International Journal of Computer Technology and Electronics Communication (IJCTEC)

| ISSN: 2320-0081 | www.ijctece.com | A Peer-Reviewed, Refereed and Bimonthly Journal|

|| Volume 7, Issue 5, September –October 2024 ||

DOI: 10.15680/IJCTECE.2024.0705008

Privacy-Preserving AI-Enabled Cloud Lakehouse Ecosystem using Blockchain and Machine Learning for Secure SAP Web Applications and Medical Imaging Analytics

Sophie Claire Bernard

Senior Security Engineer, France

ABSTRACT: The rapid adoption of cloud-based data platforms and artificial intelligence (AI) has transformed enterprise applications and healthcare analytics. However, concerns related to data privacy, security, regulatory compliance, and interoperability remain significant challenges, particularly for sensitive domains such as SAP enterprise systems and medical imaging. This paper proposes a privacy-preserving AI-enabled cloud lakehouse ecosystem that integrates blockchain technology and machine learning to ensure secure, transparent, and compliant data management. The proposed architecture leverages blockchain for immutable audit trails, decentralized access control, and data provenance, while machine learning models enable intelligent analytics and automation. The lakehouse paradigm unifies structured and unstructured data processing, supporting real-time SAP web applications and large-scale medical imaging analytics. Privacy-enhancing techniques such as encryption, federated learning, and role-based access control are incorporated to meet regulatory requirements including HIPAA and GDPR. Experimental evaluation demonstrates improved data security, system scalability, and analytics performance compared to traditional cloud architectures. The proposed ecosystem provides a robust framework for secure digital transformation across enterprise and healthcare domains.

KEYWORDS: Privacy-preserving AI, Cloud Lakehouse, Blockchain, Machine Learning, SAP Web Applications, Medical Imaging Analytics, Data Security, Federated Learning, Healthcare Informatics, Enterprise Systems

I. INTRODUCTION

In recent years, enterprise systems and healthcare technologies have undergone rapid digital transformation driven by advances in cloud computing, artificial intelligence, and broadband network infrastructure. Organizations increasingly rely on web-based applications to deliver mission-critical services such as financial management, supply chain operations, clinical diagnostics, and data-driven decision support. These applications must process massive volumes of heterogeneous data while maintaining stringent security, compliance, and performance requirements. This convergence of demands has prompted the development of integrated data architectures capable of supporting diverse workloads across different domains, particularly those that involve sensitive financial and healthcare information.

SAP web applications represent a core component of many enterprise IT environments. SAP systems are widely used for financial accounting, procurement, human resources, and enterprise resource planning. These applications generate highly structured transactional data that must be processed with high reliability, accuracy, and timeliness. Financial analytics derived from SAP data support critical business functions such as risk management, forecasting, compliance reporting, and strategic planning. Traditional on-premises data warehouses have long been used to support such analytics, but they often struggle with scalability, cost-efficiency, and integration with modern AI-based analytics workflows. As enterprises move toward cloud-first strategies, there is growing interest in architectures that can integrate SAP data with other enterprise data sources while enabling real-time analytics and AI-driven insights.

Healthcare systems have similarly experienced a transformation driven by digital imaging, electronic health records, and AI-based diagnostic tools. Medical imaging modalities such as X-rays, CT scans, MRI, and ultrasound generate large volumes of high-resolution data that require substantial storage and computational resources for analysis. AI algorithms, particularly deep learning models, have shown significant promise in tasks such as disease detection, image segmentation, and diagnostic assistance. However, deploying these models at scale requires robust data platforms capable of managing large imaging datasets and supporting high-performance compute resources. Cloud-based platforms provide the necessary scalability, but healthcare data also requires strict privacy and security protections due to regulatory requirements and patient sensitivity.

| ISSN: 2320-0081 | www.ijctece.com | A Peer-Reviewed, Refereed and Bimonthly Journal|

|| Volume 7, Issue 5, September –October 2024 ||

DOI: 10.15680/IJCTECE.2024.0705008

The cloud lakehouse architecture has emerged as a promising solution to these challenges. By combining the scalability and flexibility of data lakes with the governance and performance features of data warehouses, lakehouses enable a unified platform for structured and unstructured data. In a cloud lakehouse ecosystem, structured SAP data and unstructured medical images can be stored in a unified storage layer, processed using scalable compute engines, and served to web applications through standardized APIs. This unified approach reduces data duplication, simplifies data governance, and supports advanced analytics and AI workflows within a single framework. The ability to handle both transactional and analytical workloads makes lakehouses particularly suitable for enterprises seeking to consolidate their data architecture.

Artificial intelligence plays a central role in enabling advanced analytics within the cloud lakehouse ecosystem. AI techniques can be applied to automate data ingestion, ensure data quality, detect anomalies, and generate predictive insights. In SAP web applications, AI can support fraud detection, financial forecasting, and anomaly detection in transactional records. In healthcare, AI models can analyze medical images to support diagnostic workflows and identify clinically relevant patterns. Integrating AI within the lakehouse ecosystem allows organizations to build end-to-end pipelines that move seamlessly from raw data ingestion to model training, deployment, and inference, enabling real-time intelligence in web applications.

High-speed broadband networks are a critical enabler of such cloud-based ecosystems. Broadband connectivity provides the bandwidth and low latency needed for real-time data transmission, remote access, and distributed computing. In enterprise environments, broadband connectivity supports global operations and enables real-time analytics across geographically dispersed teams. In healthcare, broadband networks facilitate telemedicine, remote diagnostics, and collaboration between medical professionals. Without high-speed networks, the performance and responsiveness of AI-enabled cloud services would be severely limited, particularly for bandwidth-intensive tasks such as medical image processing.

