

SAP Sybase IQ 16 Unstructured Data Analytics Option Technical Overview

Andrew Neugebauer/Analytics Product Management
March 08, 2013



AGENDA

What's Happening in the Marketplace

SAP Sybase IQ Product Success

SAP Sybase IQ 16

Unstructured Data Analytics Landscape

Unstructured Data Analytics Option

Unstructured Data Analytics Option Features

Summary

Marketplace Today

What's Happening in the Marketplace...

BUSINESS INTELLIGENCE

PROFITABILITY

PREDICTIVE

LOB ANALYTIC APPLICATIONS

ADVANCED ANALYTICS

INFORMATION MANAGEMENT

COMPLIANCE MONITORING

LOB & INDUSTRY

BUSINESS ANALYTICS

PERFORMANCE MANAGEMENT

RISK MANAGEMENT

DATA WAREHOUSE

STRATEGY

Exploding Data
Volumes

The Need for
Speed

Rising IT Cost and
Complexity

Challenges customer face today?



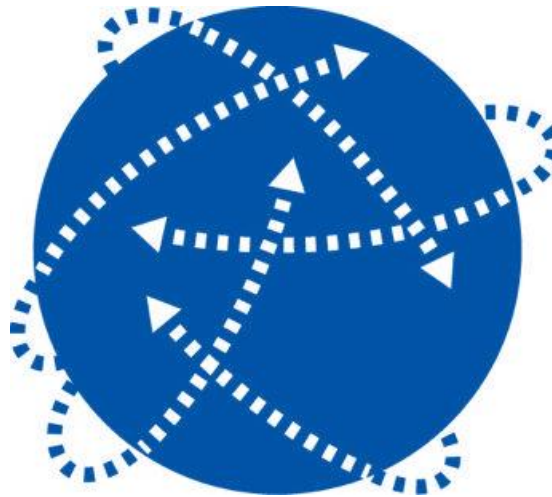
Lost revenues due to
lack of insight



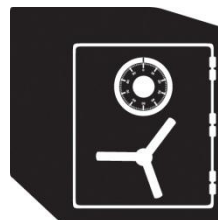
High Costs &
Complexities



Slow Performance



Data Management
Challenges

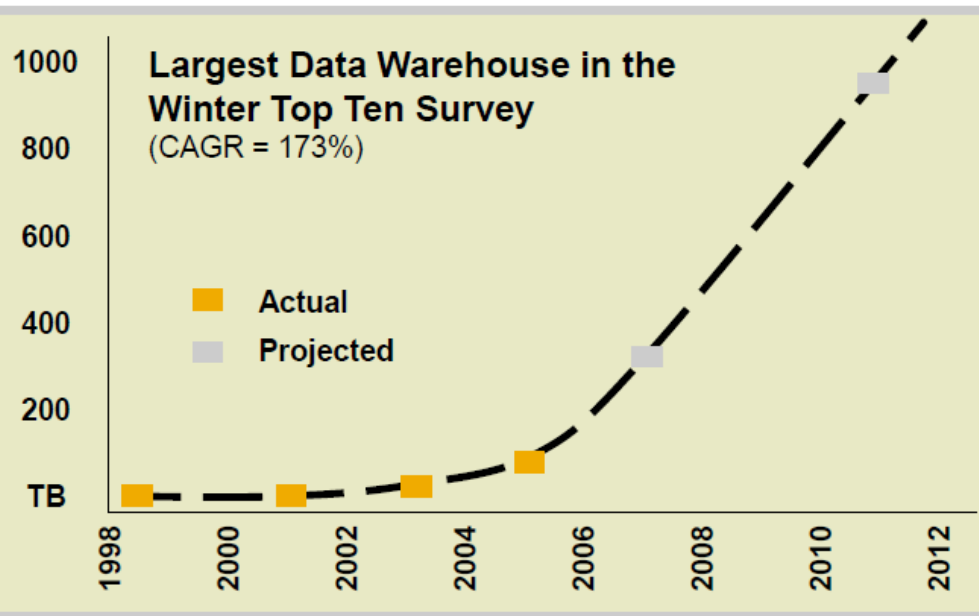


Security

SAP Sybase IQ 16 Motivators

“Petabyte is the new Terabyte” - Forbes

The data explosion continues: Data volumes in analytics environments are growing exponentially...



Source: WinterCorp

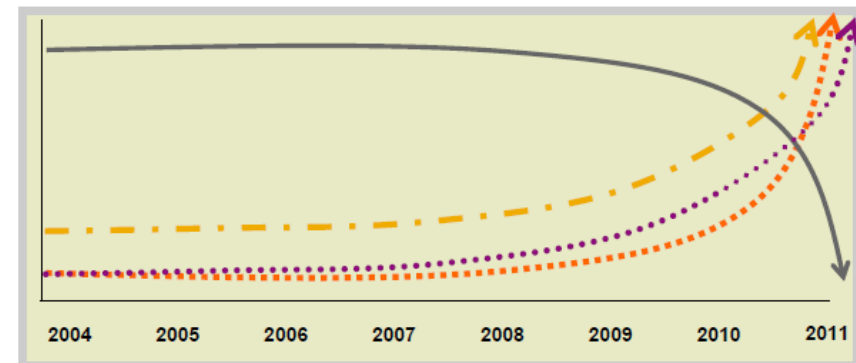
Meanwhile, what trends are you seeing...

Number of business decisions supported by data?

Number of users accessing analytics environments?

Required response time?

Resources required to maintain performance?



