SAP Solution in DetailSAP Solutions for Enterprise Information Management
SAP Data Services

Deliver Trusted Data for Accurate Analytics



Table of Contents

- 3 Quick Facts
- 4 Gain a Solid Foundation for Data Management

Establish an Information Infrastructure

- 5 Extract New Insights from Big Data with Accurate Analytics
- Gain a Complete View of Information

Access and Integrate Data

Get Unprecedented Insight from Big Data

Take Advantage of Apache Hadoop Process Text Data

8 Trust Your Information

Manage Data Quality
Profile Data and Manage Metadata

Satisfy the Information Governance Needs of Both Business and IT Users

Support for the Business User Support for IT

Manage Data the Way You Want with Greater Simplicity and Ease

Scale to Handle Very Small and Very Large Data Volumes

Reuse As Often As Possible

Optimize Integration with SAP Software and Data Sources

Integrate Structured and Unstructured Data

Optimize Implementation with SAP Services

Complementary SAP Solutions for EIM

Integrate, Transform, and Improve Your Data

Find Out More

Quick Facts

Summary

Today's data explosion forces you to remove inefficiencies from data management processes. SAP® Data Services software, along with SAP Information Steward software, helps integrate, transform, and improve your data enterprisewide, letting you cost-effectively deliver trusted information to all critical business functions. As a result, you can optimize operations, boost data quality, and gain essential access to disparate sources and applications.

Business Challenges

- Manage increasing volumes of data and numbers of applications
- Minimize application silos to enable greater collaboration across teams
- Reduce the time users spend locating data and assessing its quality
- Free IT staff from system integration tasks and encourage innovation

Key Features

- Data integration and text data processing – Access and integrate structured and unstructured data from data sources across your enterprise for deeper insights
- Data quality Make more-informed decisions and improve business process efficiencies with accurate, consistent, and complete data
- Data profiling and metadata management Understand and analyze the trustworthiness of your enterprise information and get continuous insight into the quality of your data

Business Benefits

- Streamlined business processes through a single solution to move, improve, govern, and unlock the value of data
- Increased staff productivity through a shorter learning curve and powerful data transformation tools
- Tighter cross-functional collaboration through an intuitive interface that lets business users contribute to defining data standards
- Improved quality of application data by integrating input from a wide variety of sources
- Lower costs by reducing customization and downstream data cleanup

For More Information

Call your SAP representative, or visit us online at www.sap.com/eim.

Gain a Solid Foundation for Data Management

Data is the life of your enterprise, and the health of your business depends on using trusted, accurate, and complete data. With SAP® Data Services software for data integration, data quality, and text data processing and SAP Information Steward software for data profiling and metadata management, you can gain insight across your entire business – where and when needed. As a result, you can meet the needs of "Big Data" analytics and information governance initiatives more effectively and simply.

ESTABLISH AN INFORMATION INFRASTRUCTURE

A lack of integration between lines of business and IT can cause information to be fragmented – or inaccurate. This can lead to difficulties around data quality and consistency, along with issues regarding global compliance initiatives. In fact, poor data quality is the primary reason that 40% of the anticipated value of all business initiatives is never achieved.¹

In this era of Big Data analytics and information governance, you need to take action to address these issues – before they negatively impact your business processes and decisions. The bottom line: if you want to make meaningful use of your data, you need meaningful data.

What you need to tackle these challenges is a modern information management infrastructure that can help you harness and unlock the value of your data in real time. Gaining this infrastructure is critical to your success. Through 2015, organizations that integrate high-value, diverse, and new information types and sources into a coherent information

management infrastructure are expected to financially outperform their industry peers by more than 20%.²

SAP Data Services offers a comprehensive data management foundation. It lets you boost productivity and cut costs with a single solution for data integration, data quality, and text data processing. The integration of SAP Information Steward adds data profiling and metadata management functionality.

SAP Data Services empowers you to handle all types and varieties of data from any source. And it does this in real time, allowing you to prevent data quality issues instead of fixing them after the fact. The software helps move, improve, and govern your data – letting you cost-effectively deliver trusted information to all critical business functions. Your organization can maximize its return on data by using SAP Data Services to:

- Access relevant data regardless of data type, data domain, or data source
- Deliver complete, accurate, and trusted data
- Empower both IT and business users to govern the quality of data

SAP® Data Services software is developed and designed to best meet the needs of Big Data analytics and information governance initiatives – all with the goal of simplicity of use. It offers intuitive user interfaces and provides enough flexibility to evolve with your business as it grows over time.