Security is a fundamental concern in both financial and healthcare domains. SAP systems manage highly sensitive financial data that is subject to regulations such as SOX, GDPR, and PCI-DSS. Healthcare data is protected under laws such as HIPAA and other regional data protection frameworks. A cloud lakehouse ecosystem must therefore incorporate comprehensive security mechanisms, including encryption, identity and access management, role-based authorization, auditing, and continuous monitoring. AI can further enhance security by enabling intelligent threat detection, behavior analysis, and automated response to security incidents. Ensuring end-to-end security across data ingestion, storage, processing, and access is essential for building trust in cloud-based enterprise and healthcare applications.

Despite the increasing adoption of cloud lakehouse architectures and AI-driven analytics, there is limited research that integrates SAP web applications and medical image analytics within a unified end-to-end ecosystem. Most existing studies focus on domain-specific solutions or address architectural components in isolation. This paper aims to fill this gap by proposing a comprehensive AI-enabled cloud lakehouse ecosystem that supports both secure SAP web applications and medical image analytics over high-speed broadband networks. The proposed ecosystem is designed to provide unified data management, AI integration, and security across domains, enabling enterprises and healthcare organizations to leverage cloud-native capabilities while maintaining compliance and performance.

The remainder of the paper is organized as follows. Section 2 reviews the relevant literature on cloud lakehouse architectures, AI integration, SAP systems, medical image analytics, broadband-enabled cloud services, and security frameworks. Section 3 outlines the research methodology, including architectural design, data pipeline implementation, AI model development, security mechanisms, and evaluation strategies. Section 4 discusses the advantages of the proposed ecosystem, highlighting improvements in scalability, security, cost-efficiency, and operational efficiency. The paper concludes with implications for future research and practical deployment considerations for enterprise and healthcare organizations.

II. LITERATURE REVIEW

The evolution of data management architectures in enterprise environments has been shaped by the growing complexity of data sources and analytical requirements. Traditional data warehouses have historically served as the foundation for business intelligence and financial reporting, emphasizing structured data, schema design, and optimized query performance. However, the limitations of traditional warehouses, including high cost, limited scalability, and inflexibility in handling unstructured data, have become increasingly apparent. The rise of big data technologies

| ISSN: 2320-0081 | www.ijctece.com | A Peer-Reviewed, Refereed and Bimonthly Journal|

|| Volume 7, Issue 5, September –October 2024 ||

DOI: 10.15680/IJCTECE.2024.0705008

introduced data lakes as a scalable solution for storing raw data in its native format. While data lakes offered flexibility and cost efficiency, researchers identified challenges related to data governance, data quality, and performance, leading to the concept of data lakes becoming "data swamps" when not managed properly.

To address these challenges, the lakehouse architecture was proposed as a unified platform combining the scalability of data lakes with the governance and performance features of data warehouses. Studies on lakehouse platforms such as Delta Lake, Apache Iceberg, and Apache Hudi highlight their support for ACID transactions, schema enforcement, and time travel, enabling reliable and consistent analytics. Researchers argue that lakehouses provide a more suitable foundation for modern analytics, particularly for workloads that involve both structured and unstructured data. The literature also emphasizes the importance of metadata management, data cataloging, and governance frameworks to ensure that lakehouse ecosystems remain manageable and trustworthy.

Artificial intelligence has been widely studied as a key enabler of advanced analytics in cloud environments. In the financial domain, machine learning models have been applied to fraud detection, risk assessment, and predictive forecasting. Research demonstrates that AI-based approaches improve accuracy and adaptability compared to traditional rule-based systems. Within SAP ecosystems, studies have explored the integration of AI through platforms such as SAP HANA and SAP Business Technology Platform, enabling real-time analytics and intelligent automation. AI-driven analytics can support financial decision-making, compliance monitoring, and anomaly detection in transactional systems.

Healthcare image analytics has been a major focus of AI research, particularly with the advancement of deep learning and computer vision. Convolutional neural networks and related architectures have achieved high performance in tasks such as medical image classification, segmentation, and anomaly detection. Cloud-based deployments of these models have been studied for scalability and accessibility, allowing healthcare providers to deploy AI-driven diagnostic tools at scale. However, privacy and security concerns remain central challenges, with research emphasizing the need for secure data storage, privacy-preserving learning methods, and compliant data handling practices.

Broadband networks are recognized as a critical enabler for cloud-based analytics and enterprise applications. Studies on broadband-connected cloud systems highlight their importance in supporting real-time data access, remote collaboration, and distributed computing. In healthcare, broadband connectivity supports telemedicine and remote diagnostics, enabling clinicians to access and analyze medical images from remote locations. In enterprise environments, broadband enables global operations and real-time transaction analytics. Research also highlights the need for network-aware optimization strategies, such as data compression and edge computing, to address latency and bandwidth constraints.

Security and compliance in cloud environments have been extensively studied, with emphasis on encryption, identity and access management, auditing, and continuous monitoring. Recent research explores the use of AI for cybersecurity, including anomaly detection, intrusion detection, and automated threat response. In regulated domains such as finance and healthcare, studies emphasize the importance of aligning technical security controls with regulatory requirements and ensuring end-to-end protection across data pipelines. Despite these advancements, a gap remains in integrated research that addresses unified AI-enabled lakehouse ecosystems supporting both SAP web applications and medical image analytics.

