



SAP HANA | PUBLIC

Manage Data Access, Integration, and Quality Across Your Digital Enterprise with SAP HANA®

Table of Contents

-
- 3 Simplify Information Management to Improve Analytics and Operations**

 - 5 Use Cases for SAP HANA Smart Data Integration and SAP HANA Smart Data Quality**

 - 7 Using SAP HANA Smart Data Integration for Data Provisioning**

 - 9 Using SAP HANA Smart Data Quality to Deliver Data You Can Trust**

 - 11 Gaining Powerful New Data Quality and Data Integration Functionality**

 - 12 Tallying the Benefits of Data Management with SAP HANA**

Simplify Information Management to Improve Analytics and Operations

You are immersed today in vast, rapidly increasing amounts of data from internal and external sources. Keep afloat by taking advantage of the integrated data quality, data integration, and data virtualization capabilities that SAP HANA® provides for in-memory computing. **Help simplify your IT landscape and enable successful information management** for virtually all your analytical and operational implementations.

DELIVERING TRUSTED, COMPLETE, AND RELEVANT INFORMATION

Digital transformation is a top priority in the 21st century. But how exactly do you become a digital enterprise? At SAP, we designed SAP HANA to help you harness the latest digital technologies in the service of these overarching goals:

- Faster time to insight and action with real-time analytics
- Improved operational efficiency and productivity
- Agility to respond to rapidly changing market conditions
- Simplification of the IT landscape with integrated capabilities

Managing your valuable information successfully is the key to thriving in today's business environment. With the data access, integration, and quality capabilities of SAP HANA, you can deliver immediate targeted insight, allowing stakeholders to act in the moment and make quick decisions based on complete, timely, and accurate data. You can also streamline processes to make your operations

more efficient and effective – from your inner core to the edge of your global business network, including cloud platforms, mobile devices, and the Internet of Things. And you can seize fleeting moments of opportunity, reimagine long-standing business models, and create radical new sources of value that propel you to a leadership position in your industry.

LEARNING TO RUN SIMPLER

If you're like many other companies, you're challenged to merely survive today's business challenges with an inherited web of solutions that maintain your data in numerous silos. That kind of complexity slows progress toward achieving better business results – stunting growth, reducing profits, and increasing risk. And in the information management realm, if you bring your current data problems with you, the chance to simplify and realize the long-term benefits of digital transformation is lost.

STRIVING FOR INFORMATION EXCELLENCE

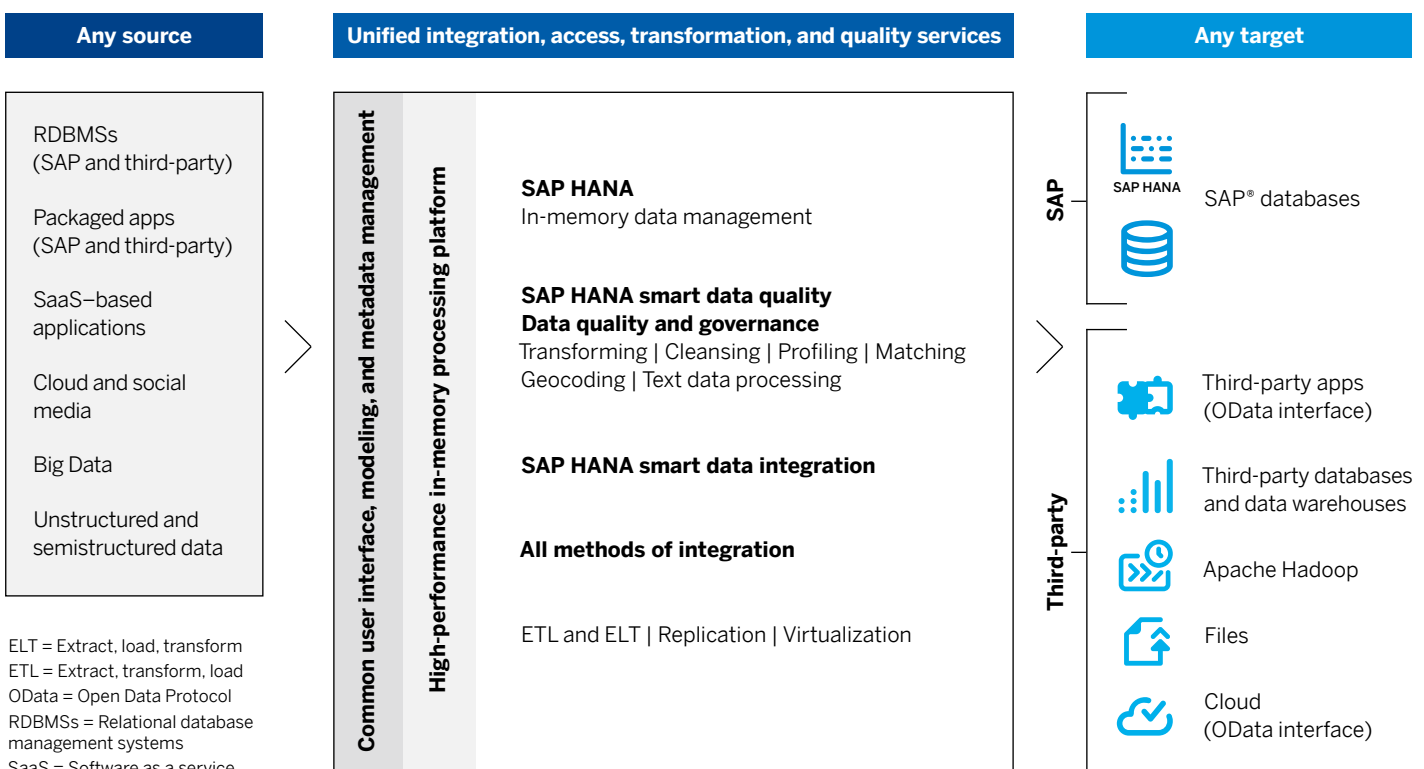
With SAP HANA, you no longer need to rely on a bundle of stand-alone data management solutions. You can radically simplify your IT landscape by using a consolidated set of data management capabilities offered in a **single unified database**, as shown in Figure 1.