Product Success

SAP Sybase IQ: Market Leader for Extreme-Scale EDW and Analytics

High performance analytics server

- Columnar RDBMS (stores data in columns-versus rows)
- Optimized for managing and accessing massive amounts of data for analytics (versus transactions)

Accelerates analytics and reporting

- Up to 1000-times faster than traditional transactional databases
- Handles structured and unstructured data
- High compression and low TCO
- Highly scalable grid architecture

SAP Sybase IQ Facts

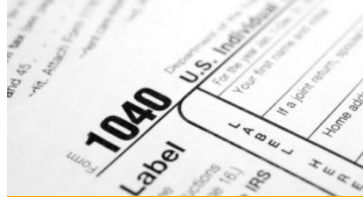
- 2200+ customers with over 4500+ installations worldwide
- Used by twice as many companies as the next leading provider
- Patented data compression dramatically reduces data storage requirement; cuts TCO
- Only column-based solution to support full text search, in-database analytics, and federated analytics
- 96%+ customer satisfaction rates
- Leader, 2013 Gartner Magic Quadrant for Data Warehouse DBMS

SAP Sybase IQ big data analytics

Pervasive across data intensive industries worldwide



Manage and analyze statistical measures for the entire nation of Canada



Analyze ALL Federal tax returns in the US



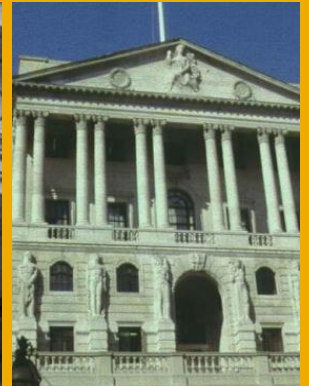
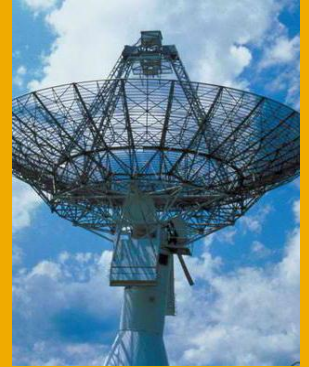
Analyze complex models in more than 200 financial institutions worldwide



Store and analyze massive amounts of industry segment data in 30 of the largest information providers in the world, including Transunion, Nielsen and Axiom



Stands out as the leading enterprise data warehouse among the largest banks, insurance agencies, and telecom operators worldwide



SAP Sybase IQ 16

Solution Overview – SAP Sybase IQ 16



SAP Sybase IQ transforms the way companies compete and win through actionable intelligence delivered at the speed of business to more people and processes.

Value of SAP Sybase IQ 16

- 1 Exploits the value of Big Data**
- 2 Transforms businesses through deeper insights**
- 3 Extends the power of analytics across the entire enterprise**

Unstructured Data Analytics Landscape

Unstructured data analytics

Definitions

Unstructured data either does not have a pre-defined data model and/or does not fit well into relational tables

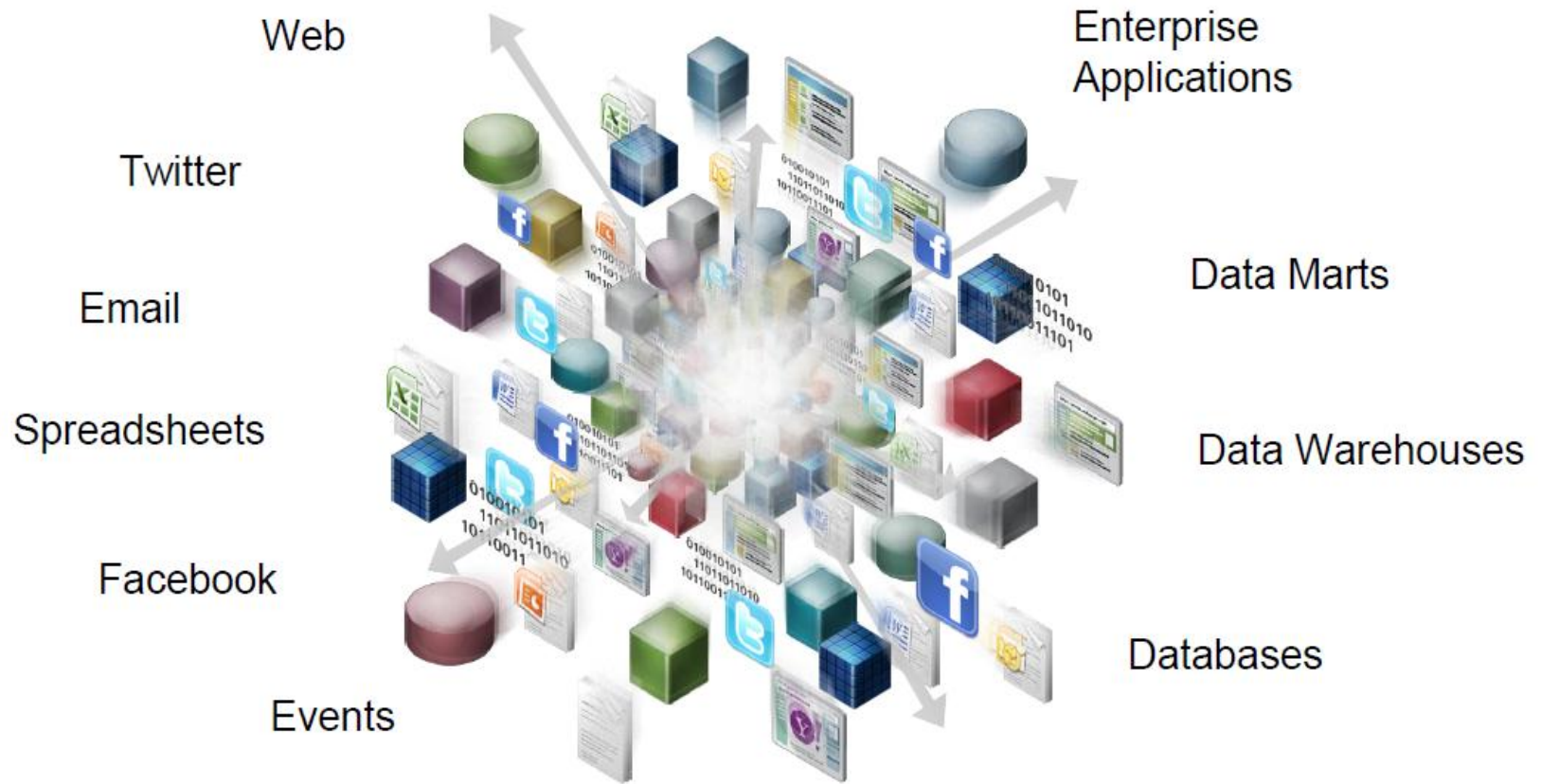
Examples: Binary documents, images, videos, blogs, telemetry data, XML

