BARC

BARC Spotlight

SAP Data Integration – Turning Challenges into Opportunities

Authors: Timm Grosser, Thomas Zeutschler

Publication: August 23, 2024

Provided by

DataTechnology

Index

BARC

SAP Data Integration – Turning Challenges into Opportunities

Introduction

In today's rapidly evolving digital landscape, businesses are increasingly recognising the urgent need to modernise their data platforms to maintain their competitive advantage. Specifically, maximising the value of SAP data is essential for businesses looking to optimise supply chains, modernise finance functions and empower employees with better analytics and insights.

"Seamless Data Pipelines should be a top priority on your agenda."

This phrase encapsulates the essence of modern data management. A seamless integration process is a prerequisite for long-term competitiveness, enabling businesses to be data-driven, data-connected, and data-intelligent. Achieving this requires a structured and well-managed approach, offering significant returns in business agility, value generation, and reduced costs. The necessary technologies for this transformation are already available, making it an achievable goal. We should modernise our SAP integration and it's worth it... then next chapter shows why.

Opportunities in SAP Data Integration

Reducing Complexity and Increasing Efficiency

The traditional setup of SAP data integration often involves multiple, redundant interfaces. Traditionally, SAP data integration setups often suffer from having numerous redundant interfaces, which complicate the overall system architecture and elevate maintenance costs. By shifting from multiple, fragmented SAP access interfaces to a more streamlined model, organisations can significantly simplify their data integration landscape. Modernisation is achieved through abstraction. 1:1 connections require a lot of effort, as a connector for each case needs to be developed and maintained. However, by using connectors that be can fed with different parameters depending on the use case, several cases can be handled with one connector.

As a result, there are fewer interfaces and changes and maintenance can be done more efficiently – thus saving man power and cutting costs. This approach encapsulates complexity within the SAP systems, making them more robust and adaptable to changes. Furthermore, eliminating these redundant interfaces not only reduces the overhead associated with managing them but also leads to substantial cost savings. The streamlined setup allows for more efficient and manageable data integration processes, improving overall system efficiency.

Improving Data Availability and Accessibility

The adoption of advanced data platforms that offer high-performance, low-cost data processing and analysis at scale can greatly enhance data availability. This is particularly advantageous compared to relying on multiple, specialised tools for different integration tasks such as ETL, data preparation, enterprise application integration, and data orchestration. Tools like SAP Data Intelligence and SAP Data Services facilitate unified data management, enabling seamless handling of data pipelines across both SAP and non-SAP systems.

Enhancing API Access and System Integration

Open interfaces (APIs) play a vital role in providing access to specific data points, thus facilitating seamless system integration and enabling comprehensive business process orchestration and automation. By introducing additional APIs that encapsulate SAP logic under a user-friendly interface, businesses can significantly broaden access to SAP data and metadata. This enhanced accessibility allows for better context and meaning to be attached to the data, enabling both humans and machines to understand and utilise it more effectively. This is also a prerequisite for AI and ML. The result is improved business agility, support for a broader range of use cases, increased trust in the data, and the establishment of a more robust data foundation.

Recommendations for Modernising SAP Data Integration

To modernise SAP data integration, businesses should focus on standardising their data provisioning and utilisation processes, ensuring a seamless and efficient flow of information across systems. This involves conducting a detailed assessment of the existing data integration landscape to identify inefficiencies and potential areas for improvement.

Additionally, by offloading data to analytical systems like data lakes or data warehouses, organisations can reduce costs and improve processing speeds and create the basis for implementing further analytical use cases by combining SAP and non-SAP data, which can be made available openly to other tools. Implementing continuous data integration, where feasible, further enhances real-time data availability, supporting more dynamic and responsive business operations. The following recommendations outline key strategies for achieving these goals and leveraging advanced integration capabilities to streamline and enhance SAP data management.

