Gain Faster, Broader Insight from Text Sources with SAP HANA®

The Best-Run Businesses Run SAP™

©2012 SAP AG. All rights reserved.

Discover the business value of text-based data

It's said that structured data tells us what and unstructured data tells us why. With approximately 80% of enterprise-relevant information originating in unstructured data, you can gain a distinct business advantage by using this data effectively. With the text search and analysis features in the SAP HANA® platform, now you can – quickly and across large volumes of data.

To make decisions that create business value, you need to extract the knowledge embedded in your data. But with the volume of data that's produced on a daily basis, gathering, distilling, and analyzing it can be challenging. Enterprise software is generating massive amounts of structured data. An even larger volume of unstructured data is being produced across call center operations, social media, blogs, forums, e-mail, and a growing number of document types.

To use this information effectively, you have to combine unstructured and structured data so you can query, analyze, and visualize it. Ideally, this needs to happen in real time to guide your business and solve critical problems.

The challenge is to accomplish all of these tasks on one integrated platform. That's now possible. With the text search, text analysis, and analytics features in SAP HANA, you now have a single platform to extract and analyze information from unstructured content.

©2012 SAP AG. All rights reserved.

Experience the power of in-memory technology

Experience the power of in-memory technology

Unlock key information with full-text searching

Perform text analysis on unstructured data

Rapidly develop search apps and deploy via browser or iOS

To be competitive, your company needs answers to specific questions. What do your customers really think about your brand? How did they respond to your last marketing campaign? Why is a particular piece of equipment failing? Is a new drug therapy causing side effects? By using SAP HANA, you can get immediate answers to really complex questions and execute like never before.

SAP HANA is a next-generation data management platform. You can instantly access huge volumes of structured and unstructured data, including text data from different sources. The platform streamlines transactions, analytics, planning, predictive analysis, and sentiment data processing.

Because the platform runs on a single in-memory database, it can process large volumes of data extremely fast, including data-intensive calculations, algorithms, and analysis. You can report on results live even as transaction data continues to be processed. And there are no limits to the amount of data or analytical processing SAP HANA can support, because the architecture can scale to the needs of your business now and in the future.

Instantly analyze huge volumes of structured and unstructured data as it's created, and speed up the processing of high-performance applications.

©2012 SAP AG. All rights reserved.

Unlock key information with full-text searching

Experience the power of in-memory technology

Unlock key information with full-text searching

Perform text analysis on unstructured data

Rapidly develop search apps and deploy via browser or iOS

Business intelligence typically harnesses structured data. SAP HANA enables native, full-text searching to exploit unstructured content. You can search unstructured text much like you would when searching for information on the Internet.

With SAP HANA studio, you can directly query data using an intuitive search definition process or you can rapidly build search apps using an information-access toolkit. You can search on single or multiple columns of almost any visible data type. In addition to standard-string search, SAP HANA also supports full-text search.

During a full-text search, the SAP HANA search engine examines structured text, such as author and date attributes, as well as unstructured text, such as body text, regardless of the sequence of words and characters. A full-text index enables this by analyzing and preprocessing the available text linguistically. This includes normalization, tokenization, word stemming, part-of-speech tagging, and "fuzzy" searching. With a fuzzy search, a database query returns records even if the search term (the user input) contains spelling errors.

Natively process and search any structured or unstructured text content in one flexible and robust database platform.

©2012 SAP AG. All rights reserved.

Perform text analysis on unstructured data

Experience the power of in-memory technology

Unlock key information with full-text searching

Perform text analysis on unstructured data

Rapidly develop search apps and deploy via browser or iOS

Unstructured text data generally does not support automated processing. There's no data model, and it's not easily understood. Structured data conforms to data models associated with databases and spreadsheets. SAP HANA gives structure to unstructured textual content so you can unlock information that drives insight.

File filtering unlocks text from binary documents. It extracts and processes unstructured text data from various file formats such as text, Web, word processing, and messages. In addition, you can load binary, flat, and other documents directly into SAP HANA for native text search and analysis. The figure on the next page shows how to build a text search and text analysis–based application.

SAP HANA exposes the linguistic mark-up for text-mining uses. You can classify entities, such as people, companies, or things. And you can identify domain facts, such as sentiments, topics, and requests. The platform supports up to 31 languages for linguistic mark-up and extraction dictionaries, 11 languages for predefined core extractions, and 4 languages for sentiment analysis.

Analyze, search, and make predictions by performing a full-text analysis on any topic related to your business.

©2012 SAP AG. All rights reserved.