Text analytics automates what researchers, writers, scholars, and all the rest of us have been doing for years. (Seth Grimes)

Text analytics

- Applies linguistic and/or statistical techniques to extract concepts and patterns that can be applied to categorize and classify documents, audio, video, images
- Transforms “unstructured” information into data for application of traditional analysis techniques
- Unlocks meaning and relationships in large volumes of information that were previously unprocessable by computer

Unstructured data



Data doubles
Every 18 months

80% of enterprise data
is unstructured

Information is a strategic
corporate asset

Where does unstructured data fit with Big Data?

Big data:

A term applied to data sets whose size is beyond the ability of commonly used software tools to capture, manage, and process the data within a tolerable elapsed time. (wikipedia)

Unstructured data fits into the 3 common characteristics of “Big data” data sets:

- **Volume:** An amount of data beyond traditional RDBMS
- **Variety:** Relational, text, and/or multimedia data types
- **Velocity:** Frequency of data generation or of data delivery

How does does unstructured data analytics generate value?

Historic

- Better analyse customer perceptions
- Identify emerging opportunities in real-time
- Improve product/ service design, delivery
- Model cross-sell and up-sell opportunities

Real-time

- Act on emerging opportunities in real-time
- Generate new insights into fraud and risk

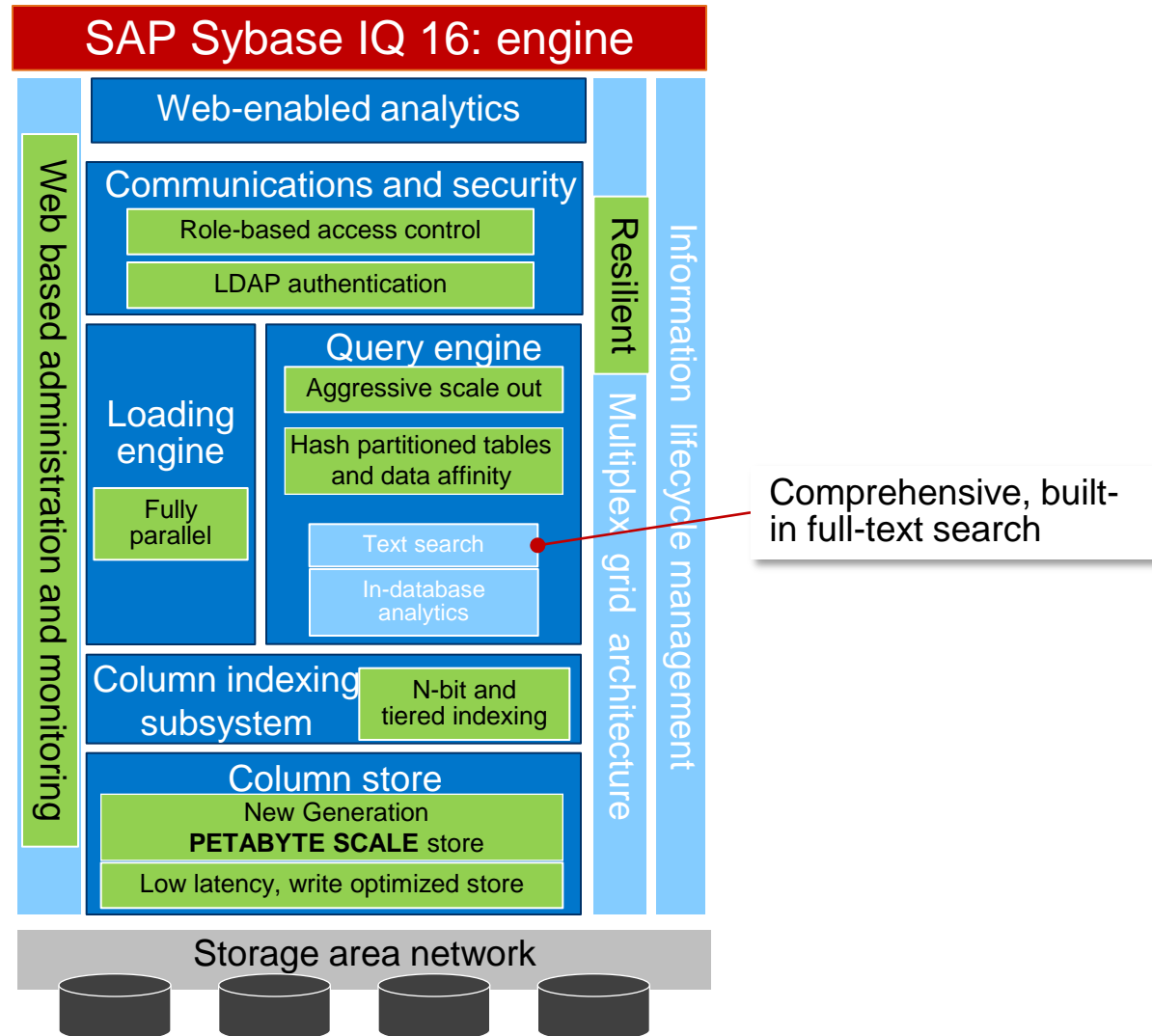
Predictive

- Improved satisfaction and retention
- Faster, more accurate processing of enquiries, claims, requests
- Identify root causes of product/service quality issues

Unstructured Data Analytics Option

SAP Sybase IQ 16

Unstructured Data Analytics



Unstructured data analytics

Getting a 360 degree view

The unstructured data analytics option

- The ability to unlock binary documents and index them
- The ability to search for words and phrases within text data
- The ability to perform Boolean and proximity searches
- The ability to score results based on relevance

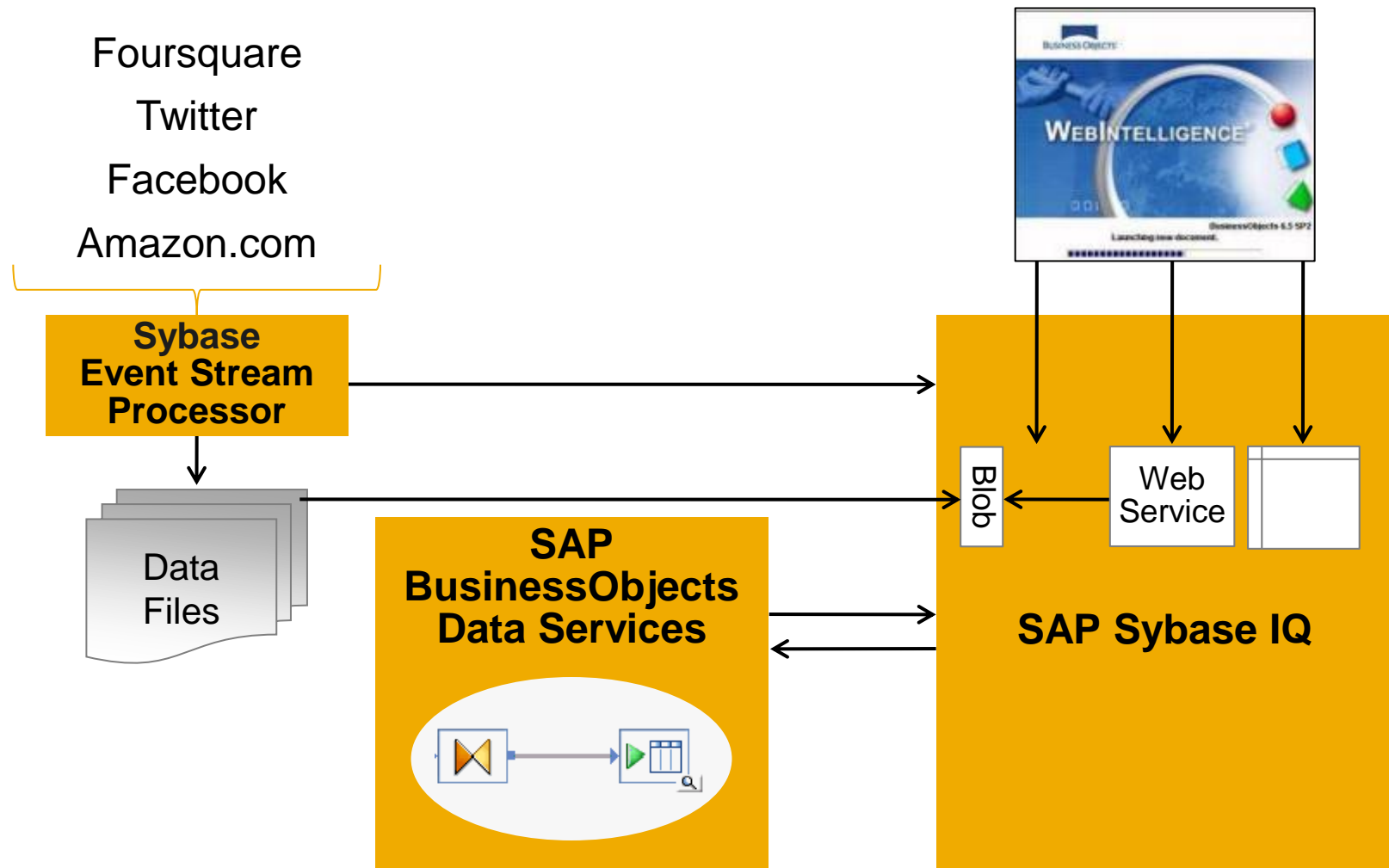
Needed for:

- Market intelligence
- Competitive research and analysis
- Voice of the customer
- Customer satisfaction
- Customer retention
- Product feedback
- Request tracking
- Employee satisfaction
- Employee retention
- Feedback to management

