Emerging Trends in Machine Intelligence and Big Data

Integrating Data from Geographically Diverse Non-SAP Systems into SAP HANA: Implementation of Master Data Management, Reporting, and Forecasting Model

Suman Shekhar[1]

[1] *Project Manager, Zimmer Biomet, https://orcid.org/0009-0001-3524-1619*

This manuscript was compiled on March 3, 2018

Abstract

Integrating data from geographically diverse non-SAP systems into SAP HANA presents numerous challenges and opportunities for organizations seeking to use data management, reporting, and forecasting capabilities. This paper explores the implementation of a Master Data Management (MDM) framework, robust reporting mechanisms, and forecasting models within SAP HANA, focusing on the integration of heterogeneous data sources. The study highlights the significance of MDM in ensuring data consistency, accuracy, and reliability across different systems. It discusses the methodologies for integrating non-SAP data into SAP HANA, emphasizing data extraction, transformation, and loading (ETL) processes. The design and implementation of reporting solutions that enable real-time data analytics and business intelligence are also discussed. The paper examines forecasting models that uses integrated data to provide actionable guides and support strategic decision-making. This paper provides a roadmap for organizations to achieve a unified and efficient data management system within SAP HANA, aiming to improve operational efficiency and support informed decision-making.

Keywords: *data integration, ETL processes, forecasting models, MDM framework, non-SAP systems, SAP HANA, real-time analytics*

Received: January 10, 2018 **Revised:** March 16, 2018 **Accepted:** April 2, 2018 **Published:** April 4, 2018

1. Introduction

1.1. Data Integration

In today's global business environment, organizations operate multiple systems across various geographical locations, a practice necessitated by the increasing complexity and scale of operations [1]. This practice is driven by the need to manage and streamline operations across diverse markets while ensuring compliance with local regulations and standards. The distribution of systems and processes is not merely a result of organic growth or expansion but represents a deliberate, strategic approach aimed at enhancing overall efficiency, improving accessibility, and increasing responsiveness to the specific demands of different markets [2].

Organizations with a global footprint must contend with many regulatory, cultural, and operational nuances that differ markedly from one region to another. This heterogeneity requires businesses to maintain distinct systems and processes tailored to the unique requirements of each locale. For example, regulatory frameworks surrounding data protection, taxation, labor laws, and environmental standards can vary significantly between countries. As a result, global organizations must implement systems that are flexible enough to comply with these diverse legal mandates without compromising operational integrity.

Cultural differences also play a critical role in shaping how organizations operate across various geographical locations. Different regions may have distinct customer expectations, business practices, and employee behaviors, all of which need to be considered when designing and implementing systems. For instance, marketing strategies that resonate with customers in one region might be ineffective or even counterproductive in another due to cultural differences in communication styles, values, and consumer behavior [3]. Operational details further complicate the management of systems across multiple locations. Factors such as supply chain logistics, workforce availability, technological infrastructure, and economic conditions can vary widely between regions, necessitating the customization of operational systems to align with local conditions. For example, an organization operating in both developed and developing markets may need to deploy different inventory management systems to account for variations in supply chain efficiency, transportation infrastructure, and market demand. The strategic approach of maintaining distinct systems across different regions also allows organizations to remain agile and adaptable in the face of varying market demands and legal requirements.

The need for responsiveness is further heightened by the increasing pace of technological change and globalization. As new technologies and business models emerge, organizations must be able to quickly integrate these innovations into their existing systems to remain competitive. For example, the rise of e-commerce and digital platforms has transformed the retail industry, requiring organizations to implement new systems for online sales, digital marketing, and customer relationship management.

In addition to technological adaptability, the responsiveness of global organizations is also shaped by their ability to navigate and respond to shifting regulatory landscapes. Governments and regulatory bodies around the world are constantly updating and revising laws and regulations to address emerging issues such as data privacy, cybersecurity, environmental sustainability, and labor rights.

Moreover, the strategic distribution of systems across different regions enables organizations to leverage local expertise and resources to enhance their operations. By decentralizing certain functions and empowering local teams to manage their own systems, organizations can tap into the unique knowledge and capabilities of their regional offices. This approach not only improves operational efficiency but also fosters innovation and creativity by allowing local teams to develop solutions that are specifically tailored to their markets. For example, a global company might allow its regional offices to develop and implement their own marketing strategies, drawing on their deep understanding of local consumer behavior and cultural trends [4].

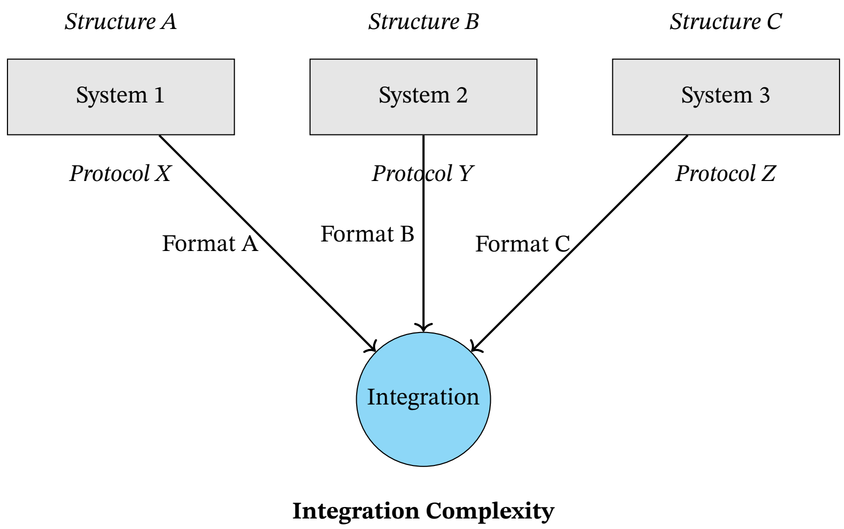

Figure 1. Illustration of Data Heterogeneity Leading to Complex Integration

This decentralized approach to system management also allows organizations to build stronger relationships with local stakeholders, including customers, suppliers, and regulatory bodies. By demonstrating a commitment to understanding and addressing the specific needs and concerns of each region, organizations can build trust and credibility with their local partners. This, in turn, can lead to increased customer loyalty, more favorable regulatory treatment, and stronger partnerships with local suppliers and service providers.

In terms of accessibility, operating multiple systems across various geographical locations ensures that organizations can provide timely and relevant support to their global operations. In a globally dispersed organization, it is critical that systems are accessible to employees, customers, and partners regardless of their location. This requires the implementation of robust IT infrastructure and communication networks that can support seamless access to systems and data from anywhere in the world. By ensuring that their systems are accessible to all stakeholders, organizations can improve collaboration, enhance decision-making, and provide better service to their customers.

The practice of operating multiple systems across different regions also enhances the organization's ability to manage risk and ensure business continuity. By distributing systems across various locations, organizations can reduce their reliance on any single system or location, minimizing the impact of disruptions. For example, if one region experiences a natural disaster, political instability, or a technological failure, the organization can continue to operate smoothly by relying on its systems in other regions. This redundancy is important for organizations that operate in volatile or high-risk environments, where the likelihood of disruptions is higher.

The ability to operate multiple systems across various geographical locations also provides organizations with greater flexibility in terms of scaling their operations. As organizations expand into new markets or regions, they can easily replicate and customize their existing systems to meet the needs of these new locations. This scalability is important in today's fast-paced business environment, where organizations must be able to quickly seize new opportunities and respond to changing market conditions.

These systems often include diverse non-SAP platforms that generate and store vast amounts of data. These platforms may encompass a wide range of software and technologies, including legacy systems, cloud-based applications, and other enterprise resource planning (ERP) solutions. The data generated by these platforms is often heterogeneous, encompassing various data types, formats, and structures, which can complicate the process of data integration and management. Despite the diversity of these platforms, integrating the data they produce into a unified system is essential for several key reasons, including achieving a holistic view of business operations, enhancing data accuracy, and improving decision-making processes.

