https://doi.org/10.70844/ijas.2025.1.19

Innovative Journal of Applied Science

Review Article

IJAS-25-051

Redefining Data Management with BDaaS-Big Data As-a-Service A Deep Dive into BDaaS Governance, Compliance and Security

Muhammad Rawish Siddiqui*

Department of Data Governance, Modern Data Management Team, Riyadh, Saudi Arabia

Corresponding Author: Muhammad Rawish Siddiqui, Department of Data Governance, Modern Data Management Team, Riyadh, Saudi Arabia, E-mail: rawishsiddiqui@yahoo.ca

Received date: 12 February, 2025, **Accepted date:** 26 February, 2025, **Published date:** 05 March, 2025

Citation: Siddiqui MR (2025) Redefining Data Management with BDaaS-Big Data As-a-Service A Deep Dive into BDaaS Governance, Compliance and Security. Innov J Appl Sci 2(2): 19.

Abstract

Big Data as a Service (BDaaS) is reshaping data management by providing scalable, cost-effective solutions that enable organizations to leverage the full potential of big data. This paper investigates the integration of BDaaS with critical aspects of data governance, compliance and protection. It explores the strategic frameworks and methodologies for effectively implementing BDaaS, highlighting its role in improving data accessibility, analytics and decision-making processes. The paper further examines real-world applications across industries and best practices. Additionally, it also identifies challenges and risks associated with BDaaS, such as data security, privacy concerns, regulatory compliance and operational complexities. *Via* an in-depth analysis, this study aims to offer an overall perspective on how BDaaS can be integrated into modern data ecosystems while ensuring compliance with data protection laws, minimizing risks and optimizing data governance strategies.

Keywords: BDaaS, Big data as-a-service, Data governance, Compliance, Data protection, Scalability, Cloud computing, Regulatory, HIPAA, GDPR, CCPA, PDPL

Introduction

As data continues to grow at an outstanding rate, organizations face distinct challenges in managing, analyzing and protecting a big amount of information. To address this, Big Data as a Service (BDaaS) has emerged as an optimal solution. BDaaS offers cloud-based platforms that enable businesses to efficiently process large volumes of data, making it possible to extract valuable insights while maintaining high performance. However, the success of BDaaS depends not only on its technological capabilities but also on the strength of the frameworks governing data management, privacy and security.

While global regulations like General Data Protection Regulation (GDPR), Health Insurance Portability and Accountability Act (HIPAA), California Consumer Privacy Act (CCPA) and even Personal Data Protection Law (PDPL) becoming stricter, it is essential that BDaaS platforms comply with these laws. This ensures that organizations handle data responsibly while maintaining transparency, protecting personal information and minimizing risks. BDaaS must, therefore, align with these regulatory standards whereas ensuring its operations are seamless, secure and scalable.

This research will explore how BDaaS can enhance data governance frameworks, facilitate compliance with evolving regulations and implement robust data protection mechanisms. The goal is to show how BDaaS can contribute to sustainable, ethical data practices, supporting organizations in navigating the complexities of data management in a highly regulated environment.

Explanation

BDaaS simplifies the complex process of managing big data by providing cloud-based platforms that reduce the reliance on costly and space-consuming on- premises solutions. Traditionally, businesses had to invest heavily in physical hardware, data centers and skilled personnel in order to maintain and operate data management systems. With BDaaS, these challenges are drastically addressed by shifting the responsibility to cloud providers who offer scalable, flexible and cost-effective solutions.

Organizations using BDaaS leverage access to a comprehensive suite of tools designed for data storage, processing and analytics. These platforms allow businesses to handle vast amounts of data without the need for managing physical servers or worrying about their maintenance. By hosting data in the cloud, companies can securely store their information in highly available environments, ensuring that it is both protected and easily accessible when needed.

Moreover, BDaaS platforms are designed to help organizations comply with a wide range of regulatory requirements, such as HIPAA, GDPR, CCPA, PDPL, PCI-DSS and industry-specific regulations. These platforms come with built-in features to ensure that data is handled according to legal standards, for instance encryption for privacy and audit trails for tracking data access to meet evolving compliance standards. By leveraging these capabilities, organizations can mitigate risks associated with data breaches, loss, or misuse, while maintaining legal and ethical standards.