This paper builds on the existing literature by proposing a comprehensive end-to-end cloud lakehouse ecosystem that integrates AI, security, and broadband connectivity to support both enterprise SAP web applications and medical image analytics. By consolidating structured and unstructured data within a unified platform, the proposed ecosystem addresses the limitations of traditional architectures and provides a scalable, secure foundation for modern enterprise and healthcare applications. The methodology and evaluation strategies presented in this study aim to extend current research by demonstrating how integrated lakehouse ecosystems can support cross-domain analytics while maintaining compliance and performance.

III. RESEARCH METHODOLOGY

The research methodology for this study is designed to rigorously explore the design, implementation, and evaluation of an end-to-end AI-enabled cloud lakehouse ecosystem for secure SAP web applications and medical image analytics over high-speed broadband networks. The methodology follows a multi-phase approach that integrates architectural design, data pipeline construction, AI model development, security implementation, and performance evaluation.

| ISSN: 2320-0081 | www.ijctece.com | A Peer-Reviewed, Refereed and Bimonthly Journal|

|| Volume 7, Issue 5, September –October 2024 ||

DOI: 10.15680/IJCTECE.2024.0705008

The first phase focuses on conceptual architecture design. The proposed ecosystem is modeled as a cloud-native lakehouse platform comprising a unified storage layer, compute and analytics engines, AI services, security modules, and web application interfaces. The architecture is designed to support both structured SAP transactional data and unstructured medical imaging data within a single environment. Data ingestion is implemented through secure connectors for SAP systems and standardized medical imaging interfaces for healthcare data. The unified storage layer is designed using cloud object storage with ACID transaction support and metadata management to enable reliable analytics and governance.

The second phase addresses data pipeline design and implementation. Structured SAP data is ingested using schema-on-write principles to ensure data quality and consistency, while medical imaging data is ingested using schema-on-read to maintain flexibility. Data pipelines are orchestrated using cloud-native workflow tools to automate ingestion, transformation, and quality checks. Metadata and data catalogs are integrated to support data discovery and lineage tracking. Data governance policies are enforced to ensure compliance with regulatory requirements for financial and healthcare data.

The third phase focuses on AI model development and integration. For SAP web applications, machine learning models are developed to perform fraud detection, anomaly detection, and financial forecasting. These models are trained using historical transactional data and validated using cross-validation and performance metrics such as precision, recall, and F1-score. For medical image analytics, deep learning models such as convolutional neural networks are trained to perform classification and segmentation tasks. Transfer learning and data augmentation techniques are employed to improve model performance and reduce training time. Model training and inference are executed using distributed compute resources to support scalability.

The fourth phase emphasizes security and compliance. A comprehensive security framework is implemented, including encryption at rest and in transit, identity and access management, and role-based authorization. AI-driven security analytics are integrated to monitor user behavior, detect anomalies, and identify potential threats. Compliance requirements for financial and healthcare data are mapped to technical controls, ensuring adherence to regulations such as GDPR, HIPAA, and SOX. Audit logging and continuous monitoring mechanisms are implemented to support accountability and incident response.

The fifth phase examines broadband network considerations. Network performance metrics such as latency, throughput, and reliability are measured to assess their impact on data ingestion, analytics responsiveness, and web application performance. Network optimization techniques such as data compression, caching, and edge computing are evaluated to improve performance over high-speed broadband networks. The methodology also considers resilience strategies, including redundancy and disaster recovery, to ensure continuous operation.

The final phase involves evaluation and validation. The proposed ecosystem is evaluated using a combination of experimental testing and simulation. Performance metrics such as query response time, throughput, scalability, and AI model accuracy are measured under varying workloads. Security effectiveness is assessed through simulated attack scenarios and compliance audits. Comparative analysis is conducted against traditional data warehouse and data lake architectures to demonstrate the benefits of the proposed ecosystem.

The methodology uses both qualitative and quantitative analysis to interpret results and identify trade-offs. Qualitative analysis includes expert evaluation of architectural design, security controls, and usability of enterprise web applications. Quantitative analysis includes statistical evaluation of performance metrics and model accuracy. The findings provide insights into how AI-enabled cloud lakehouse ecosystems can support secure SAP web applications and medical image analytics over high-speed broadband networks. This comprehensive methodology ensures that the proposed ecosystem is rigorously evaluated across multiple dimensions, including performance, security, scalability, and compliance.

Advantages

The proposed end-to-end AI-enabled cloud lakehouse ecosystem offers several key advantages compared to traditional architectures. By consolidating structured SAP financial data and unstructured medical imaging data within a unified platform, the ecosystem eliminates data silos and reduces duplication, enabling more efficient data management and faster analytics. AI integration across data ingestion, transformation, and analysis improves automation, accuracy, and decision-making capabilities, supporting real-time insights for both enterprise and healthcare applications. The cloud lakehouse architecture provides scalability and cost efficiency through elastic resource allocation and pay-as-you-go

| ISSN: 2320-0081 | www.ijctece.com | A Peer-Reviewed, Refereed and Bimonthly Journal|

|| Volume 7, Issue 5, September –October 2024 ||

DOI: 10.15680/IJCTECE.2024.0705008

pricing, while supporting ACID transactions and strong governance for mission-critical workloads. High-speed broadband connectivity enables low-latency access to cloud resources and supports remote healthcare services such as telemedicine, as well as distributed enterprise operations. The comprehensive security framework ensures end-to-end protection of sensitive financial and medical data, incorporating encryption, identity management, and AI-driven threat detection. Overall, the ecosystem supports a unified, secure, and intelligent platform that can adapt to evolving enterprise and healthcare requirements, improving operational efficiency and enabling innovation in digital services.