A holistic view of business operations is crucial for organizations that operate in complex, dynamic environments. This comprehensive perspective allows organizations to understand how different functions and processes interact with one another and impact overall performance. For example, integrating data from financial systems, supply chain management platforms, customer relationship management (CRM) systems, and human resources (HR) applications can provide a more complete picture of how various departments contribute to the organization's success. Unifying this disparate data, organizations can identify patterns, correlations, and trends that might not be apparent when data is siloed within individual platforms. This holistic approach to data management enables organizations to make more informed, strategic decisions that take into account the full spectrum of their operations [5].

Enhancing data accuracy is another critical benefit of integrating data from diverse non-SAP platforms. When data is fragmented across multiple systems, there is a risk of inconsistencies, duplications, and errors that can undermine the reliability of the information used for decision-making. These issues can arise from differences in data formats, the timing of data updates, or variations in data entry practices across different platforms. By consolidating data into a unified system, organizations can implement standardized processes for data collection, validation, and cleansing, for improving the overall quality and accuracy of the data. Accurate data is essential for effective business intelligence, as it forms the foundation upon which insights and analyses are built. With high-quality data, organizations can have greater confidence in their decisions and strategies [6].

Improving decision-making processes is perhaps one of the most significant advantages of integrating data from diverse platforms. Organizations need to be able to respond quickly and effectively to changes in the market, customer preferences, and competitive dynamics. This requires access to timely, accurate, and comprehensive data that can inform decisions at all levels of the organization. Integrating data from multiple sources, organizations can break down information silos and create a more transparent and accessible data environment. This integration enables stakeholders to access the information they need when they need it, facilitating faster and more accurate decision-making. Furthermore, integrated data supports advanced analytics and predictive modeling, which can provide deeper insights into business performance and help organizations anticipate future trends and challenges [7].

SAP HANA, an in-memory database and application platform, offers robust capabilities for integrating and managing large volumes

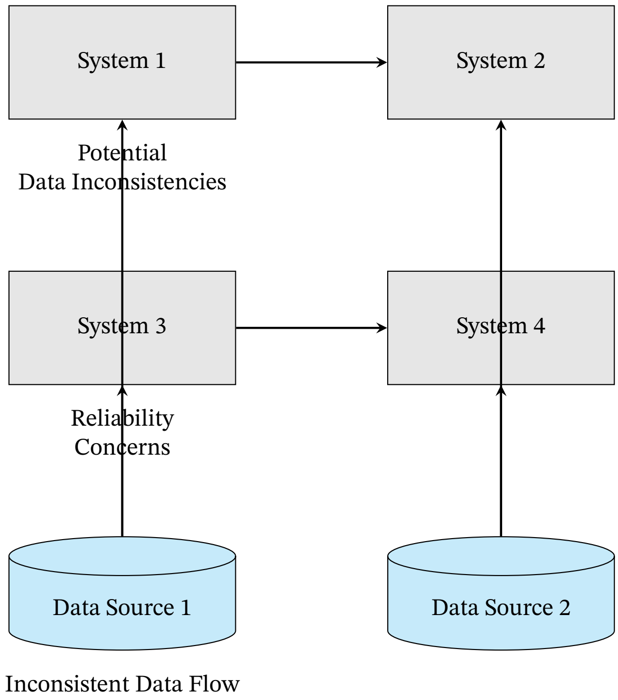

Figure 2. Data Flow and Consistency Challenges in Non-SAP Systems

of data from multiple sources. Unlike traditional databases that store data on disk, SAP HANA stores data in-memory, which allows for significantly faster data processing and retrieval times [8]. This speed is particularly advantageous for organizations that require real-time analytics and decision-making capabilities. SAP HANA's architecture is designed to handle large-scale data integration, making it an ideal solution for organizations with complex data environments that include multiple non-SAP platforms.

One of the key features of SAP HANA is its ability to handle both transactional and analytical workloads on a single platform. This means that organizations can perform real-time analytics on live transactional data without the need for data duplication or complex data extraction processes. This capability is particularly valuable in scenarios where organizations need to make immediate decisions based on the most current data available. For example, in a retail environment, SAP HANA can be used to analyze sales data in real-time, enabling retailers to adjust pricing, promotions, and inventory levels dynamically based on current demand and sales trends.

SAP HANA also supports the integration of structured and unstructured data from a variety of sources, including traditional databases, big data platforms, and cloud-based applications [9]. This flexibility allows organizations to consolidate data from diverse non-SAP platforms into a single, unified system. SAP HANA's data integration capabilities are further enhanced by its support for a wide range of data connectors and integration tools, which facilitate the seamless transfer of data from different sources into the SAP HANA environment. This integration not only streamlines data management processes but also ensures that data is readily available for analysis and reporting, providing organizations with a comprehensive view of their operations.

In addition to its data integration capabilities, SAP HANA offers powerful analytics and business intelligence tools that enable organizations to derive actionable insights from their data. SAP HANA's in-memory processing capabilities allow for the rapid execution of complex queries and data-intensive operations, making it possible to analyze large datasets in real-time. This real-time analytics capability is essential for organizations that need to monitor and respond to changes in their business environment quickly. For instance, in the financial services industry, SAP HANA can be used to perform real-time risk assessments and portfolio analysis, allowing financial institutions to manage risk more effectively and make more informed investment decisions.

Moreover, SAP HANA's advanced analytics capabilities include support for predictive analytics, machine learning, and artificial intelligence. These features enable organizations to go beyond descriptive analytics and explore predictive and prescriptive insights that can inform future business strategies. For example, by analyzing historical sales data alongside external factors such as weather patterns and economic indicators, organizations can predict future sales trends and optimize their inventory and supply chain operations accordingly [1]. The ability to leverage advanced analytics on integrated data from multiple sources gives organizations a significant competitive advantage by enabling them to make data-driven decisions with greater accuracy and foresight.

Furthermore, SAP HANA's integration and analytics capabilities are supported by its strong data security and governance features. Data security is a critical concern for organizations that handle sensitive or regulated data, and SAP HANA provides robust encryption, access control, and data masking features to protect data from unauthorized access and breaches. Additionally, SAP HANA supports comprehensive data governance frameworks that help organizations manage data quality, compliance, and lifecycle management. This ensures that data is not only integrated and accessible but also secure and compliant with relevant regulations and standards.

The ability to integrate diverse data from non-SAP platforms into a unified SAP HANA system also enhances collaboration and information sharing across the organization. When data is centralized and accessible through a unified platform, different departments and teams can collaborate more effectively, sharing insights and working together to achieve common goals. This collaborative environment fosters innovation and helps break down silos that can impede communication and decision-making [10]. For example, by integrating sales, marketing, and customer service data into a unified SAP HANA system, organizations can gain a 360-degree view of the customer journey, enabling cross-functional teams to collaborate on improving customer experience and loyalty.

Challenge	Description
Data Heterogeneity	Non-SAP systems often use different data formats, structures, and protocols, making integration complex. This can be represented by the variability in data types, $X_1, X_2, \ldots, X_n$, where X_i denotes a data type from system i.
Data Quality	Ensuring the accuracy, consistency, and reliability of data across different systems is a significant concern. Mathematically, this involves maintaining the integrity constraints, such as if $x = f(y)$ then $x = f(y')$ for all $y = y'$.
Scalability	Integrating data from geographically dispersed systems requires scalable solutions that can handle large volumes of data, denoted by $O(n \log n)$ complexity, where n is the size of the data.
Latency	Minimizing data transfer latency to enable real-time data analytics and reporting is critical. The goal is to reduce latency, τ, such that $\tau \to 0$ as $n \to \infty$.