In addition to regulatory compliance, BDaaS platforms are also structured to support robust data governance frameworks. They offer tools for monitoring, controlling and auditing data throughout its lifecycle, ensuring that data is properly categorized, stored and accessed only by authorized and concerned individuals. These governance capabilities help businesses maintain data integrity, enforce policies and ensure accountability, providing a transparent view of how data is being managed, processed and shared. This not only reduces risks but also promotes trust with customers, stakeholders and regulatory bodies [1-3].

Key Strategic Points

- **Scalability:** BDaaS provides flexible resources, making it easy for businesses to grow and adjust as needed. It lets them increase storage, computing power and analytics when required, ensuring smooth performance as data grows.
- **Cost efficiency:** BDaaS removes the need for costly infrastructure by offering a pay-as-you-go system, helping businesses save money. Since they only pay for what they use, companies can spend more on important projects and avoid wasting resources.
- **Data integration and centralization:** BDaaS collects data from different sources in one place, making it easier to manage and analyze. This helps keep data accurate, accessible and useful for making quick and smarter business decisions.
- **Regulatory compliance:** BDaaS normally has built-in tools to help businesses follow rules like GDPR, PDPL, HIPAA, PCI-DSS and CCPA. It simplifies compliance by automatically tracking and reporting data use, reducing risks and legal issues.
- **Enhanced transparency:** BDaaS helps businesses track data, review processes and ensure accountability. With clear records and monitoring, companies can see where data comes from and catch any issues in early stages, enhancing trust and better data management.

Data governance in BDaaS (big data as a service)

Data governance in BDaaS is about ensuring that the data generated, processed and stored in cloud-based environments is managed securely, accurately and in compliance with regulatory requirements. BDaaS environments mostly deal with massive amounts of diversified data across distributed systems, which demands specialized governance strategies to maintain data quality, security and compliance.

BDaaS-data governance principles

Accountability and responsibility in a cloud environment: In BDaaS, data ownership, accountability and responsibility may become complex and challenging because of the cloud infrastructure and multiple third-party vendors involvement. Clear roles and responsibilities must be defined in the beginning for managing data across various cloud platforms and services. This includes ensuring that both internal teams and cloud service providers understand their responsibilities in terms of data security, quality and compliance.

Data privacy with global access: Given the global nature of BDaaS, data privacy policies need to be very comprehensive. Data can be accessed from anywhere in the world and personal data might cross borders, making compliance with international privacy laws like PDPL, GDPR and CCPA more challenging. Data governance in BDaaS must focus on ensuring that personal and sensitive data is properly cataloged, classified, anonymized and managed to comply with these regulations, even as data moves across multiple cloud regions.

Policy enforcement across cloud services: BDaaS environments typically use multiple cloud services (e.g., AWS, Azure, Google Cloud etc.) with different data management policies. It’s important to establish consistent governance policies that span these services, enforcing rules around data access, usage, storage and sharing. This may include automating policy enforcement to prevent mismanagement of data across various cloud environments.

Data security in cloud: BDaaS environments face unique security challenges due to the cloud’s distributed nature. Ensuring data security involves setting up strong encryption protocols for data in transit as well as at rest, along with implementing Identity and Access Management (IAM) strategies to control who can access data. Data governance in BDaaS should also focus on ensuring that cloud service providers comply with the security standards and certifications necessary for the organization to comply with the regulatory requirements.

Data integrity with cloud-specific challenges: Due to its distributed nature, BDaaS introduces challenges related to data consistency and integrity. Data governance strategies must focus to ensure data accuracy and consistency across multiple cloud services and throughout its lifecycle. This involves using advanced monitoring tools to detect data discrepancies or integrity issues in real-time.

Data quality at scale: BDaaS often deals with data at an enormous scale, sourced from various locations and systems. Ensuring data quality in this context means implementing robust mechanisms for data cleansing, validation and transformation before the data is ingested into the cloud platforms. Automating data quality checks using machine learning tools is essential for large-scale data operations to maintain accuracy, consistency and reliability.