How unstructured analytics is used

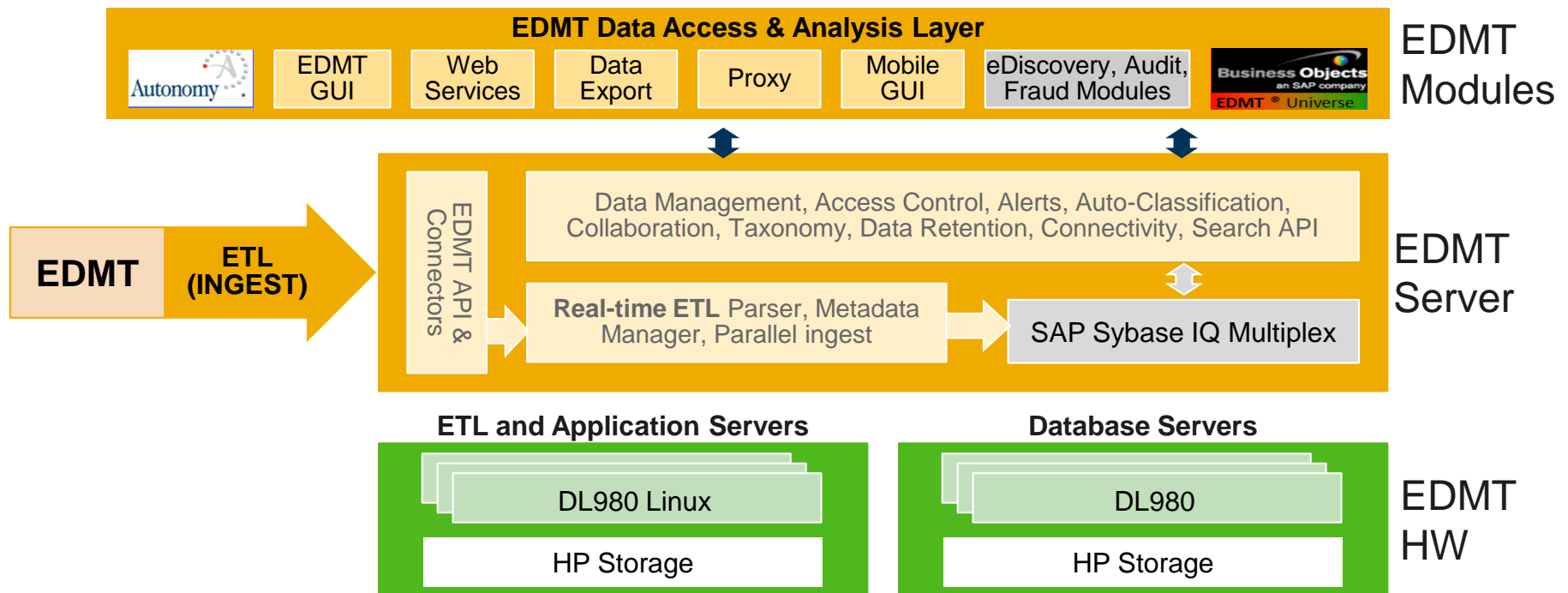
| Intelligence | Governance | Innovation | Performance | Relationships |
|--|--|--|--|--|
| <ul style="list-style-type: none">• Market research and intelligence• Social media analysis• Military intelligence• Law enforcement• National security | <ul style="list-style-type: none">• Compliance• Security (e.g. fraud detection) | <ul style="list-style-type: none">• R&D — science, biotechnology, engineering, manufacturing | <ul style="list-style-type: none">• Claims processing• Marketing effectiveness• Service delivery | <ul style="list-style-type: none">• Customer experience management• Voice of the customer |

Integrated with Sybase ESP and SAP BusinessObjects text processing



Example application

BMMSoft EDMT solution



- **BMMsoft EDMT® Solution (email / documents / multi-media / transactional data) supports enterprise electronic discovery compliance**
- **Single repository for all enterprise data**
- **Structured and unstructured data.** Almost 30% of all enterprise data is comprised of database, or structured data — and structured data is a critical part of many litigation and regulatory electronic discovery requirements. The EDMT® Solution is the only archive product that permits cross-analysis of structured and unstructured data within a single repository