Table 1. Challenges of Integrating Non-SAP Data into SAP HANA

1.2. Challenges of Integrating Non-SAP Data

The integration of data from non-SAP systems into SAP HANA is fraught with a variety of challenges that stem from the inherent complexities of managing diverse systems and ensuring seamless data flow across different platforms. These challenges can significantly impact the effectiveness and efficiency of data integration efforts, leading to disruptions in business operations and decision-making processes. Below are some of the key challenges associated with integrating non-SAP data into SAP HANA:

1.2.1. Data Heterogeneity

One of the foremost challenges in integrating non-SAP data into SAP HANA is the issue of data heterogeneity. Non-SAP systems often employ different data formats, structures, and protocols, which can vary widely depending on the technology stack, vendor specifications, and the specific use cases for which they were designed. For example, data might be stored in relational databases, NoSQL databases, flat files, or even proprietary formats unique to a particular application. The disparity in data types—such as structured, semi-structured, and unstructured data—compounds the difficulty of integration, as each type requires different handling and processing methods.

Moreover, non-SAP systems may follow different data modeling principles, resulting in varying data schemas, metadata, and relationships between data entities. This lack of uniformity complicates the process of mapping and transforming data into a format that is compatible with SAP HANA [11]. Without a standardized approach, data integration efforts can become fragmented and inconsistent, leading to errors and mismatches that can compromise the integrity of the integrated data.

1.2.2. Data Quality

Ensuring data quality during the integration of non-SAP data into SAP HANA is another significant challenge. Data quality encompasses several dimensions, including accuracy, consistency, completeness, and reliability. When data is sourced from multiple non-SAP systems, there is a risk that these quality attributes may be compromised due to discrepancies in data entry practices, differences in data validation rules, and the potential for human error [7].

For example, the same data element, such as a customer's name or product code, might be recorded differently in various systems, leading to duplication or inconsistency when the data is integrated. Additionally, data from different systems may have varying levels of completeness, with some systems capturing more detailed information than others. This can create gaps or ambiguities in the integrated dataset, making it difficult to derive accurate insights.

Furthermore, the temporal aspects of data quality—such as the timeliness and freshness of data—are critical in real-time analytics environments like SAP HANA. If data from non-SAP systems is outdated or not synchronized properly, it can lead to outdated or incorrect analysis, which in turn affects decision-making processes. Ensuring that data from multiple sources meets the required quality standards is a complex task that requires robust data governance and quality assurance mechanisms.

1.2.3. Scalability

Scalability is a critical concern when integrating data from non-SAP systems into SAP HANA, particularly for organizations with large-scale, geographically dispersed operations. As businesses grow and expand their operations, the volume of data generated by non-SAP systems can increase exponentially. This growth poses significant challenges in terms of the infrastructure and processes required to support the integration of large volumes of data into SAP HANA [2].

The scalability challenge is twofold: it involves both the technical scalability of the integration platform and the ability to manage increasing data complexity. Technically, the integration infrastructure must be capable of handling high data throughput and storage demands without performance degradation. This includes ensuring that the data pipelines, ETL (Extract, Transform, Load) processes, and network bandwidth are sufficient to accommodate the growing data volumes. Additionally, as the number of data sources and the complexity of the data increase, the integration processes must be scalable in terms of managing diverse data sets, maintaining data consistency, and ensuring that the system can continue to deliver real-time analytics capabilities.

1.2.4. Latency

Minimizing data transfer latency is another significant challenge in integrating non-SAP data into SAP HANA, especially when real-time analytics and reporting are critical to business operations. Latency refers to the delay between data generation in non-SAP systems and its availability for analysis in SAP HANA. High latency can severely impact the timeliness and relevance of business intelligence, leading to decisions based on outdated information [12].

Latency issues can arise from several factors, including the geographical dispersion of data sources, the complexity of data transformation processes, and the performance of the network infrastructure. For example, data originating from systems located in different regions may experience delays due to network congestion, bandwidth limitations, or the need to traverse multiple network hops. Additionally, complex data transformation and cleansing processes can introduce further delays, as the data must be processed before it can be loaded into SAP HANA.

The challenge is particularly acute in scenarios where real-time data integration is required, such as in financial services, supply chain management, or customer relationship management, where immediate access to up-to-date information is crucial. Ensuring low-latency data transfer requires optimizing data pipelines, reducing bottlenecks in data processing, and enhancing network performance to meet the stringent requirements of real-time analytics.

1.3. Master Data Management (MDM)

Master Data Management (MDM) is a fundamental component in the data integration process, particularly when it involves the incor-

Process	Description
Data Governance	Establishing policies and procedures for managing data quality and consistency.
Data Consolidation	Integrating data from various sources into a single, unified repository.
Data Cleansing	Identifying and rectifying data quality issues to ensure accuracy and reliability.
Data Synchronization	Ensuring that master data is consistent and up-to-date across all systems.

Table 2. Implementing MDM within SAP HANA

poration of data from multiple systems into a unified platform like SAP HANA. MDM focuses on the consistent, accurate, and reliable management of an organization's critical master data—such as customer information, product details, supplier records, and location data—across various systems and processes. Implementing MDM within SAP HANA is essential for ensuring that this master data remains consistent, high-quality, and synchronized across the entire organization. This involves several key elements: data governance, data consolidation, data cleansing, and data synchronization [13].

Data governance is the backbone of Master Data Management, providing the framework and guidelines necessary for ensuring the quality, consistency, and security of master data throughout the organization. Establishing robust data governance policies involves setting up a comprehensive structure that includes roles, responsibilities, processes, and standards for managing master data. This governance framework is critical in ensuring that all data-related activities adhere to the organization's policies, regulatory requirements, and industry best practices.

In the context of SAP HANA, data governance involves defining clear ownership of master data entities, such as assigning specific departments or roles the responsibility for maintaining and updating data related to customers, products, or suppliers. These data stewards are tasked with ensuring that the data remains accurate, complete, and up-to-date, while also monitoring compliance with established data quality standards. Governance also includes setting up data access controls and security measures to protect sensitive master data from unauthorized access or breaches.

Moreover, data governance in MDM includes the establishment of data lifecycle management processes. This ensures that data is appropriately managed from its creation through its use and eventual archival or deletion. By enforcing data governance, organizations can maintain a high level of data integrity, reduce redundancies, and ensure that all business units across the organization are working with consistent and reliable data.

Data consolidation is a critical step in the MDM process, involving the integration of master data from various sources into a single, unified repository within SAP HANA. Given that organizations often operate multiple systems—each housing different pieces of master data—consolidation is essential to create a single source of truth for all critical business entities.

The process of data consolidation typically begins with identifying and extracting master data from disparate systems, which may include legacy databases, CRM systems, ERP solutions, and other enterprise applications. Once extracted, this data needs to be transformed into a common format that is compatible with SAP HANA. This transformation process can involve standardizing data structures, aligning data fields, and harmonizing coding schemes used across different systems.

The consolidated data is then loaded into SAP HANA, where it can be managed and utilized for various business processes. By centralizing master data in a unified repository, organizations can ensure that all departments and functions have access to the same accurate and up-to-date information. This centralization not only improves the efficiency of business operations but also enhances the ability to perform cross-functional analysis and reporting.

Data Cleansing Data cleansing is an essential component of MDM, focusing on identifying and rectifying data quality issues to ensure that the master data is accurate, reliable, and fit for use. In the integration process, data cleansing involves systematically reviewing and correcting inconsistencies, errors, and inaccuracies in the master data before it is integrated into SAP HANA.

Common data quality issues that require cleansing include duplicate records, incorrect data entries, missing values, and outdated information. For example, multiple systems might contain slightly different versions of the same customer record, leading to duplications that need to be merged. Similarly, incorrect data formats, such as inconsistent date formats or improper use of currency symbols, need to be standardized [13].

Data cleansing can be both a manual and automated process. Manual data cleansing involves human intervention to review and correct data, while automated data cleansing tools can detect and rectify certain types of errors based on predefined rules and algorithms. In SAP HANA, automated data cleansing can leverage machine learning and advanced analytics to identify patterns and anomalies that suggest data quality issues.