Data lineage in complex ecosystems: In BDaaS, tracking data lineage becomes more complex due to the dynamic and often fragmented nature of cloud-based data pipelines. Tools that provide end-to-end visibility into data’s journey across various systems, applications and transformations (such as ETL processes) are crucial for maintaining transparency and auditability. This is especially important for ensuring that data flows through various stages without losing integrity.

Data retention and archiving in the cloud: In a BDaaS environment, the volume of data grows dramatically. Establishing clear data retention policies becomes an essential part to prevent unnecessary data retention and ensure compliance with regulatory requirements. Governance frameworks must ensure that data is archived or deleted in a way that is consistent across cloud services and platforms.

Data accessibility and scalability: In BDaaS, ensuring that authorized users have timely access to the right data while also maintaining security is key. Scalability must be built into data governance policies to allow for easy access to data without compromising performance. Data governance tools in BDaaS should automate data access policies and apply them in real-time as data is ingested, processed and analyzed at large scales.

Data compliance in dynamic environments: BDaaS platforms constantly evolve, with new features and services being added frequently. Ensuring compliance in such dynamic environments requires continuous monitoring and automated compliance checks. Cloud data governance frameworks must be designed to adapt to regulatory changes, automating audits and data retention processes to ensure that data always remains compliant.

BDaaS-data governance frameworks

Policy development for cloud platforms: BDaaS necessitates creating detailed policies customized to the specific needs of cloud data environments. These policies should define the acceptable use of cloud-native tools, establish protocols for integration with third-party systems and set guidelines for data storage, sharing and lifecycle management. Additionally, they must address the challenges of operating in multi-cloud or hybrid cloud setups. For consistent governance, the policies should ensure uniform standards and practices are maintained across all cloud platforms, enabling smooth and seamless interoperability and compliance.

Attribute/role-based access control in cloud systems: Implementing either ABAC or RBAC is critical in BDaaS to ensure data access aligns strictly with individual roles and responsibilities. In multi-cloud environments, this involves defining and enforcing role-specific access controls across distinct cloud systems. ABAC/RBAC limits access to only the data necessary for a user, thereby minimizing security risks and enhancing compliance. This approach not only protects sensitive information but also simplifies access management in complex cloud ecosystems.

Metadata management: Effective metadata management, in addition to proper data cataloging, is vital for BDaaS environments, as it provides a centralized repository that tracks all data assets within the cloud. This repository should include detailed metadata, such as data origins, transformation processes, ownership and usage patterns. Tools designed for metadata management must integrate seamlessly with cloud data lakes and warehouses, offering a unified view of the organization’s data. Such visibility supports in decision-making, improves data discovery and supports compliance initiatives.

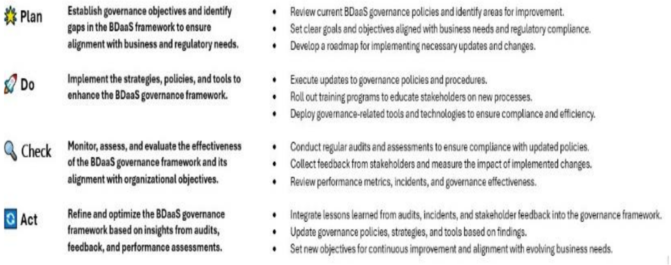
Automated auditing for cloud data: Automated auditing tools are crucial for monitoring data usage, ensuring compliance and maintaining governance in BDaaS setups. These tools should enable continuous tracking of data access and interactions across cloud environments, providing real-time insights into governance practices. By automating audit processes, organizations can quickly identify anomalies, detect policy violations and respond to compliance risks, developing a proactive governance culture.

Data stewardship for cloud-based data: Designating data stewards in BDaaS environments ensures data remains accurate, secure and compliant with policies and regulations. Data stewards are responsible for overseeing specific data domains, resolving security and quality issues, while ensuring adherence to governance standards. They also play a critical role in facilitating cross-departmental data collaboration, enabling effective data sharing and utilization across cloud systems.