FOOTNOTES

- **1.** Gartner Inc., *Measuring the Business Value of Data Quality*, October 10, 2011.
- **2.** Gartner Inc., *Predicts 2011: Enterprise Information Management Programs Connect IT and Business*, November 25, 2010.

Extract New Insights from Big Data with Accurate Analytics

Today's organizations are creating, collecting, and managing massive amounts of data. And the volume grows each day as Internet sites, social media venues, online forums, and other sources generate data at faster and faster rates. Successful businesses demand the ability to work with, understand, and manage these data sources at the speed of a real-time business application. SAP Data Services is ready to serve those needs.

SAP Data Services enables accurate analytics for Big Data by offering a flexible, open, real-time solution that can help you:

- Gain a complete view of information by accessing and integrating data at any scale and from any source with high velocity
- Get unprecedented insight from Big Data by extracting useful intelligence from unstructured content and combining it with structured content for new contextual insight
- Trust your information by governing data quality, correcting issues during data movement, and defining policies to know when data is fit for use

With SAP Data Services software, you can cleanse and enrich native data, with direct access to heterogeneous sources and applications, in a single software environment.



Gain a Complete View of Information

SAP Data Services provides you with a big picture of information by making it possible to access and combine any volume of data from any source to gain contextual insight and unlock true meaning. In fact, you can access and integrate both SAP-based and non-SAP-based data sources, as well as structured and unstructured content across your enterprise. You can even move and manage data from analytical and Big Data sources such as social media, as well as machine-generated and geographic data.

ACCESS AND INTEGRATE DATA

SAP Data Services enables you to gain deeper insight with a single, trusted view by making it possible to extract, transform, and load data quickly into a data warehouse to create a complete view. You can also access data from Apache Hadoop to combine unstructured and structured data for new insight, as well as unlock meaning from unstructured documents with native text data processing. Moreover, you can do this with many data stores, including the SAP HANA® database, SAP NetWeaver® Business Warehouse (SAP NetWeaver BW) application. SAP Rapid Marts™ packages, and SAP Sybase® IQ server. What's more, SAP Data Services is natively built into SAP HANA and SAP Sybase IQ.

USE CASE: REAL-TIME MARKETING ANALYSIS

With SAP® Data Services software, you can reduce customer churn by analyzing marketing offers made to prospects and existing customers at retail centers and Web stores or via call centers and text messages. You can integrate this information from a variety of data sources into the SAP HANA® database, which can process billions of records in hours to deliver rapid insight. As a result, you can deliver more targeted offers to customers to improve profitability and boost retention.

THE FIVE CORE FUNCTIONALITIES OF SAP® DATA SERVICES

SAP® Data Services software supports the following functionalities:

Data integration – You can gain deeper insight with a single, trusted view by accessing and integrating structured and unstructured data from data sources across your enterprise.

Data quality management – You can make more-informed decisions and improve business process efficiencies with accurate, consistent, and complete data.

Text data processing – By leveraging unstructured text data alongside regular data, you can gain significant insights into your business, identify emerging trends, and proactively respond to opportunities or potential risks.

Data profiling – SAP Information Steward software is built on and works in conjunction with SAP Data Services – enabling IT and business users to collaborate on data profiling through a special user interface.

Metadata management – IT and business users can also collaborate on metadata management by leveraging the user interface in SAP Information Steward.

Get Unprecedented Insight from Big Data

SAP Data Services makes it possible to gain newfound insight into Big Data. That's because it enables reliable, efficient, and optimized real-time analysis across all enterprise information assets, whether structured or unstructured.

TAKE ADVANTAGE OF APACHE HADOOP

SAP Data Services provides an interface that allows high-performance reading from and loading into Apache Hadoop. From there, SAP Data Services identifies, extracts, structures, and transforms meaningful information discovered in the Hadoop Distributed File System or Apache Hive database. It then provisions the data to SAP HANA, SAP Sybase IQ, or other data stores for deeper analysis.

PROCESS TEXT DATA

With SAP Data Services, you can leverage unstructured text alongside regular data to gain significant insights into your business, identify emerging trends, and proactively respond to opportunities or potential risks. Specifically, you can:

- Analyze linguistics to create base metadata (for example, lexical forms and parts of speech)
- Extract information to create semantic metadata, such as by identifying properties of groups of words (for example, person, location, or date) or by identifying relationships between groups of words
- · Analyze text data in 31 languages

You can process text within files, Web logs, surveys, content fields, social media, geolocation data, and more. And you can do so using the Hadoop Distributed File System, SAP HANA, or SAP Sybase IQ.