SAP HANA offers native real-time data replication; data virtualization; bulk and batch data movement through extract, transform, load (ETL) and extract, load, transform (ELT); data transformation; and data quality management. It provides all this functionality in a **single in-memory computing database** that consolidates data sources and accelerates data-related tasks exponentially. It dramatically reduces the number of moving parts in your information management landscape and lets you build processes once and deploy them multiple times – on premise and in the cloud.

As part of this revolutionary simplification, you can shorten the journey to information excellence and achieve the following goals:

- Acquire information from virtually any source using practically any style of data integration, including bulk and batch data movement (ETL and ELT), virtualization, and real-time replication
- Deliver the information that's fit for use in your business processes and analytics by improving data quality
- Keep information accessible and trusted for decision-making by improving data integration, making data available when and where it's needed

Figure 1: SAP HANA with Native Data Access, Integration, and Quality



Use Cases for SAP HANA Smart Data Integration and SAP HANA Smart Data Quality

SAP HANA provides both smart data integration and smart data quality software. This section describes three typical use cases for these capabilities.

DATA INTEGRATION

Few, if any, computer systems in major organizations stand in isolation. As businesses transition to the digital economy and rely on initiatives involving Internet of Things, cloud and hybrid landscapes, and business value networks, the ability to move and integrate data efficiently, securely, and reliably becomes increasingly important.

There are three major styles of data integration: bulk and batch, real-time replication, and data virtualization. SAP HANA smart data integration supports all three, making it the ideal choice for data integration tasks in SAP HANA-centric landscapes. By handling both initial loading and real-time data capture in a single flow graph, the smart data integration option simplifies the process of performing these data integration jobs.

ANALYTICS

Data quality is extremely important in analytic applications. If you are basing your analysis on faulty, incomplete, or incorrect data, you cannot get a clear view of your business.

Analytics have traditionally been used for historical reporting – that is, to tell you what happened last month or last week. With today's real-time analytics, however, this historical view sheds clear light on what happened just a few seconds ago. What's more, companies increasingly depend on data for predictive analysis (to determine what will happen) or prescriptive analysis (to decide what you should do about it). As a result, data quality is even more important, as divining the future from an incorrect picture of the past and present leads you down the wrong path.

SAP S/4HANA® breaks new ground by allowing you to run analytics on the same copy of data used for transactional processing. In the past, if you had poor-quality data in your SAP® ERP application, you could at least cleanse it on the way to the data warehouse to improve your analytics. But with just one copy of the data, all the information had better be correct from the start (or corrected as quickly as possible before negatively affecting business). SAP HANA smart data quality software is designed expressly to cleanse data for analytics within SAP HANA.



Gain **open and extensible support** for virtually any data volume, data type, and data source with SAP HANA smart data integration.