Data classification and tagging in BDaaS: Data classification and tagging are essential for managing the distributed nature of cloud data. By employing automated classification and tagging tools, organizations can categorize data, based on sensitivity, regulatory requirements and usage contexts. This classification enables the application of appropriate security controls, simplifies data discovery and ensures compliance with privacy regulations and internal policies.

Data quality management with cloud-based tools: Maintaining high data quality is a foundation of BDaaS governance. Cloud-based tools equipped with advanced features such as machine learning algorithms can perform automated data quality checks, monitor large datasets in real time and detect anomalies. These tools ensure that only high-quality data is used for analytics and decision-making, reducing the risk of errors and improving business outcomes.

Continuous improvement of cloud governance frameworks: Data governance in BDaaS must remain dynamic and flexable to the rapid advancements in cloud technology and evolving regulatory landscapes. Continuous improvement involves periodically reviewing governance policies, incorporating lessons from audits, incidents and stakeholder feedback and updating strategies to reflect new innovations and compliance requirements. This iterative approach ensures the governance framework remains relevant, effective and aligned with organizational goals (Figure 1).



Plan	Establish governance objectives and identify gaps in the BDaaS framework to ensure alignment with business and regulatory needs.	• Review current BDaaS governance policies and identify areas for improvement. • Set clear goals and objectives aligned with business needs and regulatory compliance. • Develop a roadmap for implementing necessary updates and changes.
Do	Implement the strategies, policies, and tools to enhance the BDaaS governance framework.	• Execute updates to governance policies and procedures. • Roll out training programs to educate stakeholders on new processes. • Deploy governance-related tools and technologies to ensure compliance and efficiency.
Check	Monitor, assess, and evaluate the effectiveness of the BDaaS governance framework and its alignment with organizational objectives.	• Conduct regular audits and assessments to ensure compliance with updated policies. • Collect feedback from stakeholders and measure the impact of implemented changes. • Review performance metrics, incidents, and governance effectiveness.
Act	Refine and optimize the BDaaS governance framework based on insights from audits, feedback, and performance assessments.	• Integrate lessons learned from audits, incidents, and stakeholder feedback into the governance framework. • Update governance policies, strategies, and tools based on findings. • Set new objectives for continuous improvement and alignment with evolving business needs.

Figure 1: BDaaS governance framework-continuous improvement cycle.

Compliance in BDaaS

Regulatory landscape

PDPL (Saudi Arabia): Mandates organizations to protect personal data, ensure accuracy and obtain clear consent before processing.

GDPR: Emphasizes data subject rights, requiring explicit consent, data portability and breach notifications.

CCPA: Focuses on consumer rights, including the right to opt-out of data sales and access personal data.

HIPAA: Protects sensitive health information through strict security and privacy rules.

ISO/IEC 27001: Encourages the implementation of comprehensive information security management systems.

PIPEDA (Canada): Regulates how private organizations collect, use and disclose personal information during commercial activities.

NIST privacy framework: Provides a structured approach to building privacy programs that align with regulatory requirements.

Aviation-specific compliance: Regulations such as the EU PNR Directive require airlines to secure and govern passenger data

FERPA: US Federal Law to protect the privacy of student education records.

Compliance features in BDaaS

- Data Encryption protects sensitive information by securing it both when stored (at rest) and during transfer (in transit), making it inaccessible to unauthorized users.
- Access Logs maintain detailed records of all data access, modifications and usage, which are invaluable for tracking user activity and conducting audits.
- Automated Compliance Checks monitor the platform's operations continuously, ensuring adherence to regulations such as PDPL, GDPR, HIPAA and CCPA and provide alerts for potential violations.
- Breach Response Tools allow for rapid detection, suppression and reporting in the event of a security breach, helping mitigate risks and ensuring timely communication with stakeholders.
- RBAC restricts access to sensitive data and system operations based on user roles, reducing the risk of unauthorized access and insider threats.
- Data Masking and Anonymization protect personal and sensitive information during data processing, enabling secure analytics while maintaining privacy.
- Compliance Certifications and Audits demonstrate adherence to recognized standards like ISO 27001, SOC 2 and PCI DSS, with regular audits ensuring continued compliance.
- Retention Policies enforce the secure storage and timely deletion of data after the legally required retention period, minimizing risks and reducing storage costs.
- Real-time Monitoring and Alerts provide continuous oversight of system performance and data usage, with immediate notifications of any potential compliance risks or violations.
- Data Classification organizes information based on sensitivity and applicable regulatory requirements, ensuring that the necessary safeguards are applied appropriately to different types of data. Together, these features form a comprehensive framework for compliance, enabling organizations to manage and analyze large-scale data securely and within regulatory boundaries.