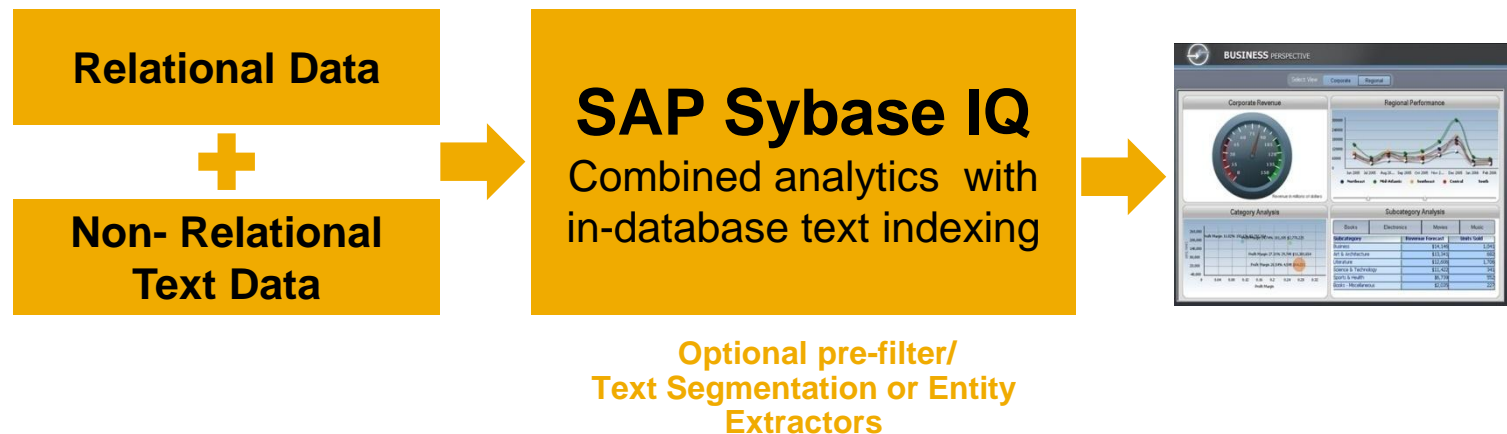
Option features

Architecture

Advanced text processing

Enable analytics on textual data + structured relational data

- Text index — SQL based on terms/phrases, prefix, proximity, scoring
- Interface to plug in 3rd party Document converters or Term Breakers
- SAP Data Services text analytics library
- 3rd party text analytics library