Experience the power of in-memory technology

Unlock key information with full-text searching

Perform text analysis on unstructured data

Rapidly develop search apps and deploy via browser or iOS

Create model	Use SAP HANA® studio to define the search data model and configure the search behavior
Create full-text index	Use SAP HANA studio to create full-text indexes for search, file filtering, and optionally run text analysis
Run text analysis	Extract salient information from text (linguistic markup, entity, and sentiment extraction)
Configure app	Use the information-access toolkit to define layout and data for the app
Consume data	Search on text and filter, analyze, and perform advanced analytics on text analysis table output

Building a text search and a text analysis–based application

©2012 SAP AG. All rights reserved.

Rapidly develop search apps and deploy via browser or iOS

Experience the power of in-memory technology

Unlock key information with full-text searching

Perform text analysis on unstructured data

Rapidly develop search apps and deploy via browser or iOS

The SAP HANA information-access toolkit provides user-interface building blocks for developing search apps for iOS- and browser-based deployments. The apps can provide real-time information access and faceted search features for huge volumes of structured and unstructured text data.

The toolkit enables a freestyle search of an SAP HANA attribute view, displaying and analyzing the result set. The toolkit provides user-interface elements (widgets) such as a search box, a result list with a detailed view, and charts for basic analytics on the result set. The widgets are interconnected and adapt in real time to user entries and mouse-over (hover) selections.

Along with the toolkit widgets, SAP offers a fully operational demo app. You can explore how the search application functions using a set of sample data. You can also use the demo app as a template or for reference when you start developing your own app.

Visualize and interact with structured and unstructured data in SAP HANA using full-text searches, filters, drill-downs, and charts.

©2012 SAP AG. All rights reserved.

Maximize your business and IT benefits

By using SAP HANA, you can extract and analyze unstructured content to gain strategic insight. It offers the flexibility to perform text search, text analysis, and analytics all in one unified platform. Because data resides in a single database, it simplifies the IT landscape and reduces the cost of ownership by reducing the need for duplication, latency, and operational overhead.

By leveraging a high-performing in-memory platform, SAP HANA enables a faster time to analysis. You no longer need to load and process transaction data from your enterprise software into your reporting and analytics solutions – no more converting, preaggregating, or indexing data and no more performance tuning.

The platform provides a unified access layer so you can quickly develop and connect applications to search and explore data in SAP HANA.

Take advantage of high-speed predictive and text analysis of "Big Data" the second it's created to get answers based on information right now.

©2012 SAP AG. All rights reserved.

www.sap.com

Summary

The SAP HANA® platform provides a powerful, unified foundation for storing, processing, and analyzing both structured and unstructured data. It runs on a single, in-memory database, eliminating data redundancy and speeding up the time for information research and analysis and search application development and deployment.

Objectives

- Learn what customers really think about your brands, products, and services
- Analyze business in real time
- Combine predictive and text data analysis for complete data exploration
- Deploy business applications without traditional data limits and complexity
- Simplify IT landscapes

Solution

- Natively process and search any structured or unstructured text content in one flexible and robust database platform
- Natively extract information from documents and perform text analysis on unstructured data
- Rapidly build and deploy search models and apps across operating platforms

Benefits

- Exploit unstructured content to inform strategic decisions
- Reduce data duplication by leveraging one infrastructure for analytical and search workloads
- Leverage intuitive tools to create search models
- Build search applications quickly with user-interface building blocks

Learn more

To find out more, call your SAP representative or visit www.sap.com/hana.

CMP23597 (12/12) ©2012 SAP AG. All rights reserved.

© 2012 SAP AG. All rights reserved.

No part of this publication may be reproduced or transmitted in any form or for any purpose without the express permission of SAP AG. The information contained herein may be changed without prior notice.

Some software products marketed by SAP AG and its distributors contain proprietary software components of other software vendors. National product specifications may vary.

These materials are provided by SAP AG and its affiliated companies ("SAP Group") for informational purposes only, without representation or warranty of any kind, and SAP Group shall not be liable for errors or omissions with respect to the materials. The only warranties for SAP Group products and services are those that are set forth in the express warranty statements accompanying such products and services, if any. Nothing herein should be construed as constituting an additional warranty.

SAP and other SAP products and services mentioned herein as well as their respective logos are trademarks or registered trademarks of SAP AG in Germany and other countries.

Please see http://www.sap.com/corporate-en/legal/copyright/index.epx#trademark for additional trademark information and notices.

The Best-Run Businesses Run SAP™