Turn "Big Data" into Business Value with Real-Time BI

Timo Elliott, March 2012
Legal Disclaimer

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 in 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’s willful misconduct or gross negligence.

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.
Executives believe companies can benefit greatly from using data, especially information generated within the company.

- **Agree: 69%**
  - Data is extremely important for competitive advantage

- **Agree: 77%**
  - Data makes an important contribution to customer relations efforts

- **Agree: 70%**
  - Business information has helped manage costs or improve operations
Top Technology Priorities of CIOs, Gartner

### 2011

<table>
<thead>
<tr>
<th>Rank</th>
<th>Technology</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Cloud computing</td>
</tr>
<tr>
<td>2</td>
<td>Virtualization</td>
</tr>
<tr>
<td>3</td>
<td>Mobile technologies</td>
</tr>
<tr>
<td>4</td>
<td>IT management</td>
</tr>
<tr>
<td>5</td>
<td>Business Intelligence</td>
</tr>
<tr>
<td>6</td>
<td>Networking, voice, and data communications</td>
</tr>
<tr>
<td>7</td>
<td>Enterprise applications</td>
</tr>
<tr>
<td>8</td>
<td>Collaboration technologies</td>
</tr>
<tr>
<td>9</td>
<td>Infrastructure</td>
</tr>
<tr>
<td>10</td>
<td>Web 2.0</td>
</tr>
</tbody>
</table>

### 2012

<table>
<thead>
<tr>
<th>Rank</th>
<th>Technology</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Analytics and business intelligence</td>
</tr>
<tr>
<td>2</td>
<td>Mobile technologies</td>
</tr>
<tr>
<td>3</td>
<td>Cloud computing (SaaS, IaaS, PaaS)</td>
</tr>
<tr>
<td>4</td>
<td>Collaboration technologies (workflow)</td>
</tr>
<tr>
<td>5</td>
<td>Legacy modernization</td>
</tr>
<tr>
<td>6</td>
<td>IT management</td>
</tr>
<tr>
<td>7</td>
<td>CRM</td>
</tr>
<tr>
<td>8</td>
<td>ERP applications</td>
</tr>
<tr>
<td>9</td>
<td>Security</td>
</tr>
<tr>
<td>10</td>
<td>Virtualization</td>
</tr>
</tbody>
</table>
Top 10 CIO Business and Technology Priorities, 2012

<table>
<thead>
<tr>
<th>Rank</th>
<th>Top 10 Technology Priorities</th>
<th>Top 10 Business Priorities</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Analytics and business intelligence</td>
<td>Increasing enterprise growth</td>
</tr>
<tr>
<td>2</td>
<td>Mobile technologies</td>
<td>Attracting and retaining new customers</td>
</tr>
<tr>
<td>3</td>
<td>Cloud computing (SaaS, IaaS, PaaS)</td>
<td>Reducing enterprise costs</td>
</tr>
<tr>
<td>4</td>
<td>Collaboration technologies (workflow)</td>
<td>Creating new products and services (innov.)</td>
</tr>
<tr>
<td>5</td>
<td>Legacy modernization</td>
<td>Delivering operational results</td>
</tr>
<tr>
<td>6</td>
<td>IT management</td>
<td>Improving efficiency</td>
</tr>
<tr>
<td>7</td>
<td>CRM</td>
<td>Improving profitability (margins)</td>
</tr>
<tr>
<td>8</td>
<td>ERP applications</td>
<td>Attracting and retaining the workforce</td>
</tr>
<tr>
<td>9</td>
<td>Security</td>
<td>Improving marketing and sales effectiveness</td>
</tr>
<tr>
<td>10</td>
<td>Virtualization</td>
<td>Expanding into new markets and geos</td>
</tr>
</tbody>
</table>
Business Analytics Continues Steady Growth

“After three decades, the business analytics market is finally reaching the mainstream”
“There are few growth inhibitors in the foreseeable future”

Dan Vessel, IDC


2005 2006 2007 2008 2009 2010 2011 2012

17.5 19.4 22.1 24.3 24.9 28.1 30.4* 33.9*
“BI spending has far surpassed IT budget growth overall for several years”
Dan Sommer, Gartner

Gartner: worldwide BI, analytics and performance management software revenue
New revenue generated from IT initiatives (enterprise innovation, context-aware computing, social networks, etc.) will become the primary factor determining CIOs compensation.

Information-smart businesses will increase recognized IT spending per head by 60%.

