Introduction to
SAP BusinessObjects Data Services XI 3.0
The information in this presentation is confidential and proprietary to SAP and may not be disclosed without the permission of SAP. This presentation is not subject to your license agreement or any other service or subscription agreement with SAP. SAP has no obligation to pursue any course of business outlined in this document or any related presentation, or to develop or release any functionality mentioned therein. This document, or any related presentation and SAP’s strategy and possible future developments, products and or platforms directions and functionality are all subject to change and may be changed by SAP at any time for any reason without notice. The information on this document is not a commitment, promise or legal obligation to deliver any material, code or functionality. This document is provided without a warranty of any kind, either express or implied, including but not limited to, the implied warranties of merchantability, fitness for a particular purpose, or non-infringement. This document is for informational purposes and may not be incorporated into a contract. SAP assumes no responsibility for errors or omissions in this document, except if such damages were caused by SAP intentionally or grossly negligent.

All forward-looking statements are subject to various risks and uncertainties that could cause actual results to differ materially from expectations. Readers are cautioned not to place undue reliance on these forward-looking statements, which speak only as of their dates, and they should not be relied upon in making purchasing decisions.
1. **Why Enterprise Information Management**

2. Data Services: One tool for data integration and data quality management

3. Metadata Management: Impact and lineage to prove trustworthiness of your data

4. Some data services use cases in an SAP environment

5. Wrap-up
The volume of data within enterprises is exploding

Application silos make it difficult for people to collaborate

Users have little understanding of the quality of available data

Excessive time is spent on integration rather than innovation

IT is struggling to address rapidly changing business requirements
90% of upper level management feel they don’t have the necessary information for critical business decisions; 50% of them are afraid they are making poor decisions because of it.”

― Economist Intelligence Unit

50% to 70% of ERP implementations are reported as “challenged” in part to data integrity and/or data accuracy problems.”

― Adaptive Growth, Inc.

90% of all businesses still do not have sufficient policies in place to meet data governance regulations.”

― IT Policy Compliance Group
Challenges to Effectively Managing Information
How Does This Impact IT?

Top issues

- How do I empower all users with the information they need to make better decisions?
- How do I help people to work across enterprise boundaries?
- How do I provide applications with accurate data to drive business operations?
- How do I implement data governance to ensure compliance and meet regulatory requirements?
- How can I minimize cost and manage complexity?
Inconsistent Data Across Different Silos
Impacts Business Results and Increases Costs

- Departments create and store their own data
- Data inconsistencies, redundancies, and errors impact business results and increase costs
Manage Information as a Strategic Asset
Reduce Inefficiencies and Deliver Trusted Information

Enterprise Information Management

- Each information asset is readily understood, available and trusted
- Formal information architecture to identify, share and govern all data
SAP Provides A Complete Approach To EIM
Encompassing People, Process & Technology
SAP Provides Best-In-Class EIM Solutions
Deliver Information That Is Complete, Accurate, and Accessible

Data Integration & Quality Management:
- SAP BusinessObjects Data Services
- SAP BusinessObjects Data Federator
- SAP BusinessObjects Text Analysis
- SAP BusinessObjects Data Insight
- SAP Data Migration services

Master Data Management:
- SAP NetWeaver Master Data Management
- SAP Master Data Governance for Financials
- SAP Data Maintenance by Vistex

Content & Information Lifecycle Management:
- SAP NetWeaver Information Lifecycle Management
- SAP Extended ECM by Open Text
- SAP Document Access by Open Text
- SAP Archiving by Open Text

Enterprise Data Warehousing:
- SAP NetWeaver Business Warehouse
- SAP NetWeaver Business Warehouse Accelerator
- SAP BusinessObjects Rapid Marts
- SAP BusinessObjects Metadata Management
1. Why Enterprise Information Management
2. **Data Services: One tool for data integration and data quality management**
3. Metadata Management: Impact and lineage to prove trustworthiness of your data
4. Some data services use cases in an SAP environment
5. Wrap-up
SAP BusinessObjects Data Services

