyed used e-learning as a MOD, and every institution had at least one online activity. This makes it a challenge for both researchers and industries to provide safe and secure data architecture in this region.

In addition, WAU are always looking for research gaps that may be addressed by academic researchers in addition to staff employment, enrolling more students, and developing staff capacity. However, the results of this study indicate that, in order to run these Universities efficiently, there is a need to optimize the utilization of data analysis results. Moreover, Table 5 shows that data analysis results for academic development have a mean of 2.74, employment has a mean of 2.70, environmental impacts have a mean of 2.68, research opportunities have a mean of 2.70, and student satisfaction has a mean of 3.11; this demonstrates the specific areas that need improvement, particularly areas with less than 3.0. Additionally, the type of data acquired for analysis before decision-making and the tools used for analysis are also causes for concern. According to the study's findings, only 7.6% of participants believed their institutions used past, present, and future-looking data for analysis, while 35.1% agreed that they used their intuition and experience instead. Furthermore, with a mean of 3.09, the majority of participants chose to use spreadsheet software for data analysis, compared to less than 2.6 for the other tools, which is worrying. This creates a vacuum for WAU to enhance the type of data and analysis tools.

Findings show that WAU are always conducting activities online, be it application, registration, lectures or facilitation, or examination, which yields data generation and are yet to get a secured means of storing their data. Only 12.6% indicated that their universities were not victims of cyberattacks. These attacks are due to several factors particularly:

- Inability to upgrade their data architecture to the newest, this study finds out that 71.8% use DIA which is the most obsolete DA in existence, followed by DD with 24.1%, and

4.1% employed DCA which is the advanced DA in existence now and it is highly secured with few security vulnerabilities (Kim, 2019)

- The technical staff maintaining learning management systems, websites, and portals lack cybersecurity knowledge and training. This study reveals that 41.1% of the staff have never taken cybersecurity training, 40.3% have done so after every 12 months, 10.2% have taken it after 6 months, and 8.1% have taken it after 3 months. The training has a mean of 1.85, indicating that the majority of participants have never taken cybersecurity training (Table 6), and the analysis of the cybersecurity skills of the participants reveals they have a mean of 3.43, demonstrating the need for frequent training and workshops.
- Lack of adequate countermeasures to efficiently prevent and detect cyberattacks. The finding of this study shows that universities use several techniques when repelling cyberattacks, 52.9% use firewalls and anti-virus software which is not efficient, while 0.1% believed their institution employed firewalls, intrusion detection systems and intrusion prevention systems.

Additionally, on a scale of yes, neutral, and no, the participants were also asked to rate their level of satisfaction with their institution's countermeasures strategy. Analysis reveals that it has a mean of 2.24 (Table 6), indicating that the majority of participants are not satisfied with their institution's countermeasures strategy.

4.3 Comparative Analysis

4.3.1 Introduction

This section serves as a comprehensive comparative analysis aimed at discerning the optimal data architecture for digital learning universities, specifically tailored to accommodate the multifaceted requirements of e-learning. To facilitate this evaluation, the study initiates the process by identifying and scrutinizing 109 distinct e-learning solution use cases. Through meticulous classification, each use case is systematically categorized based on the underlying data architectures employed.

To enrich the depth of this comparison, the study delves further into the practical insights garnered from the e-learning industry by procuring and analyzing 983 user reviews. This qualitative approach ensures a nuanced understanding of the user experience, shedding light on the nuances and intricacies of various data architectures within the e-learning domain.

The identification and classification of e-learning solution use cases lay a robust foundation for the subsequent analysis. By dissecting each use case based on the employed data architecture, the study unveils patterns, strengths, and potential limitations associated with different approaches. This intricate examination is instrumental in formulating a nuanced understanding of the diverse landscape of data architectures in the context of digital learning universities.

The inclusion of 983 user reviews amplifies the comparative assessment, providing a qualitative dimension to the evaluation process. These reviews, sourced from within the e-learning industry, encapsulate real-world experiences and perspectives, offering valuable insights into the practical implications of various data architectures. Users' feedback becomes a crucial lens through which the study gauges aspects like user satisfaction, system performance, and overall efficacy, adding a layer of authenticity to the comparative analysis.

In essence, this section not only outlines a methodology for identifying and classifying e-learning solution use cases based on data architectures but also extends its reach into the realm of user experiences. By combining quantitative data on use cases with qualitative feedback from industry users, the study aspires to present a well-rounded and informed perspective on the most suitable data architecture for digital learning universities engaged in the dynamic landscape of e-learning.

4.3.2 E-learning Solution Use cases

Figure 9 presents a detailed breakdown of e-learning use cases, with employee training emerging most frequently at fourteen (14) occurrences. Following closely are customer training at thirteen (13), compliance training, and academic learning, both registering twelve (12) each. Employee onboarding and training companies are tied at eleven (11) each, while continuing education follows with seven (7) instances. Further down the list, extended enterprise and dealer training both have six (6), and channel training is documented at four (4). Immersive learning and competency management share a count of four (4), while workforce development concludes the breakdown with three (3) instances.

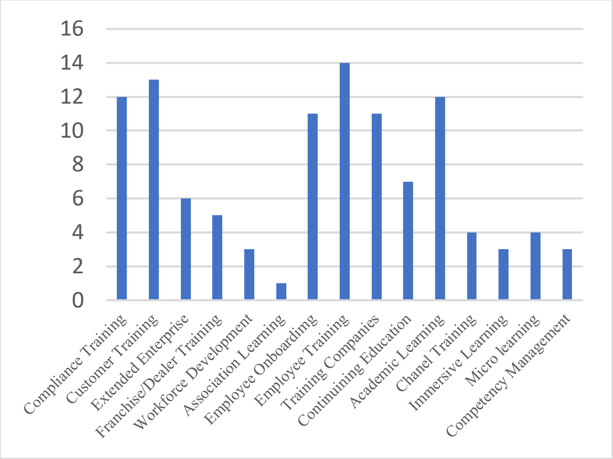


Figure 9. Use cases.

The use cases were then mapped to twenty (20) e-learning solutions (Table 2), with five (5) different solutions containing seven (7) use cases each, one (1) solution containing six (6) use cases, twelve (12) solutions containing five (5) use cases each, and two (2) solutions containing four (4) use cases.

Table 9. Use Cases

			Use Cases														
Reference(s)	E-learning solutions	Number of Use cases	Academic Learning	Association Learning	Compliance Training	Customer Training	Continuing Education	Chanel Training	Competency Management	Extended Enterprise	Employee Onboarding	Employee Training	Franchise Training	Immersive Learning	Microlearning	Training Companies	Workforce Development
(Ghatak, 2021)	Adobe Captivate Prime	5			X	X				X			X				X
(Lynch, 2021)	Absorb LMS	7	X		X	X				X	X	X				X	
(Kapadia, 2021)	GyrusAim	5			X	X						X					X
(Docebo, 2021)	Docebo	4			X	X					X	X					
(Gray, 2021)	Xperiencify	7	X				X	X			X			X	X	X	
(Jennifer, 2021)	Inquisiq	5	X		X	X						X				X	
(Brown, 2021)	Coassemble	7			X	X					X	X	X		X	X	
(Butler, 2021)	Nimble LMS	7			X	X					X	X				X	X
(Malekos, 2021)	LearnWorlds	5				X	X				X	X				X	
(Ponomarev, 2021)	Gurucan	4					X				X					X	X
(Media, 2021)	Eurekos LMS	5				X		X		X		X	X			X	
(Doust, 2021)	glo™ learn	7	X		X					X	X	X	X			X	
(Papagelis, 2021)	TalerntLMS	5			X	X				X	X	X					
(Bellaj, 2021)	Etakwin	5	X			X	X					X				X	
(Scott, 2021)	Thinkific	5	X			X	X						X			, X	
(Shodeinde, 2021)	Claned	5	X		X		X					X				X	
(Pappas, 2021)	Edysby	5	X	X	X		X		X								
(Ispring, 2021)	Ispring	5			X	X				X	X	X					
(learn upon, 2021)	LearnUpon LMS	6			X	X				X	X	X				X	
(Gogos, 2021)	Looop	5			X	X				X	X	X					

4.3.3 E-learning Solutions Mapping with Data Architectures

Table 10 shows the mapping of the e-learning solutions with data architectures, data-driven architecture accounting 70% of the mapping, including adobe captivate prime, absorb LMS, inquisiq, gurucan, eurekos LMS, glo™ learn, talerntLMS, etakwin, thinkific, claned, looop, canopyLAB, ispring, and learnUpon LMS, and data-centric architecture accounting for 30% of the mapping including gyrusAim, docebo, xperiencify, coassemble, nimble LMS, and learnWorlds.

Table 10 Data Architectures Mapping

Data architecture(s)	E-learning solutions
Data-driven architecture	Adobe Captivate Prime(Adobe, 2021; Ghatak, 2021) Absorb LMS (Inc, 2021; Lynch, 2021) Inquisiq (Jennifer, 2021; "Privacy Policy," 2021) Gurucan (Campus, 2021) Eurekos LMS (Bill, 2021) glo™ learn (GDPR, 2018; Policy, 2020) TalerntLMS (Papagelis, 2021; talent LMS, 2021) Etakwin (Bellaj, 2021) Thinkific (Scott, 2021; thinkific, 2020) Claned (claned, 2021; Shodeinde, 2021) Looop(Gogos, 2021) CanopyLAB (Hjorth, 2021) Ispring (Ispring, 2021) LearnUpon LMS (learn upon, 2021)
Data-centric architecture	GyrusAim (gyrus, 2021; Kapadia, 2021) Docebo (Docebo, 2021; docebo, 2021) Xperiencify (Gray, 2021; xper, 2020) Coassemble (Brown, 2021; pseudonymisation, 2020) Nimble LMS (Policy, 2018) LearnWorlds (CIO, 2021)

4.3.4 Comparative Analysis

In this sub-section, we present the comparison between data-driven and data-centric architectures. The data obtained for this study is presented in Table 4. Data-driven architecture received the most reviews, with 696 from 14 different e-learning solutions, while data-centric architecture received 287 from 6 other e-learning solutions.

Table 11. Users Review Data

Data architecture	Reference	Reviews	Total
Data-driven architecture	(Bellaj, 2021; claned, 2021; Doust, 2021; Ghatak, 2021; Gogos, 2021; Hjorth, 2021; Ispring, 2021; Jennifer, 2021; learn upon, 2021; Lynch, 2021; Media, 2021; Papagelis, 2021; Ponomarev, 2021; Scott, 2021)	Adobe Captivate Prime 63, Absorb LMS 20, Inquisiq 20, Gurucan 121, Eurekos LMS 16, glo™ learn 20, TalerntLMS 220, Etakwin 7, Thinkific 12, Claned 16, Looop 89, CanopyLAB 26, Ispring 29,LearnUpon LMS 37	696
Data-centric architecture	(Brown, 2021; Butler, 2021; Docebo, 2021; Gray, 2021; Kapadia, 2021; Malekos, 2021)	GyrusAim 64, Docebo 33, Xperiencify 77, Coassemble 36, Nimble LMS 40, LearnWorlds 37	287

Table 11 shows the comparison details based on e-learning requirements under data-driven architecture with 696 reviews. 69.7% of reviewers claimed it has real-time collaboration, 55.6

% claimed it gives students the ability to perform tasks anywhere, 39.7% claimed it is friendly with third-party applications, 45.8% suggested it has data integration, and 31.5% claimed it has a flexible environment. On the other hand, it has no data security, no data ownership, no data access permission, no data traceability, and data insight, according to 3.3%,10.9%, 41.5%, 2.2%, and 3.6%, respectively. Data-centric architecture received the fewest reviews (287), with 65.2 %, 37.9%, 41.4 %, 74.6 %, 93.3 %, 39.7%, 24 %, 27.2 %, 5.9%, and 5.2 % believing it has real-time collaboration, the ability for students to perform tasks anywhere, interoperability with other apps, data integration, customization, data ownership, data access permission, and data insight respectively. In comparison, 3.1% claimed it has no data security.

Table 12. Comparative Analysis

Data architecture(s)	**Data-driven architecture**			**Data-centric architecture**		
Number of reviews	696			287		
Analysis	Evaluation	Number of reviews	Percentage of the reviews	Evaluation	Number of reviews	Percentage of the review
Real-time collaboration	✓	485	69.7%	✓	187	65.2%
Ability for students to perform labs task anywhere	✓	387	55.6%	✓	109	37.9%
Interoperability with other apps	✓	276	39.7%	✓	119	41.4%
Data integration support	✓	319	45.8%	✓	214	74.6%
Flexible environment	✓	219	31.5%	✓	268	93.3%
Data security	X	23	3.3%	X	09	3.1%
Customization	✓	145	20.8%	✓	114	39.7%
Data ownership	X	76	10.9%	✓	69	24%
Data access permission	X	289	41.5%	✓	78	27.2%
Data traceability	X	15	2.2%	✓	17	5.9%
Data insight	X	25	3.6%	✓	15	5.2%

4.3.5 Discussion

This study undertakes a comprehensive examination, identifying and categorizing 109 e-learning solution use cases based on their respective data architectures. Additionally, a rich dataset of 983 user reviews from the e-learning industry was amassed, contributing valuable insights into the perceived efficacy of the identified data architectures. Subsequently, a meticulously crafted conceptual framework was developed, harnessing the wealth of user reviews to facilitate a nuanced comparison of the identified data architectures.

The framework, deployed to scrutinize both data-driven and data-centric architectures, unearthed noteworthy observations. Data-driven architecture, while exhibiting strengths, revealed limitations in critical aspects such as data security, data ownership, data access control, data traceability, and data insight. Conversely, data-centric architecture displayed a limitation primarily in the domain of data security.

This discerning analysis positions data-centric architecture as the more fitting choice for e-learning environments. The identified strengths and limitations underscore the importance of tailoring data architecture to the unique demands and sensitivities of the e-learning landscape. Considering the findings, this study advocates for a paradigm shift towards embracing data-centric architecture to optimize the overarching efficacy and security of e-learning solutions.

4.4 Data-Centric Model

Data-centric architecture is an approach to data management that places data at the center of decision-making processes (Sinan, Degila, et al., 2022b). In education, this approach involves the creation of a data-driven culture where data is treated as an asset and is leveraged to drive decision-making. The key principles of data-centric architecture in education include data governance, data quality, data integration, and data analytics (Figure 10).

Data governance involves establishing policies and procedures for data management, ensuring compliance with privacy and security regulations, and promoting data sharing across different departments (Al-Naser et al., 2013). Data quality involves ensuring that data is accurate, complete, and consistent. Data integration involves consolidating data from different sources to create a single view of student data. Data analytics involves using advanced analytics to gain insights into student behavior and improve learning outcomes.

The benefits of data-centric architecture in education are numerous. It enables personalized learning, where learning resources are tailored to the needs of individual students. It also enables early identification of at-risk students, allowing for timely interventions to be made. Furthermore, it enables the optimization of resource allocation, where resources are allocated based on student needs and performance.

Data-centric architecture also promotes collaboration among stakeholders in education. For example, teachers can collaborate with other teachers to share best practices and improve student outcomes. School administrators can collaborate with teachers to identify areas where resources need to be allocated to improve student outcomes. Additionally, data-centric architecture enables data sharing across different departments, which facilitates cross-functional decision-making.

However, there are also challenges associated with data-centric architecture in education. One challenge is ensuring data privacy and security. Another challenge is the complexity of data integration, where data is collected from different sources and in different formats. Additionally, there may be resistance to change from stakeholders who are accustomed to traditional data architectures.