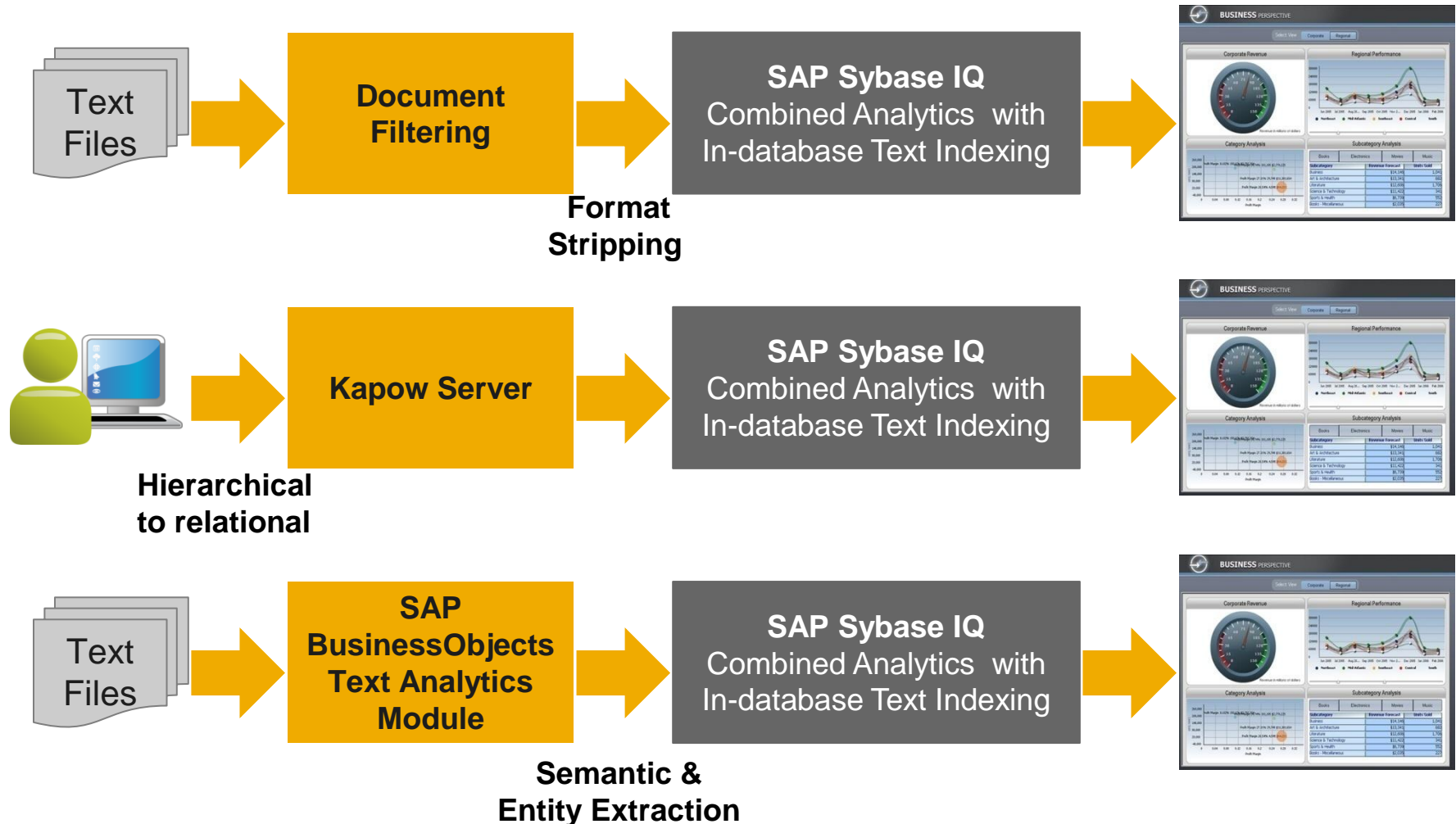


Text Search: Email-Archiving, E-discovery, E-library

Text Analytics: Fraud detection, Risk analytics, News feed analysis

Text Mining: Clustering, categorization, sentiment analysis

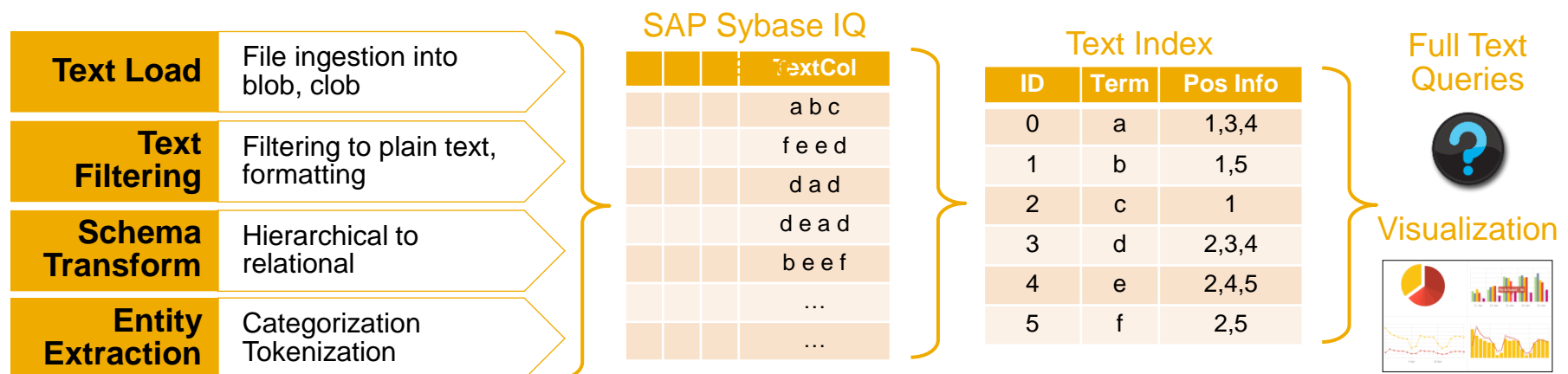
Loading unstructured data



Analyzing textual data

Four step process to start analyzing text data

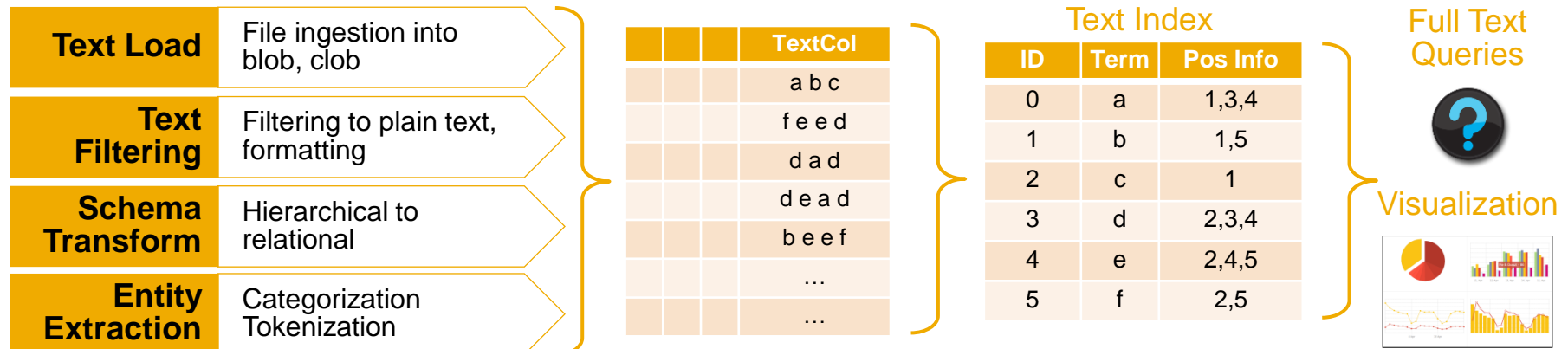
- Step I: Load text data using SAP Sybase IQ load command that may invoke file filtering
- Step II: Create a Text Configuration Object, e.g.
 - CREATE TEXT CONFIGURATION myTextCfg FROM default_char;
- Step III: Create a Text Index using the above Text Configuration Object, e.g.