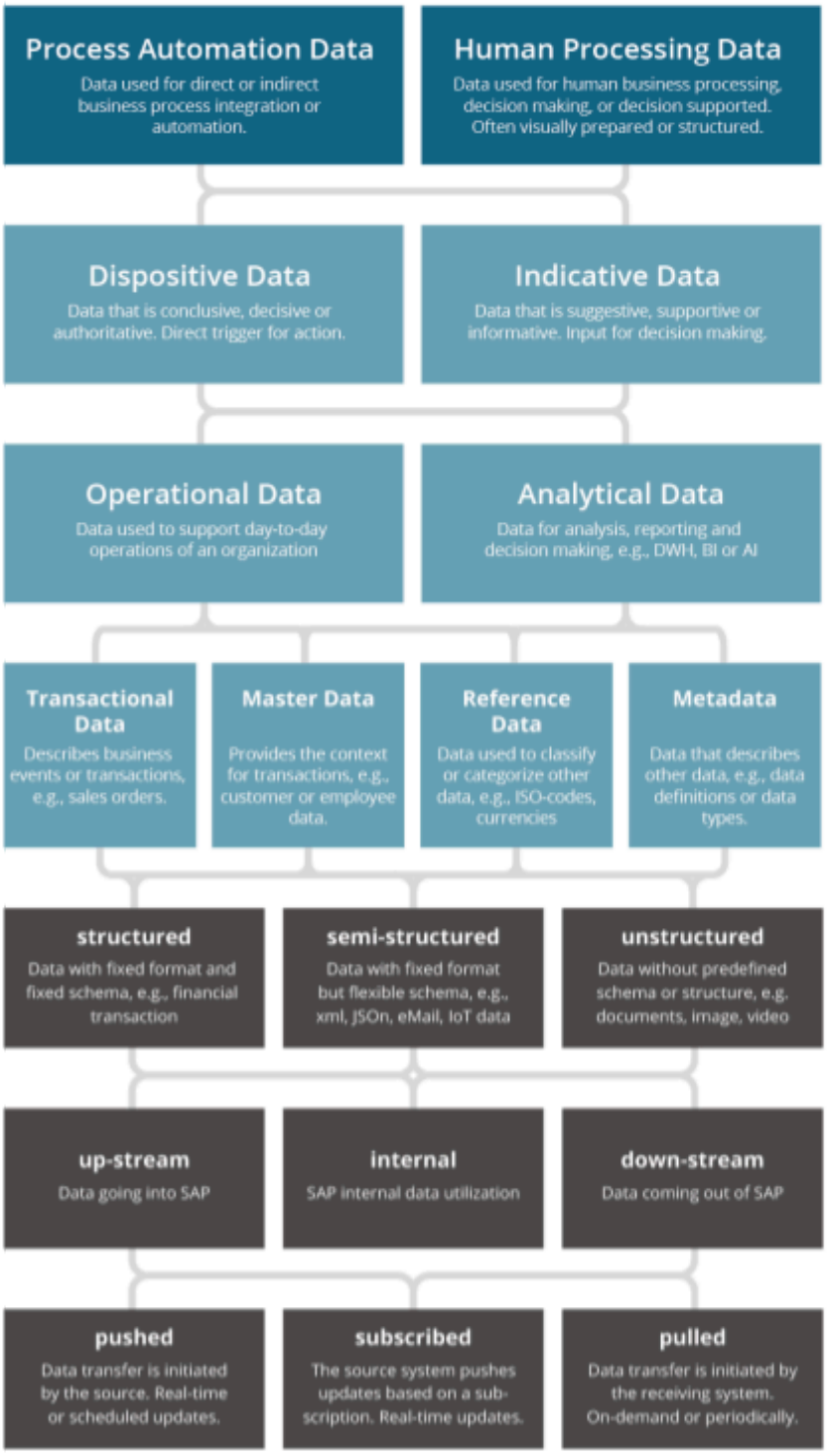
To fully leverage SAP data integration, consider the following recommendations:

- **Standardise Data Provisioning and Utilisation**: Create a clear, standardised connection between data provisioning and utilisation to minimise complexity and cost.
- **Assess and Modernise Existing Infrastructure**: Conduct a thorough assessment of the current SAP data integration landscape and identify areas for improvement.
- **Offload Data to Specialised Systems**: Moving data to specialised management systems can reduce costs and improve speed. Consider using data lakes or data warehouses for downstream applications.
- **Implement Continuous Data Integration**: Where possible, shift from batch processing to continuous data integration for real-time data availability.
- **Streamline and Eliminate Redundant Interfaces**: Identify and eliminate redundant interfaces to reduce system strain and costs.
- Leverage Modern Data Integration Capabilities: Utilise advanced features of BI systems and dedicated data integration platforms to enhance data lineage, security, and governance. Also consider modern non-SAP tools in the evaluation. Depending on the intended use, these can help to significantly improve SAP data integration and make it more efficient.

BARC

Data Type Classifications Relevant For SAP Data Integration

All directly supported by SAP through various data integration technologies and methodologies.



Modern business processes are increasingly distributed, automated and executed in real-time. In addition, business process expertise is too valuable to be spent on the execution of processes. **The future goals of data management are to feed process automation** and to provide information to humans to **shape and manage business processes**, ideally not run them.

The goal of modern data management is to **generate direct input for decision making**. For machines as for humans, the more decisive data is, e.g., a simple yes or no, the more valuable and usable it is. Modern statistics and AI helps to transform indicative data into dispositive data. Nevertheless, indicative data is crucial for context and to enable automation.

The more modern data analysis and AI is utilized for automated or semi-automated decision support and data analysis, the more the logical and physical **distinction between operational and analytical data becomes obsolete** and obstructive.

The rising need for data and business **process automation requires proper Metadata**. Therefore, data integration technologies need to be able to actively manage and use metadata. Metadata leverages to **find, understand, share and protect** data and make it accessible for machines and humans.

Structured data is still the most valuable type of data for analytical purposes and AI. **But the advances in classic and generative AI** will likely make the accessibility and utilization of unstructured data (text, images, video) even more relevant. Modern data integration technologies are going to support both at scale.

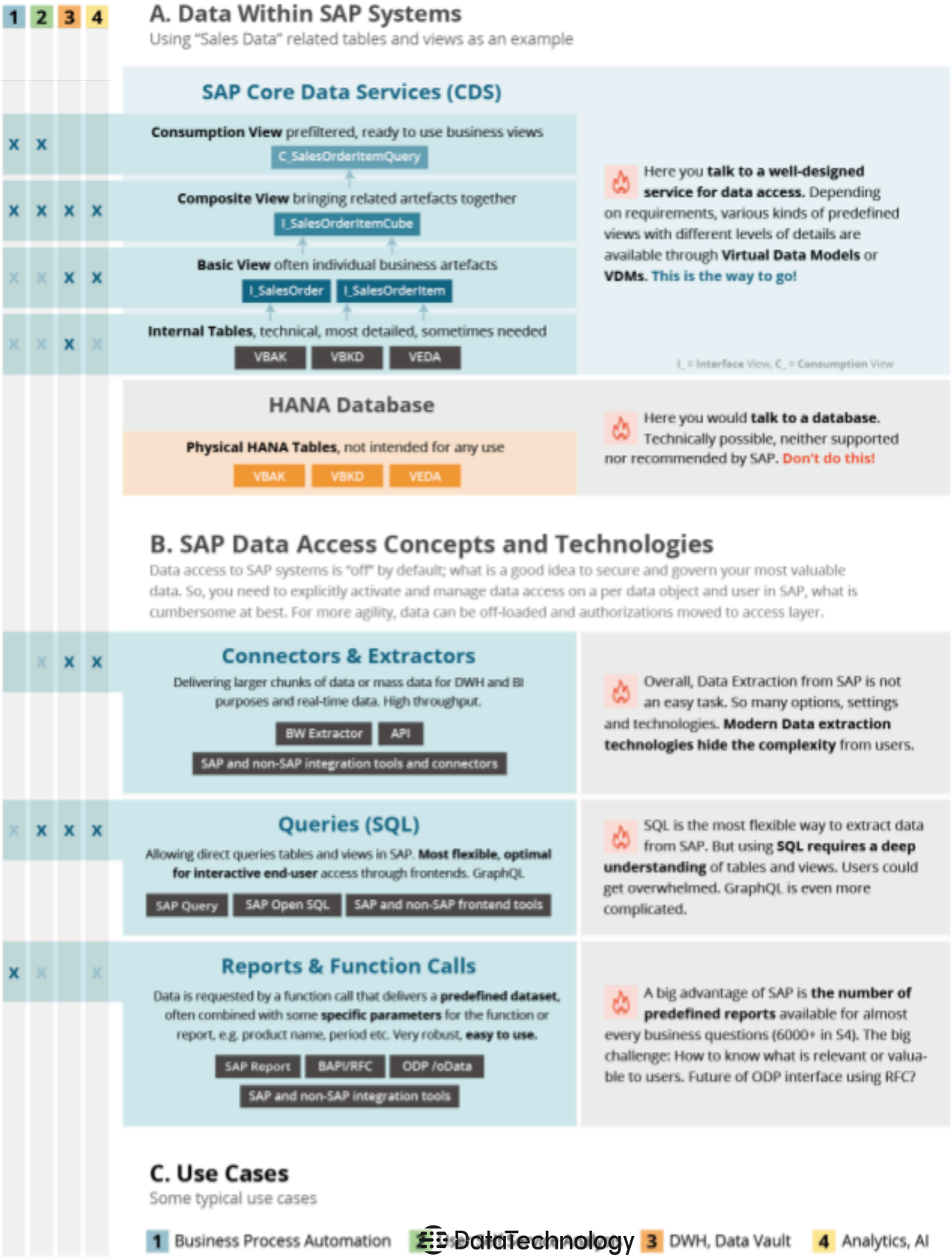
Data needs to be forwarded to as many business processes as possible to leverage its full potential and value. A trend is to simply **down stream all relevant SAP data on a granular level**. Low-cost data storage (e.g. Data Lake) is used increasingly for data offloading. Requires highly efficient and scalable data integration technologies.