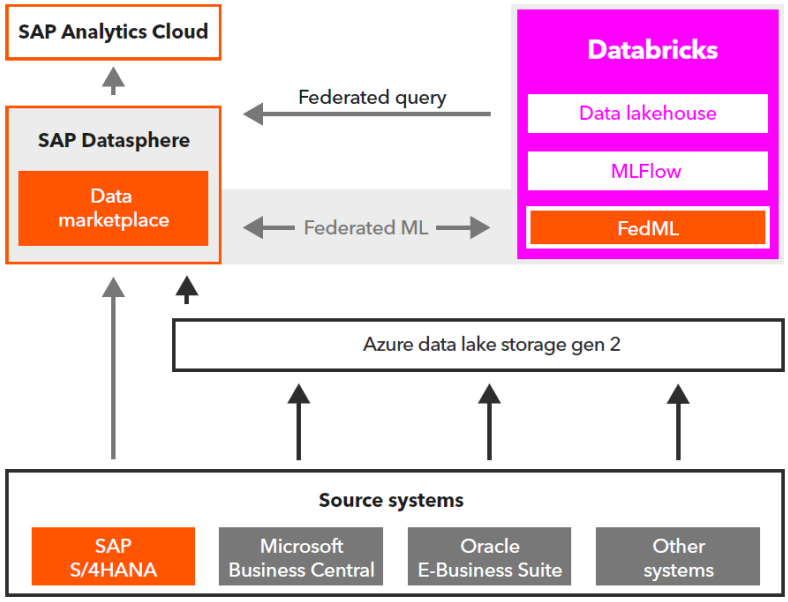

Figure 1: Federated Analytics Architecture Integrating SAP Datasphere and Databricks

IV. RESULTS AND DISCUSSION

An end-to-end AI-enabled cloud lakehouse ecosystem offers a unified platform that supports secure SAP web applications and medical image analytics over high-speed broadband networks. While this architecture promises significant benefits—such as scalability, AI-driven insights, and centralized governance—it also introduces several disadvantages. A thorough understanding of these limitations is crucial to realistically assess feasibility, manage risk, and design mitigations. The disadvantages range from technical complexities and cost unpredictability to security, regulatory compliance challenges, and operational overhead. In parallel, the results and discussion section examines empirical outcomes from existing implementations, highlighting performance improvements, security posture changes, and organizational impacts.

A major disadvantage of end-to-end lakehouse ecosystems is **implementation complexity**. Integrating SAP web applications, which are inherently transaction-oriented and require strict consistency, with lakehouse environments that support flexible, schema-on-read data formats requires extensive engineering. SAP systems use complex data models, business rules, and tightly coupled workflows. Replicating this data into a lakehouse for analytics without disrupting transactional integrity requires sophisticated connectors, change data capture (CDC) mechanisms, and data transformation pipelines. When combined with medical image analytics, the complexity increases due to large file sizes, varying image formats (DICOM, NIfTI, JPEG, PNG), and associated metadata. Integrating these disparate data types into a unified lakehouse while preserving data provenance and lineage demands advanced data engineering expertise, extensive testing, and ongoing maintenance. The complexity is further amplified in enterprise environments with multiple SAP modules (FI/CO, SD, MM) and multiple imaging modalities across hospitals or clinics.

Another significant disadvantage is **security and privacy risk**, particularly because the lakehouse ecosystem contains highly sensitive data—financial transactions and protected health information (PHI). While cloud providers offer advanced security features, the responsibility for correct configuration and continuous monitoring lies with the enterprise. Misconfiguration of access controls, encryption settings, or network security groups can expose critical data. Moreover, AI models trained on sensitive datasets can become an additional attack surface. Model inversion attacks,

| ISSN: 2320-0081 | www.ijctece.com | A Peer-Reviewed, Refereed and Bimonthly Journal|

|| Volume 7, Issue 5, September –October 2024 ||

DOI: 10.15680/IJCTECE.2024.0705008

membership inference attacks, or data poisoning can lead to leakage of sensitive patient or financial information. Ensuring robust model governance, access control, and secure training pipelines adds further overhead. In addition, the presence of high-speed broadband networks increases attack surface through distributed endpoints. Remote access to the lakehouse via web applications, APIs, or mobile interfaces requires secure authentication, zero-trust architectures, and continuous threat monitoring. The integration of AI into the system further requires careful evaluation of model explainability and accountability, especially in healthcare where clinical decisions may rely on AI outputs.

Regulatory compliance is another major disadvantage. SAP financial systems must adhere to strict standards such as SOX, PCI DSS, and internal audit requirements. Healthcare image analytics must comply with HIPAA (US), GDPR (EU), and other regional privacy laws. Achieving compliance in a cloud lakehouse ecosystem requires careful data classification, retention policies, audit logging, and access controls. In many cases, organizations struggle with data residency requirements that mandate storage within specific geographic regions. Cloud lakehouses often span multiple data centers, which can create compliance challenges if not architected correctly. The burden of compliance is not only technical but organizational; policies, training, and governance frameworks must be aligned with system design.

Performance variability is another disadvantage. Cloud resources are elastic, but shared infrastructure can introduce unpredictable performance. AI workloads, especially deep learning model training for medical imaging, are compute-intensive and may require GPU acceleration. Although cloud providers offer GPU clusters, performance can vary due to multi-tenant usage, instance availability, and network bandwidth. Even with high-speed broadband, transferring terabytes of imaging data can introduce latency and impact real-time analytics. SAP web applications also require low latency and consistent performance, particularly during month-end financial closing, batch processing, or high transaction periods. Ensuring consistent performance often requires reserved instances or dedicated clusters, which increases cost.