Figure 10. Data-centric Architecture (Arora et al., 2018)

4.4.1 Components of data-centric architecture

i. Data Sources: In the foundation of a data-centric architecture lie the data sources, diverse origins of information ranging from databases and applications to cloud services and IoT devices. These sources are instrumental in providing a comprehensive view of data, encompassing structured and unstructured formats, batch or real-time streams, and varying data quality considerations.

ii. Connectors: Acting as crucial bridges between disparate data sources and the central data hub, connectors play a vital role in facilitating the extraction, transformation, and loading processes. They ensure a smooth and seamless flow of data from source systems to the architecture, managing complexities related to data extraction, format transformation, and secure transfer.

iii. Data Hub: At the heart of the architecture lies the data hub, serving as the central repository where data from diverse sources converges. It acts as a unified storage and processing hub, providing efficient data management, organization, and accessibility. A well-designed data hub supports scalability, ensuring the architecture's ability to handle growing volumes of data.

iv. Data Governance and Security: Data governance establishes policies, procedures, and controls for managing data throughout its lifecycle, ensuring compliance with regulations. Security mechanisms guarantee the confidentiality, integrity, and availability of data. These components are essential for maintaining trust in the accuracy and privacy of stored information.

v. Data as a Service: Data as a Service (DaaS) simplifies data consumption by providing on-demand access to specific data functionalities or datasets. It abstracts the complexities of data management, allowing users and applications to consume data without in-depth knowledge of underlying storage and processing intricacies. DaaS contributes to the agility of the architecture, focusing on data consumption rather than management.

vi. User-Centric Apps: In a data-centric architecture, user-centric apps are applications designed with a primary focus on user experience. Leveraging the available data, these apps provide valuable insights and facilitate decision-making. They include dashboards, reporting tools, or specialized interfaces tailored to user needs, enhancing the usability and utility of the architecture for stakeholders.

vii. Analytic Factory: The analytic factory represents a structured environment for developing, deploying, and managing analytics within the architecture. It streamlines the analytics process, from data preparation to model deployment, offering tools and processes for analytics lifecycle management. The goal is to create a systematic approach to generating actionable insights from data.

viii. Analytics as a Service: Extending the concept of DaaS to analytics functionalities, Analytics as a Service (AaaS) provides on-demand access to analytics tools, algorithms, or platforms. This allows users to perform complex analyses without managing underlying infrastructure. AaaS enhances the scalability and accessibility of analytics within the architecture, promoting flexibility and cost-effectiveness.

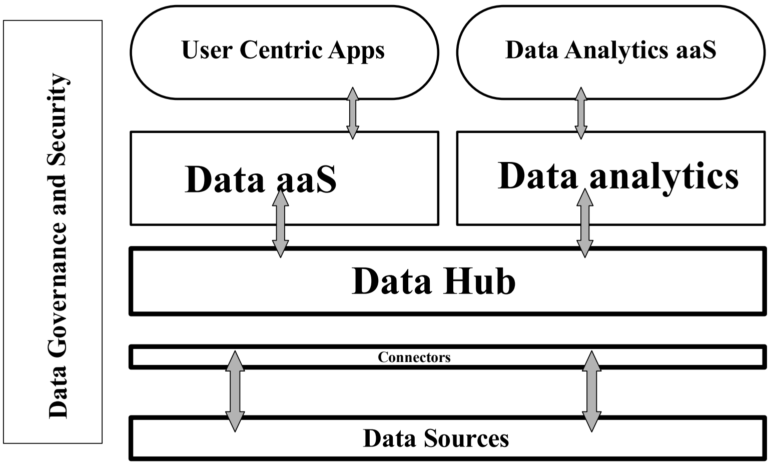


Figure 11 Components of Data-centric Model

Each of these components contributes to the cohesive and effective functioning of the data-centric architecture, ensuring it meets the demands of modern applications, particularly in contexts like digital learning universities (Figure 11).

4.4.2 Designing Data-centric Architecture Model

In the formulation of a specialized model intended for digital learning universities to effectively address the distinctive requirements prevalent in West African academic institutions, the incorporation of the following key components played a pivotal role. These components were strategically chosen to ensure a comprehensive and tailored approach to cater to the specific needs and challenges encountered within the realm of digital learning in the West African context (Figure 12).

i. Data Sources: In the intricate landscape of a digital learning university, a plethora of data sources collectively shape the foundation of a comprehensive data-centric architecture. At its core lies the Learning Management System (LMS), capturing user engagement metrics and course interactions, pivotal for assessing educational content effectiveness. Mobile applications tailored for education generate insights into user preferences and engagement patterns across diverse devices. The integration of Internet of Things (IoT) devices, administrative systems, and academic platforms contributes to a holistic understanding of the institution's operations. Social media engagement, library systems, and online collaboration tools provide additional layers of valuable data, reflecting broader trends, sentiment, and collaborative learning dynamics.

Research platforms, e-learning content platforms, and student information systems offer specialized data crucial for strategic decisions, innovation, and personalized learning approaches. Alumni engagement platforms, assessment tools, and testing platforms enrich the data landscape, providing insights into alumni contributions, academic proficiency, and continuous improvement areas.

ii. Connectors: In the realm of digital learning universities, the pivotal role of connectors within a data-centric architecture cannot be overstated. KConnect emerges as an exemplary connector, adept at seamlessly integrating diverse data sources prevalent in educational ecosystems. Its adaptability extends to Learning Management Systems, educational apps, IoT devices, administrative systems, and other academic platforms, fostering a cohesive data flow. Noteworthy for its flexibility in handling various data formats and robust security measures, KConnect ensures the synchronization of real-time data, contributing to streamlined decision-making and enhanced insights. As a linchpin in the architecture, KConnect epitomizes efficiency, reliability, and adaptability, addressing the dynamic needs of educational institutions in the digital age.

iii. Data hub: Within the data-centric architecture of digital learning universities, the Data Hub stands as a central nexus for managing and orchestrating data flow. In this context, the implementation of a hybrid storage approach is integral to the Data Hub's functionality. The hybrid storage system seamlessly integrates various storage solutions, including Network-Attached Storage (NAS), Storage Area Network (SAN), Cloud Storage, Direct-Attached Storage (DAS), and Tape Storage. This comprehensive strategy ensures optimal performance, scalability, and cost efficiency by balancing the strengths of each storage solution. Notably, all data stored in the local storage component of the hybrid model is fortified with AES encryption, bolstering data security and confidentiality. This encryption protocol serves as a robust safeguard, mitigating the risk of unauthorized access and enhancing the overall resilience of the digital learning university's data ecosystem. The Data Hub, fortified by hybrid storage and AES encryption, emerges as a strategic linchpin, fostering seamless data integration, accessibility, and security within the educational landscape.

iv. Kafka: In the data-centric architecture of digital learning universities, Kafka assumes a critical role as a distributed event streaming platform, meticulously designed by the Apache Software Foundation. Operating on a publish-subscribe model, Kafka becomes an indispensable component, orchestrating real-time data streams and fostering seamless communication across diverse applications and data sources within the university ecosystem. Serving as the backbone of a robust and scalable data pipeline, Kafka excels in handling large volumes of streaming data, ensuring the instantaneous ingestion, processing, and dissemination of information—a pivotal capability in the dynamic realm of digital education. Beyond its prowess in real-time data management, Kafka acts as a reliable connector, facilitating the integration of varied data sources like Learning Management Systems, educational apps, IoT devices, administrative databases, and academic repositories. Kafka's fault-tolerant and distributed architecture guarantees data integrity and availability, positioning it as a linchpin in constructing resilient, responsive, and data-driven educational infrastructures for digital learning universities.

v. Data governance and Security: In the intricate landscape of data-centric architecture for digital learning universities, robust data governance and security measures are imperative. Apache Ranger and Apache Atlas emerge as stalwart guardians, orchestrating a formidable shield against unauthorized access, ensuring compliance, and fostering comprehensive data management. Apache Ranger stands as a sentinel, providing fine-grained access control and centralized security policies, allowing institutions to define, enforce, and audit data access policies seamlessly. Its capabilities extend to safeguarding sensitive information, mitigating risks, and upholding regulatory compliance within the educational data ecosystem. Complementing this protective bastion, Apache Atlas takes the reins in metadata management and lineage tracking. It meticulously catalogs and classifies data entities, offering a holistic view of the data landscape. This not only enhances transparency but also fortifies the ability to trace the origins and transformations of data—an invaluable asset in maintaining data quality and integrity. Together, Apache Ranger and Apache Atlas synergize to establish an unyielding fortress, instilling confidence in the secure handling of diverse data sources, from Learning Management Systems to administrative databases, ensuring the privacy and integrity of the wealth of information within the digital learning university domain.

vi. Data as a service: In the context of data-centric architecture, Data as a Service (DaaS) emerges as a pivotal component that revolutionizes how digital learning universities handle and deliver data. DaaS acts as a cloud-based service providing on-demand access to a variety of data sources, fostering seamless integration and utilization of data across different applications and systems. In the realm of digital learning, data sources such as Learning Management Systems (LMS), educational apps, Internet of Things

(IoT) devices, administrative records, and academic databases serve as examples of the rich and diverse data offerings that can be encapsulated by DaaS. The implementation of DaaS in the data-centric architecture ensures that users within digital learning universities can easily access, retrieve, and leverage data without the constraints of physical or geographical boundaries. This streamlined accessibility not only enhances the overall efficiency of educational processes but also promotes a more personalized and adaptive learning experience for students and educators alike. Furthermore, by encapsulating data from various sources, including hybrid storage solutions, DaaS contributes to the overarching goal of creating a unified and comprehensive data ecosystem within the educational landscape. In terms of security, DaaS platforms often integrate robust access controls, encryption mechanisms, and data governance frameworks, ensuring the confidentiality and integrity of sensitive information. As digital learning continues to evolve, the incorporation of Data as a Service stands as a pivotal step towards establishing a dynamic, interconnected, and secure data environment for educational institutions.

vii. Data Analytics: In the context of data-centric architecture, the utilization of advanced data analytics plays a crucial role in extracting meaningful insights and facilitating informed decision-making. TensorFlow, a prominent open-source machine learning framework, stands at the forefront of driving data analytics within this architecture. TensorFlow empowers digital learning universities to harness the capabilities of machine learning and deep learning algorithms for processing and analyzing vast datasets. The framework provides a versatile and scalable infrastructure, enabling the development of sophisticated models that can uncover patterns, trends, and correlations within educational data. Digital learning universities can leverage TensorFlow to implement predictive analytics, allowing for the anticipation of student performance, course effectiveness, and other critical metrics. Additionally, TensorFlow facilitates the creation of intelligent applications and services that enhance the overall learning experience, such as personalized recommendations, adaptive learning paths, and automated grading systems. The integration of TensorFlow into the data-centric architecture ensures a powerful and efficient platform for data analytics, fostering innovation and optimization across various educational processes. This includes tasks such as content recommendation, student engagement analysis, and resource allocation, ultimately contributing to the continuous improvement of educational outcomes. As the field of data analytics continues to evolve, TensorFlow stands as a valuable tool for digital learning institutions seeking to unlock the full potential of their data resources through advanced machine learning capabilities.

viii. Analytics as a Service: In the realm of data-centric architecture, the concept of Analytics as a Service (AaaS) emerges as a pivotal component, enabling digital learning universities to access and deploy advanced analytical tools and capabilities without the need for extensive infrastructure investments. Analytics as a Service involves the delivery of analytical insights, data visualization, and predictive modeling functionalities through a cloud-based service model. This approach empowers educational institutions to harness the benefits of cutting-edge analytics tools without the burden of managing complex infrastructure and resources. By adopting Analytics as a Service, digital learning universities can efficiently process large volumes of educational data, gaining valuable insights into student performance, learning patterns, and overall institutional effectiveness. Cloud-based analytics platforms provide scalability, flexibility, and cost-effectiveness, allowing institutions to tailor their analytical capabilities to specific needs. Prominent cloud service providers, such as AWS, Azure, and Google Cloud, offer comprehensive Analytics as a Service solutions, providing a wide array of tools for data exploration, visualization, and machine learning. This not only streamlines the integration of analytics into educational processes but also ensures that institutions can stay at the forefront of data-driven decision-making in a rapidly evolving digital learning landscape. Analytics as a Service thus emerges as a strategic enabler, empowering digital learning universities to derive actionable insights and enhance the overall educational experience.

ix. User-Centric Apps: In the realm of data-centric architecture, the integration of User-Centric Apps plays a pivotal role, with Kibana and Elasticsearch serving as key components. User-Centric Apps are designed to provide an intuitive and tailored interface for end-users, facilitating seamless interaction with data and analytical insights. Kibana, in conjunction with Elasticsearch, forms a robust combination for developing such user-centric applications within the digital learning environment. Kibana, as an open-source data visualization platform, excels in creating dynamic and interactive dashboards, charts, and graphs. It acts as the user interface layer for Elasticsearch, enabling users to explore, analyze, and interpret data effectively. Elasticsearch, a distributed search and analytics engine, ensures high-speed data retrieval and efficient storage, making it an ideal backend for user-centric applications. Together, Kibana and Elasticsearch empower digital learning universities to build applications that offer real-time insights into various aspects of educational data, including student performance, engagement metrics, and course effectiveness.

x. User-Centric Apps using Kibana and Elasticsearch contribute to a data-driven educational environment by providing educators, administrators, and students with user-friendly interfaces to interact with complex datasets effortlessly. These applications facilitate informed decision-making, enhance the overall learning experience, and contribute to the continuous improvement of educational processes. As part of the data-centric architecture, User-Centric Apps with Kibana and Elasticsearch emerge as a powerful toolset for creating personalized, insightful, and responsive interfaces that cater to the diverse needs of stakeholders in digital learning universities.

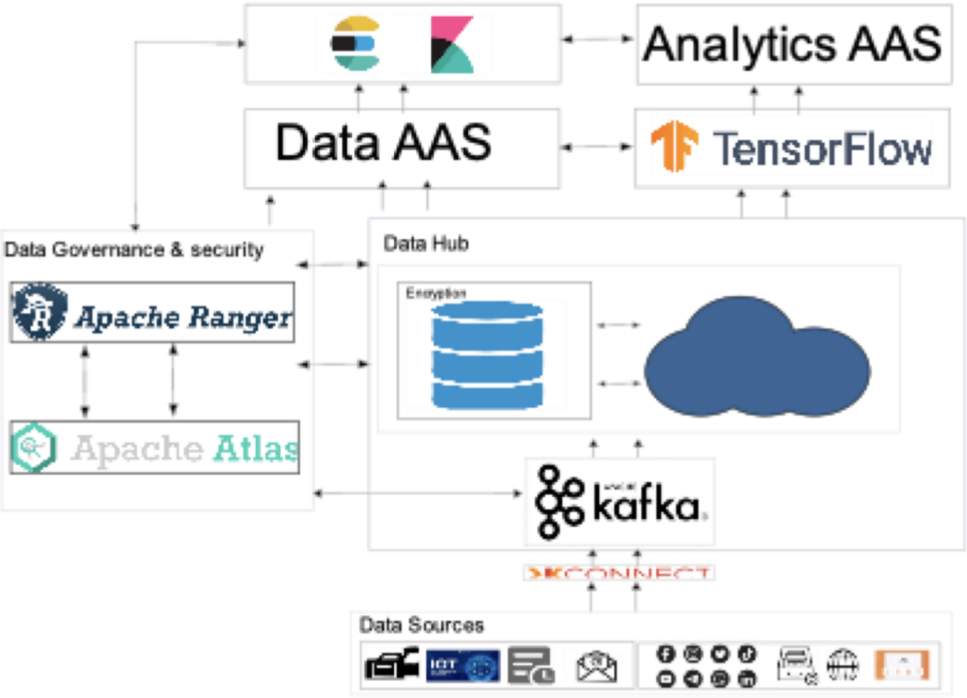


Figure 12 Data-Centric Model

4.4.3 Discussion

The developed data-centric model tailored for digital learning universities is poised for insightful discussion, shedding light on its key facets and implications. The model strategically incorporates diverse components, each contributing to the robustness and efficiency of the overarching architecture.

At the core of the model are the identified data sources, including Learning Management Systems (LMS), applications, IoT devices, administrative records, and academic databases. This diverse range ensures a comprehensive coverage of data inputs, vital for the dynamic and multifaceted nature of digital learning environments in universities.

Connectivity within the model is facilitated by KConnect, a powerful software acting as connectors. The discussion delves into the merits of KConnect in seamlessly integrating varied data sources, ensuring smooth data flow, and enhancing interoperability. This connector plays a pivotal role in harmonizing the heterogeneous data landscape present in digital learning universities.

Hybrid storage emerges as a key feature in the data hub component, incorporating a blend of network-attached storage (NAS), storage area network (SAN), and cloud storage. Notably, all local storage is encrypted using the AES algorithm, ensuring a robust security layer for stored data. This strategic integration addresses the imperative of data security and scalability in the context of digital learning.

Furthermore, the model incorporates Kafka for efficient data streaming, adding a real-time dimension to data processing. The utilization of Apache Ranger and Apache Atlas for data governance and security underscores the commitment to maintaining data integrity, confidentiality, and availability.

The discussion extends to Data Analytics as a Service (DAAS) and the incorporation of TensorFlow for advanced analytics. This ensures that the model is not only adept at handling large volumes of data but also capable of extracting valuable insights through sophisticated analytics techniques.

User-centric apps, powered by Kibana and Elastic Search, signify a user-friendly interface, ensuring accessibility and ease of use for stakeholders. Lastly, the model introduces Analytics as a Service (AaaS), providing a scalable and efficient solution for analytical needs.