Robust connectors & SAP Change Data Capture (CDC) enables constant pushing of data into DWH & BI systems. Offering always-up-to-date analytical data, reduces strain on SAP Systems and better governance.

BARC

SAP S/4 Data Integration Big Picture

With focus on for DWH, BI, analytics and AI use cases.



Final Thoughts

Effective SAP data integration is not just a technical necessity but a strategic imperative for organisations aiming to maximise the value derived from their data assets. As businesses increasingly rely on data to drive decision-making and operational processes, the complexity of managing SAP data can no longer be a bottleneck. By rethinking and optimising data integration strategies, companies can achieve several critical benefits, including significant cost savings, enhanced data accessibility and quality, and improved overall operational efficiency.

The journey towards an modernised SAP data integration landscape begins with a clear understanding of current systems and processes. Organisations must assess their existing infrastructure, identify redundant interfaces, and streamline their data flow to eliminate inefficiencies. This process involves not just technical changes but also a shift in organisational mindset towards valuing data as a key asset. By leveraging modern data integration tools and technologies, such as data lakes and advanced BI systems, businesses can offload data processing tasks, thus reducing the burden on their core SAP systems. This not only cuts costs but also frees up resources, allowing for greater innovation and agility. Moreover, adopting continuous data integration practices ensures that businesses have real-time access to critical data, empowering them to respond swiftly to market changes and emerging opportunities. In conclusion, a well-structured and efficient SAP data integration strategy is essential for organisations to stay competitive, innovate continuously, and make data-driven decisions that propel growth and success.

Terms, Facts & Insights

SAP Data Integration Tools and Interfaces

SAP offers various technologies for data integration, each suited for different purposes.
Some of the key tools include:

- **SAP Data Intelligence**: Primarily used for building and orchestrating data pipelines within SAP Business Technology Platform (BTP).
- **SAP Data Services**: Manages ETL processes for SAP and non-SAP systems.
- **SAP HANA Smart Data Integration**: Handles data pipelines within SAP HANA.
- **SAP Integration Suite**: Integrates on-premise and cloud applications in BTP.
- **SAP Process Integration / Orchestration**: Supports service-oriented architectures.
- **SAP System Landscape Transformation Replication Server**: Facilitates real-time data replication from SAP or third-party systems into SAP HANA.