BDaaS Real World Use Cases and Scenarios

Use cases-at a glance

- **Aviation:** Route optimization, passenger data analysis and predictive maintenance aligned with data protection and regulatory standards.
- **Education:** Analysis of student performance while adhering to FERPA regulations.
- **Entertainment:** Streaming platforms utilize BDaaS for personalized recommendations and compliance with copyright laws.
- **Financial sector:** Real-time fraud detection, compliance reporting and audit readiness.
- **Healthcare:** Secure patient data management, HIPAA compliance and advanced analytics for patient outcomes.
- **Hospitality:** Enhancing guest experiences through data insights while adhering to privacy regulations.
- **Logistics:** Real-time package tracking and predictive route optimization, ensuring compliance with international shipping regulations.
- **Public sector:** Open data initiatives with strict governance controls.
- **Retail:** Personalized customer insights, inventory optimization and adherence to CCPA regulations.
- **Telecommunications:** Network performance monitoring and regulatory compliance for customer data privacy.

Real-world scenarios

- **Aviation industry:** Airlines are leveraging BDaaS to analyze passenger data for enhanced customer experience and optimizing flight routes. Airlines use BDaaS for predictive maintenance and passenger safety analysis, ensuring compliance with international aviation standards such as ICAO guidelines, GDPR for passenger data privacy and PNR (Passenger Name Record) standards.
- **Education sector:** Universities and educational institutions utilize BDaaS to analyze student performance, track resource utilization and improve administrative processes, adhering to regulations like FERPA.
- **Energy sector:** An energy provider used BDaaS to analyze consumption patterns and monitor renewable energy production, ensuring compliance with environmental standards and improving grid reliability.
- **Financial services:** A global bank implemented BDaaS to centralize customer transaction data. By leveraging automated metadata management and policy enforcement tools, they improved fraud detection by 40% while ensuring compliance with GDPR and Basel III.
- **Government sector:** A national statistics office used BDaaS for census data analysis. Tools for data lineage and access logs facilitated compliance with transparency and auditability requirements.
- **Healthcare:** A hospital group used BDaaS for real-time patient monitoring and health record management. The platform ensured data quality and security, allowing seamless HIPAA compliance.
- **Hospitality sector:** Hotels leverage BDaaS for guest behavior analysis, enhancing personalized services while ensuring compliance with local and international privacy regulations.
- **Manufacturing:** A multinational manufacturing company adopted BDaaS for predictive maintenance. Real-time analytics on machine performance data reduced downtime by 25%, while adherence to local compliance regulations ensured secure cross-border data transfers.
- **Pharmaceutical industry:** A pharmaceutical company leveraged BDaaS for managing clinical trial data. The

Copyright © 2025 | ijas.meteorpub.com

system ensured FDA compliance by automating reporting and ensuring data integrity.

- **Retail industry:** A multinational retailer engaged BDaaS to consolidate customer data from multiple regions. Automated compliance checks ensured alignment with regional privacy laws, including PDPL, GDPR and CCPA.
- **Telecommunications:** A telecom company implemented BDaaS to manage network performance and detect anomalies in customer usage data. Compliance with PDPL and/or GDPR etc., ensured customer privacy while boosting service quality.

Enablement Methodology

Framework development

Establish a comprehensive governance framework that smoothly integrates Big Data as a Service (BDaaS) capabilities with organizational goals and objectives. This involves:

- **Defining policies and standards:** Based on organization Data Strategy, create policies that address data quality, security, access and usage in line with business priorities.
- **Scalability and flexibility:** Design the framework to accommodate evolving BDaaS capabilities and future organizational needs.
- **Technology integration:** Ensure compatibility between BDaaS solutions and existing infrastructure to optimize performance and enable smooth implementation.
- **Key Performance Indicators (KPIs):** Develop metrics to measure the framework's effectiveness in supporting strategic goals.