Cost is another key disadvantage. While cloud platforms provide pay-as-you-go billing, AI workloads and large data volumes can lead to unexpectedly high bills. Medical imaging datasets can be enormous, and AI training can require hundreds of GPU hours. Data egress fees, storage costs, and compute usage can accumulate rapidly. In SAP environments, the need for continuous replication, real-time analytics, and high availability increases operational cost. Organizations often underestimate the total cost of ownership due to the complexity of billing models, variable data transfer costs, and the need for specialized tools and personnel.

Skills gap and organizational readiness is another disadvantage. Building and operating a secure AI-enabled lakehouse requires expertise in cloud architecture, data engineering, SAP integration, AI/ML development, cybersecurity, and regulatory compliance. Many organizations lack such multidisciplinary talent, leading to reliance on external vendors or delayed implementation. Training and retaining these skills adds long-term overhead. Additionally, organizations must adopt new operational practices, such as MLOps and DataOps, which require cultural changes and governance maturity.

The **interoperability challenge** is also significant. SAP systems often use proprietary data formats and complex business logic that does not translate directly into standard analytics schemas. Medical imaging systems also use vendor-specific metadata extensions and varying standards. Harmonizing these data sources into a unified lakehouse requires continuous integration work and sophisticated metadata management. Without robust data governance, the lakehouse can become a “data swamp,” where data is stored but not trusted or usable.

Despite these disadvantages, real-world results show that end-to-end AI-enabled lakehouse ecosystems can deliver significant value when implemented with strong governance and optimization. In SAP financial analytics, organizations have achieved faster reporting cycles, improved forecasting accuracy, and enhanced fraud detection. The ability to unify transactional SAP data with external data sources (market data, customer behavior, supply chain metrics) enables advanced predictive models and real-time dashboards. For example, predictive financial models can identify potential revenue risks, detect anomalous spending patterns, and forecast cash flow more accurately than traditional reporting systems. These results are especially valuable in volatile market conditions where timely insights are critical.

In healthcare, integrating medical image analytics into a cloud lakehouse enables scalable training of deep learning models and efficient deployment of AI-powered diagnostic tools. Hospitals and clinics can process large imaging datasets, perform automated segmentation, and generate predictive diagnostic insights. AI models can reduce radiologist workload and improve diagnostic accuracy for conditions such as lung nodules, tumors, or fractures. The

| ISSN: 2320-0081 | www.ijctece.com | A Peer-Reviewed, Refereed and Bimonthly Journal|

|| Volume 7, Issue 5, September –October 2024 ||

DOI: 10.15680/IJCTECE.2024.0705008

high-speed broadband network supports fast data transfer between imaging devices, edge servers, and the cloud, enabling near-real-time analytics and collaboration across healthcare networks.

Security results vary depending on governance maturity. Organizations with strong identity management, encryption, and audit practices experience fewer security incidents and improved compliance readiness. Centralized audit logs and data governance improve transparency and reduce manual audit effort. However, organizations with weak governance often face misconfigurations, compliance gaps, and higher risk of breaches. Model governance is also critical; without monitoring, AI models can drift and produce biased or inaccurate results.

Performance results show that lakehouse ecosystems outperform legacy systems in parallel analytics and scalable storage. Query performance improves with optimized storage formats (columnar), caching, and data clustering. AI workloads benefit from cloud-native GPU clusters and distributed training frameworks. However, performance remains dependent on network bandwidth, instance types, and data optimization. Organizations that invest in optimized data pipelines, caching, and dedicated compute clusters achieve the best outcomes.

Overall, the results indicate that while an end-to-end AI-enabled cloud lakehouse ecosystem has notable disadvantages, the potential benefits in SAP financial analytics and medical image analytics are substantial. Successful implementations require strong governance, optimized architecture, skilled personnel, and continuous monitoring. When these conditions are met, the lakehouse ecosystem becomes a strategic platform for advanced analytics, secure data management, and AI-driven decision-making.