In summation, the developed data-centric model exhibits a thoughtful integration of components, addressing the unique challenges of digital learning universities. Its comprehensive nature, coupled with a focus on security, scalability, and analytics, positions it as a robust framework for optimizing data management in the evolving landscape of educational technology.

4.5 Security Model Design

4.5.1 Development of the Security Model

Developing a robust model for enhanced data security is pivotal to safeguarding sensitive information within digital learning universities. This section outlines the systematic process undertaken in crafting the framework for enhanced data security, ensuring resilience against potential risks and threats.

4.5.1.1 Data Storage

The study identifies all the available storage solutions and select the most suitable for digital learning setting.

Storage Solutions:

1. Optical Discs: Optical discs, such as CDs, DVDs, and Blu-ray discs, utilize lasers for data reading and writing. They have been extensively used for data distribution, software installation, and media storage. Despite their widespread use, optical discs have limited storage capacity compared to other solutions.
2. Flash Drives: Also known as USB drives or thumb drives, flash drives are compact, portable devices using flash memory. Commonly employed for data transfer, portable storage, and backups, flash drives offer varying capacities, ranging from a few gigabytes to multiple terabytes.
3. Memory Cards: Utilized in devices like digital cameras and smartphones, memory cards come in various formats, including SD cards and microSD cards. They offer storage capacities ranging from a few gigabytes to hundreds of gigabytes.
4. Hard Disk Drives (HDD): Traditional mechanical storage devices, HDDs use rotating platters for data storage. They provide varying capacities, suitable for both consumer and enterprise-grade drives.
5. Solid State Drives (SSD): Faster and more reliable than HDDs, SSDs use flash memory. They offer storage capacities ranging from a few hundred gigabytes to several terabytes.

6. Direct-Attached Storage (DAS): Involves connecting storage devices directly to a single server or computer. It includes internal and external hard drives and SSDs connected via interfaces like USB, Thunderbolt, or eSATA.

7. Network-Attached Storage (NAS): A dedicated file-level storage solution connected to a network, NAS offers centralized storage accessible by multiple clients.

8. Storage Area Network (SAN): A high-performance storage solution providing block-level storage accessed through a dedicated network.

9. Cloud Storage: Involves storing data on remote servers accessed over the internet, offering virtually unlimited storage capacity.

10. Object Storage: A scalable solution storing data as objects rather than traditional file hierarchies.

11. Hybrid Storage: Combines on-premises storage infrastructure with cloud storage, optimizing performance, cost, and data availability.

12. Tape Storage: Involves storing data on magnetic tape cartridges for long-term archival and backup purposes.

4.5.1.2 Selecting Suitable Storage Solutions for Digital Learning Universities

In evaluating storage solutions for digital learning universities, it's crucial to consider factors such as storage capacity, performance, cost, and data availability. Optical discs, flash drives, and memory cards, while useful for certain applications, may have limitations in handling the vast amounts of data generated by digital learning platforms.

Therefore, focusing on solutions like Network-Attached Storage (NAS), Storage Area Network (SAN), Cloud Storage, Object Storage, Direct-Attached Storage (DAS), Hybrid Storage, and Tape Storage becomes essential. These solutions provide higher capacities, scalability, and better performance to manage the substantial data requirements of digital learning universities.

4.5.1.3 Hybrid Storage

Hybrid storage emerges as a comprehensive and suitable approach for digital learning universities, offering a balanced integration of various storage solutions. This section outlines the advantages of hybrid storage in the context of digital learning universities, showcasing its ability to harness the strengths of different storage technologies while mitigating their limitations.

1. Performance Optimization: Hybrid storage optimizes performance by leveraging the strengths of different storage technologies. Network-Attached Storage (NAS) provides centralized storage for efficient data sharing, while Direct-Attached Storage (DAS) offers fast, direct access. Tape storage ensures cost-effective long-term archival and a combination of Hard Disk Drives (HDDs) and Solid State Drives (SSDs) balances capacity and speed.

2. Cost Efficiency: Hybrid storage ensures cost efficiency by balancing storage requirements. It combines cost-effective solutions like NAS, DAS, and tape storage for large-scale data storage while utilizing HDDs for high capacity and SSDs for faster performance. This enables digital learning universities to allocate resources based on cost considerations, achieving an optimal balance between performance and budget.
3. Data Availability and Resilience: Enhancing data availability and resilience, hybrid storage utilizes NAS and DAS for immediate access and tape storage for reliable backup and long-term archival. Redundant storage across multiple technologies reduces the risk of data loss, ensuring high data availability critical for uninterrupted learning experiences.

By filtering out less suitable solutions and focusing on the strengths of hybrid storage, digital learning universities can establish a robust and flexible storage infrastructure that meets the demands of extensive data generation, performance optimization, and cost efficiency. Figure 4.1 visually illustrates the integration of various storage solutions in a hybrid storage model for digital learning universities.

4.5.1.4 Data Classification:

Academic Data:

1. Course Catalog: Information about the courses offered by the university, including course titles, descriptions, prerequisites, credit hours, and learning outcomes.

2. Course Materials: Educational resources provided to students for a specific course, such as textbooks, lecture notes, presentations, readings, multimedia content, and online learning materials.
3. Assignments and Assessments: Details about assignments, projects, quizzes, exams, and other forms of assessments given to students as part of their coursework. This includes submission deadlines, grading criteria, rubrics, and feedback provided by instructors.
4. Academic Calendar: Important dates and deadlines related to the academic year, including semester start and end dates, holidays, registration periods, add/drop deadlines, and examination schedules.
5. Student Registration Data: Information about student enrollment and registration, including course selections, schedule preferences, waitlists, and changes to enrollment status.
6. Academic Advising Records: Documentation of student-advisor interactions, academic plans, course recommendations, and progress towards degree completion.
7. Degree Requirements: Information on the requirements for different academic programs and degrees, including core courses, electives, major/minor requirements, credit hours, and any additional program-specific criteria.
8. Transcripts: Official records of a student's academic performance, including courses taken, grades earned, cumulative GPA, and degree(s) conferred.
9. Grading Records: Data related to student grades and assessments, including individual assignment grades, midterm and final exam grades, and overall course grades.
10. Graduation and Degree Audit Data: Information regarding students' progress towards graduation, degree audit reports, and requirements for degree completion.
11. Faculty and Staff Profiles: Profiles of academic faculty and staff members, including their educational background, research interests, areas of expertise, contact information, and office hours.
12. Academic Policies and Procedures: Documentation of institutional policies and procedures related to academic matters, such as grading policies, academic integrity policies, transfer credit policies, and academic appeals processes.

4.5.1.5 Risk Assessment

To establish a robust security framework tailored to the specific needs of West African universities, a comprehensive risk assessment has been conducted. The assessment identifies potential risks associated with various categories of data, each critical to the functioning of academic institutions. The following risk categories have been analyzed:

1. Personal Identifiable Information of Students and Staff:
   - Confidentiality Risks: Potential data breaches leading to privacy violations.
   - Integrity Risks: Risks associated with unauthorized data manipulation.
   - Availability Risks: Threats leading to limited access to crucial information.
2. Research Data:
   - Confidentiality Risk: Potential data breaches posing privacy violations.
   - Integrity Risks: Risks related to unauthorized data manipulation.
   - Availability Risks: Threats leading to potential data loss.
3. Financial Data:
   - Confidentiality Risks: Exposure to data breaches and privacy violations.
   - Integrity Risks: Risks associated with unauthorized data manipulation.
   - Availability Risks: Threats leading to limited access to financial information.
4. Academic Records:
   - Confidentiality Risks: Risks of data breaches compromising the confidentiality of academic records.
   - Integrity Risks: Potential threats related to unauthorized data manipulation.
   - Availability Risks: Risks leading to limited access to academic records.
5. Library Data:
   - Confidentiality Risks:Exposure to data breaches and privacy violations.
   - Integrity Risks: Risks associated with unauthorized data manipulation.
6. Administrative Records:
   - Confidentiality Risks: Risks of data breaches compromising the confidentiality of administrative records.
   - Integrity Risks: Potential threats related to unauthorized data manipulation.
7. Data from Digital Devices, Websites, LMS, Portals, etc.
   - Confidentiality Risks: Exposure to data breaches and privacy violations.

- Availability Risks: Risks leading to downtime and service disruptions.

Each identified risk is associated with specific potential consequences, enabling a focused approach to the development of mitigation strategies. The resulting risk assessment framework is crucial for informed decision-making and the establishment of proactive security measures. Table 13 visually represents the risk assessment matrix, offering a clear overview of the identified risks and their potential impacts on data security within the university context.

Table 13 Risk Assessment

Data	**Risk**	**Likelihood of Occurrence**	**Description**
Personal Identifiable information of students and staff	Data Breach	Moderate	The threat is **somewhat likely** to occur
	Privacy Violation	Moderate	The threat is **somewhat likely** to occur
	Data Manipulation	Low	The threat is **unlikely** to occur
	Limited Access	Very High	The threat is **almost certain** to occur
Research data	Data Breach	Low	The threat is **unlikely** to occur
	Privacy Violation	High	The threat is **highly likely** to occur
	Data Manipulation	Low	The threat is **unlikely** to occur
	Data Loss	Moderate	The threat is **somewhat likely** to occur
Financial Data	Data Breach	Very High	The threat is **almost certain** to occur
	Privacy Violation	Very High	The threat is **almost certain** to occur
	Data Manipulation	Low	The threat is **unlikely** to occur
	Limited Access	Moderate	The threat is **somewhat likely** to occur
Learning Management Systems Data	Data Breach	Low	The threat is **unlikely** to occur
	Data Manipulation	Very Low	The threat is **highly unlikely** to occur
	Downtime and Service Disruptions	Very High	The threat is **almost certain** to occur
Academic Data	Data Breach	Very High	The threat is **almost certain** to occur
	Privacy Violation	Very High	The threat is **almost certain** to occur
	Data manipulation	Very High	The threat is **almost certain** to occur
Library Data	Data Breach	Very High	The threat is **almost certain** to occur
	Privacy Violation	Very High	The threat is **almost certain** to occur
	Data manipulation	Very High	The threat is **almost certain** to occur
Administrative Records	Data Breach	Low	The threat is **unlikely** to occur
	Data manipulation	Very Low	The threat is **highly unlikely** to occur
Data from digital devices, websites, LMS, Portals etc	Data Breach	Low	The threat is **unlikely** to occur
	Privacy Violation	Very Low	The threat is **highly unlikely** to occur
	Downtime and service disruptions	Very High	The threat is **almost certain** to occur

4.5.1.6 Encryption:

In fortifying the security model, encryption stands as a paramount element, safeguarding sensitive data from unauthorized access. Among various encryption algorithms, the Advanced Encryption Standard (AES) emerges as the most fitting choice for the security model designed for digital learning universities.

AES, renowned for its cryptographic strength and widespread adoption, offers a high level of security in data protection. Its symmetric key encryption approach ensures that the same key is used for both encryption and decryption, streamlining the process without compromising security. The flexibility of AES in supporting key sizes of 128, 192, or 256 bits enhances its adaptability to varying security requirements.

The choice of AES aligns with the model's emphasis on robust data protection, particularly within the context of digital learning where confidentiality and integrity of academic, administrative, and user-related information are paramount. The algorithm's track record of resistance against various cyber threats and its compliance with industry standards make it a dependable choice for encrypting sensitive data within the digital learning environment.

Moreover, AES accommodates the hybrid storage approach integrated into the model, ensuring that all data stored locally undergoes encryption. This safeguards data at rest, reinforcing the security layers in place and aligning with the model's commitment to comprehensive data protection.

In conclusion, the adoption of AES encryption within the security model attests to the meticulous consideration of cryptographic principles and industry standards. Its robustness, versatility, and alignment with the overarching security goals make AES the optimal choice for fortifying the security measures within the designed data-centric architecture for digital learning universities.

4.5.1.7 Access Control Model

Access control within the context of a data-centric architecture in digital learning universities is a critical aspect of ensuring the security and integrity of sensitive information. Various access control models offer distinct methodologies for regulating access to data based on different criteria. The following access control models have been identified and evaluated for their applicability in the context of digital learning universities:

1. Attribute-Based Access Control (ABAC):ABAC assesses access decisions based on attributes associated with persons, objects, and the environment. It offers fine-grained control and dynamic access decisions.
2. Context-Based Access Control (CBAC): CBAC makes access decisions based on contextual information like time, location, and user behavior. It provides adaptive access control and enhanced security through real-time context consideration.
3. Graph-Based Access Control (GBAC): GBAC employs a graph structure to express access control interactions and dependencies, supporting complex access control situations
4. Lattice-Based Access Control (LBAC): LBAC defines security layers and access control regulations using a lattice structure, offering a mathematical foundation for layered security.
5. Mandatory Access Control (MAC): MAC implements access choices based on subject and object security labels, ensuring strict access rule enforcement and robust protection in high-security contexts.
6. Organization-Based Access Control (OrBAC): OrBAC involves access control rules within a company, utilizing organizational roles, connections, and hierarchies.
7. Role-Based Access Control (RBAC): RBAC grants access rights based on preset roles, simplifying access control administration and eliminating administrative complexity.
9. Rule-Based Trust Management (RTM): RTM establishes trust rules regulating resource access based on trust connections between entities.
10. Attribute-Based Encryption (ABE): ABE encrypts data based on attributes, enabling fine-grained access control over encrypted data.
11. Rule-Set-Based Access Control (RSBAC): RSBAC specifies access control rules for determining access choices based on various parameters.

12. Capability-Based Security (CBS): CBS gives access based on specified capabilities or tokens, providing decentralized control over resource access.
13. Discretionary Access Control (DAC): DAC enables resource owners to provide access to other users or groups, allowing control over resource access.
14. Hierarchical Attribute-Based Access Control (HABAC): HABAC adds attribute hierarchies to ABAC, offering more flexible and scalable access management.
15. Discretionary Mandatory Access Control (DMAC): DMAC combines DAC and MAC features, providing discretionary authority over some resources while imposing required access limits on others.

4.5.1.8 Selecting a Suitable Access Control Model for Data-centric Architecture in Digital Learning Universities:

When selecting an access control model for digital learning universities, careful consideration of specific criteria is crucial. The evaluation is based on three key factors:

1. Implementation Difficulties: RBAC simplifies access control administration through role-based organization, reducing complexity compared to fine-grained models.
2. Ability to handle large amounts of data: RBAC's hierarchical organization facilitates easier management and scalability compared to fine-grained models, ensuring efficient access control operations.
3. Maintenance Difficulties: RBAC's role-based approach allows for more straightforward policy modifications, enhancing flexibility in adapting to changes in user roles or data access requirements.

Considering these factors, RBAC emerges as the preferred choice for a data-centric architecture in digital learning universities, offering simplicity, scalability, and ease of maintenance in ensuring effective access control. This conclusion is drawn through a comparative analysis of RBAC against fine-grained access control models, taking into account implementation difficulties, data-handling capabilities, and maintenance challenges.

4.5.2 The Model

The Security Model is an integrated framework designed to fortify the data-centric architecture, ensuring robust protection against potential threats and vulnerabilities. It comprises interconnected components that collaboratively contribute to the overall security posture, emphasizing confidentiality, integrity, and availability of data (Figure 13). The key components include:

1. Role-Based Access Control (RBAC): RBAC serves as a pivotal component, offering a structured approach to access control. Users, including staff and students, undergo authentication processes, after which they are authorized based on their roles. Distinct roles such as student, staff, and examination officer are assigned specific spaces, delineating access to student spaces, course spaces, collaborative spaces, and data analytics spaces. This granular access control ensures that each user operates within predefined boundaries, enhancing data security and maintaining the principle of least privilege.
2. Data Encryption Protocols: All data stored locally is encrypted using the Advanced Encryption Standard (AES), ensuring end-to-end encryption. This comprehensive encryption strategy safeguards data at rest, in transit, and during processing, thwarting unauthorized access attempts effectively.
3. Continuous Monitoring and Surveillance: The security model features a dynamic surveillance mechanism that continuously monitors data activities, user interactions, and potential security incidents. This real-time monitoring capability enhances the model's responsiveness to emerging threats, enabling swift identification and mitigation of security breaches.
4. Incident Response Framework: In the event of a security incident, the model integrates an incident response framework. This component outlines predefined strategies and procedures to be executed promptly, ensuring a swift and effective response to security breaches. By minimizing response time, the model aims to mitigate potential damages and restore normalcy swiftly.
5. Security Awareness and Training Initiatives: Acknowledging the human factor in data security, the model incorporates educational initiatives and training programs. These aim to enhance the awareness and cybersecurity literacy of users within the digital learning university environment, fostering a culture of security consciousness.