Compliance alignment

Map BDaaS capabilities to specific regulatory requirements to ensure adherence to applicable laws and standards. This includes:

- **Regulatory analysis:** Conduct a thorough assessment of regulations, such as PDPL, GDPR, CCPA, or local laws, that impact your organization.
- **Risk management:** Identify and mitigate potential compliance risks through effective data handling, storage and processing practices.
- **Audit readiness:** Establish mechanisms for real-time monitoring, reporting and auditing to demonstrate compliance during assessments.
- **Policy updates:** Periodically review and update policies to align with changing regulatory landscapes.

Continuous improvement

Leverage analytics and performance metrics to refine data governance and compliance strategies iteratively. This may include:

- **Feedback loops:** Gather insights from operational performance, compliance audits and stakeholder feedback to identify gaps and opportunities.
- **Predictive analytics:** Use data trends to forecast potential issues and proactively implement improvements.
- **Automation:** Implement automated tools to monitor data governance practices, ensuring consistency and reducing manual effort.
- **Benchmarking:** Compare governance strategies against industry standards to stay competitive and effective.

Stakeholder engagement

Motivate inter-departmental collaboration to ensure alignment and create a unified approach to data governance. This involves:

- **Cross-functional teams:** Form dedicated teams with representatives from IT, legal, operations and business units to drive initiatives.
- **Training and awareness:** Conduct regular training sessions and workshops to build organizational awareness and capability in leveraging BDaaS.
- **Change management:** Communicate the value and importance of BDaaS and governance initiatives to gain buy-in and minimize resistance.
- **Regular communication:** Establish channels for ongoing dialogue, updates and feedback to keep all stakeholders informed and engaged.

General Activation Steps

Assessment: Understanding organizational needs and compliance requirements

Begin by conducting a thorough assessment of your organization’s specific needs and objectives. Identify the types of data your organization handles, the volume of data processed and any existing challenges in managing it. Pay close attention to compliance requirements, such as PDPL, GDPR, CCPA, or industry-specific regulations for instance ICAO, to ensure your BDaaS implementation aligns with legal obligations. This evaluation will help you define clear goals and set a foundation for the implementation.

Vendor selection: Choosing the right BDaaS provider

Research and select a BDaaS provider that aligns with your organization’s technical and compliance requirements. Prioritize vendors that offer robust compliance certifications, such as ISO 27001, SOC 2, HIPAA, PCI/DSS, PDPL or GDPR (if applicable). Additionally, look for features like built- in governance and privacy tools, advanced data analytics capabilities, scalability and strong customer support. Vendor evaluations should include reviewing case studies, conducting demos and seeking customer testimonials.

Integration: Connecting BDaaS with existing infrastructure

Once a vendor is selected, focus on faultless and smooth integration between the BDaaS platform and your organization’s existing IT systems. This includes connecting databases, applications and tools while ensuring minimal disruption to ongoing operations. Work closely with the provider to configure APIs, CDC, data pipelines and authentication protocols to enable smooth data flow and compatibility with your current architecture.

Policy implementation: Automating governance and compliance measures

Enforce governance policies that align with organizational objectives and regulatory standards. Leverage the automated tools provided by the BDaaS platform to monitor and manage data access, security and usage. Implement safeguards for example, ABAC/RBAC, data encryption and real-time monitoring as well as observability to ensure data integrity and protect sensitive information.

Training: Empowering teams for effective usage and management

Plan and launch comprehensive training programs to ensure teams understand the BDaaS system's functionality and features. Customize training sessions to address specific roles, such as data analysts, IT staff and compliance officers. Equip them with the skills to analyze data efficiently, troubleshoot issues and ensure ongoing compliance. Continuous education and support will help maximize the value derived from BDaaS and enhance organizational data capabilities.

Keep in mind that data security and compliance must be implemented with great care, as these are complex areas, which will be discussed further in the section titled "Challenges and Risks in DBaaS Implementation".