SAP also provides various interfaces such as RFC, BAPI, IDoc, and CDS Views, which can be accessed through protocols like oData, SOAP, REST, and HTTP/S. The choice of interface depends on the specific use case, spanning batch to real-time scenarios, and the performance and complexity of required transformations. This diversity of tools and interfaces is a significant driver of complexity in SAP Data Integration, alongside the enormous data model complexity in the SAP source systems themselves.

Glossary

Data Lakehouse
A data lakehouse is an architectural concept that combines a data warehouse with a data lake. Raw data is stored permanently and cost-effectively in the data lake. Only structured data that is required for certain analyses is loaded into the data warehouse. The data lakehouse concept is now recommended by all cloud hyperscalers with different names.

Data Warehouse
A Data Warehouse (DWH), or Enterprise Data Warehouse (EDW), is a central data repository for reporting and analysing integrated data from multiple sources. It stores current and historical data to support BI applications like reporting and OLAP. DWHs may include Data Marts and extract data from operational systems such as marketing and sales. While DWHs can be used as data sources for other systems, they should not be misused as data hubs, as varying service levels can complicate adjustments.

SAP data cataloging & lineage
Context information about data is needed to find, understand, analyse and protect data for both humans and machines. Catalogs help to inventory disparate SAP data and make it much more intuitive to navigate. It can provide useful information that helps to get a better understanding about data provenence, applied logic and rules, usage of the data, dependencies and also quality of information. This information help to streamline different use cases from migration (e.g. understand table dependencies), bi or analytical use cases (find right data or queries to build analytical applications) to governance use cases (e.g. monitor data usage, observe policies, understand data quality). If a catalog encompasses SAP and non-SAP systems metadata it can be a good way to integrate both worlds for better decision making.

SAP Connector
A connector provides performant and easy access to a source system and ideally encapsulate all technical complexity. Why mention it here: Connectors can differ in various ways like performance of data extraction, supported integration styles and extraction methods (e.g. read/write, ACID compliance, support of bulk load or mass data extraction), profiling capabilities, sample extraction, depth of access to source systems data objects, and more. To have a reliable, performant connectors that encapsulate SAP data and provides robustness against changes in source system as well as supports needed integration styles and depth of data access is crucial. Finally user need trusted data, in the right format, in the right time at scale.

ETL/ELT
ETL is a classic data warehouse loading process in which the data is retrieved from the source system (extract), checked and converted in a second step (transform) and loaded into the data warehouse in a final step. In contrast to the ETL process, the data is loaded directly into the data warehouse after extraction. The transformation takes place within the DBMS. Particularly relevant in the context of big data and data lakes as well as data warehouse automation (DWA).

Data Virtualisation
Data virtualisation enables users to access, query, and manipulate data from multiple sources as a unified database without moving the data. It offers advantages such as unified access, real-time integration, and cost savings by reducing physical data storage. Challenges include potential performance issues and complexity in integrating diverse data sources.

A use case example is a financial services company using data virtualisation to combine customer data from CRM, transaction systems, and external credit scores, providing a real-time, 360-degree view of customer profiles for better decision-making and personalised service.

Data Preparation
Data preparation is the data preprocessing for data discovery and data science, typically performed by business users or data scientists in a decentralised manner. It primarly provides low-code/no-code interfaces with lightweight or well-guided transformation and mapping functionality. It can also support data quality remediation use cases to prepare a data set for further usage. The goal is to provide data for answering specific questions. This distinguishes data preparation from the generic processing in data warehousing (ETL).

Data Streaming
A data stream is a continuous flow of data records of the same type in a fixed chronological order, potentially with varying data input rates. Using data streaming, sensor data, for example, can be stored long-term as a time series at any desired resolution, enabling comprehensive analyses based on synchronised timestamps and comparable events.

