BI²

Business Intelligence and
A Center of Excellence in Higher Education

Presented By: Christina Rouse

at

Ohio Association of Institutional Research

March 19, 2010
Presenter

Christina Rouse
Chief Architect
chris.rouse@incisiveanalytics.com
Mobile: 216 849-8237
Symptoms: Common Themes
Why is a BI Center of Excellence needed?

- Poor data quality
- Non-standard calculations
- Poor and inconsistent access to data
- Manual, laborious and untimely data manipulation
Data Quality State-of-the-Union
Survey: What is your data quality?

My organization thinks our data quality is?
  • Excellent, Good, OK, Poor

What is really the data quality?
  • Better Than Everyone Thinks
  • Worse Than Everyone Thinks
  • The Same As Everyone Thinks

What is the status of your data quality initiative?
  • No Plans
  • Under Construction
  • Design or Implementation Phase
  • Deployed
Since 2001, data quality has decreased

More organizations are aware of data quality issues

More data quality initiatives are being launched

Source: TDWI, Data Quality Survey, 2001 and 2005
Based on 647 respondents in 2001 and 750 in 2005
Data Quality
What’s the difference between reliability and validity?

- **Validity**
  - A data source value is true and is what it says it is. If a data element source is valid it is also reliable

- **Reliability**
  - A data source value consistently returns the same value which may or may not be valid. A reliable data source is not necessarily a valid data source
Need: A Common Statement

Your Institution:

- Desires comprehensive, campus-wide analytics to support decisions that:
  - Are proactive and focused at multiple management levels
  - Improve institutional performance and accountability
  - Establish repeatable compliance reports with ease
  - Enhance access and service for faculty and staff
  - Leverage technology assets for expanded access and service
  - Unify the digital campus

- Seeks advice on Business Intelligence tools
- Wants to explore available BI offerings from software vendors like SunGard Banner, PeopleSoft, Datatel and others
- Establish a business intelligence center of excellence with institution-wide governance
What is Business Intelligence? Isn't it the same as Data Warehousing?

**Business Intelligence** (BI) is the use of mathematical facts to improve business decisions. These facts typically focus on company performance across time and throughout the organization. BI is the use of a data warehouse.

**Data Warehousing** (DW) is the methodology and technology used to capture data from the company’s operational systems and then present the data in a meaningful way.

Business

'biz-n&s, -n&z
A usually commercial or mercantile activity engaged in as a means of livelihood

Intelligence

in-'te-l&-j&n(t)s
The ability to apply knowledge to manipulate one's environment or to think abstractly as measured by objective criteria (as tests)

**DW** is the technology. **BI** is the use of the technology to make better decisions that improve company performance.

Facts: Tuition, Lab Expenses, Facilities Costs, Enrollment Headcount by Course by Department by College, etc.
What is institutional intelligence?
Specific to Higher Education

• Academic Analytics
  – Reports
  – Analytics
  – Graphics
  – Projections

• With one…
  – Source of data
  – Calculation method
Institutional Intelligence

- A campus wide BI environment
- Built on world class BI tools
- Governed by an established council
BI Leverage

What does a BI solution enable us to do better?

- Know your student (and faculty) customer
  - Cross-offer
  - New student on-boarding
  - Retention
  - Loyalty

- More quickly respond to enrollment changes (i.e. enrollment management)
- Manage curricula to market demand
- Improve admission, registration and other process efficiencies
- Seek additional grants via better measurable objectives
- Make compliance reports routine; minimize time
- The process starts with **data** from many business sources
- **Data** is translated into **information** by adding business calculations and metrics
- Users (VP, Dean, etc.) look at **information** and turns it into **knowledge**
- This **knowledge** worker takes an action to improve operations
- Knowledge applied in context is **wisdom**
BI Strategy
Gaps, Strengths and Opportunities for Leverage

A comprehensive BI solution involves people, process, technology and data

- **People** – Create a BI Center of Excellence
  - Institutional research analysts
  - Information technology report writers
  - Financial analyst

- **Process** – Design a database that includes daily updates

- **Technology** – Microsoft, Business Objects, SAS, Cognos and many more

- **Data** – Turn data into information
  - Banner, PeopleSoft, Datatel, or homegrown transactional data
  - Institutional goals data
  - External sources like IPEDS, job market, voter registration, etc.
People: BI Center of Excellence
What is it? What does it do?

- Collection of people where the whole is greater than the sum of the parts
  - Institutional research analysts
  - Information technology report writers
  - Financial analysts
- Center of knowledge and understanding of institutional data
- Incubator of data into information, knowledge and ultimately wisdom
- Lead the culture of analytics!

- Create and publish executive dashboards
- Study the analytics of drivers and success
- Author operational reports
- Standardize reporting and data definitions
- Support “measurable objectives” in grant writing
- Produce IPEDS, State and other compliance reports
- Lead ad hoc research
- Organize the institution’s intelligence
Process: Data Warehouse
What is the best design?