Challenges Risks in BDaaS Implementation

Data security

One of the core challenges when implementing BDaaS is ensuring the security of centralized data, as this architecture inherently poses a higher risk of security breaches. The more data that is stored in a single, centralized cloud environment, the more attractive it becomes as a target for cyber-attacks.

Key risks

- **Data access control:** In a centralized system, securing data access becomes more complex. If improper controls are implemented, unauthorized users could gain access to sensitive business information.
- **Data in transit:** When large volumes of data are transferred across networks, there is an increased risk that it could be intercepted or tampered with, especially if proper encryption protocols are not enforced.
- **Compliance risks:** Poor data security measures could lead to violations of compliance frameworks, like PDPL, GDPR, CCPA or HIPAA, that mandate strong data protection controls.

Suggestions

- **Identity and Access Management (IAM):** Use robust IAM mechanism such as AWS IAM and Azure AD etc. to control and monitor access to data.
- **Multi-Factor Authentication (MFA):** Enforce MFA for all users accessing BDaaS platforms to add an extra layer of security.
- **Encryption:** Implement end-to-end encryption for data in transit and at rest.

Compliance complexity

As organizations adopt BDaaS, they are often required to comply with a variety of regulatory and legal frameworks. With data often being stored and processed across multiple countries and jurisdictions, maintaining compliance becomes a major challenge.

Key risks

- **Jurisdictional issues:** Regulations such as PDPL, GDPR, CCPA and others may require data to remain within specific geographic boundaries. In BDaaS, where data may be distributed across different cloud providers globally, organizations must ensure compliance with data residency requirements.
- **Dynamic regulations:** Data privacy regulations are constantly evolving. This creates a challenge for organizations in terms of staying up- to-date and ensuring that their BDaaS solution remains compliant with the latest laws.

Suggestions

- **Compliance automation tools:** Leverage tools like Collibra and Alation for data governance and compliance management, ensuring that data usage adheres to legal standards.
- **Audit trails:** Implement comprehensive audit trails and logging using solutions like Splunk, AWS CloudTrail and Azure Monitor to track data access and usage.
- **Data privacy frameworks:** Utilize Data Masking and Anonymization tools, like IBM InfoSphere and Privitar, to protect sensitive data while ensuring compliance.

Vendor lock-in

BDaaS solutions often come with pre-configured compliance and security features. However, relying too heavily on a single BDaaS provider for critical functionalities can lead to vendor lock-in, making it difficult to switch providers or scale the platform without facing compatibility issues.

Key risks

- **Cost overruns:** If an organization becomes too dependent on one vendor, it may find itself facing higher costs due to the lack of flexibility in negotiating contracts.
- **Limited customization:** Vendors often provide standardized solutions that may not fully address specific regulatory or operational requirements. This could limit an organization's ability to innovate and customize its solution.
- **Data transfer complexity:** Migrating large datasets from one provider to another can be resource-intensive, requiring significant time and effort to ensure data integrity and consistency during the transfer.

Suggestions

- **Multi-cloud strategy:** Adopt a multi-cloud or hybrid-cloud strategy, using multiple BDaaS providers (e.g., AWS, Azure, Google Cloud) to avoid heavy reliance on a single vendor.
- **Containerization:** Leverage containerization tools like Kubernetes and Docker, which can provide a more portable and flexible infrastructure that is less tied to specific cloud vendors.
- **Interoperability:** Ensure that the BDaaS platform supports industry-standard integration protocols, such as RESTful

APIs to facilitate easier migrations and data exchange between providers.

Skill gaps

The implementation and ongoing management of BDaaS require highly specialized skills in several areas, including data governance, data security, cloud computing and regulatory compliance. Unfortunately, these skills may be in short supply within many organizations, leading to operational inefficiencies and compliance risks.