Ensuring clean data is critical because poor data quality can lead to significant business inefficiencies, such as incorrect reporting, poor customer experiences, and faulty decision-making. By implementing rigorous data cleansing processes within MDM, organizations can significantly improve the quality of their master data, leading to more reliable and accurate business insights.

Data Synchronization Data synchronization is the process of ensuring that master data is consistent and up-to-date across all systems within the organization. This is particularly important in environments where data is constantly being updated across multiple platforms, and there is a need to ensure that these updates are reflected uniformly across the entire enterprise.

In an SAP HANA environment, data synchronization involves establishing processes and tools that continuously monitor and update master data across different systems. This can include real-time data integration solutions that automatically propagate changes made in one system to all other connected systems, ensuring that data remains consistent and synchronized. For example, if a change is made to a customer's address in a CRM system, this change should be automatically reflected in the SAP HANA repository and any other system that relies on this data.

Synchronization also involves resolving data conflicts that may arise when different systems have different versions of the same data entity. This requires establishing rules and mechanisms to determine which data version takes precedence and how conflicts are resolved. In SAP HANA, data synchronization can be managed using replication technologies, middleware solutions, or data integration platforms that facilitate seamless data flow and consistency across systems.

Maintaining synchronized data across an organization is crucial for ensuring that business processes run smoothly and that decision-makers are working with the most current and accurate data. Inconsistencies in data synchronization can lead to errors, delays, and a lack of trust in the data, all of which can have negative consequences for business operations.

2. Integrating Non-SAP Data into SAP HANA

Integrating Non-SAP Data into SAP HANA requires meticulous planning and execution, given the complexity and diversity of data sources involved. SAP HANA, an in-memory, column-oriented, relational

ETL Step	Description
Data Extraction	Extracting data from various non-SAP systems, which may include databases, flat files, APIs, and other sources.
Data Transformation	Converting data into a format compatible with SAP HANA, which may involve data cleaning, filtering, aggregation, and enrichment.
Data Loading	Importing the transformed data into SAP HANA, ensuring it is accurately stored and accessible for analysis.

Table 3. ETL Process for Integrating Non-SAP Data into SAP HANA

database management system, offers high-performance data processing capabilities that can significantly enhance business operations when integrated with data from non-SAP systems. This integration process hinges on the effective execution of the Data Extraction, Transformation, and Loading (ETL) process, which serves as the backbone of the integration effort.

2.1. Data Extraction, Transformation, and Loading (ETL)

Data Extraction is the first step in the ETL process and involves identifying and extracting relevant data from various non-SAP systems. These systems may include traditional relational databases such as Oracle, MySQL, or SQL Server; NoSQL databases like MongoDB or Cassandra; flat files such as CSV or XML; as well as data streams from APIs or message queues. The extraction phase must be designed to handle the heterogeneity of these sources, including differences in data models, formats, and protocols. In particular, attention must be paid to ensuring that the extraction process is both efficient and minimally disruptive to the source systems, especially when dealing with large volumes of data or systems that are operationally critical [3].

The extracted data often contains inconsistencies, redundancies, and errors that must be addressed during the Data Transformation phase. This phase involves converting data into a format that is compatible with SAP HANA, which may require various operations such as data cleaning, filtering, aggregation, and enrichment. Data cleaning involves removing or correcting erroneous data entries, which can be caused by human error, data corruption, or outdated information. Filtering is applied to exclude irrelevant data that does not need to be loaded into SAP HANA optimizing storage and processing resources. Aggregation involves summarizing detailed data, which is particularly useful for reducing the volume of data that needs to be stored and processed. Enrichment is the process of enhancing the data by adding new attributes or by integrating data from multiple sources to provide a more comprehensive view. During transformation, special attention must be paid to maintaining data integrity and consistency, particularly when dealing with relational data, where preserving relationships between different entities is crucial.

The final step, Data Loading, involves importing the transformed data into SAP HANA. This step must ensure that the data is accurately stored and readily accessible for analysis. The loading process must be optimized for performance, as SAP HANA is designed to handle large datasets and perform complex queries in real-time. Various loading strategies can be employed depending on the data volume and update frequency. For instance, initial bulk loading might be followed by incremental loading to update the dataset with new or changed data. Additionally, parallel loading techniques can be used to further optimize performance. It is also essential to implement robust error-handling mechanisms during the loading phase to ensure that any issues are quickly identified and resolved, minimizing the risk of data corruption or loss [4].

2.2. ETL Tools and Techniques

To facilitate the integration process, several ETL tools and techniques are available, each offering distinct advantages and capabilities that can be leveraged depending on the specific requirements of the integration project.

SAP Data Services is a comprehensive ETL tool that is particularly well-suited for environments where SAP HANA is the target system. It offers robust support for data integration, transformation, and cleansing, providing a graphical interface for designing ETL workflows, which can significantly streamline the development process. SAP Data Services also integrates seamlessly with other SAP tools and technologies, enabling a more cohesive and integrated approach to data management within SAP ecosystems. The tool supports a wide range of data sources and formats, making it versatile enough to handle diverse integration scenarios. Moreover, it offers advanced data quality features, such as data profiling, deduplication, and validation, which are critical for ensuring the accuracy and reliability of the integrated data [14].

Apache Kafka is a distributed streaming platform that can handle real-time data ingestion and processing, making it ideal for scenarios where timely data integration is crucial. Kafka can be used to stream data from non-SAP systems into SAP HANA in real-time, providing near-instantaneous access to the latest data for analysis. This is particularly valuable in use cases where businesses need to react quickly to changing conditions, such as in financial trading or real-time monitoring of industrial processes. Kafka's architecture is designed for scalability and fault tolerance, allowing it to handle high-throughput data streams with low latency. Integration with SAP HANA can be achieved using connectors and stream processing frameworks like Apache Flink or Apache Spark, which can transform and load data into SAP HANA as it is ingested by Kafka.

Custom ETL Scripts represent a more tailored approach to data integration, where scripts are written in programming languages such as Python or Java to address specific integration requirements. This approach provides maximum flexibility, allowing developers to implement custom logic and optimizations that may not be possible with off-the-shelf ETL tools. For instance, custom scripts can be designed to handle complex data transformations, perform specific data validations, or integrate with proprietary systems that are not supported by standard ETL tools. However, this approach also requires a higher level of expertise and maintenance effort, as custom scripts must be thoroughly tested and kept up-to-date with any changes in the source or target systems.

2.3. Data Integration Architectures

To effectively integrate non-SAP data into SAP HANA, various data integration architectures can be implemented, each offering different trade-offs in terms of performance, flexibility, and complexity.

Centralized Architecture involves consolidating data from different sources into a central repository within SAP HANA. In this architecture, all data is physically moved into SAP HANA, where it can be stored, managed, and analyzed in a single, unified environment. This approach offers several advantages, including simplified data management, as all data resides in one location, and improved query performance, as data does not need to be retrieved from external sources during query execution. However, centralized architecture can also present challenges, particularly in terms of data volume and storage costs, as all data must be physically loaded into SAP HANA. Additionally, it may introduce latency if the data must be frequently updated or synchronized with the source systems. To mitigate these issues, data archiving and partitioning strategies can be employed to