USE CASE: CLEAR INSIGHT FROM UNSTRUCTURED DATA

SAP® Data Services software lets you uncover valuable insight from text messages you receive from customers across 31 languages. After key points are extracted from these messages, industry-leading cleansing techniques are used to match the newly parsed text and reply back to the customer. The result? Top-notch customer services that are quick, reliable, and cost-effective.

Your data quality rules are applied whenever data is created, updated, or moved, so you can trust your data to deliver accurate and complete information anywhere, anytime.



Trust Your Information

SAP Data Services enables you to improve data quality for more-effective decision making and business operations. It allows you to correct data issues as they arise and prevent quality issues before they occur. The software prevents and automatically fixes data redundancies, inconsistencies, and errors within any application. It applies your data quality rules whenever data is created, updated, or moved, so you can deliver accurate and complete information anywhere, anytime.

MANAGE DATA QUALITY

With SAP Data Services, you can make more-informed decisions and improve business process efficiencies with accurate, consistent, and complete data. You can use the software to:

- Improve data quality by parsing, standardizing, and cleansing data from any source, domain, or type
- Enhance data with internal or external sources to maximize its value
- Consolidate data to uncover hidden relationships and provide a single version of the truth

Moreover, you can perform any of this in real time or using batch processing. Plus, because the software is built on Web services, you can use its functionality in a wide variety of applications, platforms, and databases – including those from SAP and third parties, along with proprietary software.

To prevent troubles integrating data, SAP Data Services allows you to define and enforce business rules for data standardization, helping to ensure that everyone has the same understanding and insight into the data. You can assess and improve the quality of your data within any industry, locality, or data domain, including customer, product, and material.

The software handles cleansing, fuzzy matching, auto-merge, and survivorship to optimize data quality. Plus, it can

validate and enrich data to enable compliance with company standards. It also offers extensive global data quality coverage with support for over 230 countries. On top of that, you can deploy SAP Data Services in multiple target sources across your enterprise. You can also use different deployment options, such as batch or real-time processing, to fit your needs. For example, you can fix data in motion when loading it into SAP HANA using batch processing. You can also fix data at the source and prevent data problems at point of entry using real-time processing.

PROFILE DATA AND MANAGE METADATA

Using SAP Data Services and its native integration with SAP Information Steward, you can easily trace the source of specific data for information governance. SAP Information Steward is built on and works with SAP Data Services, enabling you to understand and analyze the trustworthiness of your enterprise information and get continuous insight into the quality of your data. You can monitor how well data adheres to governance rules and spot quality trends through functions driven by industry-leading dashboards. By analyzing data levels over time and tracking trends in quality, you increase reliability, leaving your staff free to apply its expertise to initiatives that have greater strategic value. Through built-in data quality dashboards, you see the impact of data quality problems on all downstream systems and applications.

Moreover, SAP Data Services and SAP Information Steward serve as a central repository for data asset analyses, metadata, a business glossary, validation rules, and data profiles. They enable all members of your data management team to work together, eliminating the need to share reports via spreadsheets or import data into applications.

USE CASE: INFORMATION YOU CAN TRUST

With SAP® Data Services software, you can load and cleanse data from very large external data sources – and deliver trustworthy information. Business rules are applied at each data source, so all files are managed consistently. You can then use the SAP HANA® database to rapidly process this large volume of data and glean key insights.

Satisfy the Information Governance Needs of Both Business and IT Users

Information governance empowers business users to serve their customers and accomplish their tasks with the greatest benefit and impact. It also ensures that all deliverables are based on accurate information that can stand the test of time. SAP Data Services supports information governance by offering the core data management foundation that can help you:

- Extend the value of information Improve business process efficiency and accuracy and deliver reliable and relevant information directly into business applications
- Transform information into business value – Consolidate and enhance data across all sources with greater accuracy, so you can deliver trusted information for confident decisions
- Know the value of information –
 Meet governance and compliance
 regulations by measuring, monitoring,
 and managing the lifecycle of all data
 types that are integrated across business and analytic applications

SAP Data Services and SAP Information Steward, both SAP solutions for enterprise information management (EIM), enable all members of your data management team to collaborate. As a result, you eliminate the need for employees to share reports via spreadsheets or import data into applications.