“Enterprise leaders and stakeholders must change their way of thinking that “lower is better” for IT spending per employee”

Gartner Predicts 2011
User Growth

Percentage of Potential Users That Analytics Will Reach

- 100% (2020)
- 75% (2014)
- 50% (2011)
- 25%
Collaborative, real-time, and predictive rose most, while ease of implementation, fast exploration, and ease of use remained highly important.
What’s Hard?

Poor organization of data is a challenge

Agree: 88%

Poor processes for sharing data between departments and employees is a hurdle

Agree: 81%

Technical issues—such as data silos or incompatible systems—represent at least some challenge

Agree: 77%
What’s Hard?

Big Data!
Data Scientists Wanted
2.5 PB
7.9 ZB
BIG DATA
REAL TIME
PREDICTIVE
Worldwide digital content will double in 18 months, and every 18 months thereafter.

In 2005, humankind created 150 exabytes of information. In 2011, **1,200 exabytes will be created.**

80% of enterprise data will be unstructured, spanning traditional and non-traditional sources.
Data Quality is a Real Problem

88% See data as a strategic asset

33% Lost a customer or new business deal as a result of missing data
30% Lost business due to the mishandling of data
30% Of operating revenue is wasted through process failure and information rework
4% Rate data quality as excellent
63% Have no idea how much data quality might be costing
17% Have no plans to start a data quality initiative

Sources: Vanson Bourne; Larry English; The State of Data Quality Today
Applications for Data Stewards
Present    Big Data Applications?

Process    Azkaban    Oozie    Pig    Hive    Hadoop    MapReduce    S4    Storm

Store    Voldemort    Cassandra    Hbase

Ingest    Kafka    Flume    Scribe
The New Data Warehousing

```
The New Data Warehousing

External data
Cloud

Massively Parallel
Hadoop, NoSQL

New Analytic Platforms

Complex Event Processing
In-Memory
Calculation Engine
Columnar Databases

Sensors, Real-time
Web, Unstructured

 "Extreme data performance"

Everybody in the industry is doing some combination

ERP, Replication
```
Understand Big Data through Information Management Lifecycle
Technologies supporting enterprise requirements naturally co-exist and co-evolve

**Structured**
- Replication/synchronization
- Extract-transform-load
- Event stream processing
- Database
- Many instance formats
- Many storage formats
- ACID properties
- Transactional
- Versatile data queries using SQL/OLAP
- Scripting and UDF to process In-DB
- Federate queries across DB and DFS
- Low-latency processing
- Mature connectors to transport data to variety of visualization tools

**Ingest**
- File transfer
- Extract-load data from variety data sources

**Store**
- Distributed File System (DFS)
- Files stored in native format
- Software-driven reliability model
- BASE properties
- Lack of BI tool support

**Process**
- Transform to pre-process or at time of request using map-reduce
- Huge amount of efforts to define the SQL
- Scheduled chains to create consumable information
- Simple SQL-like connectivity
- High-latency processing

**Present**
- Need new storage formats for high speed and high concurrency
- Emerging visualization tools

**Unstructured**
- Replication/synchronization
- Extract-transform-load
- Event stream processing
- Versatile data queries using SQL/OLAP
- High-latency processing
SAP Big Data Processing Framework

Mobile

Big Data Analytics

Big Data Applications

Present

Sybase ESP
- Stream & event processing

Sybase ASE
- Transaction Processing

SAP HANA
- In-memory Technology

SAP Replication Server, SAP BusinessObjects Data Services
  (Integrate / synchronize data across deployment options)

Sybase IQ
- Analytic Grid

Sybase ESP Monitor / filter streaming events

Hadoop
- MapReduce
- Batch Compute Framework
- Hive/HDFS

Ingest

Semi-structured Data

Structured Data

Unstructured Data
New Analytic Infrastructures

We now have the possibility to dramatically simplify our analytic infrastructures.

SPEED & FLEXIBILITY
Blur Operations and Analytics

Use a single platform for both analytics and applications
The New Analytics Experience

Extranets

Collaborative Decision Support

Action-oriented

Mobile

Predictive

Data Discovery

Vizualization, Geolocation

Social

Context

“Actionable Insights”

Specialization vs coverage and business process
What Causes Decision Problems? People!