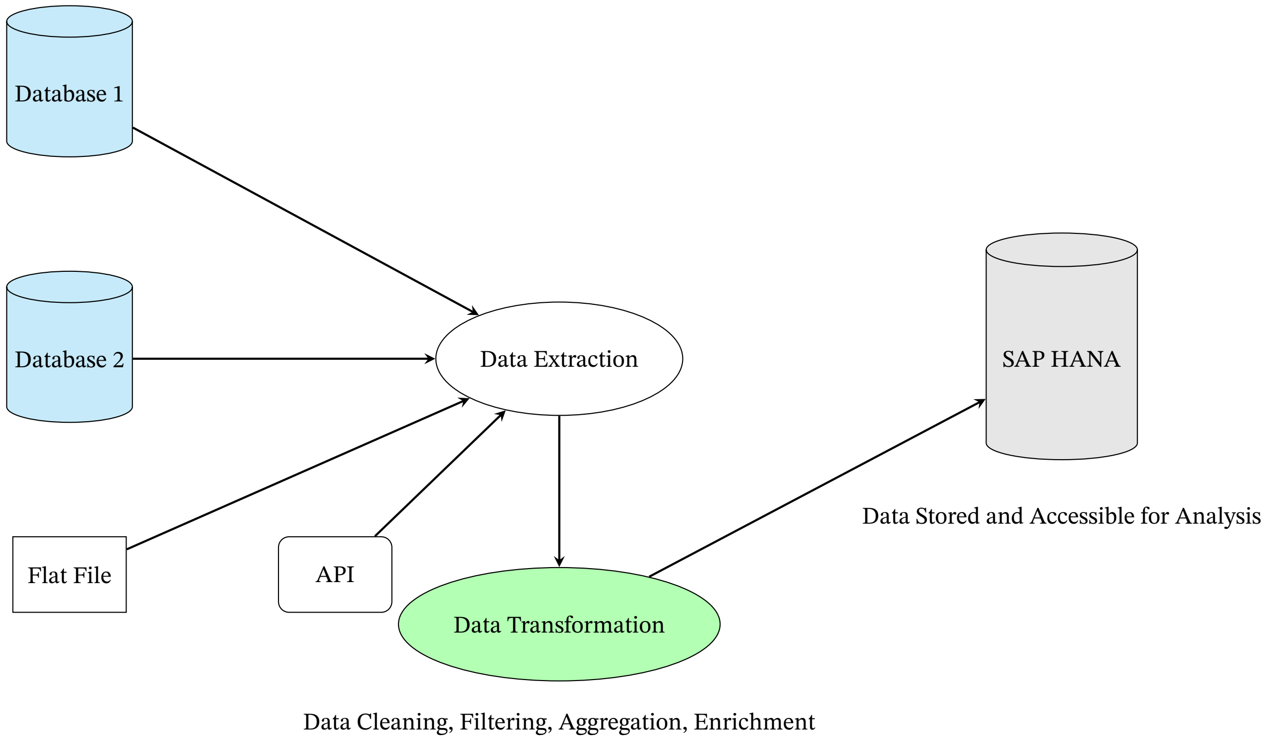

Figure 3. ETL Process: Extracting data from various non-SAP systems, transforming it for compatibility with SAP HANA, and loading it into SAP HANA for analysis.

ETL Tool/Technique	**Description**
SAP Data Services	A comprehensive ETL tool that supports data integration, transformation, and cleansing.
Apache Kafka	A distributed streaming platform that can handle real-time data ingestion and processing.
Custom ETL Scripts	Tailored scripts written in programming languages such as Python or Java to address specific integration requirements.

Table 4. ETL Tools and Techniques for Integration into SAP HANA

manage storage requirements and optimize query performance [15].

Federated Architecture allows data to remain in its original location, and SAP HANA queries the data in real-time as needed. This architecture is particularly useful when integrating data from systems that are not suitable for full data replication, such as large-scale operational databases or third-party systems that cannot be fully migrated to SAP HANA. In a federated architecture, SAP HANA acts as a virtual data layer, providing a unified view of the data across multiple sources without the need to physically move the data. This approach offers greater flexibility, as it allows organizations to integrate and query data from diverse sources without disrupting the existing systems. However, federated architecture can introduce performance challenges, as query execution may be affected by the performance of the external systems. To address these challenges, techniques such as caching, query optimization, and parallel processing can be employed to improve query performance and ensure timely access to the data.

Hybrid Architecture combines elements of both centralized and federated architectures to balance performance and flexibility. In a hybrid architecture, some data is consolidated into SAP HANA, while other data remains in its original location and is accessed in real-time as needed. This approach allows organizations to optimize the storage and processing of different types of data based on their specific requirements. For example, frequently accessed or critical data can be loaded into SAP HANA to ensure fast query performance, while less critical or less frequently accessed data can be left in external systems and accessed on-demand. Hybrid architecture provides a flexible and scalable solution for integrating diverse data sources, allowing organizations to leverage the strengths of both centralized and federated approaches. However, it also requires careful planning and management to ensure that data is properly synchronized and that queries are optimized to access the data efficiently.

Data Governance and Security are critical aspects of integrating non-SAP data into SAP HANA, particularly when dealing with sensitive or regulated data. Organizations must ensure that data integration processes comply with relevant data protection regulations, such as GDPR or HIPAA, which may impose strict requirements on how data is collected, stored, and accessed. Implementing robust data encryption, access control, and auditing mechanisms is essential to protect data integrity and prevent unauthorized access. Additionally, data governance frameworks should be established to define data ownership, data quality standards, and data stewardship roles, ensuring that data integration efforts are aligned with organizational

Architecture Type	**Description**
Centralized Architecture	Data from different sources is consolidated into a central repository within SAP HANA.
Federated Architecture	Data remains in its original location, and SAP HANA queries the data in real-time as needed.
Hybrid Architecture	Combines centralized and federated approaches to balance performance and flexibility.