SUPPORT FOR THE BUSINESS USER

Unique support for the business user includes:

- Business-user interface Define standards, profile data, and perform impact analysis using an intuitive business-user interface
- Reporting and metrics Monitor and assess the quality of any data source against business metrics to meet compliance
- Scorecards Integrate data quality scorecards with other analytical

- applications to allow better analysis; expose a scorecard as a URL to any application, and then zoom in to see quality trends and get more granular details
- Metadata management Provide metadata context for data such as definitions, business terms, and use
- Business rule organization Flexibly organize data validation rules based on a data steward's own data categories and hierarchies; modify and add quality dimensions to the rule definitions and scorecards of SAP Information Steward
- Match review user interface –
 Participate in the match review process by leveraging an accessible and intuitive user interface to review match groups, resolve conflicts, and set the master record

SUPPORT FOR IT

SAP Data Services provides a single development environment, metadata repository, data connectivity layer, runtime environment, and management console designed specifically to help locate, categorize, cleanse, and catalog information. You need only to maintain one data management application rather than a jumble of point-to-point solutions. And to save time, you can write completely reusable transformations, as well as customize them if the need arises. Additional support for IT includes:

- Data quality firewall You can prevent and automatically fix data redundancies, inconsistencies, and errors within any application.
- Integration with a range of data sources – To bring users quick access to information, SAP Data Services supports all leading relational database management solutions. These include generic open database connectivity (ODBC) products, IBM DB2 Universal Database (UDB), and offerings from

Hewlett-Packard Development Company, Oracle Corporation, Netezza Corporation, and Teradata Corporation. SAP Data Services also supports fixedwidth and delimited flat files, Microsoft Excel spreadsheets, COBOL copybooks, and XML documents. The software supports HTTP, HTTPS, electronic data interchange, and Simple Network Management Protocol (SNMP) technology standards. And it offers rich connectivity to packaged SAP applications and to Oracle applications such as E-Business Suite, OneWorld, and PeopleSoft.

- Unstructured data SAP Data Services lets you interpret, transform, and distribute the huge portion of enterprise data that arrives in unstructured formats.
- Easy migration to SAP software You can migrate data into SAP solutions and establish master data such as material, customer, and supplier data using prebuilt models and templates.
- Optimized integration with SAP solutions With SAP Data Services, you can use a wide range of supplementary data management offerings from SAP. These include SAP HANA appliance software, SAP NetWeaver BW, SAP NetWeaver BW Accelerator software, SAP Information Steward, SAP Rapid Marts, and SAP Test Data Migration Server software.
- Metadata integration You can collect additional metadata and lineage details for objects stored in SAP NetWeaver BW, SAP BusinessObjects™ Web Intelligence® software, and the SAP BusinessObjects Business Intelligence (BI) platform (including its dashboard and launch-pad functionality). You can also integrate a link to the auto doc and metadata explorer functionalities of SAP Data Services to show the details of data flow, transforms, and mappings.

Manage Data the Way You Want with Greater Simplicity and Ease

SCALE TO HANDLE VERY SMALL AND VERY LARGE DATA VOLUMES

SAP Data Services supports native bulk and batch loading of data by fully integrating with loaders supplied with leading relational database management software. For example, you can trickle-feed a data warehouse. You can create a request-reply message broker to collect message requests in real time, route them to a 24x7 service, and deliver a reply within a user-specified time frame. You can respond to Web service requests, listen in on messaging systems, and manage the capture of both log- and trigger-based changed data. You can optimize parallelism and grid computing through a scalable, high-performance extract, transform, and load (ETL) engine. This mechanism uses the same repository, user interface, and engine to perform both classic loading and data quality management.

REUSE AS OFTEN AS POSSIBLE

With full support for third-generation language in the software, you can write completely customized and reusable transformations. If you've already created custom transformations as stored procedures in a relational database, for example, you can simply import them. In addition, since SAP Data Services was built with reusability in mind, all objects are reusable, including jobs, workflows, and data flows. Typically, you design once and reuse many times, with changes affected globally, and you can copy your work to make duplicates. A difference viewer lets you compare similar objects for example, a local version and a central version – and list the differences between them. Built-in staging transforms and reusable data-store connections let you break apart logic and reuse components at will. And finally, you can create objects from scratch with an XML-based tool kit.