Data Services is the first single tool for data integration and data quality

Data Integrator XI R2
- Development User Interface
- Metadata Repository
- Runtime Architecture
- Administration and Connectors

Data Quality XI R2
- Development User Interface
- Metadata Repository
- Runtime Architecture
- Administration and Connectors

Data Services
- One Development User Interface
- One Metadata Repository
- One Runtime Architecture
  - Access
  - Transform
  - Improve
  - Deliver
- One Administration Environment
- One Set of Connectors
Data Services Architecture

SAP ERP, SAP CRM, SAP Master Data Management (MDM), SAP NetWeaver BI, ...

SOA

Data profiling

Data Services Engine

Data Cleansing

Data Validation

Data Auditing

Real Time

Batch

SAP NetWeaver BI

Data Migration, Synchronization, ...

Query, Reporting, Analysis, and Dashboards

Shared Metadata

Data Lineage

Impact Analysis

Files, XML, Mainframe, Excel, etc.

PeopleSoft, Oracle Apps, Siebel, etc.

Oracle, SQL, DB2, etc.

SAP R/3, SAP ERP, SAP NetWeaver BI

Data Services Architecture

Files, XML, Mainframe, Excel, etc.

PeopleSoft, Oracle Apps, Siebel, etc.

Oracle, SQL, DB2, etc.

SAP R/3, SAP ERP, SAP NetWeaver BI
Enterprise-Wide Data Access

Support for structured and unstructured data
Broad connectivity to databases, applications, legacy, file formats, and unstructured data

<table>
<thead>
<tr>
<th>Databases</th>
<th>Applications</th>
<th>Files/Transport</th>
<th>Mainframe (with partner)</th>
<th>Unstructured Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oracle</td>
<td>JD Edwards</td>
<td>Text delimited</td>
<td>ADABAS</td>
<td>Any text file type</td>
</tr>
<tr>
<td>DB2</td>
<td>Oracle Apps</td>
<td>Text fixed width</td>
<td>ISAM</td>
<td>32 languages</td>
</tr>
<tr>
<td>Sybase &amp; IQ</td>
<td>PeopleSoft</td>
<td>EBCDIC</td>
<td>VSAM</td>
<td></td>
</tr>
<tr>
<td>SQL Server</td>
<td>Siebel</td>
<td>XML</td>
<td>Enscribe</td>
<td></td>
</tr>
<tr>
<td>Informix</td>
<td>Salesforce.com</td>
<td>Cobol</td>
<td>IMS/DB</td>
<td></td>
</tr>
<tr>
<td>Teradata</td>
<td>SAP NetWeaver BI</td>
<td>Excel</td>
<td>RMS</td>
<td></td>
</tr>
<tr>
<td>ODBC</td>
<td>SAP R/3</td>
<td>HTTP</td>
<td>Both direct and change data</td>
<td></td>
</tr>
<tr>
<td>MySQL</td>
<td>– ABAP</td>
<td>JMS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Netezza</td>
<td>– BAPI</td>
<td>SOAP (Web Services)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HP NeoView</td>
<td>– IDoc</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Data Integrator: Enterprise-class Data Integration platform

Explore, transform, and move data anywhere, at any frequency

- **Deliver trusted information**
  - Market leading data quality functionality within the data integration environment
  - Built-in features for validating data against business rules and auditing data movement execution
  - End-to-end metadata impact analysis and lineage

- **Agility and ease-of-use**
  - Single easy to use development interface to build, test, and deploy
  - Web-based administration and management
  - Collaboration features for team development

- **Deliver extreme ETL scalability**
  - Parallelized processing down to the transform level
  - Grid computing for high availability and throughput
  - Services-based architecture enabling right-time data delivery
  - Powerful prepackaged transformations
Increase Value of Data Assets with Data Quality