6. Collaborative Data Governance Policies: The model emphasizes the establishment of comprehensive data governance policies that promote collaboration between different stakeholders. These policies define roles, responsibilities, and best practices to ensure a cohesive and well-coordinated approach to data security.

7. Hybrid Storage Integration: Hybrid storage is seamlessly integrated into the security model, offering a balanced approach to data storage. This includes local storage, encrypted using AES, and cloud-based solutions. The hybrid storage configuration optimizes performance, scalability, and data availability while maintaining a robust security posture.

8. Regulatory Compliance Framework: Aligning with relevant data protection regulations and industry standards, the security model incorporates a compliance framework. This ensures that the digital learning university adheres to legal requirements, fostering trust and accountability in handling sensitive data.

In essence, the Security Model's components collectively create a resilient and adaptive security infrastructure, integrating RBAC and hybrid storage to tailor access controls and storage solutions based on user roles within the digital learning university. The inclusion of encryption, monitoring, incident response, education, collaborative governance, and regulatory compliance forms a comprehensive security framework that safeguards the integrity and confidentiality of institutional data.

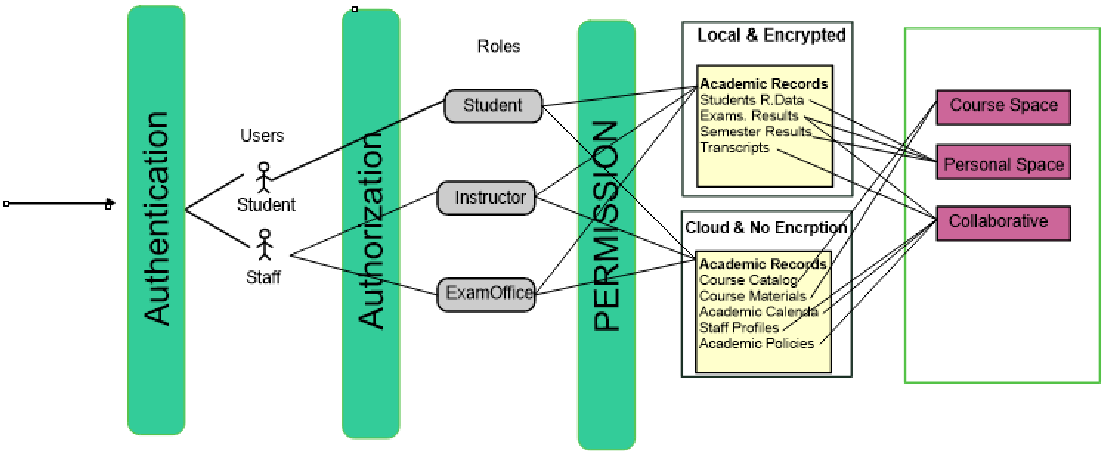


Figure 13 Security Model

4.6 Secure Grade Distribution Scheme

The “Secure Grade Distribution Scheme” will implement the student grade protection part from the security model in the previous section and will increase the security and privacy of grade data on the Moodle platform, and its effective deployment and thorough assessment will yield significant outcomes. Our research’s main findings and conclusions are shown in this section.

4.6.1 Key Management and Hardware Security Modules (HSMs)

Key management is essential to this technique as a fundamental component of data security. Hardware Security Modules (HSMs), specialized equipment made to generate and secure cryptographic keys, are used in the scheme. Grade data is encrypted and decrypted using these keys. Key management gains additional security and reliability with the integration of HSMs.

4.6.1.1 Key Generation

Mathematically, key generation within the scheme can be represented as follows:

Let K be the set of cryptographic keys used within the scheme, where $K =$*K1, K2, ..., Kn}.*
For each encryption session, a unique cryptographic key, K_i*, is generated using HSMs:*
$K_i = HSM.GenerateKey()$,

4.6.1.2 Key Storage

The generated cryptographic keys are securely stored within the HSMs. This can be mathematically represented as:

HSMs ensure the tamper-resistant storage of keys, which can be denoted as:
$HSM.StoreKey(K_i)$,

4.6.1.3 Key Retrieval

The cryptographic keys are safely retrieved from the HSMs for encryption or decryption. Mathematically, this can be represented as:

To obtain a specific key for an encryption or decryption session:

$K_i = HSM.RetrieveKey()$,

4.6.1.4 Nonce Integration

Nonces, arbitrary integers created for every encryption session, are key to improving security and preventing replay attacks. They are made securely and integrated into the generation of keys. The representation of nonce integration mathematically is as follows:

Let N be the set of nonces used within the scheme, where N = {N1, N2, ..., Nn}.

For each encryption session, a unique nonce, Ni, is generated:

$N_i = HSM.GenerateNonce()$,

The nonce is securely combined with the cryptographic key to create a session-specific key, denoted as Ki_nonce, ensuring unique keys for each session:

. $Ki_{nonce} = Ki \; XOR \; Ni$,

4.6.2 Advanced Encryption Standard (AES) Implementation

One of the scheme's main components for maintaining data security and secrecy is using the Advanced Encryption Standard (AES). AES is a well-known symmetric encryption technique known for its high security. To secure grade data against unwanted access, our system carefully integrates AES for encryption and decryption.

4.6.2.1 AES Encryption Process

AES operates on data in fixed-size blocks, applying a series of transformation rounds using a specific encryption key. In the context of the "Secure Grade Distribution Scheme," we employ AES-256, which operates a 256-bit encryption key for maximum security.

Mathematical Representation:

1. Data Division: Grade data, denoted as G, is divided into fixed-size blocks, represented as G1, G2, ..., Gn.
2. AES Encryption Rounds: AES performs a series of transformation rounds using the encryption key, K. The number of rounds depends on the critical size, with AES-256 using 14 rounds.

3. Block Encryption: AES encrypts each data block, Gi, using the encryption key, K, and the respective round, Ri. Mathematically, this can be represented as $C_i = AES\_Encrypt\ (G_i, K, R_i)$
4. Ciphertext Concatenation: The ciphertext blocks, Ci, are concatenated to form the complete Ciphertext, C.

4.6.3 AES Decryption Process:

On the recipient's end, AES decryption is applied to retrieve the original grade data from the Ciphertext. The decryption process is the reverse of encryption and is mathematically represented.

1. Ciphertext Division: The Ciphertext, C, is divided into blocks, represented as C1, C2, ..., Cn.
2. AES Decryption Rounds: AES decryption employs the same number of rounds and the decryption key, K, denoted as Rn, Rn-1, ..., R1, where n is the number of rounds.
3. Block Decryption: Each ciphertext block, Ci, is decrypted using the decryption key, K, and the corresponding round, Ri, yielding the original data block, Gi. Mathematically: $G_i = AES\_Decrypt\ (C_i, K, R_i)$
4. Data Concatenation: The decrypted data blocks, Gi, are concatenated to obtain the original grade data, G.

4.6.4 Diffie-Hellman Key Exchange Protocol

The Diffie-Hellman Key Exchange Protocol is a critical component of this scheme that enables secure key exchange between users, notably instructors and students. This protocol helps create a secure channel within the Moodle platform, allowing secure communication without requiring pre-shared keys.

4.6.4.1 Diffie-Hellman Key Exchange Process:

The Diffie-Hellman protocol allows two parties, in this case, instructors and students, to generate a shared secret key over an unsecured channel without explicitly sharing it. This process can be mathematically represented as follows:

1. Parameter Setup: A set of parameters, including a large prime number (p) and a primitive root (g), is chosen, and made publicly available. Both instructors and students use these parameters.
2. Key Generation (Instructors): Instructors generate their private keys (InstructorPrivateKey) and corresponding public keys (InstructorPublicKey) using the chosen parameters:
    a. InstructorPrivateKey = a (a randomly chosen secret integer)
    b. InstructorPublicKey = g^amod p
3. Key Generation (Students): Similarly, students generate their private keys (StudentPrivateKey) and corresponding public keys (StudentPublicKey) using the same parameters:
    a. StudentPrivateKey = b (b randomly chosen secret integer)
    b. StudentPublicKey = g^bmod p
4. Key Exchange:
    a. Instructors and students exchange their public keys (InstructorPublicKey and StudentPublicKey) over the unsecured channel.
5. Shared Secret Key: Both parties independently compute the shared secret key (SharedSecretKey) using the received public keys and their own private keys. Mathematically:
    a. Instructors calculate: SharedSecretKey = $studentPublicKey^a \ mod \ p$
    b. Students calculate: SharedSecretKey = $studentPublicKey^b \ mod \ p$

4.6.5 Message Integrity Code (MIC) Verification

To guarantee the integrity of grade data throughout transmission, the scheme includes Message Integrity Code (MIC) checking as a crucial security mechanism. It provides an effective way to find any unauthorized changes or tampering with the grade data.

4.6.5.1 MIC Generation Process:

Mathematically, the MIC generation process can be represented as follows:

1. MIC Generation (Instructors): When an instructor prepares to distribute grade data, a unique MIC is generated for each data packet (MIC_Instructor1, MIC_Instructor2, etc.). This is achieved by hashing the grade data and a secret key known only to the instructor.

   Mathematically:

   $$MIC\_Instructor\_i = Hash\ (GradeData\_i + InstructorSecretKey)$$

2. MIC Generation (Students): When students receive the data packets, they generate their own MICs for the received data. This ensures that they can verify data integrity and detect any unauthorized changes:

   $$MIC\_Student\_i = Hash(ReceivedData\_i + StudentSecretKey)$$

4.6.6 Key Metrics

In the context of key management, the scheme necessitates the generation and distribution of cryptographic keys for encryption and decryption processes. The key metrics include:

1. Instructor Keys:
   a. Private Key: Each instructor possesses a private key generated securely within the Moodle environment.
   b. Public Key: The corresponding public key is derived from the private key using the Diffie-Hellman key exchange protocol.
2. Student Keys:
   a. Private Key: Each student has a unique private key generated within the Moodle environment.
   b. Public Key: Similarly, the student's public key is generated through the Diffie-Hellman key exchange protocol.
3. Staff Keys:
   a. Private Key: Staff members also have a private key generated securely within Moodle.
   b. Public Key: The public key for staff is generated using the same Diffie-Hellman key exchange process.

4.6.6.1 Number of Keys Calculation

The number of keys required can be calculated based on the number of participants within the Moodle system. If there are 'n' instructors, 'm' staff members, and 'p' students, the total number of keys can be expressed as:

Total Keys = n(Instructor Keys) + m(Staff Keys) + p(Student Keys)

This formula accounts for the unique keys associated with each role within the educational environment. The integration ensures that each participant has the necessary cryptographic keys to engage in secure grade distribution.

4.6.7 Justification of Each Element

This section presents a detailed justification for each element incorporated into the "Secure Grade Distribution Scheme" based on the results of experiments conducted to assess their effectiveness.

4.6.7.1 Key Management and Hardware Security Modules (HSMs)

- Security Enhancement: The integration of Hardware Security Modules (HSMs) is justified by their ability to securely generate, store, and retrieve cryptographic keys. HSMs provide a dedicated and tamper-resistant environment, enhancing the overall security of the key management process.
- Reliability: The secure storage of cryptographic keys within HSMs ensures their reliability and protection against unauthorized access. This reliability is crucial for maintaining the confidentiality and integrity of grade data.

4.6.7.2 Advanced Encryption Standard (AES) Implementation

- High-Level Security: The use of the Advanced Encryption Standard (AES), specifically AES-256, is justified by its reputation for providing a high level of security. AES is widely recognized for its resistance to various cryptographic attacks, making it suitable for safeguarding sensitive grade data.

- Symmetric Encryption Efficiency: AES's symmetric encryption approach is efficient for bulk data encryption and decryption, ensuring that the process is both secure and computationally feasible within the Moodle environment.

4.6.7.3 Diffie-Hellman Key Exchange Protocol

- Secure Key Exchange: The Diffie-Hellman Key Exchange Protocol is justified by its ability to facilitate secure key exchange between instructors, staff, and students. It eliminates the need for pre-shared keys, enhancing the overall security of communication channels within Moodle.
- Public and Private Key Generation: The use of public and private keys in the Diffie-Hellman protocol allows entities to securely share public keys while maintaining the confidentiality of their private keys. This ensures a secure and efficient key exchange process.

4.6.7.4 Message Integrity Code (MIC) Verification

- Tamper Detection: MIC verification is crucial for detecting any unauthorized changes or tampering with grade data during transmission. This element ensures the integrity of the data, preventing malicious alterations.
- Hashing for Integrity: The use of hash functions for MIC generation provides a reliable and efficient method for verifying data integrity. Hashing ensures that even minor changes to the data result in significantly different MIC values.

4.7 Case Scenario

In this scenario, we will examine the processes and steps an instructor takes to transmit grades to staff and students safely. The instructor broadcasts the ciphertexts and Message Integrity Codes (MICs) during transmission via a public channel, allowing staff and students to view the encrypted grades (refer to Figure 14).

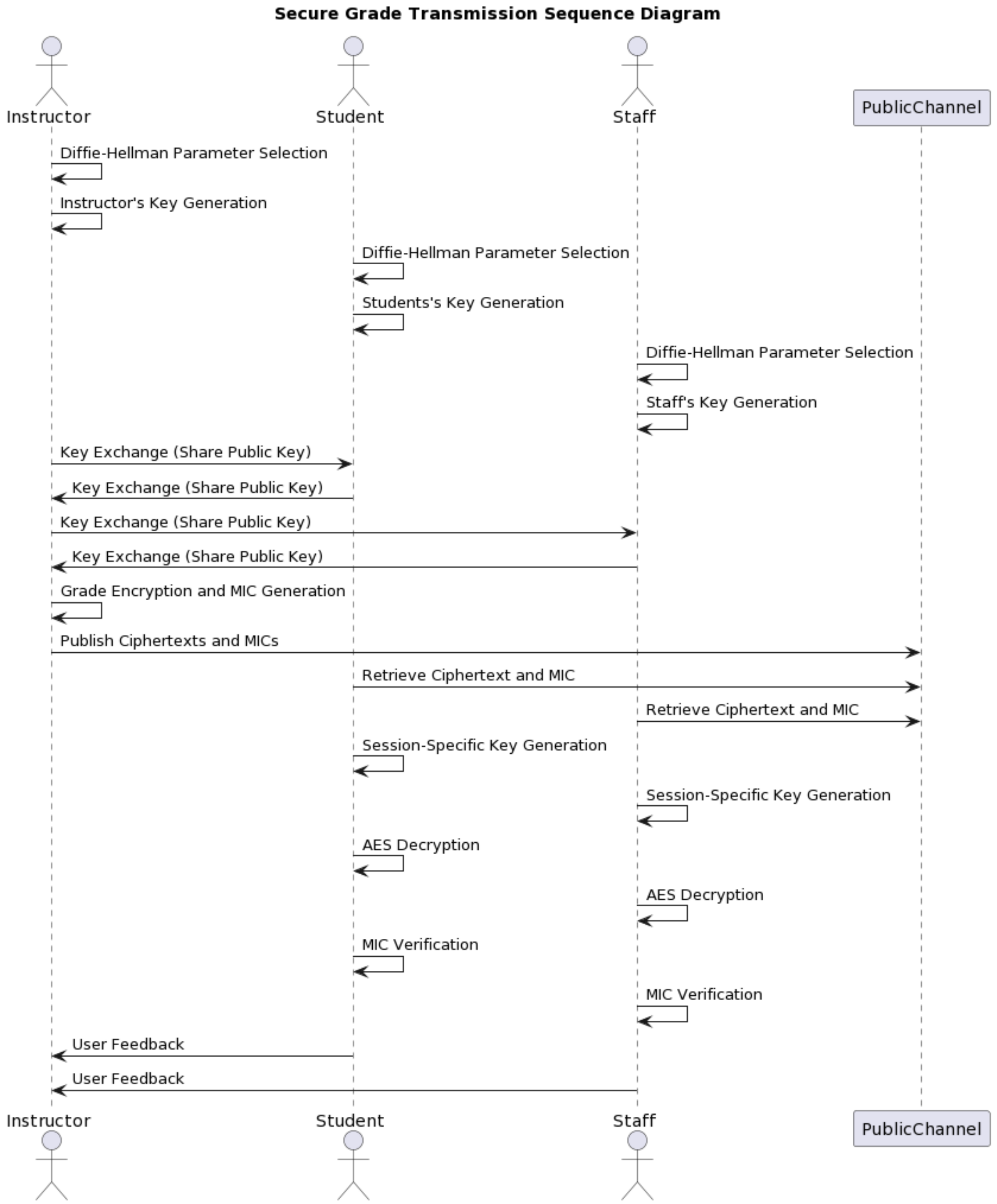

Secure Grade Transmission Sequence Diagram
Instructor
Student
Staff
PublicChannel
Diffie-Hellman Parameter Selection
Instructor's Key Generation
Diffie-Hellman Parameter Selection
Students's Key Generation
Diffie-Hellman Parameter Selection
Staff's Key Generation
Key Exchange (Share Public Key)
Key Exchange (Share Public Key)
Key Exchange (Share Public Key)
Key Exchange (Share Public Key)
Grade Encryption and MIC Generation
Publish Ciphertexts and MICs
Retrieve Ciphertext and MIC
Retrieve Ciphertext and MIC
Session-Specific Key Generation
Session-Specific Key Generation
AES Decryption
AES Decryption
MIC Verification
MIC Verification
User Feedback
User Feedback
Instructor
Student
Staff
PublicChannel

Figure 14: Use case diagram of the case scenario

Step 1: Instructor’s Initial Setup

1. Diffie-Hellman Parameter Setup:

The instructor selects a large prime number, p.