CUSTOM DEVELOPMENT

Perhaps the most exciting thing about SAP HANA smart data integration and SAP HANA smart data quality is that they're created from the ground up as **built-in database services in SAP HANA**. As a result, developers are now able to build data integration and data quality directly into their applications. They can both ingest and improve data using the same development tools that they use to build SAP HANA applications. This is an entirely new breakthrough in the development world. Imagine developing a custom analytic application, but rather than hoping that users will deduplicate the data, instead actually taking care of that within your application.

Consider how much better and more accurate your application would be with this approach. Developers think of application development and integration development as separate domains, but with SAP HANA, they're now merging into a single set of competencies. And developers can perform them both with an intuitive and unified tool set that allows them to focus on what they want to do without worrying about how to do it.



Deliver **immediate and targeted insight** with the powerful capabilities in SAP HANA for managing data access, integration, and quality.

Using SAP HANA Smart Data Integration for Data Provisioning

Data provisioning is the process of providing data to users (and to applications) in an orderly and secure way. In the context of this section, think of data provisioning as simply getting data into the SAP HANA database from external sources as well as moving it out to other targets.

SAP HANA smart data integration consists of built-in bidirectional data movement capabilities that allow the database to ingest data from a wide variety of sources and move data out to a large number of targets. You can move data using bulk and batch or real-time replication, or you can simply virtualize the data. Data virtualization is particularly useful if you want to create a quick proof of concept or construct test scenarios to prove the value of certain data before fleshing out permanent data movement processes. It is also valuable for accessing and exploring Big Data sources.

SAP HANA smart data integration is available preconfigured with a host of built-in adapters for common data sources, including SAP and third-party databases, Big Data from applications and transactions, text documents, e-mail, and instant messaging. You can easily customize adapters with either on-premise or Web-based design tools and develop custom adapters with an open software development kit.

APPLYING MULTIPLE STYLES OF DATA PROVISIONING IN SAP HANA

You may need to use several different styles of data provisioning in one SAP HANA system. That's why SAP HANA smart data integration supports three of the major methods of data provisioning, as described below.

Bulk and Batch Data Integration

Bulk and batch data integration is used to move large quantities of data in a single job (or in multiple related jobs). Even for real-time replication projects, the first step is usually a bulk load.

Bulk data loading typically involves transformation, the process of changing the data from the source format to the target format. This can include structural transformation (for example, converting field names, joins, data type changes, and so on) and functional conversions (such as value mappings and lookups).

Real-Time Replication

SAP HANA smart data integration supports both single-table and mass-data replication. It lets you feed the process with an initial load or with the real-time capture of change data. You can set configuration options to add, edit, or remove target columns and filter for the records you need. You can manage load behavior to identify changed records for consuming applications or create a history table. And you can also create logical partitions to decrease initial load times.

Classically, replication moves data as is from one system to another. But SAP HANA smart data integration combines replication with transformation to change the data while in transit from the source system to the target system. You can also use the replication functionality for high availability and disaster recovery (HA/DR) use cases to provide ready back-up for load balancing or in case of system failure. In addition, you can use replication when data needs to be current in two separate systems.

SAP HANA smart data integration combines the real-time change data capture from SAP and third-party sources with complex data transformation capabilities. You can turn raw operational data into enriched information that fuels real-time analytic and governance applications.

Depending on the source, you can run replication as a change data capture (CDC) process, in which changes made to a table are automatically replicated to the target system.

Data Virtualization

In virtualization, data does not physically move from the source system(s) to the target system(s). Rather, virtualization presents a unified view of data. Updates to data are effected on the source system without changing the target system. Data virtualization is provided by SAP HANA smart data access technology, which is incorporated into SAP HANA smart data integration.



Using SAP HANA Smart Data Quality to Deliver Data You Can Trust

SAP HANA smart data quality lets you simplify data quality operations by parsing, standardizing, validating, correcting, and enhancing individual, company, and address data in a single transformation. These capabilities help you gain and maintain confidence in the data that your organization uses to run core processes and make strategic decisions.

PRESERVING AND PROTECTING DATA QUALITY

All organizations have data quality problems. Duplicate customer master records, incomplete and incorrect material data, incorrect vendor addresses, and many more quality issues affect profitability, customer satisfaction, marketing and production costs, returns, regulatory compliance, and operational efficiency.