- **Measure and analyze** data through data assessment and continuous monitoring

- **Cleanse and enhance** customer and operational data anywhere across the enterprise

- **Match and consolidate** data at multiple levels within a single pass for individuals, households, or corporations

- **Improve and automate** the delivery of direct mail and goods
Data Cleansing examples: Customer Data

Input record
Maggie.kline@future_electronics.com
Margaret Smith-Kline phd
FUTURE Electronics
5/23/03
101 6th ave
manhattan
ny
10012
001124367

Output record
Salutation: Ms.
First name: Margaret
Last name: Smith-Kline
Postname: Ph. D.
Match standards: Maggie, Peg, Peggy
Gender: Strong Female
Company name: Future Electronics
Address 1: 101 Avenue of the Americas
City: New York
State: NY
ZIP+4: 10013-1933
Email: maggie.kline@future_electronics.com
SSN: XXX-XX-XXXX
Date May 23, 2003
## Data Cleansing examples: Product Data (Universal Data Cleanse)

<table>
<thead>
<tr>
<th>Input Description</th>
<th>Parsed Output</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Description</strong></td>
<td><strong>Product</strong></td>
</tr>
<tr>
<td>Kallkyle screw</td>
<td>screw</td>
</tr>
<tr>
<td>test steel plate 20 x 35 mm</td>
<td>plate</td>
</tr>
<tr>
<td>wire 23.33 x 40.50 cm</td>
<td>wire</td>
</tr>
<tr>
<td>34 x 60 mm steel plate</td>
<td>plate</td>
</tr>
<tr>
<td>steel plate 34,0 60 mm</td>
<td>plate</td>
</tr>
<tr>
<td>34.0 x 60,0 mm steel plate</td>
<td>plate</td>
</tr>
<tr>
<td>34 x 60 mm steel plate?</td>
<td>plate</td>
</tr>
<tr>
<td>plate</td>
<td>plate</td>
</tr>
<tr>
<td>steel plate</td>
<td>plate</td>
</tr>
</tbody>
</table>
Matching and Consolidation example

Input records

Ms. Margaret Smith-Kline Ph.D.
Future Electronics
101 Avenue of the Americas
New York NY 10013-1933
maggie.kline@future_electronics.com
May 23, 2003

Maggie Smith
Future Electronics Co. LLC
101 6th Ave.
Manhattan, NY 10012
maggie.kline@future_electronics.com
001-12-4367

Ms. Peg Kline
Future Elect. Co.
101 6th Ave.
New York NY 10013
001-12-4367
(222) 922-9922
5/23/03

Consolidated record

Name: Ms. Margaret Smith-Kline Ph.D.
Company name: Future Electronics Co. LLC
SSN: 001-12-4367
Purchase date: 5/23/2003
Address: 101 Avenue of the Americas
New York, NY 10013-1933
Latitude: 40.722970
Longitude: -74.005035
Fed code: 36061
Phone: (222) 922-9922
Email: maggie.kline@future_electronics.com
Data Services in action

Access

Cleanse

Validate

Deliver

Profile
1. Our vision of information management
2. Data Services: One tool for data integration and data quality management
3. Metadata Management: Impact and lineage to prove trustworthiness of your data
4. Some data services use cases in an SAP environment
5. Wrap-up
Data Lineage Helps Users Make Confident Decisions

Where did this number come from?

Data lineage provides information on how a number in your BI report is calculated during the ETL process and its origin.
**Data Lineage**: End users are able to trace back data in their reports to the original source through all different layers and transformation and presentation steps.

**Impact Analysis**: IT needs to know which reports and users will be affected when a change is made to one of the source systems.
Metadata Management Architecture

- **Consolidate** metadata from multiple systems
- **Integrate** into one central, open metadata repository
- **Audit** impact, usage, and lineage using the Metadata Explorer
- **Trust** your metadata with added business content
1. Our vision of information management
2. Data Services: One tool for data integration and data quality management
3. Metadata Management: Impact and lineage to prove trustworthiness of your data
4. **Some data services use cases in an SAP environment**
5. Wrap-up
Scenario 1: Data Quality for SAP CRM

Data Quality Management for SAP provides a prepackaged native integration of data quality best practices within the SAP environment using the BusinessObjects Data Services platform.