The instructor selects a primitive root, g (where g is a primitive root modulo p).

2. Key Generation (Instructor):

The instructor generates a private key:

InstructorPrivateKey (a randomly chosen secret integer).

The instructor computes the corresponding public key:

InstructorPublicKey using the equation: InstructorPublicKey = $g^{InstructorPrivateKey} mod\ p$

Step 2: Student and Staff Setup

3. Key Generation (Student and Staff):

Students generate their private keys, StudentPrivateKey (randomly chosen secret integers).

Students compute their public keys, StudentPublicKey, using the equation: StudentPublicKey = $g^{StudentPrivateKey} mod\ p$

Staff generate their private keys, StaffPrivateKey (randomly chosen secret integers).

Staff compute their public keys, StaffPublicKey, using the equation:

StaffPublicKey = $g^{StaffPrivateKey} mod\ p$

Step 3: Diffie-Hellman Key Exchange

4. Key Exchange:

Instructor and students exchange their public keys (InstructorPublicKey and StudentPublicKey).

Instructor and staff exchange their public keys (InstructorPublicKey and StaffPublicKey).

5. Shared Secret Key Calculation**:**

The instructor calculates SharedSecretKey using the equation:

$$SharedSecretKey = StudentPublickey^{InstructorPrivateKey} mod\ p$$

The instructor calculates SharedSecretKey using the equation:

$$SharedSecretKey = StaffPublickey^{InstructorPrivateKey} mod\ p$$

Students and staff calculate SharedSecretKey similarly. Both parties derive the same SharedSecretKey.

Step 4: Grade Encryption and MIC Generation

6. Grade Data: The instructor has grade data, GradeData.

7. Nonce Generation: A unique nonce, Ni, is generated for this session.

8. Session-Specific Key Creation:

The instructor creates a session-specific key by combining the SharedSecretKey with the nonce: Ki_instructor = SharedSecretKey XOR Ni.

9. AES Encryption: The instructor encrypts the grade data (GradeData) using the session-specific key (Ki_instructor) and obtains the Ciphertext Ciphertext.

10. MIC Generation: The instructor calculates a Message Integrity Code (MIC) for the encrypted grades using a cryptographic hash function:
MIC = Hash(Ki_instructor || Ciphertext).

Step 5: Publication on a Public Channel

11. Public Channel Transmission: The instructor publishes the Ciphertext and MIC on a public channel accessible to staff and students.

Step 6: Decryption and MIC Verification (Student and Staff):

12. Decryption:
Students and staff retrieve the Ciphertext and MIC from the public channel.

13. Session-Specific Key Creation (Student):
The student calculates the session-specific key: Ki_student = SharedSecretKey XOR Ni.
The staff calculates the session-specific key: Ki_staff = SharedSecretKey XOR Ni.

14. AES Decryption:

The student decrypts the Ciphertext using Ki_student and retrieves the grade data (GradeData).
The staff decrypts the Ciphertext using Ki_staff and retrieves the grade data (GradeData).

15. MIC Verification (Student and Staff):

The student calculates a Message Integrity Code (MIC) using the received encrypted data (Ciphertext) and the shared key (Ki_student) and compares it to the received MIC.
The staff calculates a Message Integrity Code (MIC) using the received encrypted data (Ciphertext) and the shared key (Ki_staff) and compares it to the received MIC.

Step 7: User Feedback

16. User Feedback: Both students and staff provide feedback on the grade distribution experience, security, and any issues or suggestions for improvement.

4.7.1 Discussion

Integrating a secure grade distribution mechanism inside the Moodle platform is essential to guaranteeing the security and privacy of sensitive academic data. The system provides a reliable solution for grade data security by integrating cutting-edge cryptographic methods, including Diffie-Hellman key exchange, AES encryption, and Message Integrity Code (MIC) verification. Using the Diffie-Hellman key exchange protocol is one of this system’s advantages.

Additionally, via the rapid and safe establishment of encryption keys made possible by this protocol, staff, students, and instructors may securely interact without directly trading sensitive information. The data is well safeguarded during transmission because of robust session-specific keys derived from shared secrets and nonces. Another feature of the system is its use of the Advanced Encryption Standard (AES) for high-grade data encryption. Since AES is a well-used and reliable encryption technology, its applicability for protecting educational data is demonstrated by its mathematical form. Users are reassured about the system's security by the openness with which the encryption and decryption procedure is explained. Adding Message Integrity Code (MIC) checking improves the system's security. It guarantees the received data's integrity, enabling the detection of any unauthorized alteration. To protect cryptographic keys, a crucial step is the adoption of a Hardware Security Module (HSM) for secure key management; by offering a safe and specialized setting for key storage and cryptographic operations, HSMs lower the possibility of key compromise.

4.8 Proof of Concept

The proof of concept serves as a pivotal phase in this research, offering a hands-on demonstration of the practical implementation of the designed security model within the context of a digital learning university. Through the establishment of a controlled lab environment using virtual machines, the model will be tested and validated to showcase its effectiveness in safeguarding data integrity, confidentiality, and availability.

This practical demonstration will utilize key components of the security model, Data-centric model and grade distribution scheme including Apache Atlas, Apache Ranger, Kafka, and Kibana with Elasticsearch as well as Moodle. The implementation will focus on academic data from digital learning universities using RBAC roles (Table 14). By meticulously configuring and deploying these components, the proof of concept aims to illustrate the seamless integration of the security model and its ability to mitigate potential cyber threats and enhance overall data governance.

Table 14 RBAC Roles

Permissions	Instructor	Student	Exermination officer/ Staff
Grade Preparation	X		
Encryption	X		
Key Management	X		
Diffie-hellman key Exchange	X		
HSM Integration	X		
Decryption		X	X
Secure Communication	X	X	X
Data analytics			X

The stepwise procedure will be accompanied by detailed explanations and visual representations, providing a comprehensive understanding of how the security model operates within the digital learning environment. This hands-on approach ensures that the research findings are not only theoretical but also practically applicable, offering a valuable resource for digital learning universities seeking robust data-centric security solutions."

4.8.1 Application of the Data-Centric Architecture Model

To assess the feasibility and applicability of the proposed data-centric architecture model for digital learning universities, a proof of concept (PoC) was conducted in a controlled lab environment. The objective was to validate the model's effectiveness in handling the unique demands of digital learning. The following outlines the stepwise process undertaken in the PoC:

*4.8.1.1 Component Deployment and Configuration***

In this phase, key components of the data-centric architecture were deployed and configured. Apache Atlas was installed initially, followed by the subsequent installation of Apache Ranger and Kafka. The computer storage served as the designated data hub for streamlined data processing.

Stepwise Procedure:

1. Apache Atlas Installation: Apache Atlas was installed on the VM to facilitate metadata management (Figure 15).

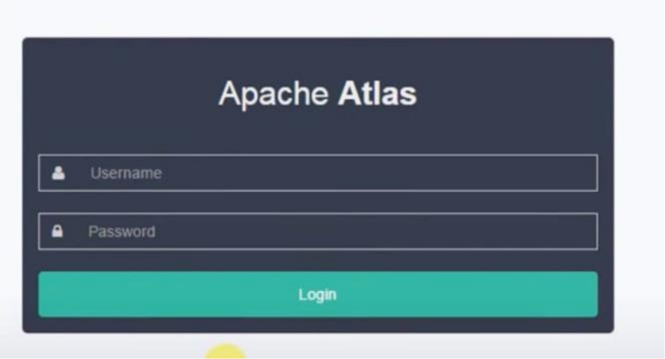
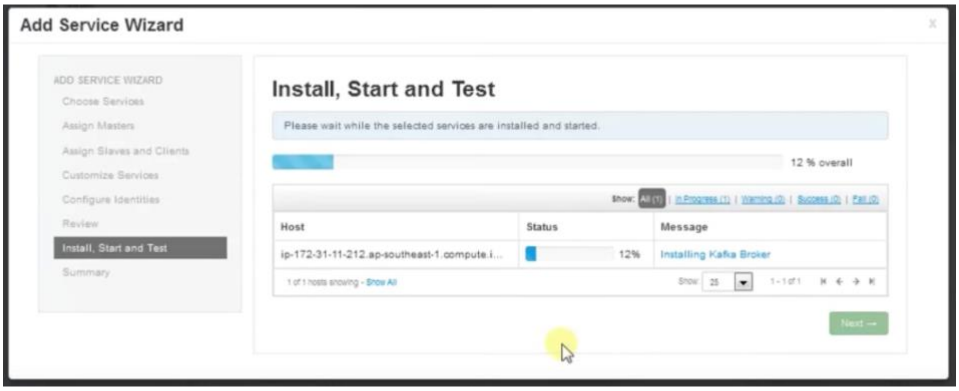


Figure 15. Apache Atlas Login Page

2. Apache Ranger Installation: Apache Ranger was installed and configured in tandem with Apache Atlas for testing the developed security model.

3. Kafka Integration: Kafka was integrated into the system to optimize data processing and distribution (Figure 16).



Install, Start and Test

Please wait while the selected services are installed and started.

12 % overall

Show: All (1) | In Progress (1) | Warning (0) | Success (0) | Fail (0)

Host	Status	Message
ip-172-31-11-212.ap-southeast-1.compute.i...	12%	Installing Kafka Broker

1 of 1 hosts showing - Show All   Show: 25   1 - 1 of 1

Figure 16. Kafka Integration with Apache Atlas

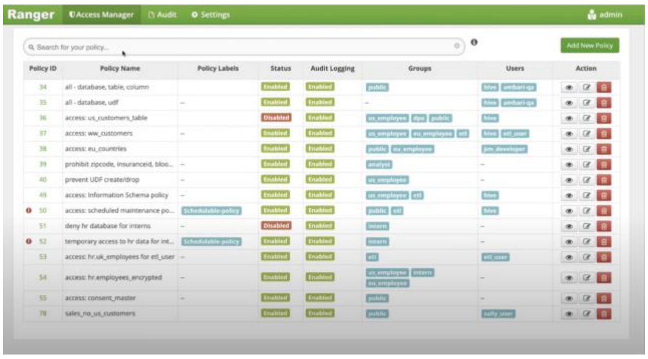
4. Data-Hub Configuration: The computer storage was designated as the central data hub for efficient data organization.

4.8.2.3 Testing and Validation

This crucial phase involved thorough testing and validation of the data-centric model to ensure its robustness and effectiveness.

Stepwise Procedure:

1. Security Model Testing: Apache Ranger and Apache Atlas were rigorously tested to validate the efficacy of the developed security model.



Policy ID	Policy Name	Policy Labels	Status	Audit Logging	Groups	Users	Action
34	all - database, table, column		Enabled	Enabled	public	hive ambari-qa	
35	all - database, udf	–	Enabled	Enabled	–	hive ambari-qa	
36	access: us_customers_table		Disabled	Enabled	us_employee dpo public	hive	
37	access: ww_customers	–	Enabled	Enabled	us_employee eu_employee etl	hive etl_user	
38	access: eu_countries		Enabled	Enabled	public eu_employee	jim_developer	
39	prohibit zipcode, insuranceid, bloo...	–	Enabled	Enabled	analyst	–	
40	prevent UDF create/drop	–	Enabled	Enabled	us_employee	–	
49	access: Information Schema policy	–	Enabled	Enabled	us_employee etl	hive	
50	access: scheduled maintenance po...	Schedulable-policy	Enabled	Enabled	public etl	hive	
51	deny hr database for interns	–	Disabled	Enabled	intern	–	
52	temporary access to hr data for int...	Schedulable-policy	Enabled	Enabled	intern	–	
53	access: hr.uk_employees for etl_user	–	Enabled	Enabled	etl	etl_user	
54	access: hr.employees_encrypted	–	Enabled	Enabled	us_employee intern eu_employee	–	
55	access: consent_master	–	Enabled	Enabled	public	–	
78	sales_no_us_customers		Enabled	Enabled	public	sally_user	

Figure 17. Apache Ranger configuration Page

2. Kafka Integration Testing: The integration of Kafka was tested to ensure seamless data processing.

4.8.2.4 Data Flow and Connectivity Testing

A critical aspect of the proof of concept involved assessing the flow of data within the architecture and ensuring seamless connectivity between components.

Stepwise Procedure:

1. Data Flow Analysis: The flow of data within the architecture was analyzed to identify potential bottlenecks.
2. Connectivity Testing: The connectivity between Apache Atlas, Apache Ranger, Kafka, and the data hub was tested for optimal performance (Figure 18).

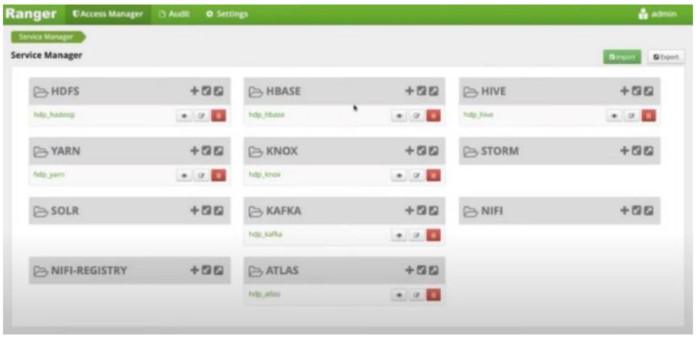


Figure 18. Apache Atlas, Apache Ranger and Kafka Configuration

4.8.2.5 Visualization and Analysis Setup

To enhance data visualization and analysis, Elasticsearch and Kibana were integrated into the architecture using Logstash and Docker.

Stepwise Procedure:

1. Elasticsearch and Kibana Integration: Elasticsearch and Kibana were configured to complement Kafka for streamlined data visualization (Figure 19).

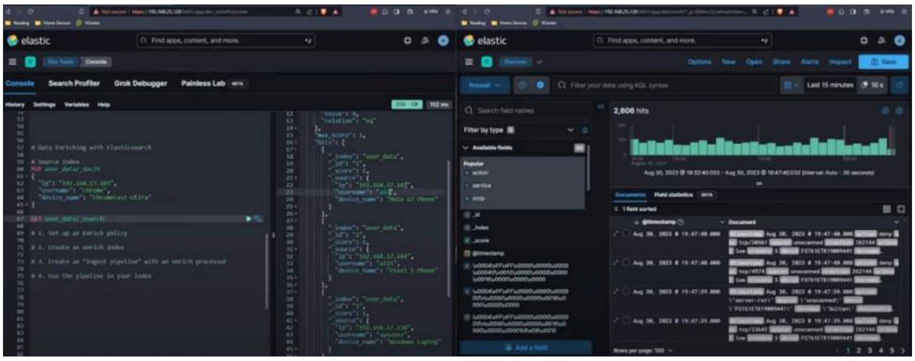
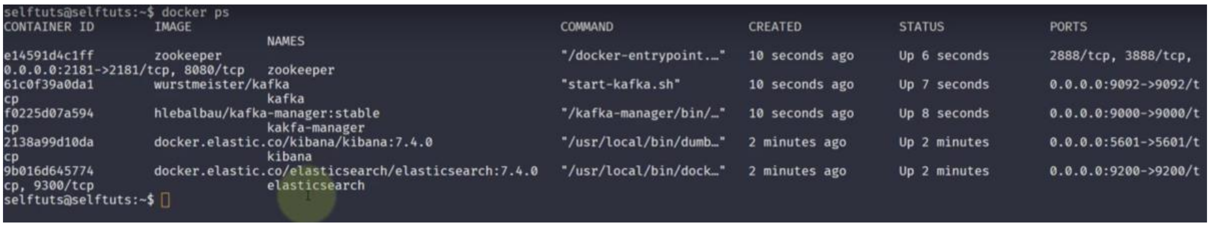


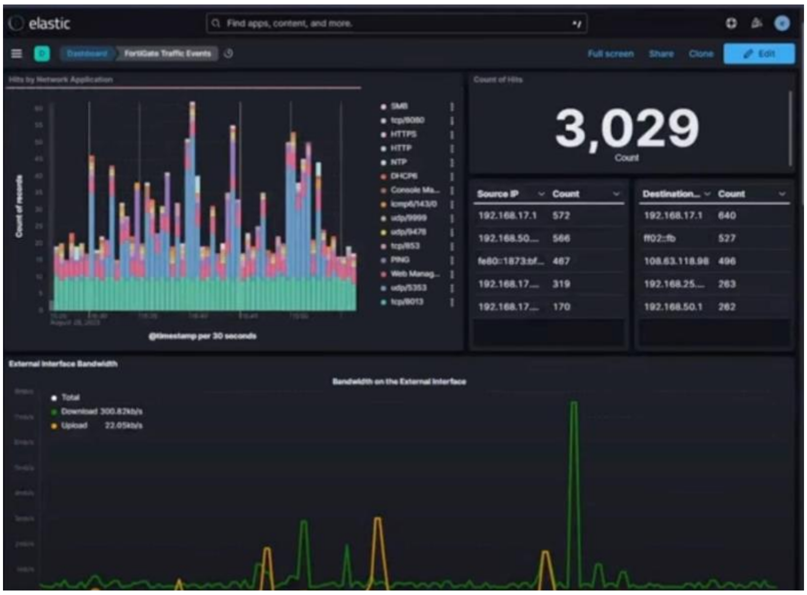
Figure 20. Data Integration to Elastic Search using Kafka.