What is typically the cause for delay between collecting data and using them to inform decisions?
(% respondents)

- No formal processes around data management: 31%
- Validating and scrubbing the data: 27%
- Lack of organisational urgency in viewing/using the data: 24%
- Lack of technology: 9%
- Other: 4%
- Don’t know: 5%

Source: Economist Intelligence Unit
Major Strategy Directions

- **DEEPER INTEGRATION**
- **RICHER INSIGHT**
- **SOCIAL**
- **MOBILE**

Methods and Templates
Ad-hoc Teams
Structure to Teamwork
Integration

HANA
Sybase IQ
SAP
BW Hierarchies
BICS
ERP InfoSets
ABAP Functions…

Microsoft Sharepoint

HANA
In-Memory Computing

Sybase IQ
Insight to Action

Supported source documents:

- SAP Crystal Reports
- SAP BusinessObjects Dashboards
Accessing Hadoop Data
Information Design Tool on Hadoop Hive

A Data Foundation against a Hive schema

- One can draw joins between the Hive tables
- We support Hive tables, aliases, derived tables, Hive views and Hive partitioned tables
Examples

National League Baseball Salaries (2005)

Year 2011

Speech: Martin_Luther_King(1963-08-28)
Richer Insight
Richer Insight
Mobile
Try the new “Experience SAP” Mobile Application
Recalls Plus
Paul Predicts
Afaria Analytics

- **Devices**
  - Number of Devices by Operating system (OS)
  - Number of Devices by Carrier
  - Number of Devices by Manufacturer
  - Number of new devices added each month by OS
  - Ownership of Devices [Corporate/Personal]

- **Activity**
  - Roaming activity of Devices [Data/Messaging/Voice]
  - International roaming activity of Devices for current billing period
  - Number of devices that exceed the defined activity threshold
  - Data, Voice and Messaging Usage by Carrier

- **Applications**
  - Number of Enterprise Applications installed by month
  - Top 10 Enterprise Applications by installation status
  - Volume Licensing status for Enterprise applications
  - Top 10 Enterprise Applications by platform & installation status
  - Number of Enterprise applications that are out of version

- **Compliance**
  - Devices that are compromised (iOS & Android)
  - Number of Devices that have not connected in time
  - Number of iOS Devices without a password policy
  - Number of Devices that are out of compliance (by platform)
  - Number of Devices that have access violations (by platform)
Afaria Analytics
Voice Recognition is Getting Siri-ous

AA Flights today from JFK to PDX

Here's a list of AAL direct flights from JFK to PDX for today from this morning to tonight:
Next Up: Brain Waves!
StreamWork: SAP’s Collaboration Foundation
Collaboration in Analytics
<table>
<thead>
<tr>
<th>Title</th>
<th>Type</th>
<th>Last Run</th>
<th>Instances</th>
<th>SAP StreamWork</th>
</tr>
</thead>
<tbody>
<tr>
<td>Charting Samples</td>
<td>Web Intelligence</td>
<td>Sep 8, 2011 7:41 AM</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>Fold Unfold Samples</td>
<td>Web Intelligence</td>
<td></td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Formatting Samples</td>
<td>Web Intelligence</td>
<td></td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Input Controls</td>
<td>Web Intelligence</td>
<td></td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Input Controls and variables</td>
<td>Web Intelligence</td>
<td></td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
SAP Social Intelligence

Expertise location — Relationship Mining — Social Network Analysis
Business Analytics Solutions from SAP

Data Sources

Analytic Capabilities

- Analytic Applications
- Business Intelligence
- Enterprise Information Management
- Data Warehousing
- Collaboration
- Governance, Risk, and Compliance
- Enterprise Performance Management

Access

Services and Best Practices
SAP BusinessObjects Links

- Virtual Event [www.sapvirtualevents.com](http://www.sapvirtualevents.com)
- Visit our [Business Analytics resource center](http://www.sap.com/analytics)
- Learn the 4.0 solutions with [4.0 tutorials](http://www.sap.com/analytics)
- Join our official fan page on [Facebook](http://www.facebook.com) and follow us on [Twitter](http://www.twitter.com)
Conclusion

Big data is not just about big data
SAP Big Data Framework for volume, velocity, variety, validity

Revolution in analytic platforms
Back end: simplification, new real-time opportunities
Front end: mobile first, collaboration and consumerization

Innovation without disruption
SAP BusinessObjects: radically change your business without having to radically change your infrastructure…
Thanks!

You Should Follow Me on Twitter: @timoelliott

Email: timo.elliott@sap.com

Analytics Blog: timoelliott.com