OPTIMIZE INTEGRATION WITH SAP SOFTWARE AND DATA SOURCES

With SAP Data Services, existing SAP customers gain greater IT efficiencies as a result of deep integration with a range of SAP software. SAP Data Services is the data integration software for loading data into suitable formats for in-memory computing.

SAP Data Services offers native access to content extractors in SAP Business Suite software. As a result, customers with SAP NetWeaver BW can use a single ETL tool to define all extraction, validation, and cleansing rules for loading data into the warehouse regardless of source. Other customers get native access to SAP Business Suite without going through the warehouse.

Customers who use SAP Data Services along with SAP BusinessObjects BI solutions, SAP solutions for enterprise performance management (EPM), or SAP solutions for governance, risk, and compliance (GRC) benefit from shared services. These shared services cover user administration and security, offering a scheduling framework and support for access management across the entire infrastructure from a single point.

SAP Data Services integrates with the SAP Solution Manager application management solution to help users monitor jobs for CPU usage and memory utilization and check elapsed time with batch jobs.

INTEGRATE STRUCTURED AND UNSTRUCTURED DATA

What if you could automatically assess and improve data related to any industry, locality, or data domain, including customer, product, supplier, and material data? SAP Data Services supports literally all data – including free-form, structured, and unstructured data – that can

be integrated, processed, and cleansed. You can harness deep linguistic extraction for both unstructured documents and partially structured data. This powerful text-processing feature, based on 15 years of experience with semantic extraction, helps you identify key entities, facts, relationships, and inflections, along with tokens and other linguistic artifacts.

SAP Data Services allows you to incorporate traditional business information and new social information in a way that builds a framework of trust and context. Your constituent organizations can leverage both historical results and current sentiment to add the "why" and "who" to every piece of information. This is all delivered and integrated with management reports, analytic dashboards, and mobiledevice displays, so everyone can understand where the information comes from and how best to apply it. Wherever they are, your stakeholders can quickly and efficiently make the well-informed decisions that drive business success.

OPTIMIZE IMPLEMENTATION WITH SAP SERVICES

Are you looking to optimize the success of your data management initiative? The SAP Services organization is at your service.

SAP Services offers a range of targeted services designed to support you at every stage of the data management lifecycle to optimize the benefits of SAP Data Services. We can help you understand data health issues on various data quality dimensions like completeness, uniqueness, consistency, and timeliness. We can also quickly enable you to get up and running on SAP Data Services with our installation and configuration services, along with templates and best practices for software usage and administration. And finally, you can take advantage of numerous offerings designed to

jump-start your data migration using a mature information management infrastructure and proven implementation methods. Whatever your needs, SAP Services can support you with critical master data management services that deliver real value, minimize your risk, and reduce your cost.

COMPLEMENTARY SAP SOLUTIONS FOR EIM

SAP offers additional solutions for enterprise information management (EIM) that you can use in conjunction with SAP Data Services.

SAP Master Data Governance

The SAP Master Data Governance application enables you to create, maintain, and replicate master data through a centralized, governed process prior to the data being loaded into your line-of-business applications. SAP Master Data

Governance supports line-of-business processes and business rules across partner, customer, supplier, material, finance, and custom domains. Moreover, SAP Master Data Governance is integrated with SAP Data Services and SAP Information Steward to support governance from start to finish.

SAP Sybase PowerDesigner®

SAP Sybase PowerDesigner® software is our flagship modeling and architecture platform that fully supports data, information, and enterprise architecture environments. As a modeling platform, this software provides a pictorial representation of the complexity of the environment to simplify and streamline communication, thus fostering a greater understanding among all stakeholders. By improving communication and collaboration, it enables enterprises to translate business requirements into design and development deliverables.

INTEGRATE, TRANSFORM, AND IMPROVE YOUR DATA

Harnessing the full value of all the data at your disposal is no simple task. But with SAP Data Services, you gain a single solution to streamline business processes and move, improve, govern, and unlock the value of your data. The software's powerful and intuitive data transformation tools make it possible to boost staff productivity while encouraging tighter collaboration between the business and IT. Moreover, you can improve the quality of application data by integrating input from a variety of sources. Just as important, you can lower costs by reducing data customization and downstream cleanup.

FIND OUT MORE

To learn more, call your SAP representative or visit us on the Web at www.sap.com/eim.

For both unstructured documents and partially structured data, a semantic extraction feature helps you identify key entities, facts, relationships, and inflections, along with tokens and other linguistic artifacts.