Table 5. Architectures for Data Integration into SAP HANA

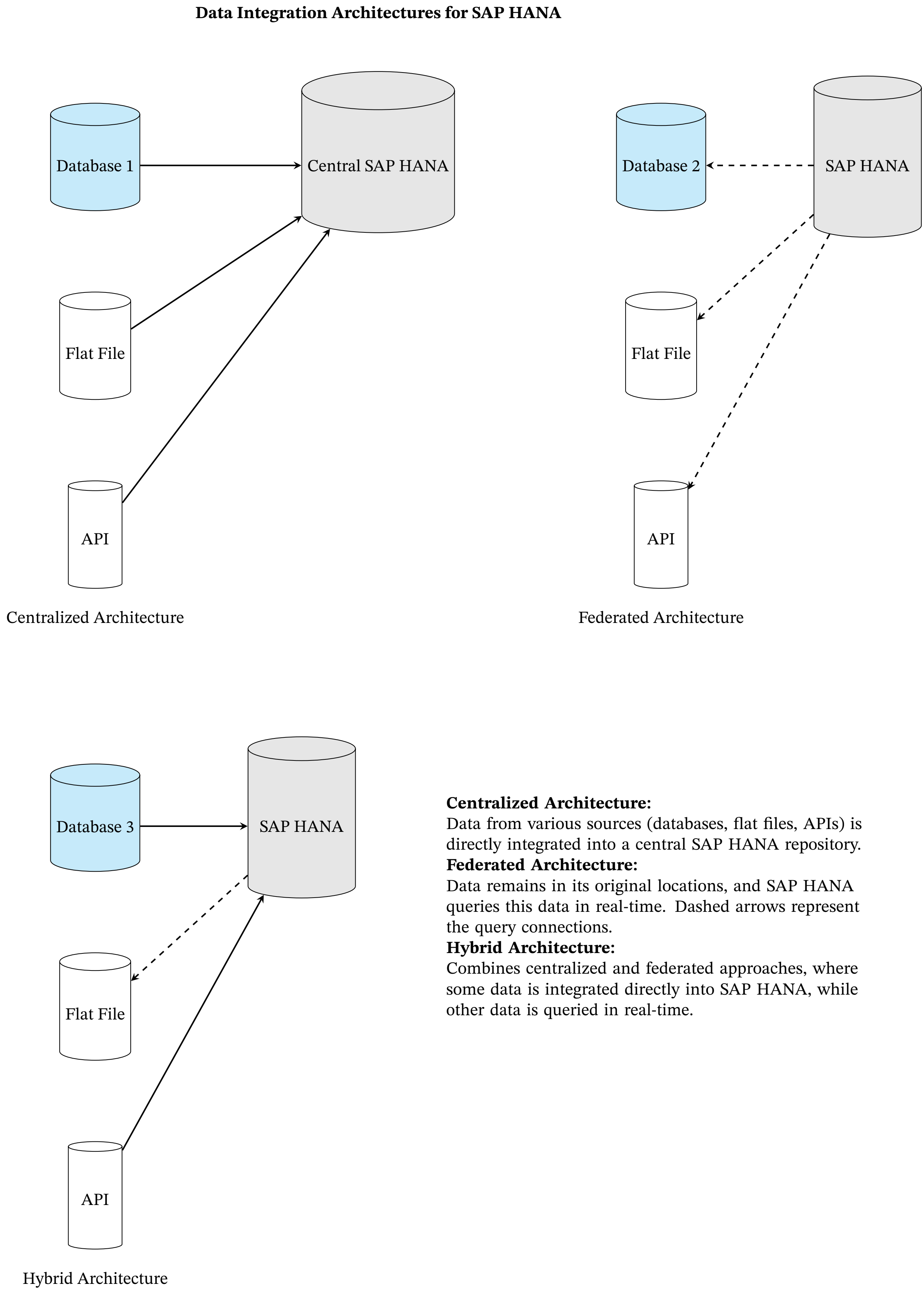

Figure 4. Data Integration Architectures: Centralized, Federated, and Hybrid. Centralized architecture consolidates data into a central SAP HANA repository, Federated architecture queries data in its original locations, and Hybrid architecture combines both approaches.

objectives and regulatory requirements.

Performance Optimization is another key consideration, as the success of data integration efforts often hinges on the ability to efficiently process and analyze large volumes of data. To optimize performance, organizations should employ techniques such as data partitioning, indexing, and query optimization to reduce the time and resources required for data processing and analysis. Additionally, leveraging SAP HANA's in-memory processing capabilities can significantly enhance performance, enabling real-time analytics and reducing latency. Organizations should also consider the use of data compression techniques to minimize storage requirements and improve query performance.

Scalability and Flexibility are also important factors to consider when designing data integration architectures. As organizations grow, their data integration needs are likely to change, requiring scalable solutions that can accommodate increasing data volumes and diverse data sources. Implementing modular and extensible architectures, such as microservices-based approaches, can provide the flexibility needed to adapt to changing requirements and integrate new data sources as they emerge. Additionally, organizations should consider leveraging cloud-based solutions, which offer scalable storage and processing capabilities, enabling them to scale their data integration efforts without the need for significant upfront investments in infrastructure.

Monitoring and Maintenance are essential to ensuring the success of data integration efforts. Data integration processes must be continuously monitored to detect and resolve any issues that may arise, such as data inconsistencies, errors, or performance bottlenecks. Implementing automated monitoring and alerting tools can help organizations quickly identify and address these issues, minimizing the impact on data processing and analysis.

Furthermore, regular maintenance activities, such as data refreshes, re-indexing, and updating ETL workflows, are necessary to keep

the integration process running smoothly. These activities should be scheduled and automated where possible to reduce the burden on IT staff and ensure that data remains accurate and up-to-date.

In addition to the technical aspects of integration, effective communication and collaboration between different teams, such as IT, data management, and business users, are crucial. Clear documentation of the ETL processes, data models, and integration architecture helps ensure that all stakeholders have a shared understanding of the integration efforts and can work together to resolve any issues that arise.

Finally, organizations should be prepared to adapt their integration strategies as new technologies and best practices emerge. Staying informed about the latest developments in data integration, ETL tools, and SAP HANA capabilities can help organizations continuously improve their integration processes and maintain a competitive edge in their data management practices.

3. Reporting Solutions

Integrating non-SAP data into SAP HANA significantly enhances an organization's ability to perform real-time data analytics, enabling immediate and actionable insights into their operations. The process of integration is vital, as it allows the consolidation of disparate data sources into a single, coherent platform. This consolidation is not merely a technical endeavor; it transforms how organizations approach decision-making by providing a unified view of data across various functions and systems.

Real-time data analytics is a transformative capability that shifts the paradigm from reactive to proactive management. Organizations can move from making decisions based on historical data to leveraging current, up-to-the-minute information. This capability is crucial in dynamic industries where timely and accurate data can make the difference between seizing an opportunity or suffering a loss. The integration of non-SAP data into SAP HANA is accomplished through various techniques such as ETL (Extract, Transform, Load) processes, real-time data replication, and the use of connectors and APIs that bridge different systems. This process ensures that data from legacy systems, third-party applications, and external sources is available within the SAP HANA environment for comprehensive analysis.

3.1. SAP HANA Live: Real-Time Operational Reporting

SAP HANA Live plays a pivotal role in this integrated environment by providing pre-configured content for real-time operational reporting. SAP HANA Live is essentially a library of pre-defined calculation views and data models tailored to specific business processes. These views are optimized for SAP HANA's in-memory computing capabilities, allowing for rapid query execution and real-time data retrieval. The calculation views in SAP HANA Live are designed to be modular and reusable, enabling organizations to customize and extend them according to their specific requirements. This flexibility is crucial in adapting the reporting framework to the unique needs of an organization, whether they are in manufacturing, finance, or customer service.

Moreover, SAP HANA Live supports a broad range of reporting and analytics tools within the SAP ecosystem, including SAP BusinessObjects and SAP Lumira. By leveraging these tools in conjunction with SAP HANA Live, organizations can create a comprehensive reporting environment that supports both operational and strategic decision-making. The real-time nature of SAP HANA Live ensures that reports reflect the most current data, providing stakeholders with up-to-date information that is essential for effective decision-making.

3.2. SAP BusinessObjects: Comprehensive Reporting, Analysis, and Data Visualization

SAP BusinessObjects is a suite of tools designed to offer comprehensive capabilities in reporting, analysis, and data visualization. It plays an integral role in translating the raw data stored in SAP HANA into actionable insights. BusinessObjects includes tools like Web Intelligence, Crystal Reports, and Analysis for Office, each serving different reporting and analytical needs within an organization.

Web Intelligence allows users to create ad-hoc queries and reports, providing the flexibility to explore data from multiple angles. This tool is particularly useful for business users who need to generate their own reports without relying on IT support. Crystal Reports is renowned for its pixel-perfect reporting capabilities, making it ideal for producing highly formatted documents such as invoices, statements, and regulatory reports. Analysis for Office integrates SAP data with Microsoft Excel, offering a familiar interface for users who prefer to analyze data within spreadsheets.

The strength of SAP BusinessObjects lies in its ability to provide a unified platform for diverse reporting needs. It supports both traditional static reports and dynamic dashboards that can be customized and interactive. Furthermore, BusinessObjects integrates seamlessly with other SAP tools and data sources, making it a versatile solution for organizations that need to report on both SAP and non-SAP data.

3.3. SAP Lumira: Self-Service Data Visualization

SAP Lumira is a self-service data visualization tool that empowers users to create interactive dashboards and reports without needing extensive technical knowledge. Lumira allows users to connect to various data sources, including SAP HANA, and transform raw data into meaningful visual representations. This capability is critical in enabling non-technical users to participate in the data analysis process, democratizing access to data insights across the organization.

The tool's drag-and-drop interface and intuitive design make it accessible to users with varying levels of expertise. Lumira supports a wide range of visualization options, from simple charts and graphs to complex heat maps and geospatial analyses. This flexibility allows users to tailor their visualizations to the specific needs of their audience, whether it be executives requiring high-level dashboards or analysts needing detailed, data-rich visual reports.

Component	Description
SAP HANA Live	Pre-configured content for real-time operational reporting.
SAP BusinessObjects	A suite of tools for comprehensive reporting, analysis, and data visualization.
SAP Lumira	A self-service data visualization tool that allows users to create interactive dashboards and reports.