Key Risks

- **Lack of expertise in cloud architecture:** Many professionals may lack the specific expertise needed to design and maintain complex BDaaS systems, especially in terms of scaling infrastructure and managing distributed data pipelines.
- **Governance and data stewardship:** Effective data governance is critical to ensure that data is accurate, consistent and secure. Insufficient knowledge of governance best practices can lead to poor data management practices, which may result in compliance violations or inefficiencies.
- **Difficulty with compliance:** Without a deep understanding of the various compliance frameworks, teams may inadvertently fail to meet the requirements, exposing the organization to legal and financial risks.

Suggestions

- **Training and certification:** Invest in specialized training for your team through platforms like Udemy, Coursera, or Pluralsight, particularly on topics like Data Governance, Compliance and BDaaS
- **Data governance platforms:** Leverage platforms such as Collibra, TrustArc, Privitar, LogicGate, OneTrust, or Talend to help enforce governance practices and ensure that data stewardship is a priority.
- **Consultancies managed services:** Consider working with third-party consultancies or managed services providers who specialize in BDaaS to fill any skills gaps during implementation and maintenance.

Latency issues

While BDaaS platforms are designed to scale and handle large volumes of data, performance issues related to latency can arise, especially in operations that involve large-scale real-time data processing or analytics. High latency can significantly reduce the speed at which insights are generated, affecting business decision-making.

Key Risks

- **Real-time processing delays:** In scenarios where, immediate insights are needed (e.g., fraud detection, real-time analytics), any delays in data processing can lead to missed opportunities or delayed reactions to critical events.
- **Cross-region latency:** For organizations using multi-cloud or hybrid-cloud environments, latency issues may occur when data needs to be transferred across different geographical regions, especially in areas where internet connectivity is limited or inconsistent.
- **Increased costs:** To overcome latency issues, organizations may need to provision additional resources (e.g., higher-performance instances or additional data processing capacity), which can lead to increased costs.

Suggestions

- **Edge computing:** Use edge computing techniques and services like AWS Greengrass or Azure IoT Edge to process data closer to the source and reduce latency.
- **Optimized data pipelines:** Design efficient data pipelines using technologies like Apache Kafka to minimize delays in data ingestion and processing.
- **Content Delivery Networks (CDNs):** For global operations, use CDNs like Akamai or Cloudflare to reduce latency by caching and delivering content closer to the user or endpoint [4,5].

Conclusion

Big Data as a Service (BDaaS) has become a game-changer for businesses dealing with large amount of diversified dataset. It provides scalable and high- performance platforms to process and analyze information efficiently. Beyond technology, BDaaS empowers data governance by ensuring compliance with strict regulations like PCI-DSS, GDPR, HIPAA, CCPA and PDPL. This helps organizations maintain transparency, accuracy and security while avoiding legal risks and building trust with stakeholders.

However, implementing BDaaS comes with challenges, especially in maintaining privacy, security and regulatory compliance. As data protection laws evolve, businesses must ensure that BDaaS platforms integrate strong security measures such as encryption, strict access controls and privacy-by-design principles. Additionally, continuous monitoring, risk assessments and compliance checks are essential to prevent data breaches and unauthorized access. Without these safeguards, companies risk legal penalties, reputational damage and operational disruptions.

Despite these challenges, a well-structured BDaaS approach allows businesses to turn regulatory demands into opportunities. By proactively addressing data protection requirements, organizations can enhance their competitive advantage, drive innovation and build a strong foundation for secure and responsible data management. In an increasingly regulated digital world, BDaaS serves as a vital tool for balancing compliance, efficiency and growth.

Conflicts of Interest

The authors declare no conflicts of interest in this research.

References

1. Batko K, Ślęzak A (2022) The use of big data analytics in healthcare. Journal of Big Data 9(1): 3. [Crossref] [Google Scholar] [Pubmed]
2. Wang X, Yang L, Liu H, Deen MJ (2017) A big data-as-a-service framework: State-of-the-art and perspectives. IEEE Transactions on Big Data 4(3): 1-1. [Crossref] [Google Scholar]
3. Marr B (2021) A big data-as-a-service: How to choose the best provider.
4. Barlow B, Greene J (2018) Data as a service: The what, why, how, c when.

5. Abdalla HB (2022) A brief survey on big data: Technologies, terminologies and data-intensive applications. Journal of Big Data 9(1): 107. [Crossref] [Google Scholar]