

Background and Experience

Throughout TSI's Higher Education experience (see higher-education.transforming.com) there has been a consistent need to bring disparate sources of data together in a coherent manner to create education data warehouses that improve decision making quality and capabilities for a host of constituents. TSI is in a unique position to assist Higher Education in these efforts as we couple our higher education experience along with our technical experience in ETL (Extract Transform Load), data modeling, data warehousing and analytics/decision support. Provided below are a few recent examples of our experience in this area:

- *Rapidly-growing, Study Abroad Organization - Chicago, IL* - As part of a process improvement and technology re-engineering effort, TSI worked with various departments (Academics-Program Deans, Enrollment Management, Student Services, Marketing, Finance and IT) and a 3rd party OLAP vendor to architect and manage a solution whereby data was extracted and combined on a monthly basis into a SQL Server data warehouse using: key Student Information System (SIS) Oracle data from 33 campuses, General Ledger SQL Server data from the accounting application and budget data from various Excel spreadsheets. Using the 3rd party OLAP application, users were able to make key strategic budget decisions and dynamically update their budgets.
- *Large Public University - Urbana, IL* - TSI played a key role on this critical project to develop an Advancement Technology Strategy emphasizing the sharing of data across its four campuses. This project identified important process and policy changes that needed to occur to achieve a 360 degree view of constituents while identifying important security considerations and policies. During this project, TSI facilitated workshops on each campus and for each college and many departments. This project resulted in the implementation of a number of technology recommendations that enhanced and extended the current DW (Data Warehouse) and BI (Business Intelligence) solutions and leveraged the current investment in technology.
- *Large Healthcare Professional Education and Accreditation Organization, Northbrook, IL* - TSI led a project to improve the efficiency and utility of the IT Department including the technology processes, people and infrastructure that supports this \$150M organization. TSI quantified the number of vendors and disparate software and database products at over 100 companywide and analyzed how to improve the software capabilities, processes and data quality by adopting more of a vendor centric approach. TSI quantified the true technology spend and by implementing the recommendations, this client will reduce its technology spend by 50%. This will also allow the professional education and accreditation departments to spend less time pulling data from several different applications/systems and spend more time on strategic decision making.

In addition to Higher Education, TSI has a wealth of experience in food and drug industries (over 20 years) concerning strategic decision support. The food and drug industries have been leaders in strategic decision support and their industry leading best practices are increasingly being used in Higher Education. TSI staff has:

- Managed major strategic decision support applications for large clients involving data integration, data aggregation, data warehousing, and business intelligence reporting/graphics against point of sale (POS) and customer data
- Worked with leading clients such as: Abbott, Coke, D&B, Dole, Colgate, IMS, Kraft, Miller, Nestlé, P&G, Quaker, RJR, Unilever and Walgreens.

With TSI's mix of Higher Education experience and other industry experience (e.g. food and drug) on the next few pages are 4 steps we take for each strategic decision support project, namely:

1. Planning the Education Data Warehouse for Strategic Decision Making
2. Planning ETL Processes and Data Warehouse Creation
3. Detailed Analysis
4. Project Planning / Execution / Implementation

Approach:

1. Planning the Education Data Warehouse for Strategic Decision Making

TSI's approach is collaborative to involve key stakeholders to understand their current and future information needs and, at times, their underlying business process issues. This knowledge is the foundation that enables us to assist in the solution architecture and define how data sources are populated, collected, maintained and analyzed.

For an education data warehouse we frequently work with the following types of data sources:

- Student Information (e.g. Recruit to Enroll, Curriculum, Scheduling, Class Assignment)
- Advancement (e.g. Alumni, Donors, Foundations, Government Agencies, Corporations)
- Financial (e.g. Student Payments for tuition/housing/meals, Financial Aid/Awards/Scholarships)
- Budget (e.g. Forecast and Actuals across Departments/Campuses)
- Human Resources (e.g. Faculty, both in-house and external)
- 3rd Party information (e.g. Government Census data, IPEDS, competition data from other universities/higher education facilities)
- Other (e.g. Grants)

Our next step is to create a "360 view" of all the data sources. It is important to get this understanding with the stakeholders prior to commencing any ETL work so as to avoid re-work later on. To do this "360 view", TSI typically takes a *dimensional* approach when creating an education data warehouse. Here are some *examples* of these dimensions:

- Constituents - Student, Alumni, Donors, Demographics
- Location - Campus, Colleges, Schools (onsite and remote)
- Time/Period - Year, Semester, Term
- Program - Majors, Courses, Disciplines
- Previous Education - High School, Other Institutions
- Fact/Measures - GPA/Grades, Tuition, Fees, etc.