Data cleansing is the process of removing duplicate records, creating a best-quality “golden record,” correcting addresses (with the help of regularly

updated postal directories), and enriching data with geocoding or bureau data (such as that provided by Dun & Bradstreet). By so doing, you can ensure that data is correct and complete. With SAP HANA smart data quality, we deliver unified cleansing, enrichment, deduplication, matching, geocoding, and other data management capabilities as callable services through one user-friendly interface. You can:

- Cleanse individual, company, and address data
- Identify duplicates within a single source or across multiple sources of data
- Select a best record from a group of duplicates
- Enrich address data with geocode information
- Search for addresses by latitude and longitude

You can further support geodependent business operations by identifying nearby addresses and returning U.S. census data where applicable.



SAP HANA offers native real-time data replication, data virtualization, bulk and batch data movement through ETL and ELT, data transformation, and data quality management on a **single in-memory computing database**.

DESIGNING DATA FLOWS

With SAP HANA smart data integration and SAP HANA smart data quality, you can design simpler or more-complex data flows. Available structured query language (SQL) transformations support basic aggregate, filter, join, sort, and union processes, along with advanced case, lookup, pivot, and unpivot operations. Data lifecycle management capabilities include data generation, history-preserving mapping, row generation, and table comparison. You can cleanse, match, and geocode data to preserve and enhance quality, and you can execute programs in the leading data management scripting languages. Best of all, enabling data flows for real time is as straightforward as checking a box on a screen.

TAKING ADVANTAGE OF ITS REST API

The REST API for SAP HANA smart data integration is a RESTful API built for supporting execution and monitoring of SAP HANA smart data integration and SAP HANA smart data quality tasks. Developers and integrators can use the REST API for SAP HANA smart data integration to programmatically execute and monitor flow graphs, to process data for interactive data transformation within consuming applications, and to create, modify, and delete

virtual tables. In order to use the REST API, design time object flow graphs need to be created, saved, and activated in the Web-based development workbench. Upon activation, the software automatically creates a runtime object called “Task,” which can be used by the REST API to execute and monitor. The REST API uses the term “Task” to represent the concept or object that is being executed in the workbench based on a particular flow graph model. Since the REST API for SAP HANA smart data integration is exposed as a REST API, the client application does not need to be running on the engine for SAP HANA extended application services.

REVERSE GEOCODING IN SAP HANA

A reverse-geocoding function in SAP HANA smart data quality can pinpoint a location with latitude and longitude and display the addresses within a radius of that point. This is particularly useful in scenarios such as targeting marketing campaigns in mobile apps or determining the households at risk in a natural disaster or public health crisis.

Gaining Powerful New Data Quality and Data Integration Functionality

At SAP, we understand that the data in your SAP Customer Relationship Management, SAP ERP, and SAP Master Data Governance applications – as well as in your SAP S/4HANA landscape – is vital to making sound business decisions. By enhancing the data integrity within your SAP solutions, data quality helps to establish and maintain trust in your data. To achieve this goal, it is very important to tackle on-entry validation, duplicate prevention, and periodic data cleansing to meet accuracy and completeness requirements.

The version of SAP Data Quality Management software that is available for SAP solutions delivers prepackaged capabilities to enforce data accuracy for customer, vendor, and business partner data with a smoother look and feel to your SAP software landscape. This integration brings the benefits of SAP HANA smart data quality directly into your

SAP Business Suite applications and provides immediate value without the need for custom development. The add-on provides prebuilt data quality configurations that support real-time address validation, geocoding, batch data cleansing, and batch duplicate record detection.

SAP Master Data Governance provides built-in, domain-specific governance functionality to centrally create, change, distribute, or consolidate master data across enterprise system landscapes. It uses SAP HANA smart data quality and SAP HANA smart data integration to integrate and load master data from many different sources and to standardize, validate, and enrich address data. It also matches data to detect duplicates based on customer-specific matching rules and calculates best records.



Turn raw operational data into **enriched information** that fuels real-time analytic and governance applications.