- Enforces data discipline directly within SAP CRM or SAP ERP systems
  - No extracts for external processing required
- Virtually undetectable presence that provides:
  - Global address correction and standardization
  - Comprehensive duplicate detection
  - Convenient record searching
- Supported versions
  - SAP CRM: Releases 4.0 and 5.0
  - SAP ERP Central Component: Release 5.0 and 6.0
  - SAP R/3 4.6C
Examples

Potential Duplicates Identified!

Review the Corrected and Fielded Information
Scenario 2: Load SAP NetWeaver BW

Get easy access to data stored in non-SAP sources
- Replace database connectivity with Rapid Mart application understanding

Include data cleansing and validation rules before loading data
- Standardize, correct, and enhance
- Match and consolidate
- Validate against validation rules before loading into SAP NetWeaver BI
Scenario 3: Data Migration for SAP

Traditional Steps

Let’s cross the fingers and hope it loads …

But attention: Data migration is far from easy.

Analysts

More than 80% of data migration projects run over time and/or over budget. Cost overruns average 30%. Time overruns average 41%.

Bloor, 2007

80% of organizations … will underestimate the costs related to the data acquisition tasks by an average of 50 percent

Gartner, Data Warehouse Infrastructure for Growth and Change

Why?

- Lack of experience
- Migration project plan is unrealistic
- Data quality issues/lack of confidence in the data
- You need to extend the consultants time for the SAP ERP project

It is crucial to identify the right partner with the right expertise for your data migration project.
Data migration is a critical element in the success of the ERP implementation.

- Ensure accountability and a robust project approach to data migration from a single strategic vendor.

- Enable data governance best practices that live on after the project.

- Industry-leading BusinessObjects Data Services platform benefits data migration and delivers many additional benefits in information management.

- Pre-built framework and tools reduce the need for custom scripts and deliver data insights more quickly and consistently.

- Engage business users early with better access to data quality and metrics.
SAP Data Migration Framework
Based on BusinessObjects Data Services

Data Governance Visualization

Project Management Summary Metrics
Track % quality and progress over time against project milestones

Business users/Data Steward reports
Show detailed data exceptions requiring resolution for each SAP Business Objects, e.g., Customer, including reconciliation exceptions

Data Exceptions & DQ Metrics

Migration Framework Staging
Cleanse, Harmonize, Reconcile

SAP Business Object Validation Library
Library of DS objects to validate the data pre-load.
The metrics created from this process drive the Data Governance Engine

SAP Load: XML, IDOC, File

Extract & transform to SAP structures
Oracle Apps, PeopleSoft, JDEdwards, Siebel

Legacy

XML, flat file

Post load reconciliation
1. Our vision of information management
2. Data Services: One tool for data integration and data quality management
3. Metadata Management: Impact and lineage to prove trustworthiness of your data
4. Some data services use cases in an SAP environment
5. Wrap-up
7 Key Points to Take Home

- Data is spread throughout systems and applications in the enterprise
  - You need a broad connectivity to all source systems
  - Don’t forget the unstructured data!
- Make sure you understand your data well before starting any project
  - Profiling your data will give you insight into your source data and identify potential data quality issues
- The quality of data will vary among source systems
  - Standardize, cleanse, and enrich data before loading into the target applications
- Avoid duplicate data
  - Find matches and consolidate to get one “golden record”
- Use a “data quality firewall”
  - Define validation rules to prevent incorrect data to get loaded
- Deliver data in “right-time”
  - Sometimes you need real-time services, in other cases you need to bulk load data overnight. One architecture should be able to serve all needs.
- Don’t underestimate the importance of metadata
  - Metadata gives insight in where the data comes from and how it was transformed