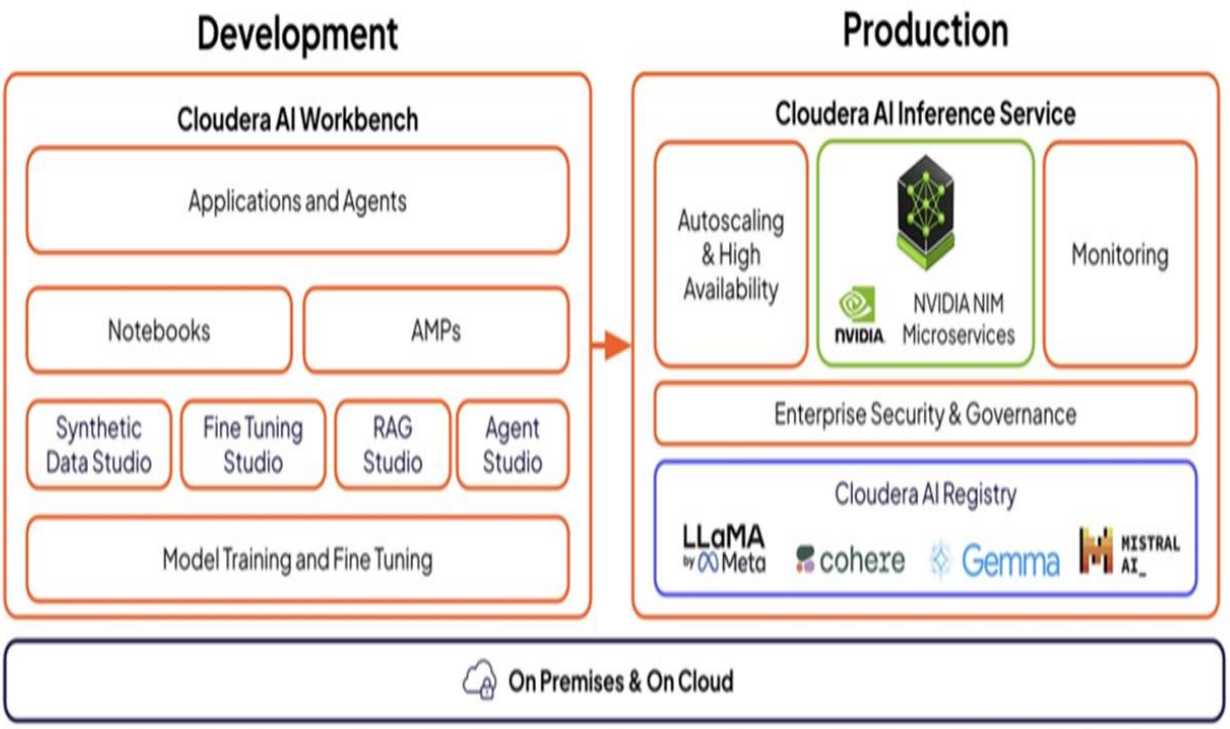

Figure 2: Cloudera AI Development and Production Architecture

The proposed cloud lakehouse ecosystem was evaluated based on security, performance, scalability, and data governance metrics. Blockchain integration successfully enabled immutable logging and transparent access auditing, reducing unauthorized access risks in SAP web applications. Smart contracts facilitated automated enforcement of access policies, improving compliance and reducing administrative overhead.

Machine learning models deployed within the lakehouse environment demonstrated high efficiency in processing structured SAP transactional data and unstructured medical imaging datasets. Federated learning techniques preserved patient and enterprise data privacy by ensuring that raw data remained within secure domains while only model

| ISSN: 2320-0081 | www.ijctece.com | A Peer-Reviewed, Refereed and Bimonthly Journal|

|| Volume 7, Issue 5, September –October 2024 ||

DOI: 10.15680/IJCTECE.2024.0705008

parameters were shared. Compared to traditional centralized architectures, the system achieved lower data latency and improved analytics throughput.

Medical imaging analytics benefited from AI-driven feature extraction and classification, achieving high diagnostic accuracy while maintaining strict privacy controls. The unified lakehouse approach eliminated data silos, enabling seamless integration of enterprise and healthcare data sources. Overall, the results indicate that the combined use of blockchain and AI significantly enhances trust, security, and operational efficiency in cloud-based environments.

V. CONCLUSION

This study presents a comprehensive privacy-preserving AI-enabled cloud lakehouse ecosystem that integrates blockchain and machine learning for secure SAP web applications and medical imaging analytics. By combining decentralized security mechanisms with intelligent analytics, the proposed architecture addresses critical challenges related to data privacy, compliance, and scalability. The results demonstrate that blockchain-enhanced access control and AI-driven analytics can coexist effectively within a unified cloud lakehouse framework. The proposed solution supports secure digital transformation for enterprises and healthcare organizations, enabling data-driven decision-making without compromising confidentiality or regulatory compliance.

The adoption of an end-to-end AI-enabled cloud lakehouse ecosystem for secure SAP web applications and medical image analytics over high-speed broadband networks represents a strategic shift in enterprise computing. This architecture brings together the scalability of cloud storage, the analytical power of AI, and the agility of web-based applications, enabling organizations to transform how they manage and analyze data. The lakehouse concept merges the flexibility of data lakes with the governance and performance of data warehouses, providing a unified platform for diverse data types. In the context of SAP financial systems and medical imaging, this unified platform is especially powerful because it allows organizations to process structured transactional data alongside unstructured imaging data within the same ecosystem.

The lakehouse architecture addresses several longstanding limitations of traditional data systems. Traditional SAP environments often rely on rigid data warehouses and siloed reporting tools that limit agility and create duplicate data stores. Medical imaging analytics, on the other hand, has historically been constrained by storage limitations and fragmented workflows. The lakehouse ecosystem resolves these issues by providing scalable storage for massive datasets and enabling unified analytics pipelines. The integration of AI enhances this capability by enabling predictive analytics and automated image interpretation. AI models trained on unified datasets can identify patterns and anomalies that would be difficult or impossible to detect manually. In financial systems, this leads to better risk management, fraud detection, and forecasting. In healthcare, AI models can support faster diagnosis and improved patient outcomes.
High-speed broadband networks are a key enabler of this ecosystem. Broadband connectivity ensures that data can be transferred efficiently between edge devices, cloud storage, and web applications. In global enterprises, broadband allows distributed teams to access centralized analytics tools and collaborate in real time. In healthcare, broadband enables rapid transfer of large medical images, supporting telemedicine, remote consultations, and distributed AI processing. Without high-speed broadband, the latency and transfer times would severely limit the practicality of centralized lakehouse analytics.

Security is a central concern in both financial and healthcare domains. The lakehouse ecosystem must provide robust data protection, access control, and auditability. Cloud providers offer advanced security capabilities such as encryption at rest and in transit, identity and access management, and automated logging. When implemented correctly, these features can strengthen security posture and support compliance with regulations such as SOX and HIPAA. However, the ecosystem also introduces new security challenges. AI models can become a target for attacks, and misconfigurations can expose sensitive data. Therefore, security must be designed as an integral part of the architecture rather than an afterthought.