Tallying the Benefits of Data Management with SAP HANA

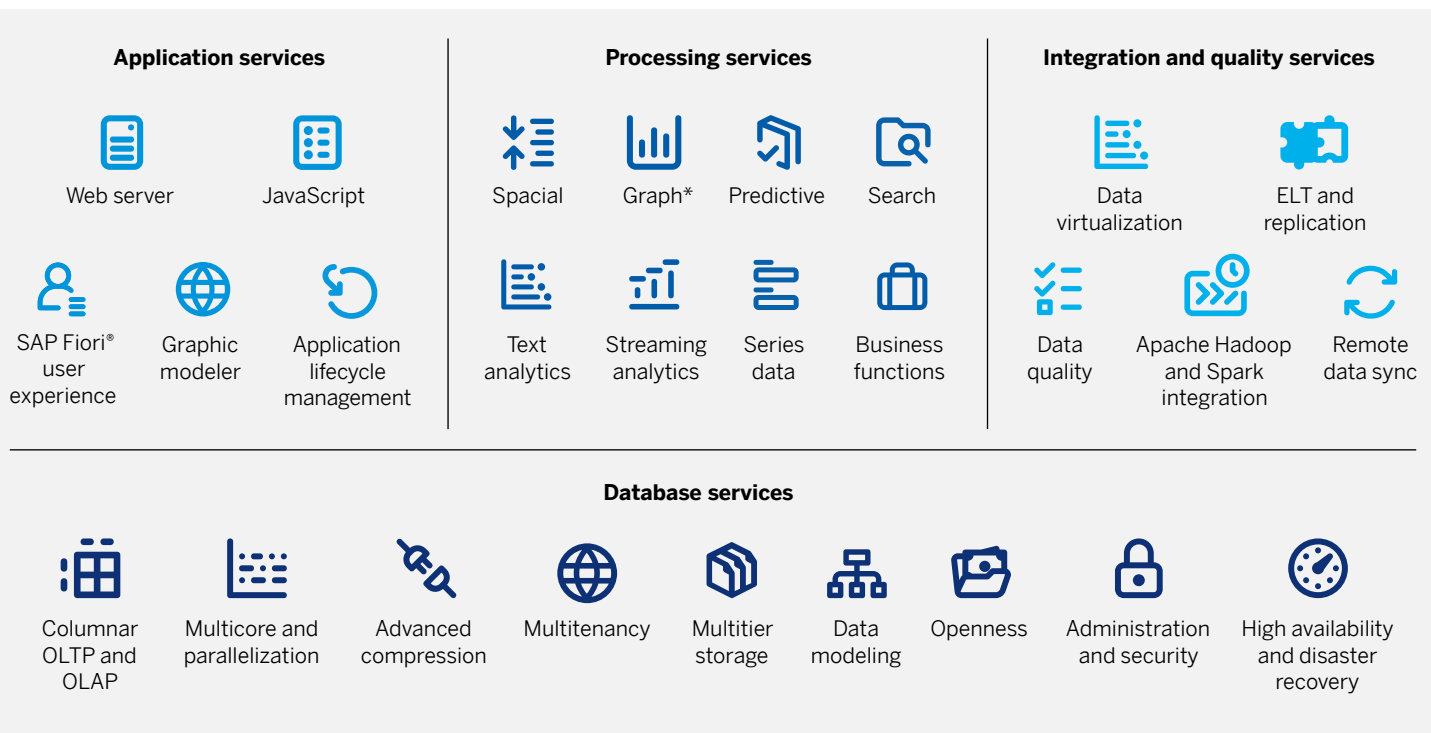
The data management capabilities native to SAP HANA help you realize a 21st-century framework for managing your valuable information assets to help ensure information excellence for your digital enterprise. SAP HANA simplifies your IT landscape and speeds operations and analytics by combining transaction and application processing on a single in-memory database. It offers you revolutionary power for smart data integration, smart data quality, and smart data access.

SAP HANA smart data integration provides open and extensible support for virtually any data volume, data type, and data source. It delivers structured,

text, social, and spatial data from SAP and third-party applications on premise, on device, or in cloud or hybrid environments. SAP HANA smart data quality helps you unify, cleanse, enrich, match, and geocode data through callable services with a single user interface. SAP HANA smart data access lets you access remote data virtually, without copying and storing it in SAP HANA.

As Figure 2 shows, SAP HANA can be the single database that supports the functionality you need to run a successful modern business.

Figure 2: SAP HANA – A Comprehensive Database to Support Your Digital Enterprise



ELT = Extract, load, transform
 OLAP = Online analytical processing
 OLTP = Online transaction processing

*Graph is in controlled availability.

Because you can see where data comes from, what transformations it has undergone, and how well it conforms to your in-house standards, you increase trust in your data. And higher data quality and deeper data context help you improve the accuracy of modeling, planning, and analysis. The bottom line is that you increase everybody's productivity and streamline business processes, ultimately lowering operational costs and driving revenues.

LEARN MORE

For more on the features and benefits of the data management capabilities in SAP HANA, talk to your SAP representative, or visit us [online](#).