2. Logstash and Docker Integration: Logstash and Docker were employed to facilitate the efficient flow of data for visualization and analysis (Figure 20 and Figure 21).



```
selftuts@selftuts:~$ docker ps
CONTAINER ID        IMAGE                                                   COMMAND                  CREATED             STATUS              PORTS
                              NAMES
e14591d4c1ff        zookeeper                                               "/docker-entrypoint.…"   10 seconds ago      Up 6 seconds        2888/tcp, 3888/tcp,
0.0.0.0:2181->2181/tcp, 8080/tcp   zookeeper
61c0f39a0da1        wurstmeister/kafka                                      "start-kafka.sh"         10 seconds ago      Up 7 seconds        0.0.0.0:9092->9092/t
cp                            kafka
f0225d07a594        hlebalbau/kafka-manager:stable                          "/kafka-manager/bin/…"   10 seconds ago      Up 8 seconds        0.0.0.0:9000->9000/t
cp                            kakfa-manager
2138a99d10da        docker.elastic.co/kibana/kibana:7.4.0                   "/usr/local/bin/dumb…"   2 minutes ago       Up 2 minutes        0.0.0.0:5601->5601/t
cp                            kibana
9b016d645774        docker.elastic.co/elasticsearch/elasticsearch:7.4.0     "/usr/local/bin/dock…"   2 minutes ago       Up 2 minutes        0.0.0.0:9200->9200/t
cp, 9300/tcp                  elasticsearch
selftuts@selftuts:~$ 
```

Figure 21. Kafka integration with Kibana and Elasticsearch



Source IP	Count
192.168.17.1	572
192.168.50....	566
fe80::1873:bf...	467
192.168.17....	319
192.168.17....	170

Destination...	Count
192.168.17.1	640
ff02::fb	527
108.63.118.98	496
192.168.25....	263
192.168.50.1	262

Figure 22. Analytics Dashboard

These stepwise procedures and corresponding diagrams provide a comprehensive view of the proof of concept conducted in the virtualized lab environment, offering clarity on the deployment, testing, and validation processes within the data-centric architecture.

4.8.2 Application of the Secure Grade Distribution Scheme

The feasibility and practicality of the proposed scheme for enhancing security and privacy in educational environments, particularly within the context of Moodle integration, were assessed through a proof of concept (Appendix B). This section outlines the steps involved in developing, implementing, and evaluating the scheme within a controlled lab environment. The proof of concept, however, does not employ HSM; instead, all cryptographic keys are kept in a file inside Moodle.

4.8.2.1 Lab Environment Setup

A virtualized lab environment was created, consisting of three distinct Virtual Machines (VMs) to represent instructors, staff, and students. These VMs were configured to operate on the same network, allowing for seamless communication.

4.8.2.2 Moodle Installation and Configuration

Moodle, the widely used open-source Learning Management System, was installed on each of the VMs, mimicking a real educational environment. The installation and configuration encompassed the web server setup, database creation, and Moodle initialization. A shared folder was established on the instructor's VM for resource sharing.

4.8.2.3 Scheme Plugin Development

A custom scheme plugin was developed to implement secure grade distribution features within Moodle. The plugin integrated Advanced Encryption Standard (AES) encryption, Diffie-Hellman key exchange, and Message Integrity Code (MIC) verification. The development environment included PHP tools and a code editor.

4.8.2.4 Plugin Installation and Activation

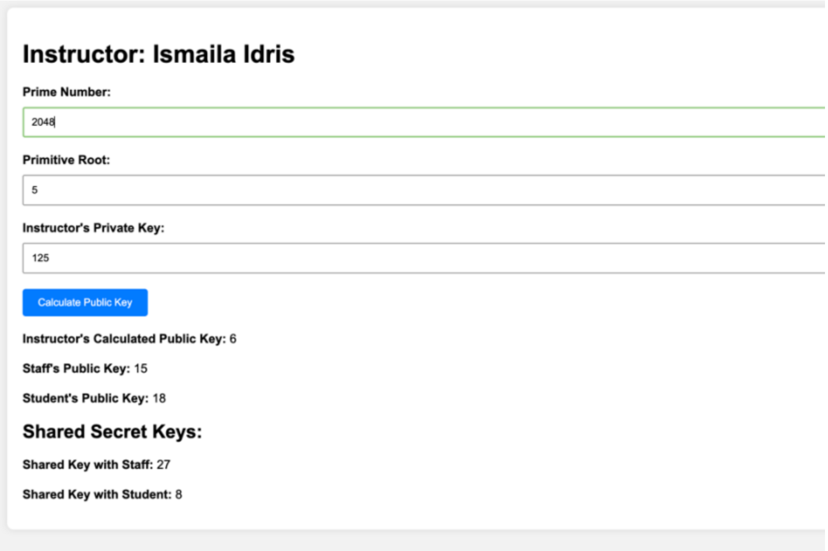
The developed scheme plugin was uploaded and activated on each of the Moodle instances representing instructors, staff, and students. This enabled the secure grade distribution features across the roles.

4.8.3 Testing and Validation

A series of tests were conducted to verify the functionality of the scheme plugin:

4.8.3.1 Key Exchange

The Diffie-Hellman key exchange was initiated between the instructor and each student to establish a shared secret key (Figure 23).



Instructor: Ismaila Idris

Prime Number:

2048

Primitive Root:

5

Instructor's Private Key:

125

Instructor's Calculated Public Key: 6

Staff's Public Key: 15

Student's Public Key: 18

Shared Secret Keys:

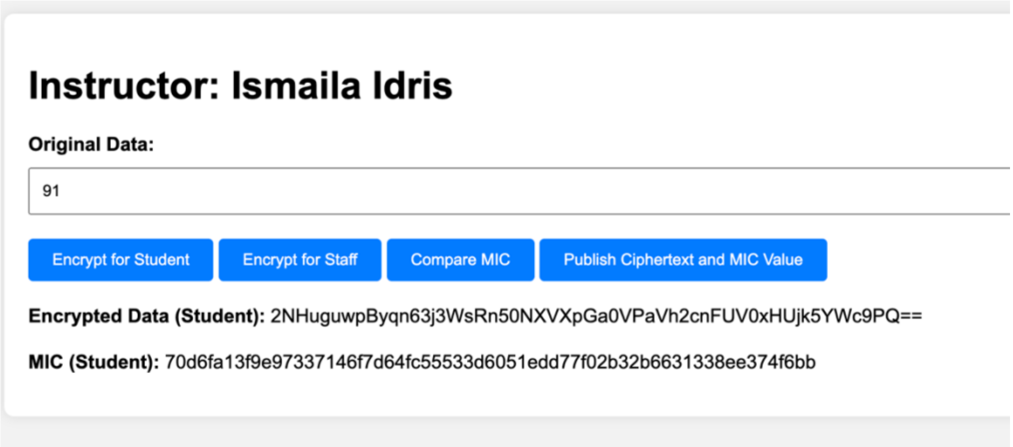
Shared Key with Staff: 27

Shared Key with Student: 8

Figure 23. Key Exchange using Diffie-hellman

4.8.3.2 AES Encryption and Decryption

Instructors encrypted grades and shared them with staff and students. Recipients successfully decrypted the ciphertext using the shared secret key (Figure 22).



Instructor: Ismaila Idris

Original Data:

91

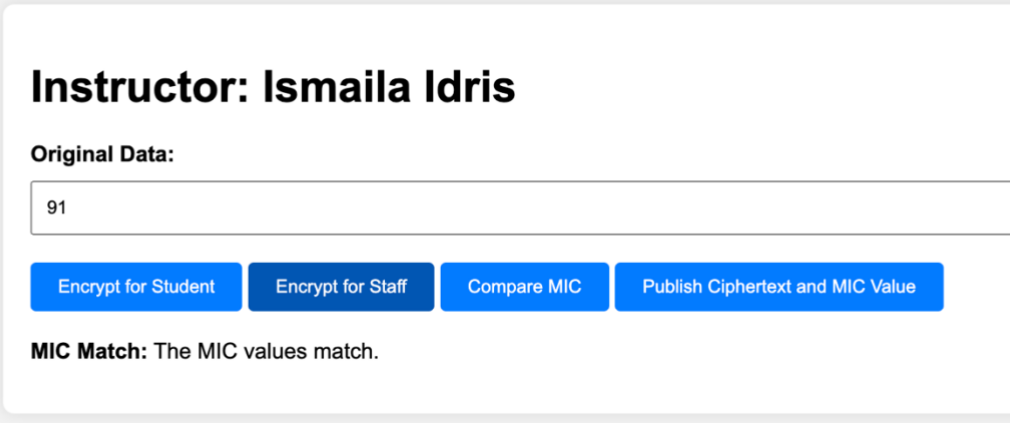
Encrypted Data (Student): 2NHuguwpByqn63j3WsRn50NXVXpGa0VPaVh2cnFUV0xHUjk5YWc9PQ==

MIC (Student): 70d6fa13f9e97337146f7d64fc55533d6051edd77f02b32b6631338ee374f6bb

Figure 23. AES Encryption and MIC generation

4.8.3.3 MIC Verification

Recipients verified the integrity of the received grades using Message Integrity Code (MIC) validation (Figure 24).



Instructor: Ismaila Idris

Original Data:

91

MIC Match: The MIC values match.

Figure 24. MIC Verification

4.8.3.4 Evaluation and Feedback

Feedback was gathered from participants who represented the roles of instructors, staff, and students in the lab environment. The feedback aimed to evaluate the ease of use, security, and effectiveness of the scheme.

CHAPTER FIVE: SUMMARY, CONCLUSION AND RECOMMENDATIONS

5.1 Summary

This research embarked on a comprehensive exploration of data-centric architecture tailored to the context of digital learning universities, particularly within the West African region. The study initiated with a meticulous survey methodology designed to capture insights from technical staff actively engaged with diverse university systems and portals. Employing a combination of survey queries and case studies, the research undertook a robust comparative analysis. The data collection process was methodically crafted, involving the development of survey questions, rigorous validity testing, wide-scale dissemination through established channels, and subsequent analysis employing both a four-point Likert scale and the statistical tool SPSS.

The participant pool, consisting primarily of technical staff, ensured the generation of focused and pertinent data. The study impressively secured responses from 93 universities, surpassing the initially set target of 70%, thereby demonstrating a robust and representative sample. The research design effectively addressed key research questions, encompassing critical aspects such as data architectures, cyber threats, countermeasures, and the role of data in decision-making within university environments. This comprehensive approach to data collection and analysis facilitated the extraction of valuable insights and trends.

Moving forward, the study undertook a meticulous comparison of data architectures, unveiling specific criteria for evaluation. This involved scrutinizing methodologies, understanding data sources, and appraising various storage solutions. The research identified and classified 109 e-learning solution use cases, offering a detailed breakdown based on the data architectures they employed. The findings from this analysis critically underscored the limitations associated with data-driven architecture in terms of security, leading to a compelling recommendation for the adoption of data-centric architecture within e-learning environments.

Continuing the exploration, the study delved into the intricate realm of access control models, providing a nuanced comparison and evaluation. A particular emphasis was placed on the selection of a suitable model for digital learning universities, with Role-Based Access Control

(RBAC) emerging as a pragmatic choice for simplifying access control administration and optimizing performance within the educational context.

The research then navigated through the design and conceptualization of a data-centric model specifically tailored to the unique demands of digital learning universities. This involved a systematic consideration of components such as data sources, connectors, data hubs, governance and security frameworks, data as service, user-centric apps, analytic factories, and analytics as a service. Each component was intricately examined to ensure seamless integration and efficiency within the designed model.

Security stood as a paramount concern throughout the study, prompting a detailed exploration of security models and the subsequent design of a robust framework. The risk assessment meticulously scrutinized various data categories, including personal identifiable information, research data, financial data, academic records, library data, administrative records, and data from digital devices. The comprehensive evaluation led to the incorporation of a security model bolstered by Apache Ranger and Apache Atlas, aimed at fortifying data governance and security within the digital learning environment.

To validate the proposed data-centric architecture and security model, a proof of concept was meticulously executed. This involved the identification of the most suitable storage solutions, particularly focusing on the unique demands of digital learning universities. The study advocated for a hybrid storage approach, leveraging both local and cloud storage, with a crucial emphasis on encrypting all data in local storage using the AES encryption standard. This multifaceted approach aimed to optimize performance, ensure cost efficiency, and enhance data availability and resilience.

In the final stages of the research, the study strategically introduced Kafka, a high-performance data streaming platform, to enhance the overall architecture. The inclusion of Kafka underscored the commitment to leveraging cutting-edge technologies for seamless data processing and distribution within the digital learning landscape.

In summation, this research unfolds as a comprehensive and meticulously executed endeavor, navigating through the intricacies of data-centric architecture in digital learning universities. From the nuanced survey methodology to the design of a tailored data-centric model, the study contributes valuable insights, recommendations, and a concrete proof of concept to the evolving landscape of educational technology and data management within the West African context.

5.2 Conclusion

In conclusion, this research has traversed the intricate landscape of data-centric architecture in the context of digital learning universities, with a specific focus on the unique demands and challenges within the West African region. The study was initiated with a robust survey methodology, engaging technical staff from numerous universities to garner rich insights and perspectives. The overwhelming response from 93 universities not only exceeded the set target but also ensured a comprehensive and representative dataset for analysis.

The comparative analysis of data architectures shed light on the limitations of data-driven architecture within the e-learning domain, paving the way for a compelling recommendation in favour of data-centric architecture. The research identified and classified 109 e-learning solution use cases, offering a detailed breakdown based on the data architectures they employed. This classification not only contributes to the academic understanding of prevailing trends but also provides practical insights for decision-makers in the educational technology landscape.

A critical aspect of the research focused on the meticulous design of a data-centric model tailored to the specific needs of digital learning universities. Components such as data sources, connectors, data hubs, governance and security frameworks, data as service, user-centric apps, analytic factories, and analytics as a service were intricately examined, ensuring a holistic and seamlessly integrated architecture.

Security remained a paramount concern throughout the study, prompting an in-depth exploration of security models and the subsequent design of a robust framework. The risk assessment, covering various data categories, facilitated the development of a security model fortified by Apache Ranger and Apache Atlas. This model, designed to enhance data governance and security, stands as a pivotal contribution to the evolving field of cybersecurity within educational institutions.

The validation of the proposed data-centric architecture and security model through a proof of concept further strengthens the practical relevance of the research. The advocacy for a hybrid storage approach, coupled with encryption standards for local storage, reflects a nuanced understanding of the balance required between performance, cost efficiency, and data resilience within digital learning environments.

The strategic introduction of Kafka, a high-performance data streaming platform, adds a layer of sophistication to the overall architecture, emphasizing the research's commitment to leveraging cutting-edge technologies for seamless data processing and distribution.