 - CREATE TEXT INDEX myTextIndex ON myTable (TextCol) CONFIGURATION myTextCfg
- Step IV: Select queries to search and analyze, e.g.
 - SELECT * FROM myTable CONTAINS (TextCol, 'd'); — returns rows and scoring



Creating text index

Text index creation

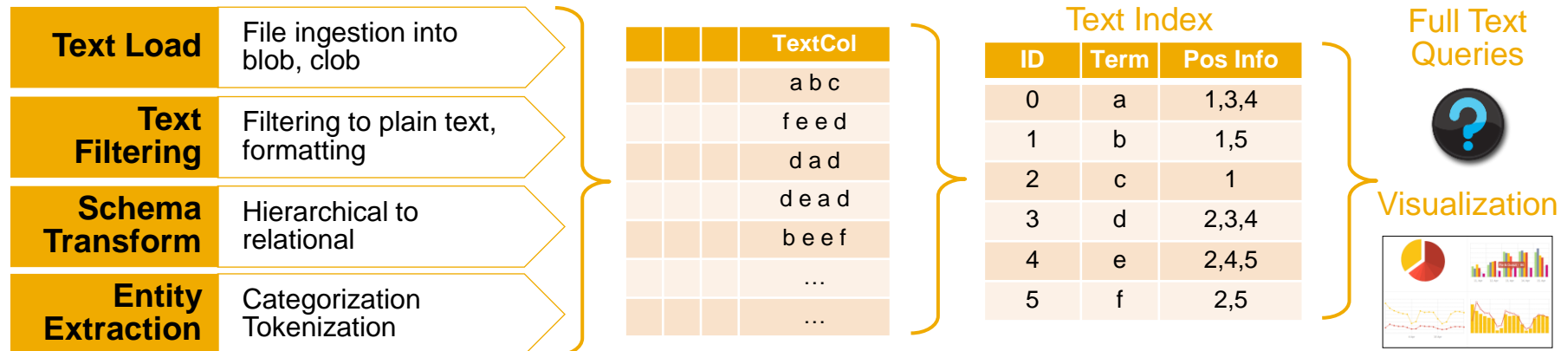
- Create a Text Configuration Object with Minimum Term Length of 4
`CREATE TEXT CONFIGURATION myTextCfg FROM default_char;`
`ALTER TEXT CONFIGURATION myTextCfg MINIMUM TERM LENGTH 4;`
- Create a Text Index using the above Text Configuration Object
`CREATE TEXT INDEX myTextIndex ON myTable (TextCol) CONFIGURATION myTextCfg;`