Table 6. Key Components of a Robust Reporting Solution

Technique	Description
Time Series Analysis	Utilizing historical data, $\{x_t\}_{t=1}^{T}$, to identify patterns and trends for future predictions. This involves techniques such as ARIMA (AutoRegressive Integrated Moving Average) where future values are predicted as $x_{t+h} = \phi_1 x_t + \phi_2 x_{t-1} + \cdots + \theta_h \epsilon_{t-h}$.
Machine Learning Algorithms	Applying advanced algorithms such as regression, neural networks, and decision trees to enhance forecasting accuracy. For instance, a linear regression model can be represented as $\hat{y} = \beta_0 + \beta_1 x_1 + \cdots + \beta_p x_p$, where $\hat{y}$ is the predicted value.
Predictive Analytics	Combining statistical methods and machine learning to predict future outcomes based on historical data. This can be formalized as $P(Y\|\mathbf{X})$, where Y is the future outcome and $\mathbf{X}$ represents historical data and other features used in the model.

Table 7. Key Techniques for Developing Forecasting Models Using Integrated Non-SAP Data in SAP HANA

SAP Lumira also facilitates collaboration and sharing within organizations. Users can publish their dashboards and reports to the SAP Lumira Server, where they can be accessed by other users. This sharing capability ensures that insights generated by one user can be leveraged by others, fostering a data-driven culture within the organization.

3.4. Designing Effective Reports

Designing effective reports is a critical aspect of maximizing the value of integrated data. Effective report design requires careful consideration of various factors to ensure that the reports not only deliver accurate information but also present it in a way that is easily understood and actionable by stakeholders.

Data relevance is paramount in report design. It is essential to ensure that reports focus on key performance indicators (KPIs) and metrics that are directly aligned with the organization's strategic objectives. This alignment ensures that the data being reported on is meaningful and that the insights derived from it are actionable. For instance, a financial report might focus on metrics such as revenue growth, profit margins, and cash flow, while a supply chain report might highlight metrics such as inventory turnover, lead times, and supplier performance.

Ensuring data relevance also involves filtering out extraneous information that could clutter the report and distract from the key insights. This requires a deep understanding of the business context in which the report will be used, as well as the specific information needs of the report's intended audience. By focusing on the most critical data points, report designers can create reports that provide clear and actionable insights.

User accessibility is another crucial consideration in report design. Reports must be designed with the end-user in mind, ensuring that they are easy to navigate and interpret. This involves creating user-friendly interfaces that allow stakeholders to access the information they need quickly and efficiently. For example, reports can be structured to include summary sections or executive dashboards that provide a high-level overview of key metrics, with the option to drill down into more detailed data for those who require it.

Accessibility also extends to the technical aspects of report distribution. Reports should be available in formats that are easily accessible to users, whether they are working from a desktop computer, a tablet, or a smartphone. This may involve leveraging web-based reporting tools or mobile applications that allow users to access reports on the go. Additionally, reports should be designed to accommodate users with varying levels of data literacy, providing clear explanations of complex metrics and offering guided analysis where appropriate.

Visual clarity is a fundamental element of effective report design. The goal of any report is to communicate information clearly and concisely, and visual design plays a key role in achieving this. Effective use of charts, graphs, and other visual elements can help to highlight trends, patterns, and outliers in the data, making it easier for stakeholders to grasp the key insights.

One important aspect of visual clarity is the use of appropriate chart types. Different types of data are best represented by different types of charts, and selecting the right chart type can make a significant difference in how easily the data is understood. For example, line charts are ideal for showing trends over time, while bar charts are effective for comparing quantities across different categories. Heat maps can be used to show the distribution of values across two dimensions, and scatter plots are useful for identifying correlations between variables.

In addition to selecting the right chart types, it is important to ensure that visualizations are designed with clarity in mind. This includes using a consistent color scheme, avoiding clutter, and providing clear labels and legends. The goal is to create visualizations that are not only aesthetically pleasing but also functional, making it easy for users to extract the information they need at a glance.

Tool/Capability	Description
SAP Predictive Analytics	An integrated suite that supports the entire predictive modeling process, from data preparation to model deployment.
SAP HANA Predictive Analysis Library (PAL)	A collection of predictive algorithms that can be used directly within SAP HANA for real-time analytics.
SAP Leonardo Machine Learning Foundation	A platform that enables the development and deployment of machine learning models within SAP environments.

Table 8. Tools and Capabilities for Building and Deploying Forecasting Models in SAP HANA

4. Conclusion

In the increasingly interconnected global business landscape, organizations often manage numerous systems across various geographical regions. These systems, encompassing both SAP and non-SAP platforms, are vital for generating and storing vast amounts of data. However, the existence of these disparate systems can lead to data silos, where critical business information is isolated, resulting in challenges in achieving a cohesive view of business operations. The ability to integrate these diverse data sources into a unified system is essential for ensuring data accuracy, fostering enhanced decision-making, and enabling comprehensive analytics. SAP HANA, a high-performance in-memory database and application platform, offers robust capabilities for integrating and managing large volumes of data from multiple sources. By facilitating real-time data aggregation, SAP HANA enables organizations to gain instantaneous insights into business processes and supports the development of sophisticated analytics. This capability is crucial in modern enterprises, where the speed and accuracy of decision-making are directly linked to competitive advantage.

Data integration goes beyond the mere aggregation of data; it involves transforming and aligning data to ensure consistency and reliability across the enterprise. A well-integrated system allows businesses to break down data silos, leading to improved collaboration, better customer insights, and more agile responses to market dynamics. Moreover, it enhances operational efficiency by automating data flows and reducing the need for manual data reconciliation minimizing the risk of errors and discrepancies. However, integrating data from non-SAP systems into SAP HANA presents several technical and operational challenges that must be addressed to ensure a successful implementation.

One significant challenge is data heterogeneity. Non-SAP systems are often built using various technologies, data formats, and communication protocols, leading to differences in the structure, semantics, and representation of data across systems. For instance, one system might store date values in a different format than another or use different schemas to represent similar entities such as customers or products. This heterogeneity complicates the integration process, as it requires extensive mapping, transformation, and standardization of data to achieve consistency within SAP HANA. Additionally, ensuring data quality is a significant concern during integration. Data from different systems might vary in accuracy, consistency, and reliability, posing risks to the integrity of business processes and analytics. Integrating data of varying quality without appropriate cleansing and validation can lead to misleading insights and flawed decision-making. As a result, rigorous data quality management practices, including cleansing, de-duplication, and validation, are essential components of the integration process.

Scalability is another critical challenge. Integrating data from geographically dispersed systems involves handling large volumes of data that grow as organizations expand. The integration solution must be scalable to accommodate this growth without compromising performance. SAP HANA's in-memory architecture offers advantages in processing large datasets; however, the integration process must be designed to efficiently manage data transfer, storage, and processing loads. Latency is also a crucial factor to consider when integrating non-SAP data into SAP HANA. In many cases, organizations require real-time data analytics and reporting to make timely and informed decisions. High latency in data transfer and processing can hinder the ability to perform real-time analytics, reducing the responsiveness of the system. Therefore, minimizing data transfer latency is vital for ensuring that integrated data is available for analysis as quickly as possible.

Master Data Management (MDM) plays a pivotal role in addressing some of these challenges by focusing on the consistent and accurate management of master data across the organization. Master data, which includes critical business entities such as customers, products, suppliers, and locations, must be consistently defined and maintained across all systems to ensure data integrity. Implementing MDM within SAP HANA involves establishing robust data governance practices to manage data quality and consistency. Data governance ensures that policies, standards, and procedures are in place for the proper handling and maintenance of master data reducing the risk of errors and inconsistencies during integration. Data consolidation is another key aspect of MDM, involving the integration of data from various sources into a single, unified repository within SAP HANA. This process helps eliminate data silos and provides a single source of truth for master data, which is essential for accurate reporting and analytics.

Data cleansing is also a critical component of MDM, focusing on identifying and rectifying data quality issues such as duplicates, inaccuracies, and incomplete records. By ensuring that master data is clean and reliable, organizations can enhance the accuracy and trustworthiness of their analytics and decision-making processes. Data synchronization ensures that master data remains consistent and up-to-date across all systems, further enhancing the integrity of integrated data within SAP HANA.