Data Ingestion
Data ingestion is the process of collecting and importing data from various sources into a storage system, such as a data warehouse or data lake, that resides typically in the cloud. It involves extracting, transforming, and loading (ETL) data to ensure it is accessible for analysis and decision-making. Data ingestion can be done in real-time or batch processing, and it supports multiple data formats and sources, including databases, files, and streaming data, to provide a unified view of information. It differs from ETL as it is dedicated into loading data into a target system (mainly in the cloud) and as it does not focus on applying complex transformation logic to mass-data.

Data Orchestration
Data pipeline orchestration is the automated process of managing and scheduling data pipelines across multiple systems and tasks. It involves coordinating the sequential execution of data ingestion, transformation, and processing stages to ensure data flows smoothly from source to destination. Orchestration tools provide monitoring, error handling, and dependency management, allowing data engineers to automate repetitive tasks and optimise performance. This ensures data pipelines run efficiently and reliably, supporting timely data delivery for analytics and business intelligence.

Delta Loading
Refers to an essential capability of SAP BW and S4/HANA systems, that supports incremental data loads, which only contain new or changed records since the last load. Prerequisite for data warehousing at scale from within (BW, Datasphere) and/or outside the SAP Ecosystem (BI-Systems, DWH, Data Lakehouse), as full-loads, especially daily, may take much to long (days) and create high system strain from such prolonged workloads.

CDC (Change Data Capture)
Modern pattern for continues data integration, aka “streaming”. Changes, e.g., new postings, are automatically captured by the SAP system/database and immediately forwarded to subsequent processes within or outside an SAP system. This could be for operational or analytical purposes (to trigger some process) or for continuous data extraction into BI or DWH systems. This approach helps to reduce peak system strain on SAP side caused by larger delta or full loads. In addition, some receiving data management systems are able present changes from the underlying SAP system in near-real-time, data is always in sync.

Unlock Your SAP Data

In today's rapidly evolving digital landscape, businesses are increasingly recognising the urgent need to modernise their data platforms to maintain their competitive advantage. Specifically, maximising the value of SAP data is essential for businesses looking to optimise supply chains, modernise finance functions and empower employees with better analytics and insights.

With SAP ECC approaching obsolescence, many organisations are being forced to plan for a future-ready environment and have to make critical decisions on whether to leverage SAP's own cloud solutions or opt for alternatives like Snowflake, Azure, AWS, and others. While no doubt a daunting task, it does provide an opportunity for forward-thinking organisations to design and build a scalable, resilient, and competitive infrastructure that will support their business goals for years to come.

Why Modernise Your Data Platform?

The argument for moving to the cloud is clear. With cloud computing, AI, machine learning, and big data analytics, outdated data platforms can't keep up. Legacy systems often lag in performance and scalability, while modern solutions handle increased workloads effortlessly with faster processing speeds. Staying compliant, secure and reliable is also non-negotiable, and new cloud solutions are built to meet these ever-changing demands. And with no more hefty investments in infrastructure or replacing old servers, the cost savings are obvious.

Businesses are also demanding increasingly innovative uses for their SAP data in order to stay competitive and agile. This includes integrating advanced analytics and AI to derive deeper insights and streamline operations. Some examples include:

Supply Chain Optimisation

Enhanced access to SAP data allows for real-time tracking of inventory levels, supplier performance, and shipment statuses. By leveraging this data, companies can optimise their supply chain operations, reduce lead times, and respond more swiftly to disruptions.

Financial Process Automation

Better access to SAP financial data facilitates the automation of accounts payable and receivable processes. By utilising near real-time data, businesses can automate invoice matching, streamline payment approvals, and enhance cash flow management. This not only reduces manual errors and administrative costs but also improves financial forecasting and decision-making capabilities, providing a clearer picture of the company's financial health.

Human Resources Management

Improved data access in SAP can significantly impact HR processes, such as talent management and employee performance tracking. By integrating SAP with advanced analytics, HR departments can gain insights into employee productivity, turnover rates, and training needs. This enables data-driven decisions for hiring, employee development, and retention strategies, ultimately fostering a more engaged and efficient workforce. In order to implement these types of initiatives, businesses require rapid access to a mix of SAP and other business-critical, reliable data sources, enabling near real-time decision-making and fostering a data-driven culture.