In essence, this research not only advances academic knowledge in the realm of data-centric architecture but also offers actionable insights and recommendations for digital learning universities in West Africa. The comprehensive and meticulously executed nature of this study positions it as a valuable contribution to the evolving landscape of educational technology and data management, with implications for policymakers, educators, and technology practitioners in the region.

5.3 Recommendations

Based on the comprehensive findings and insights derived from this research, several recommendations are put forth to guide digital learning universities, policymakers, and technology practitioners in enhancing their data-centric architecture and cybersecurity frameworks:

1. Adoption of Data-Centric Architecture: Digital learning universities in West Africa are encouraged to transition towards a data-centric architecture. This shift should be underpinned by a thorough assessment of the specific needs and challenges within each institution. Embracing a data-centric model will contribute to improved scalability, flexibility, and efficiency in managing the vast amounts of data generated in the e-learning ecosystem.
2. Integration of Hybrid Storage Solutions: Considering the diverse data types and storage requirements in digital learning environments, the adoption of hybrid storage solutions is recommended. This approach allows for a balanced and cost-effective allocation of storage resources, leveraging the strengths of both local and cloud

storage. Implementation should also prioritize robust encryption standards for local storage to enhance data security.

3. Implementation of Apache Ranger and Apache Atlas: To fortify data governance and security, universities are advised to implement robust frameworks such as Apache Ranger and Apache Atlas. These tools provide a comprehensive set of features for access control, policy enforcement, and metadata management. Customization based on specific institutional requirements is recommended to ensure a tailored and effective security model.

4. Leveraging Kafka for Streamlining Data Processing: The incorporation of Kafka as a high-performance data streaming platform is recommended to streamline data processing and distribution. This technology can enhance the real-time capabilities of digital learning platforms, facilitating efficient data flow and analysis. Universities should consider the scalability and adaptability of Kafka to meet the evolving demands of e-learning.

5. Continuous Security Training and Awareness: Recognizing the dynamic nature of cybersecurity threats, universities should invest in continuous training and awareness programs for staff and students. This proactive approach will foster a culture of cybersecurity awareness and responsible data handling, reducing the risk of security breaches.

6. Regular Updates and Audits: It is imperative for institutions to prioritize regular updates of software, security protocols, and data management policies. Conducting periodic security audits and assessments will help identify vulnerabilities and ensure that the implemented data-centric architecture remains resilient against emerging threats.

7. Collaboration and Information Sharing: Digital learning universities are encouraged to foster collaboration and information sharing within the academic community. Establishing forums or consortia where institutions can share insights, best practices, and challenges related to data-centric architecture and cybersecurity will contribute to collective resilience.

8. Research and Development Initiatives: Investing in research and development initiatives specific to data-centric architecture and cybersecurity in the context of West African digital learning environments is recommended. This proactive stance will contribute to the evolution of tailored solutions that address regional nuances and challenges.

By embracing these recommendations, digital learning universities can not only strengthen their data-centric architecture but also enhance the overall cybersecurity posture, ensuring a secure and efficient learning environment for students and faculty.

5.4 Contributions to Knowledge

This research makes significant contributions to the field of data-centric architecture and cybersecurity within the context of digital learning universities, particularly in West Africa. The key contributions include:

a. Comprehensive Survey and Comparison: The research provides a thorough survey and comparison of data architectures employed in digital learning universities, identifying 109 e-learning solution use cases. This comprehensive analysis sheds light on the diversity of data architectures, paving the way for informed decision-making in selecting the most suitable model.

b. Tailored Data-Centric Model Design: A data-centric architecture model specifically designed for digital learning universities in West Africa is proposed. The model considers the unique challenges and requirements of the region, offering a practical blueprint for institutions aiming to enhance their data management capabilities.

c. Robust Security Model: The research contributes a detailed design of a security model tailored to the data-centric architecture in digital learning universities. This includes the integration of Apache Ranger and Apache Atlas for data governance and security, providing a robust framework to safeguard sensitive information.

d. Proof of Concept: A practical proof of concept is presented, demonstrating the viability and effectiveness of the proposed data-centric architecture and security model. The implementation of Kafka for streamlined data processing adds a practical dimension to the research, showcasing real-world applicability.

e. Insights into Storage Solutions: The research offers insights into storage solutions tailored to the needs of digital learning universities. The discussion on hybrid storage, encryption practices, and the selection of suitable storage options provides valuable guidance for institutions grappling with the management of vast and diverse datasets.

f. Recommendations for Cybersecurity Practices: The research contributes actionable recommendations for enhancing cybersecurity practices in digital learning universities. The emphasis on continuous security training, regular updates, and collaboration underscores the proactive measures needed to mitigate evolving cybersecurity threats.

g. Contextualization for West Africa: The research contextualizes its findings and recommendations within the specific socio-economic and technological landscape of West Africa. This regional focus ensures that the proposed solutions align with the unique challenges faced by digital learning institutions in this geographical context.

Overall, these contributions advance the understanding and implementation of data-centric architecture and cybersecurity practices in digital learning universities, with implications extending beyond the specific region to benefit educational institutions globally.

5.5 Future Research Directions

The research conducted in this study opens avenues for future investigations in several key areas related to data-centric architecture and cybersecurity in digital learning universities. Some potential directions for future research include:

1. Advanced Security Measures: Future research could delve deeper into innovative security measures and technologies that can further fortify data-centric architectures in digital learning universities. Exploring emerging encryption techniques, biometric authentication, and anomaly detection systems could enhance the resilience of cybersecurity frameworks.
2. Adaptability to Emerging Technologies: As technology continues to evolve, future research can explore the adaptability of data-centric architectures to emerging technologies such as blockchain, artificial intelligence, and edge computing. Investigating how these technologies can be integrated to enhance data security and processing efficiency would be valuable.
3. Global Comparative Studies: Conducting comparative studies on data-centric architectures and cybersecurity practices across various regions and continents could provide insights into the contextual differences and common challenges faced by

digital learning universities globally. This comparative approach can inform best practices that are adaptable to diverse environments.

4. User-Centric Design: Future research can focus on refining user-centric applications within data-centric architectures. Exploring ways to improve the user experience, accessibility, and personalization of educational platforms can contribute to more effective learning environments.
5. Longitudinal Studies on Security Efficacy: Conducting longitudinal studies to assess the long-term efficacy of implemented security measures would be beneficial. Tracking the performance of security models over time can reveal potential vulnerabilities and inform the development of adaptive cybersecurity strategies.
6. Integration of Ethical Considerations: Future research can explore the ethical implications of data-centric architectures in digital learning universities. Investigating issues such as data privacy, consent, and responsible use of educational data can contribute to the development of ethically sound frameworks.
7. Scalability and Resource Optimization: As digital learning platforms continue to grow, scalability and resource optimization become critical. Future research can focus on developing strategies for optimizing data-centric architectures to handle increasing volumes of data while ensuring efficient resource utilization.
8. Collaboration and Information Sharing: Exploring collaborative approaches and information-sharing mechanisms among digital learning universities can enhance collective cybersecurity efforts. Research in this area could investigate models of collaboration, threat intelligence sharing, and joint response strategies.

By pursuing these future research directions, scholars and practitioners can contribute to the ongoing evolution of data-centric architectures and cybersecurity practices in digital learning universities, fostering a more secure and adaptable educational landscape.

REFERENCES:

1. Abdelsalam, M., Idrees, A. M., & Shokry, M. (2023). A Proposed Model for Improving the Reliability of Online Exam Results Using Blockchain. *IEEE Access*. https://ieeexplore.ieee.org/abstract/document/10216975/
2. Aborode, A., Anifowoshe, O., Ayodele, T. I., Iretiayo, A. R., & David, O. O. (2020). *Impact of COVID-19 on education in sub-Saharan Africa*. https://www.preprints.org/manuscript/202007.0027
3. Adnan, M. (2020). Online learning amid the COVID-19 pandemic: Students perspectives. *Journal of Pedagogical Sociology and Psychology*, *1*(2), 45–51. https://doi.org/10.33902/JPSP.2020261309
4. Adobe. (2021). *Adobe Privacy Centre*. https://www.adobe.com/africa/privacy/policy.html
5. Aissaoui, K., & Azizi, M. (2017). El-security: E-learning systems security checker plug-in. *Proceedings of the 2nd International Conference on Big Data, Cloud and Applications*, 1–6.
6. Alassery, H. A. A. F. (2021). Securing fog computing for e-learning system using integration of two encryption algorithms. *Journal of Cybersecurity*, *3*(3), 149.
7. Alfonso, F. (2018, February 13). Data-driven versus data-centric. *Stratio*. https://blog.stratio.com/datadriven-versus-datacentric/
8. Allen, I. E., & Seaman, J. (2017). Digital Compass Learning: Distance Education Enrollment Report 2017. *Babson Survey Research Group*. https://eric.ed.gov/?id=ed580868
9. Al-Malah, D. K. A.-R., Aljazaery, I. A., Alrikabi, H. T. S., & Mutar, H. A. (2021). Cloud computing and its impact on online education. *IOP Conference Series: Materials Science and Engineering*, *1094*(1), 012024. https://iopscience.iop.org/article/10.1088/1757-899X/1094/1/012024/meta
10. Al-Naser, A., Rasheed, M., Irving, D., & Brooke, J. (2013). *A Data-Centric Approach to Data Provenance in Seismic Imaging Data*. cp. https://doi.org/10.3997/2214-4609.20130200
11. Al-Sherideh, A. S., Maabreh, K., Maabreh, M., Al Mousa, M. R., & Asassfeh, M. (2023). Assessing the Impact and Effectiveness of Cybersecurity Measures in e-Learning on Students and Educators: A Case Study. *International Journal of Advanced Computer Science and Applications*, *14*(5). https://search.proquest.com/openview/175afa53a183a601946e5c20a9abae52/1?pq-origsite=gscholar&cbl=5444811
12. Arabi, A. A. M. (2021). A zero-trust model-based framework for managing of academic dishonesty in institutes of higher learning. *Turkish Journal of Computer and Mathematics Education (TURCOMAT)*, *12*(6), 5381–5389.
13. Ascend. (2020). *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
14. Aulakh, K., Roul, R. K., & Kaushal, M. (2023). E-learning enhancement through educational data mining with Covid-19 outbreak period in backdrop: A review. *International Journal of Educational Development*, *101*, 102814. https://doi.org/10.1016/j.ijedudev.2023.102814
15. Bates, B. (2019). Learning theories simplified: And how to apply them to teaching. *Learning Theories Simplified*, 1–384.

16. Bellaj, S. (2021). *Etakwin*. ELearning Industry. https://elearningindustry.com/directory/elearning-software/etakwin
17. Bervell, B., & Umar, I. N. (2017). A Decade of LMS Acceptance and Adoption Research in Sub-Sahara African Higher Education: A Systematic Review of Models, Methodologies, Milestones and Main Challenges. *EURASIA Journal of Mathematics, Science and Technology Education*, *13*(11). https://doi.org/10.12973/ejmste/79444
18. Bezovski, Z., & Poorani, S. (2016). The evolution of e-learning and new trends. *Information and Knowledge Management*, *6*(3), 50–57.
19. Bhatia, M., & Maitra, J. K. (2018). E-learning Platforms Security Issues and Vulnerability Analysis. *2018 International Conference on Computational and Characterization Techniques in Engineering & Sciences (CCTES)*, 276–285. https://doi.org/10.1109/CCTES.2018.8674115
20. Bill, B. (2021). *Blackboard Inc. [NASDAQ:BBBB]: The Digital Learning Champions*. CIOReview. https://education.cioreview.com/vendor/2017/blackboard_inc._[nasdaq:bbbb]
21. Boh Podgornik, B., Dolničar, D., Šorgo, A., & Bartol, T. (2016). Development, testing, and validation of an information literacy test (ILT) for higher education. *Journal of the Association for Information Science and Technology*, *67*(10), 2420–2436.
22. Brown, C. (2021). *Coassemble*. ELearning Industry. https://elearningindustry.com/directory/elearning-software/coassemble
23. Butler, J. (2021). *Nimble LMS*. ELearning Industry. https://elearningindustry.com/directory/elearning-software/nimble-lms
24. Campus. (2021). *Award Winning Cloud ERP System for University, College, Cloud ERP for Small Companies*. https://campuslabs.in/campus-erp/
25. Carol, D. (2021, February 4). *The Difference Between Data-centric and Data-driven*. Applied Software. https://www.asti.com/the-difference-between-data-centric-and-data-driven/
26. Carvalho Ota, F. K., Augusto Meira, J., Frank, R., & State, R. (2020). Towards Privacy Preserving Data Centric Super App. *2020 Mediterranean Communication and Computer Networking Conference (MedComNet)*, 1–4. https://doi.org/10.1109/MedComNet49392.2020.9191550
27. CIO, Re. (2021). *SharedBook: You Can Have It Both Ways*. CIOReview. https://education.cioreview.com/vendor/2017/sharedbook
28. claned. (2021). *Privacy policy*. Claned. https://claned.com/privacy-policy/
29. Crompton, H. (2013). The benefits and challenges of mobile learning. *Learning and Leading with Technology*, *41*. https://digitalcommons.odu.edu/cgi/viewcontent.cgi?article=1136&context=teachinglearning_fac_pubs
30. Cyoy, R. B. (2022). *Framework for Effective Management of Cyber Security on E-learning Platforms in Public Universities in Kenya* [PhD Thesis, university of nairobi]. http://erepository.uonbi.ac.ke/handle/11295/161726
31. Dabbagh, N., & Kitsantas, A. (2012). Personal Learning Environments, social media, and self-regulated learning: A natural formula for connecting formal and informal learning. *The Internet and Higher Education*, *15*(1), 3–8.
32. Daniel, J. (2012). Making sense of MOOCs: Musings in a maze of myth, paradox and possibility. *Journal of Interactive Media in Education*, *2012*(3). https://jime.open.ac.uk/article/10.5334/2012-18/
33. Dave, M. (2020). The Data-Centric Revolution: Data-Centric vs. Data-Driven. *TDAN.Com*. https://tdan.com/the-data-centric-revolution-data-centric-vs-data-driven/20288

34. Djeki, E., Dégila, J., & Alhassan, M. H. (2023). E-Learning Challenges and Opportunities in West Africa During COVID-19 Pandemic. *2023 IEEE 12th International Conference on Engineering Education (ICEED)*, 159–164. https://doi.org/10.1109/ICEED59801.2023.10264039
35. Docebo. (2021). *Docebo*. ELearning Industry. https://elearningindustry.com/directory/elearning-software/docebo
36. docebo. (2021). Docebo—Privacy Policy. *Docebo*. https://www.docebo.com/company/privacy-policy/
37. Doust, S. (2021). *Glo^{TM} learn*. ELearning Industry. https://elearningindustry.com/directory/elearning-software/glo
38. Elmaghrabi, A. Y., & Eljack, S. M. (2019). Enhancement of Moodle learning management system regarding quizzes security and stability problems. *2019 2nd International Conference on Computer Applications & Information Security (ICCAIS)*, 1–7. https://ieeexplore.ieee.org/abstract/document/8769530/
39. Ertmer, P. A., Ottenbreit-Leftwich, A. T., Sadik, O., Sendurur, E., & Sendurur, P. (2012). Teacher beliefs and technology integration practices: A critical relationship. *Computers & Education*, *59*(2), 423–435.
40. Fatokun, F. B., Hamid, S., Norman, A., & Fatokun, J. O. (2019). The impact of age, gender, and educational level on the cybersecurity behaviors of tertiary institution students: An empirical investigation on Malaysian universities. *Journal of Physics: Conference Series*, *1339*(1), 012098.
41. Ferguson, R., Coughlan, T., Egelandsdal, K., Gaved, M., Herodotou, C., Hillaire, G., Jones, D., Jowers, I., Kukulska-Hulme, A., & McAndrew, P. (2019). *Innovating pedagogy 2019: Open university innovation report 7*. https://oro.open.ac.uk/59132/1/innovating-pedagogy-2019.pdf
42. Firat, M., & Bozkurt, A. (2020). Variables affecting online learning readiness in an open and distance learning university. *Educational Media International*, *57*(2), 112–127.
43. GDPR. (2018). *GDPR Compliance | LMS by Mindflash*. https://mindflash.com/gdpr
44. Ghatak, S. R. (2021). *Adobe Captivate Prime*. ELearning Industry. https://elearningindustry.com/directory/elearning-software/adobe-captivate-prime
45. Gogos, R. (2021). *Looop*. ELearning Industry. https://elearningindustry.com/directory/elearning-software/looop
46. Gray, M. (2021). *Xperiencify*. ELearning Industry. https://elearningindustry.com/directory/elearning-software/xperiencify
47. Guo, S., & Zeng, D. (2020). Pedagogical Data Federation toward Education 4.0. *Proceedings of the 2020 The 6th International Conference on Frontiers of Educational Technologies*, 51–55. https://doi.org/10.1145/3404709.3404751
48. gyrus. (2021). *LMS Privacy Policy*. https://www.gyrus.com/privacy-policy
49. Hasan, M. R., Rahman, R., & Zaman, K. (2022). Design an Information Security Framework for University Automation System. *2022 25th International Conference on Computer and Information Technology (ICCIT)*, 454–459. https://ieeexplore.ieee.org/abstract/document/10054997/
50. Hjorth, S.-J. (2021). *CanopyLAB - Social Learning Powered by AI*. ELearning Industry. https://elearningindustry.com/directory/elearning-software/canopylab-social-learning-powered-by-ai
51. Hodges, C. B., Moore, S., Lockee, B. B., Trust, T., & Bond, M. A. (2020). *The difference between emergency remote teaching and online learning*. https://vtechworks.lib.vt.edu/handle/10919/104648
52. Inc, A. L. S. (2021). *Absorb Privacy Policy—LMS Data Security—Absorb LMS Software*. https://www.absorblms.com/support/privacy-policy