The process of integrating non-SAP data into SAP HANA typically involves a sequence of steps known as Extraction, Transformation, and Loading (ETL). The ETL process is fundamental to ensuring that data from diverse sources is accurately and efficiently brought into SAP HANA for further use. The first step, data extraction, involves retrieving data from various non-SAP systems, which may include databases, flat files, APIs, and other data sources. This step is crucial because it requires handling different data formats, protocols, and access methods, which can be complex and require specialized tools and techniques. Once the data is extracted, it undergoes data transformation, where it is converted into a format compatible with SAP HANA. This transformation process may involve a range of tasks, including data cleaning, filtering, aggregation, and enrichment. The goal is to ensure that the data is standardized, accurate, and ready for analysis within the SAP HANA environment. Finally, the transformed data is loaded into SAP HANA, where it is stored and made accessible for reporting and analytics. The loading process must ensure that data is accurately mapped to the appropriate tables and structures within SAP HANA to maintain the integrity and usability of the integrated data.

Several ETL tools and techniques can be employed to facilitate the integration of non-SAP data into SAP HANA. SAP Data Services is a comprehensive ETL tool that supports data integration, transformation, and cleansing. It provides a range of functionalities for managing data from different sources and formats, making it a popular choice for organizations looking to integrate non-SAP data into SAP environments. Apache Kafka, a distributed streaming platform, is another tool that can be used for real-time data ingestion and processing. Kafka is particularly well-suited for scenarios where organizations need to handle large volumes of data in real time, such as in streaming analytics or real-time monitoring applications. Additionally, custom ETL scripts written in programming languages such as Python or Java can be developed to address specific integration requirements that are not fully met by out-of-the-box ETL tools. These scripts allow for greater flexibility and customization, enabling organizations to tailor the ETL process to their unique needs.

Different data integration architectures can be implemented to support the integration of non-SAP data into SAP HANA, each with its own set of advantages and trade-offs. In a centralized architecture, data from different sources is consolidated into a central repository within SAP HANA. This approach simplifies data management and provides a single source of truth for analytics, but it may require significant data movement and storage resources. In a federated architecture, data remains in its original location, and SAP HANA queries the data in real time as needed. This approach reduces the need for data movement and allows organizations to leverage existing data stores, but it may introduce latency and performance challenges depending on the complexity and distribution of the data sources. A hybrid architecture combines elements of both centralized and federated approaches, allowing organizations to balance performance and flexibility by centralizing critical data while leaving less critical data in its original location.

Integrating non-SAP data into SAP HANA enables organizations to perform real-time data analytics, allowing them to gain immediate insights into their operations and make data-driven decisions more quickly. Real-time analytics are increasingly important in today's fast-paced business environment, where the ability to respond to changing market conditions and customer demands in real time can provide a significant competitive advantage. SAP HANA supports real-time analytics through various tools and solutions. SAP HANA Live provides pre-configured content for real-time operational reporting, enabling organizations to quickly deploy analytics solutions that leverage integrated data. SAP BusinessObjects offers a suite of tools for comprehensive reporting, analysis, and data visualization, allowing users to create detailed reports and dashboards that provide deep insights into business performance. SAP Lumira is a self-service data visualization tool that enables users to create interactive dashboards and reports, making it easier for stakeholders to access and interpret data.

Designing effective reports is essential for maximizing the value of integrated data. When designing reports, it is important to ensure that the data presented is relevant to the organization's key performance indicators (KPIs) and business objectives. Reports should focus on the metrics that are most critical to decision-making, providing actionable insights that can drive business outcomes. Additionally, reports should be designed with user accessibility in mind, featuring user-friendly interfaces that allow stakeholders to easily access and interpret the data. Visual clarity is also crucial, as clear and concise visualizations help users understand the data and identify trends, patterns, and anomalies more easily.

Integrating non-SAP data into SAP HANA also provides a comprehensive dataset for developing accurate forecasting models. These models can be used for various purposes, including demand forecasting, inventory management, and financial planning. Leveraging integrated data allows organizations to develop more accurate and reliable forecasts, which are essential for strategic planning and resource allocation. Time series analysis is one of the key techniques used in forecasting, involving the analysis of historical data to identify patterns and trends that can be used to predict future outcomes. Machine learning algorithms, such as regression, neural networks, and decision trees, can also be applied to enhance the accuracy of forecasting models by identifying complex relationships and patterns in the data. Predictive analytics combines statistical methods and machine learning to predict future outcomes based on historical data, providing organizations with valuable insights that can inform decision-making.

SAP HANA offers a range of tools and capabilities for building and deploying forecasting models. SAP Predictive Analytics is an integrated suite that supports the entire predictive modeling process, from data preparation to model deployment. It enables organizations to develop, validate, and deploy forecasting models that leverage integrated data from SAP and non-SAP sources. The SAP HANA Predictive Analysis Library (PAL) provides a collection of predictive algorithms that can be used directly within SAP HANA for real-time analytics, enabling organizations to perform forecasting and predictive analysis in real time. Additionally, the SAP Leonardo Machine Learning Foundation provides a platform for developing and deploying machine learning models within SAP environments, further enhancing the organization's ability to leverage integrated data for advanced analytics and forecasting.

References

[1] U. Steger, *Corporate diplomacy: The strategy for a volatile, fragmented business environment*. John Wiley & Sons, 2003.

[2] J. Braithwaite and P. Drahos, *Global business regulation*. Cambridge university press, 2000.

[3] A. Cuervo-Cazurra, "Global strategy and global business environment: The direct and indirect influences of the home country on a firm's global strategy," *Global Strategy Journal*, vol. 1, no. 3-4, pp. 382–386, 2011.

[4] W. A. Dymsza, "Trends in multinational business and global environments: A perspective," *Journal of International Business Studies*, vol. 15, pp. 25–46, 1984.

[5] M. Andrei, C. Lemke, G. Radestock, *et al.*, "Sap hana adoption of non-volatile memory," *Proceedings of the VLDB Endowment*, vol. 10, no. 12, pp. 1754–1765, 2017.

[6] B. Berg and P. Silvia, *Einführung in SAP HANA*. Galileo Press Bonn, 2013.

[7] C. Bornhövd, R. Kubis, W. Lehner, H. Voigt, and H. Werner, "Flexible information management, exploration and analysis in sap hana.," in *DATA*, 2012, pp. 15–28.

[8] V. Sikka, F. Färber, W. Lehner, S. K. Cha, T. Peh, and C. Bornhövd, "Efficient transaction processing in sap hana database: The end of a column store myth," in *Proceedings of the 2012 ACM SIGMOD International Conference on Management of Data*, 2012, pp. 731–742.

[9] S. H. P. SPS, "Sap hana master guide," 2015.

[10] M. Rudolf, M. Paradies, C. Bornhövd, and W. Lehner, "The graph story of the sap hana database," 2013.

[11] N. May, A. Böhm, M. Block, and W. Lehner, "Managed query processing within the sap hana database platform," *Datenbank-Spektrum*, vol. 15, pp. 141–152, 2015.

[12] R. Brunel, J. Finis, G. Franz, *et al.*, "Supporting hierarchical data in sap hana," in *2015 IEEE 31st International Conference on Data Engineering*, IEEE, 2015, pp. 1280–1291.

[13] Z. Caklovic, P. Expert, O. Rebholz, *et al.*, "Bringing persistent memory technology to sap hana: Opportunities and challenges," *Annual SNIA Persistent Memory Summit*, 2017.

[14] F. Faerber, J. Dees, M. Weidner, S. Baeuerle, and W. Lehner, "Towards a web-scale data management ecosystem demonstrated by sap hana," in *2015 IEEE 31st International Conference on Data Engineering*, IEEE, 2015, pp. 1259–1267.

[15] S. HANA, "Sap hana® platform–technical overview," *SAP Whitepaper*, 2012.