Full text queries

Full text queries

- `SELECT * FROM myTable WHERE CONTAINS (TextCol, 'd');` – returns rows
- `SELECT * FROM myTable CONTAINS (TextCol, 'd');` – returns rows and scoring
- `SELECT * FROM myTable WHERE CONTAINS (TextCol, 'a AND NOT b');` – Boolean
- `SELECT * FROM myTable WHERE CONTAINS (TextCol, 'a NEAR b');` – proximity



Summary

Store, retrieve and analyze unstructured data as part of the same repository as transactional or analytical data. Perform full text search capabilities including:

- Searching for words and phrases within text data
- Performing Boolean and proximity searches
- Scoring results from text queries based on relevance
- Integrated with SAP Text Data Processing

Learn more

- Visit: <http://www.sap.com/iq>
- Call: 1-877-727-1127 FREE ext. 11001



Thank you

© 201' 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.

Microsoft, Windows, Excel, Outlook, PowerPoint, Silverlight, and Visual Studio are registered trademarks of Microsoft Corporation.

IBM, DB2, DB2 Universal Database, System i, System i5, System p, System p5, System x, System z, System z10, z10, z/VM, z/OS, OS/390, zEnterprise, PowerVM, Power Architecture, Power Systems, POWER7, POWER6+, POWER6, POWER, PowerHA, pureScale, PowerPC, BladeCenter, System Storage, Storwize, XIV, GPFS, HACMP, RETAIN, DB2 Connect, RACF, Redbooks, OS/2, AIX, Intelligent Miner, WebSphere, Tivoli, Informix, and Smarter Planet are trademarks or registered trademarks of IBM Corporation.

Linux is the registered trademark of Linus Torvalds in the United States and other countries.

Adobe, the Adobe logo, Acrobat, PostScript, and Reader are trademarks or registered trademarks of Adobe Systems Incorporated in the United States and other countries.

Oracle and Java are registered trademarks of Oracle and its affiliates.

UNIX, X/Open, OSF/1, and Motif are registered trademarks of the Open Group.

Citrix, ICA, Program Neighborhood, MetaFrame, WinFrame, VideoFrame, and MultiWin are trademarks or registered trademarks of Citrix Systems Inc.

HTML, XML, XHTML, and W3C are trademarks or registered trademarks of W3C®, World Wide Web Consortium, Massachusetts Institute of Technology.

Apple, App Store, iBooks, iPad, iPhone, iPhoto, iPod, iTunes, Multi-Touch, Objective-C, Retina, Safari, Siri, and Xcode are trademarks or registered trademarks of Apple Inc.

IOS is a registered trademark of Cisco Systems Inc.

RIM, BlackBerry, BBM, BlackBerry Curve, BlackBerry Bold, BlackBerry Pearl, BlackBerry Torch, BlackBerry Storm, BlackBerry Storm2, BlackBerry PlayBook, and BlackBerry App World are trademarks or registered trademarks of Research in Motion Limited.

Google App Engine, Google Apps, Google Checkout, Google Data API, Google Maps, Google Mobile Ads, Google Mobile Updater, Google Mobile, Google Store, Google Sync, Google Updater, Google Voice, Google Mail, Gmail, YouTube, Dalvik and Android are trademarks or registered trademarks of Google Inc.

INTERMEC is a registered trademark of Intermec Technologies Corporation.

Wi-Fi is a registered trademark of Wi-Fi Alliance.

Bluetooth is a registered trademark of Bluetooth SIG Inc.

Motorola is a registered trademark of Motorola Trademark Holdings LLC.

Computop is a registered trademark of Computop Wirtschaftsinformatik GmbH.

SAP, R/3, SAP NetWeaver, Duet, PartnerEdge, ByDesign, SAP BusinessObjects Explorer, StreamWork, SAP HANA, 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.

Business Objects and the Business Objects logo, BusinessObjects, Crystal Reports, Crystal Decisions, Web Intelligence, Xcelsius, and other Business Objects products and services mentioned herein as well as their respective logos are trademarks or registered trademarks of Business Objects Software Ltd. Business Objects is an SAP company.

Sybase and Adaptive Server, iAnywhere, Sybase 365, SQL Anywhere, and other Sybase products and services mentioned herein as well as their respective logos are trademarks or registered trademarks of Sybase Inc. Sybase is an SAP company.

Crossgate, m@gic EDDY, B2B 360° , and B2B 360° Services are registered trademarks of Crossgate AG in Germany and other countries. Crossgate is an SAP company.

All other product and service names mentioned are the trademarks of their respective companies. Data contained in this document serves informational purposes only. National product specifications may vary.

The information in this document is proprietary to SAP. No part of this document may be reproduced, copied, or transmitted in any form or for any purpose without the express prior written permission of SAP AG.