53. Ispring. (2021). *ISpring Learn*. https://elearningindustry.com/directory/elearning-software/ispring-learn
54. Jennifer. (2021). *Inquisiq*. ELearning Industry. https://elearningindustry.com/directory/elearning-software/inquisiq
55. Joksimović, S., Kovanović, V., & Dawson, S. (2019). The journey of learning analytics. *HERDSA Review of Higher Education*, *6*, 27–63.
56. Jusas, V., Butkiene, R., Venčkauskas, A., Grigaliūnas, Š., Gudoniene, D., Burbaite, R., & Misnevs, B. (2022). *Sustainable and Security Focused Multimodal Models for Distance Learning. Sustainability 2022, 14, 3414*. s Note: MDPI stays neutral with regard to jurisdictional claims in published …. https://www.academia.edu/download/89627685/pdf.pdf
57. Kale, A. W., Narawade, V. E., & Kothoke, P. M. (2023). 7 A Study on Online Learning Systems' Identification with Security Schemes and Applications. *Online Learning Systems: Methods and Applications with Large-Scale Data*, 73–80.
58. Kampakis, D. S. (2018, October 22). What are the differences between data-driven, data-informed and data-centric? *The Data Scientist*. https://thedatascientist.com/data-driven-data-informed-data-centric/
59. Kapadia, V. (2021). *GyrusAim*. ELearning Industry. https://elearningindustry.com/directory/elearning-software/gyrusaim
60. Kim, H. (2019). Research issues on data centric security and privacy model for intelligent internet of things based healthcare. *ICSES Trans. Comput. Netw. Commun*, *5*, 1–3.
61. Korać, D., Damjanović, B., & Simić, D. (2022a). A model of digital identity for better information security in e-learning systems. *The Journal of Supercomputing*, *78*(3), 3325–3354. https://doi.org/10.1007/s11227-021-03981-4
62. Korać, D., Damjanović, B., & Simić, D. (2022b). A model of digital identity for better information security in e-learning systems. *The Journal of Supercomputing*, *78*(3), 3325–3354. https://doi.org/10.1007/s11227-021-03981-4
63. learn upon. (2021). *LearnUpon LMS - eLearning Industry*. https://elearningindustry.com/directory/elearning-software/learnupon-lms
64. Lewis, N. J., & Orton, P. (2000). The Five Attributes of I Innovative E-Learning. *Training &amp; Development*, *54*(6), 47–47.
65. Li, C., Guo, J., Zhang, G., Wang, Y., Sun, Y., & Bie, R. (2019). A blockchain system for E-learning assessment and certification. *2019 IEEE International Conference on Smart Internet of Things (SmartIoT)*, 212–219.
66. Lu, Y., & Da Xu, L. (2018). Internet of Things (IoT) cybersecurity research: A review of current research topics. *IEEE Internet of Things Journal*, *6*(2), 2103–2115.
67. Lynch, M. (2021). *Absorb LMS*. ELearning Industry. https://elearningindustry.com/directory/elearning-software/absorb-lms
68. Mackey, T. P., & Jacobson, T. E. (2011). Reframing information literacy as a metaliteracy. *College & Research Libraries*, *72*(1), 62–78.
69. Malekos, N. (2021). *LearnWorlds*. ELearning Industry. https://elearningindustry.com/directory/elearning-software/learnworlds
70. Maurice, D. H., Ke, H., Ahmad, F., Wang, Y., Chung, J., & Manganiello, V. C. (2014). Advances in targeting cyclic nucleotide phosphodiesterases. *Nature Reviews Drug Discovery*, *13*(4), 290–314.
71. Means, B., & Neisler, J. (2021). Teaching and learning in the time of COVID: The student perspective. *Online Learning*, *25*(1).
72. Media, S. (2021). *Eurekos LMS*. ELearning Industry. https://elearningindustry.com/directory/elearning-software/eurekos

73. Mihailescu, M. I., Nita, S. L., & Corneliu, P. V. (2020). Applied cryptography in designing e-learning platforms. *The International Scientific Conference ELearning and Software for Education*, *2*, 179–189. https://search.proquest.com/openview/eb907a3aa93f0d1dcf65e585b349dac1/1?pq-origsite=gscholar&cbl=1876338&casa_token=NTS59L-3ZY0AAAAA:Xsps3o2s9q9DnH74w-D338iGOPrIFjxdxyvXtZudCtjEUqwXOeHoMaXbIF5YAXzqB3djw-WHJHw
74. Modesti, P. (2020). Integrating Formal Methods for Security in Software Security Education. *Informatics in Education-An International Journal*, *19*(3), 425–454.
75. Mustofa, M., Ahmadi, R., & Karimullah, I. W. (2020). Islamic Character Education in E-Learning Model: How Should It be Implemented? *Jurnal Sains Sosio Humaniora*, *4*(1), 89–93.
76. Nakagawa, H., Iwasawa, Y., & Matsuo, Y. (2019). Graph-based knowledge tracing: Modeling student proficiency using graph neural network. *IEEE/WIC/ACM International Conference on Web Intelligence*, 156–163. https://dl.acm.org/doi/abs/10.1145/3350546.3352513
77. Papagelis, A. (2021). *TalentLMS*. ELearning Industry. https://elearningindustry.com/directory/elearning-software/talentlms
78. Pappas, G. (2021). *Edsby LMS*. ELearning Industry. https://elearningindustry.com/directory/elearning-software/edsby-lms
79. Pérez, S. O., Díez, C. H., & García, J. A. M. (2017). Applying Security to Moodle Grades. *Proceedings of the International Conference on Security and Management (SAM)*, 117–123. https://www.proquest.com/docview/2139471781/abstract/BB7AE83FF8544693PQ/1
80. Plyer, L., Marcou, G., Perves, C., Schurhammer, R., & Varnek, A. (2022). Implementation of a soft grading system for chemistry in a Moodle plugin. *Journal of Cheminformatics*, *14*(1), 72. https://doi.org/10.1186/s13321-022-00645-0
81. Policy. (2018). *Privacy Policy • DigitalChalk Continuing Education Solutions*. DigitalChalk Continuing Education Solutions. https://digitalchalkeu.wpengine.com/about-digitalchalk/privacy-policy
82. Policy. (2020). *Privacy Policy | LMS by Mindflash*. https://mindflash.com/privacy-policy
83. Ponomarev, D. (2021). *Gurucan*. ELearning Industry. https://elearningindustry.com/directory/elearning-software/gurucan
84. Privacy Policy. (2021). *Inquisiq*. https://inquisiq.com/privacy/
85. pseudonymisation. (2020). *Privacy Policy | Coassemble*. Coassemble ELearning Software. https://coassemble.com/privacy-policy
86. Ramanauskaitė, S., Urbonaitė, N., Grigaliūnas, Š., Preidys, S., Trinkūnas, V., & Venčkauskas, A. (2021). Educational organization’s security level estimation model. *Applied Sciences*, *11*(17), 8061.
87. Scott, S. (2021). *Thinkific*. ELearning Industry. https://elearningindustry.com/directory/elearning-software/thinkific
88. Selwyn, N. (2016). Minding our language: Why education and technology is full of bullshit … and what might be done about it. *Learning, Media and Technology*, *41*(3), 437–443. https://doi.org/10.1080/17439884.2015.1012523
89. Setiawan, R., Arif, F. A. S., Putro, J. O., Princes, E., Silalahi, F. T. R., Geraldina, I., Julianti, E., & Safitri, J. (2023). E-Learning Pricing Model Policy for Higher Education. *IEEE Access*, *11*, 38370–38384. https://doi.org/10.1109/ACCESS.2023.3266954
90. Shodeinde, M. (2021). *Claned*. ELearning Industry. https://elearningindustry.com/directory/elearning-software/claned

91. Siemens, G., & Long, P. (2011). Penetrating the fog: Analytics in learning and education. *EDUCAUSE Review*, *46*(5), 30.
92. Sinan, I. I., Degila, J., Nwaocha, V., & Onashoga, S. A. (2022a). Data Architectures' Evolution and Protection. *2022 International Conference on Electrical, Computer and Energy Technologies (ICECET)*, 1–6. https://doi.org/10.1109/ICECET55527.2022.9872597
93. Sinan, I. I., Degila, J., Nwaocha, V., & Onashoga, S. A. (2022b). Data Architectures' Evolution and Protection. *2022 International Conference on Electrical, Computer and Energy Technologies (ICECET)*, 1–6.
94. Sinan, I. I., Nwoacha, V., Degila, J., & Onashoga, S. A. (2022). A Comparison of Data-Driven and Data-Centric Architectures using E-Learning Solutions. *2022 International Conference Advancement in Data Science, E-Learning and Information Systems (ICADEIS)*, 1–6.
95. Skinner, B. F. (1957). The experimental analysis of behavior. *American Scientist*, *45*(4), 343–371.
96. Sun, N., Ding, M., Jiang, J., Xu, W., Mo, X., Tai, Y., & Zhang, J. (2023). Cyber Threat Intelligence Mining for Proactive Cybersecurity Defense: A Survey and New Perspectives. *IEEE Communications Surveys & Tutorials*. https://ieeexplore.ieee.org/abstract/document/10117505/
97. talent LMS. (2021). *Data Security in LMS - Secure Online Learning System—TalentLMS*. https://www.talentlms.com/security
98. thinkific. (2020). *Security Overview Thinkific Website*. Thinkific. https://www.thinkific.com/security-overview-thinkific-website/
99. Vista. (2021). *Data-centric Architecture—A Different Way of Thinking | Vista Projects*. https://www.vistaprojects.com/blog/data-centric-architecture/
100. Warschauer, M., & Matuchniak, T. (2010). New Technology and Digital Worlds: Analyzing Evidence of Equity in Access, Use, and Outcomes. *Review of Research in Education*, *34*(1), 179–225. https://doi.org/10.3102/0091732X09349791
101. xper. (2020). *Privacy Policy*. https://howto.xperiencify.com/article.php?article=88
102. Zheng, X., Li, Q., & Kong, L. (2010). A Data Storage Architecture Supporting Multi-level Customization for SaaS. *2010 Seventh Web Information Systems and Applications Conference*, 106–109. https://doi.org/10.1109/WISA.2010.18

APPENDIX

7.1 Appendix A

A Survey on Cyber-attacks Faced by Data Architectures in West African Institutions During the COVID-19 Era

Data architecture is a collection of models, policies, rules, and standards used by institutions and organizations to manage which data is collected and how it is kept, processed, and integrated. Each institution has a data architecture that is either data-informed, data-driven, or data-centric.

- Data-informed architecture: Data are collected from many sources, such as external and internal hard drives of computers, flash drives, and so on. A dashboard or excel is used to analyze the data, and the results are used as part of inputs in decision-making.
- Data-driven architecture: In this design, algorithms are used to make decisions based on data collected from various data silos such as the cloud, data lakes, and so on.
- Data-centric architecture: Here, the institution builds a single data model utilized by all information systems in the institution, data science is used as the core in decision making, and all data are integrated and connected using a graph database, removing data redundancy and silos.

ACETEL and ACE-SMIA in conjunction with DSTN, ACE-Partner, IRD, AFD, AAU, and the World Bank are conducting research to provide a secure data architecture that will aid in achieving safer learning environment for west African institutions.

Consequently, this survey was initiated to gain a better understanding of the types data architectures these institutions employ and the types of cyber-attacks/threats they faced during the COVID-19 pandemic; your input will contribute immensely to this research. This survey is strictly for the purpose of research and your responses remain confidential.

1. Please select your gender *
   - Male
   - Female
2. Please select your country*
   - Bénin
   - Burkina Faso
   - Cabo Verde
   - Cote d’Ivoire
   - Gambia
   - Ghana
   - Guinea
   - Guinea-Bissau
   - Liberia
   - Mali

- Mauritania
- Niger
- Nigeria
- Senegal
- Sierra Leone
- Togo
- Other :

3. Please select the type of your institution*

- Public
- Private
- Other:

4.Please select the mode of delivery in your institution*

- Face-to-face
- E-learning
- Blended

5.Please select the age group you belong *

- Below 25 years
- 26 – 35 years
- 36 – 45 years
- 46 – 55 years
- 56 and above

6.Please indicate your current academic level *

- Diploma
- Bachelor degree
- Master’s degree
- PhD / Doctorat
- Other:

7.Please indicate, if your institution conducted any training/ workshop during the period of COVID-19 pandemic*

- Yes
- No

8.Please indicate, if your institution conducts any of the following online*

- Application
- Registration
- Lectures
- Examination

9.Please indicate the type of data architecture employ by your institution*

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

10. To what extent does your institution use the following tools to ANALYSE and REPORT on data you collect and store? *

	Extensively	Moderately	A Little	Not at all
1. Spreadsheets (Charts, counts, pivot tables)	O	O	O	O

2. Website analytics (e.g., Google Analytics).	O	O	O	O
3. Database Database (CRM analytics and reports)	O	O	O	O
4. Specialist tools (e.g., SAS, R, Stata, Python, SPSS, GIS Mapping)	O	O	O	O

11. Which of these best describes your institution’s use of data for decision-making?*

- We do not use data at all for decision-making, we rely on gut feeling and experience
- We use data about what happened in the recent past (e.g., last quarter or last year)
- We use past and recent data, including some longer-term trends analysis
- We monitor what’s happening now, in real-time, as well as past trends
- We use past, present, and forward-looking data (e.g. forecasting, modelling, and optimization)

12. To what extent has your institution’s staff use data and analysis to influence or inform their activities and decisions in the following areas?*

	Extensively	Moderately	A Little	Not at all
• Students’ satisfaction with their teaching & learning experience	O	O	O	O
• Need for student and/or staff engagement	O	O	O	O
• Academic development and performance review	O	O	O	O
• Research opportunities and potential research partners	O	O	O	O
• Environmental impacts from institutional activities	O	O	O	O
• Other societal impacts from institutional activities	O	O	O	O
• Mid and long-term strategic planning	O	O	O	O

13. Which of these best describes how your institution is planning for improvement in data?

- There is no plan and no intention to make one
- There is no plan but we are thinking we should have one
- We are actively creating a plan
- We have a plan and are implementing it
- There is a regular cycle of planning, implementation and review

14.Please indicate if your institution is a victim of cyber-attacks*

- Yes
- No

15. If Yes in the above, please indicate; the type of cyber-attack your institution faced SQL Injection

- Denial of service (DOS)
- Ransomware
- Virus
- Worm
- Phishing
- Other:

16.Please indicate; the data protection technique(s) employed by your institutions*

- Firewall
- Anti-virus
- Intrusion detection system
- Intrusion prevention system
- Other:

17.Are you satisfied with your institution data protection technique(s) / Êtes-vous satisfait de la ou des techniques de protection des données de votre établissement ?

- Yes / Oui
- Neutral / neutre
- No / Non

18.How would you rate the level of your institution cyber countermeasures / Comment évaluez-vous le niveau des contre-mesures informatiques de votre institution*

- Poor
- Fair
- Good
- Excellent
- Don't Know

19. How often do you attend workshop/ training in cybersecurity?*

- Not at all / Pas du tout
- Every 3 months
- Every 6 months
- Every 12 months
- Other:

20.How would you rate the level of your cybersecurity knowledge and skill? *

- Poor
- Fair
- Good
- Excellent
- Don't Know

Thank you

Back

Submit

Clear form

7.2 Appendix B

Secure grade distribution code is available at: https://github.com/iisinan/grade-distribution-scheme/tree/main/aes_encryption 2