- “One source of the truth”
- Star schema
  - Atomic level data
  - Conformed dimensions across stars
  - Single joins from fact tables to dimensions
  - Estimate 15 stars
- Move data once; update data at least daily
- Store data in business terms
P.I.T. Analysis of Course registrations

- By Student, Faculty
- By Course, Section
- By Campus
- By College of Course
- By Registration Status
- By Day in Term, Term
- By NCR Status
- Use a good modeling tool
- Adopt a naming scheme for databases, tables and columns
- Use surrogate keys
- Include the joins
- Keep model current
Q. If a star schema has 3 dimensions and one fact:
   - Class with 1000 rows
   - Time with 365 rows
   - Faculty with 2000 rows
   - Registration transactions fact table of 50 rows,

   • How many rows are returned in a query of select * with all dimensions joined?

A. 50
   • What else can you tell about the institution?
Data: Measures
Cumulative and Non-Cumulative

- Non-cumulative measures cannot be stored in the physical database; they must be calculated at the OLAP presentation point
- Non-cumulative measures typically have division, %, or are ratios like:
  - Completion Rate %
  - GPA
  - Retention Rate
Data: Classification
Data Type Determination

**NOIR**

- **Nominal** – Red, Yellow and Blue
  Just a name with no order or magnitude

- **Ordinal** – 1\textsuperscript{st}, 2\textsuperscript{nd} and 3\textsuperscript{rd}
  An order, but no magnitude

- **Interval** – 1\textquoteleft, 2\textquoteleft and 3\textquoteleft
  An order, a magnitude but absolute zero

- **Ratio** – -2\degree, -1\degree, 0\degree, +1\degree, +2\degree C
  Below zero possibilities

Nominal and Ordinal data type elements are dimensions and Interval and Ratio data type elements are facts in the data model.
Data: Behavior Statements
Slowly Changing Dimensions

- Slowly Changing Dimensions:
  - (Type One), doesn’t preserve history
  - (Type Two), preserve a version of history
  - (Type Three), Hybrid of Type One and Two

- Users typically want a “Type Two” methodology of SCD
- A Type Two change writes a record with the new attribute information and preserves a record of the old dimensional data.
- Type Two changes let you preserve historical data.
- Implementing Type Two changes, after the fact, will require significant analysis and development.
- Type Two changes accurately partition history across time more effectively than other types.
- Because Type Two changes add records, they can significantly increase the database’s size.
Balanced Hierarchies
• All branches of the hierarchy descend to the same level, and each member’s logical parent is the level immediately above the member

Unbalanced Hierarchies
• Branches of the hierarchy descend to different levels. For example, an Organization dimension contains a member for each employee in a company

Ragged Hierarchies
• The logical parent member of at least one member is not in the level immediately above the member. This can cause branches of the hierarchy to descend to different levels.
# Technology: Tool Options

What are the best tools for us?

<table>
<thead>
<tr>
<th>Data Modeling</th>
<th>Target Database</th>
<th>ETL (Extract, Transform &amp; Load)</th>
<th>OLAP (On-Line Analytical)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AllFusion ERwin Data Modeler</td>
<td>SYBASE</td>
<td>INFORMatica</td>
<td>COGNOS</td>
</tr>
<tr>
<td>ORACLE</td>
<td>Ascential</td>
<td>ProClarity</td>
<td></td>
</tr>
<tr>
<td>Teradata</td>
<td>Business Objects</td>
<td>Business Objects</td>
<td></td>
</tr>
<tr>
<td>IBM DB2</td>
<td>SAS</td>
<td>SPSS</td>
<td></td>
</tr>
</tbody>
</table>

© Incisive Analytics LLC
Data: Subject Matter
What data are included in institutional intelligence?

<table>
<thead>
<tr>
<th>Transactional</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Student</strong></td>
<td><strong>Employee</strong></td>
<td><strong>Financial</strong></td>
<td><strong>Grant</strong></td>
<td><strong>Alumni</strong></td>
<td></td>
</tr>
<tr>
<td>• Prospects and Admissions</td>
<td>• Employee Foundation</td>
<td>• General Ledger</td>
<td>• Grant Management</td>
<td>• Gifts and Donations</td>
<td></td>
</tr>
<tr>
<td>• Student Course Registration</td>
<td>• Employee Application</td>
<td>• Accounts Receivable</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Prospect Financial Aid</td>
<td>• Employee Degree</td>
<td>• Budget Ledger</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Student Financial Aid</td>
<td>• Employee Position</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Completions and Degrees</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>External</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Voter</strong></td>
<td><strong>Marketing</strong></td>
<td><strong>Economy</strong></td>
<td></td>
</tr>
<tr>
<td>• Registration Trends</td>
<td>• Database America</td>
<td>• Consumer Price Index</td>
<td></td>
</tr>
<tr>
<td>• Voter volume</td>
<td>• Socioeconomic Profiles</td>
<td>• Job Market</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Higher Ed</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Compliance</strong></td>
<td></td>
</tr>
<tr>
<td>• Federal IPEDS</td>
<td></td>
</tr>
<tr>
<td>• State Reporting</td>
<td></td>
</tr>
<tr>
<td>• Others</td>
<td></td>
</tr>
</tbody>
</table>
Dashboard and Reports
What are some examples?
Timeline

Implementation – Phases for Success

January 2010 – February 2011

Phase I
Student, 23 Weeks

Phase II
Employee, 10 Weeks

Phase III
Financial, 7 Weeks

Phase IV
Grants, 7 Weeks

Phase V
Compliance, 5 Weeks

Training
Adopt Vision and Knowledge Transfer