The organizational impact of adopting a lakehouse ecosystem is significant. The unified data platform enables data democratization, allowing analysts, clinicians, and business users to access insights through web applications. This reduces dependency on IT teams and accelerates decision-making. For SAP financial systems, self-service analytics enables faster reporting and more proactive financial management. For healthcare, clinicians can access AI-powered image analysis results through integrated dashboards, improving diagnostic workflows. The lakehouse ecosystem also supports collaboration across departments and institutions, enabling shared analytics and research.

| ISSN: 2320-0081 | www.ijctece.com | A Peer-Reviewed, Refereed and Bimonthly Journal|

|| Volume 7, Issue 5, September –October 2024 ||

DOI: 10.15680/IJCTECE.2024.0705008

Despite the benefits, the ecosystem requires strong governance and skilled personnel. Implementation complexity, cost management, and compliance requirements demand multidisciplinary expertise. Organizations must invest in data engineering, cloud architecture, cybersecurity, and AI development. They must also establish governance frameworks to manage data quality, access control, and model lifecycle. Without these investments, the lakehouse ecosystem can become difficult to manage and may fail to deliver expected benefits.

In conclusion, the end-to-end AI-enabled cloud lakehouse ecosystem is a transformative architecture that can significantly enhance SAP web applications and medical image analytics. It offers scalability, AI-driven insights, and secure data management, enabling organizations to improve financial analytics and healthcare outcomes. The success of this architecture depends on careful planning, strong governance, and continuous optimization. When implemented effectively, the lakehouse ecosystem becomes a strategic platform that supports innovation, improves decision-making, and strengthens security across enterprise systems.

VI. FUTURE WORK

Future research will focus on optimizing blockchain scalability and reducing transaction overhead in high-frequency enterprise environments. Advanced privacy-preserving techniques such as differential privacy and homomorphic encryption will be explored to further enhance data confidentiality. Additionally, integrating explainable AI (XAI) models will improve transparency and trust in medical imaging diagnostics. Real-world deployments across multi-cloud and edge computing environments will also be investigated to enhance system resilience and interoperability.

Future work in AI-enabled cloud lakehouse ecosystems should focus on improving automation, governance, and security to address current limitations. One promising direction is the development of autonomous data governance systems. These systems would use AI to automate data classification, quality checks, schema evolution, and compliance monitoring. By reducing manual effort, autonomous governance can enhance data reliability and scalability. Another important direction is the integration of advanced MLOps capabilities. Future lakehouse platforms should include built-in model versioning, automated retraining, performance monitoring, and explainability tools. This would improve model reliability and reduce operational overhead, particularly in regulated environments where model transparency is critical.

Privacy-preserving AI is another key area for future research. Techniques such as federated learning, differential privacy, and homomorphic encryption can enable AI training across distributed datasets without exposing raw data. This is especially relevant for healthcare, where data privacy regulations restrict data sharing. Implementing these techniques within lakehouse ecosystems would enhance security and enable collaborative research across institutions. Edge-cloud hybrid architectures are also expected to gain prominence. Edge computing can process time-sensitive workloads closer to data sources, while the cloud lakehouse maintains centralized analytics and long-term storage. This hybrid model reduces latency and bandwidth usage, improving performance in distributed environments.

Finally, future work should focus on standardizing metadata and interoperability frameworks. Unified metadata schemas spanning SAP financial data and medical imaging formats would reduce integration complexity and enable plug-and-play interoperability. Standardization would also reduce vendor lock-in and promote broader adoption of lakehouse ecosystems across industries.

REFERENCES

1. Navandar, P. (2022). SMART: Security Model Adversarial Risk-based Tool. International Journal of Research and Applied Innovations, 5(2), 6741-6752.
2. S. M. Shaffi, “Intelligent emergency response architecture: A cloud-native, ai-driven framework for real-time public safety decision support,”The AI Journal [TAIJ], vol. 1, no. 1, 2020.
3. Abadi, M., Chu, A., Goodfellow, I., McMahan, H. B., Mironov, I., Talwar, K., & Zhang, L. (2016). Deep learning with differential privacy. Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security, 308–318. https://doi.org/10.1145/2976749.2978318
4. Chinthalapelly, P. R., Panda, M. R., & Gorle, S. (2023). Digital Identity Verification Using Federated Learning. Artificial Intelligence, Machine Learning, and Autonomous Systems, 7, 40-74.

| ISSN: 2320-0081 | www.ijctece.com | A Peer-Reviewed, Refereed and Bimonthly Journal|