The Path Forward

While it's evident that migrating to the cloud is essential, determining the best path forward isn't as straightforward, especially with the impending need to transition from SAP ECC. By the end of 2027, free maintenance support for SAP ECC will end, with extended support available only at a high cost until 2030, after which all support ceases.

Your organisation might already be progressing with a SAP modernisation effort, moving to S/4HANA, SAP's latest ERP suite. Yet, according to Gartner, nearly half of SAP ECC users won't have upgraded to S/4HANA by 2027. Whether you're just beginning your transformation, in the middle of it, or already using S/4HANA, a vital aspect to consider is how to most effectively manage your SAP data - particularly, selecting the most appropriate data warehousing, ingestion, governance and analytics tools.

As part of this consideration process, you should be asking:

- What computational elasticity, variety, and compute capabilities are required?
- Is seamless integration and data transfer between SAP and other systems required?
- Is support for unstructured data required?
- Is there a requirement to share data with external third parties via APIs?
- What security and governance capabilities are required?
- Is near real-time access to data important?
- Is low-code data pipeline automation required?
- Has the cost of hardware, licensing, and specialised talent for running SAP systems been considered?

Unfortunately, SAP's cloud solutions fall short in many of these areas, including computational elasticity, data variety, and integration with non-SAP systems. Additionally, SAP's high costs, limited self-service capabilities, and lack of advanced cloud-based data sharing make it less competitive compared to modern, cloud-native solutions.

Given these limitations, a more considered strategy involving different technologies and integrations is essential. Data warehousing solutions like AWS, Azure, and Snowflake provide greater flexibility and scalability versus SAP's own cloud platforms. These cloud-native platforms offer robust computational capabilities, seamless integration with various data sources, and advanced data-sharing features.

By implementing these solutions with the right supporting infrastructure, businesses can enhance performance, reduce costs, and avoid vendor lock-in. This approach not only future-proofs operations but also ensures that companies can adapt quickly to changing demands and leverage the full potential of their data. To fully capitalise on these advanced data warehousing solutions, it's crucial to establish a supporting infrastructure that ensures efficient data ingestion, robust governance, and powerful analytics capabilities.

Implementing a comprehensive data value chain begins with reliable data ingestion processes capable of handling both structured and unstructured data from various sources in near real-time. Effective governance frameworks must be put in place to maintain data quality, ensure compliance, and safeguard sensitive information through advanced security measures. Finally, leveraging sophisticated analytics tools will enable your organisation to derive actionable insights, drive strategic decision-making, and foster a data-driven culture.

How can Data Technology help?

With over 30 years of experience, particularly in working with SAP data, Data Technology is a long-standing partner with Qlik, boasting a proven track record of driving data transformation. We help businesses migrate their SAP data onto the major hyperscaler platforms, such as Microsoft, AWS and Google, where it becomes connected, readily available and analytics-ready.

By taking a systems-agnostic approach, we align business problems to the right technical solutions and then pair them with the right toolsets. Our tailored implementation and managed support services help to make sure our clients benefit from the convenience, security and reliability afforded by cloud solutions.

Data Technology will work with your business to:

- Create clear vision and goals for data management and use, removing data silos and streamlining data collection and sharing.
- Create the foundations for analytics and predictive capabilities that will optimise your business processes
- Build a best-in-class data culture with world-leading governance, security and compliance.
- Choose the right cloud partner that matches your business requirements and satisfies any cost constraints.

We work with our technology partners, such as Qlik, through the implementation phase to:

- Migrate key finance and supply chain SAP processes to the cloud – like order to cash and inventory management – at speed and with confidence via our Accelerators
- Enable near real-time data ingestion (CDC) from SAP to your preferred cloud platform
- Replace ELT scripting with data pipeline automation, accelerating data replication and orchestration by over 90%
- Reduce deployment cost and managing on-premise SAP infrastructure, plus reduce your TCO by up to 40%

Speak to one of our experts: https://datatechnology.co.uk/contact