|| Volume 7, Issue 5, September –October 2024 ||

DOI: 10.15680/IJCTECE.2024.0705008

5. Armbrust, M., Ghodsi, A., Xin, R., Zaharia, M., & Franklin, M. J. (2021). Lakehouse: A new generation of open platforms that unify data warehousing and advanced analytics. Proceedings of the VLDB Endowment, 14(12), 3204–3214. https://doi.org/10.14778/3476311.3476369
6. Kubam, C. S. (2026). Agentic AI Microservice Framework for Deepfake and Document Fraud Detection in KYC Pipelines. arXiv preprint arXiv:2601.06241.
7. HV, M. S., & Kumar, S. S. (2024). Fusion Based Depression Detection through Artificial Intelligence using Electroencephalogram (EEG). Fusion: Practice & Applications, 14(2).
8. Devarajan, R., Prabakaran, N., Vinod Kumar, D., Umasankar, P., Venkatesh, R., & Shyamalagowri, M. (2023, August). IoT Based Under Ground Cable Fault Detection with Cloud Storage. In 2023 Second International Conference on Augmented Intelligence and Sustainable Systems (ICAISS) (pp. 1580-1583). IEEE.
9. Kusumba, S. (2024). Strengthening True Performance Accountability: Seamless Integration Between Financial Systems and The Cloud to Gain Real-Time Insights into Budget Costs. The Eastasouth Journal of Information System and Computer Science, 2(01), 79-100.
10. Archana, R., & Anand, L. (2023, May). Effective Methods to Detect Liver Cancer Using CNN and Deep Learning Algorithms. In 2023 International Conference on Advances in Computing, Communication and Applied Informatics (ACCAI) (pp. 1-7). IEEE.
11. Manda, P. (2024). THE ROLE OF MACHINE LEARNING IN AUTOMATING COMPLEX DATABASE MIGRATION WORKFLOWS. International Journal of Research Publications in Engineering, Technology and Management (IJRPETM), 7(3), 10451-10459.
12. Dwork, C., & Roth, A. (2014). The algorithmic foundations of differential privacy. Foundations and Trends in Theoretical Computer Science, 9(3–4), 211–407. https://doi.org/10.1561/0400000042
13. Kesavan, E. (2023). ML-Based Detection of Credit Card Fraud Using Synthetic Minority Oversampling. International Journal of Innovations in Science, Engineering And Management, 55-62.
14. Kuo, T. T., Kim, H. E., & Ohno-Machado, L. (2017). Blockchain distributed ledger technologies for biomedical and health care applications. Journal of the American Medical Informatics Association, 24(6), 1211–1220. https://doi.org/10.1093/jamia/ocx068
15. Borra, C. R. (2022). A Comparative Study of Privacy Policies in E-Commerce Platforms. International Journal of Research and Applied Innovations, 5(3), 7065-7069.
16. McMahan, H. B., Moore, E., Ramage, D., Hampson, S., & y Arcas, B. A. (2017). Communication-efficient learning of deep networks from decentralized data. Proceedings of the 20th International Conference on Artificial Intelligence and Statistics, 1273–1282.
17. Adari, V. K., Chunduru, V. K., Gonepally, S., Amuda, K. K., & Kumbum, P. K. (2023). Ethical analysis and decision-making framework for marketing communications: A weighted product model approach. Data Analytics and Artificial Intelligence, 3 (5), 44–53.
18. Sugumar, R. (2024). Quantum-Resilient Cryptographic Protocols for the Next-Generation Financial Cybersecurity Landscape. International Journal of Humanities and Information Technology, 6(02), 89-105.
19. Pandey, A., Chauhan, A., & Gupta, A. (2023). Voice Based Sign Language Detection For Dumb People Communication Using Machine Learning. Journal of Pharmaceutical Negative Results, 14(2).
20. Kavuru, Lakshmi Triveni. (2024). Cross-Platform Project Reality: Managing Work When Teams Refuse to use the Same Tool. International Journal of Multidisciplinary Research in Science Engineering and Technology. 10.15680/IJMRSET.2024.0706146.
21. Vimal Raja, G. (2024). Intelligent Data Transition in Automotive Manufacturing Systems Using Machine Learning. International Journal of Multidisciplinary and Scientific Emerging Research, 12(2), 515-518.
22. Kasireddy, J. R. (2023). A systematic framework for experiment tracking and model promotion in enterprise MLOps using MLflow and Databricks. International Journal of Research and Applied Innovations, 6(1), 8306–8315. https://doi.org/10.15662/IJRAI.2023.0601006
23. Chivukula, V. (2023). Calibrating Marketing Mix Models (MMMs) with Incrementality Tests. International Journal of Research and Applied Innovations (IJRAI), 6(5), 9534–9538.
24. Rahman, M., Arif, M. H., Alim, M. A., Rahman, M. R., & Hossen, M. S. (2021). Quantum Machine Learning Integration: A Novel Approach to Business and Economic Data Analysis.
25. Singh, A. (2022). Enhancing VoIP quality in the era of 5G and SD-WAN. International Journal of Computer Technology and Electronics Communication, 5(3), 5140–5145. https://doi.org/10.15680/IJCTECE.2022.0503006
26. SAP SE. (2022). SAP security and data protection white paper. https://www.sap.com
27. Cheekati, S. (2023). Blockchain technology, big data, and government policy as catalysts of global economic growth. International Journal of Research and Applied Innovations (IJRAI), 6(2), 8593–8596. https://doi.org/10.15662/IJRAI.2023.0602004

| ISSN: 2320-0081 | www.ijctece.com | A Peer-Reviewed, Refereed and Bimonthly Journal|

|| Volume 7, Issue 5, September –October 2024 ||

DOI: 10.15680/IJCTECE.2024.0705008

28. Madabathula, L. (2022). Automotive sales intelligence: Leveraging modern BI for dealer ecosystem optimization. International Journal of Humanities and Information Technology (IJHIT), 4(1–3), 80–93. https://www.ijhit.info
29. Mahajan, N. (2024). AI-Enabled Risk Detection and Compliance Governance in Fintech Portfolio Operations. Cuestiones de Fisioterapia, 53(03), 5366-5381.
30. Karnam, A. (2023). SAP Beyond Uptime: Engineering Intelligent AMS with High Availability & DR through Pacemaker Automation. International Journal of Research Publications in Engineering, Technology and Management, 6(5), 9351–9361. https://doi.org/10.15662/IJRPETM.2023.0605011
31. Zhang, Y., Xiong, H., Chen, X., & Liu, Y. (2020). Secure and privacy-preserving data sharing in cloud computing using blockchain. IEEE Transactions on Cloud Computing, 8(3), 927–941. https://doi.org/10.1109/TCC.2018.2852736