Within each dimension it is important to identify the types of attributes (also known as characteristics) that need to be collected. For example, some of these for the Student dimension might be: Gender, Major, Birth City, Ethnicity, Income level and Age. These attributes serve the basis for analysis that lead to strategic decision making. Next we work with the key stakeholders to determine the hierarchies for each dimension (e.g. for the Time dimension: weeks, months, semester, year might be the hierarchy). Special business rules also need to be reviewed to ensure any rollups are done properly (e.g. cumulative GPA is the mean not the sum of the GPA from all the terms for a given year). There may also be special business rules for not showing or suppressing certain data elements (e.g. Student SSN number).

Next, TSI will data model the warehouse creating both logical and physical data models. DDL (Data Definition Language) will be generated that can then be used to create a prototype data warehouse to be reviewed with the key stakeholders (this is often populated with some sample data). TSI also has experience with DDL tools such as CA ERwin Data Modeler and Embarcadero ER/Studio.

2. Planning ETL Processes and Data Warehouse Creation

Once the data sources and dimensionality have been identified and agreed to with the stakeholders, TSI then formulates the ETL (Extract Transform Load) routines to create the education data warehouse from the base data sources (often referred to as the operational database stores).

Provided below are some key points we consider during the ETL process:

- Confirm the availability of all data sources with technical staff and/or vendors.
- Review security requirements with technical staff and/or vendors.
- Use the appropriate toolset for extraction including API programs, web services, scripts/languages (e.g. Oracle PL/SQL or SQL Server Information Services (SSIS)) or third party tools (e.g. Informatica)
- Determine extract timing (e.g. run daily at midnight) and extract trigger events
- Decide on how to handle/notify missing files, missing data rows, missing fields
- Filter out non-necessary records
- Confirm extract/load performance (e.g. can it be handled in a timely manner)
- Plan for re-extraction routines if partial/total extract failures are encountered
- Perform various transformation/calculations
- Perform cross-referencing (i.e. coming up with common lookups, e.g. different discipline codes across departments)
- Store and index into the education data warehouse
- Create data marts of the education data warehouse (optional)

There are also some unique considerations for education data, for example:

- FERPA rules (e.g. who is allowed to see the student's record information)
- SSN privacy/masking for student data
- Handling cross discipline courses (MATH101 = COMPSCI203)

3. Detailed Analysis

Once the data warehouse has been created, there may be further procedures to load and manipulate data for the business intelligence applications being used. TSI has worked with a variety of business intelligence application vendors and analysis tools including:

- Microsoft Analysis Server (OLAP)
- PowerOLAP
- Cognos
- Microstrategy
- Business Objects
- Oracle-OBIEE

From the business intelligence applications, generally ad-hoc queries are run or scheduled queries are run to "feed" syndicated reports or dashboards.

It should be noted that TSI has no affiliation with any vendor and concentrates on working with the key stakeholders to define the best data mining and analytic techniques with the given applications. TSI can also lead unbiased reviews of the various business intelligence applications and analysis tools with the stakeholders so the right choice is made.

4. Project Planning / Execution / Implementation

Since the creation of the data warehouse and use in business intelligence applications crosses many organizational boundaries, ownership issues, security policies/maintenance and reporting processes in a higher education environment, TSI's collaborative approach helps to understand and coordinate these related issues across the various campuses, colleges and departments. TSI is experienced in project execution including the entire testing lifecycle and is often called upon to direct and assist in Use Cases/Test Cases/QA/UAT, training, and implementation.

Conclusion

In conclusion, TSI has the experience in the following areas of strategic decision support:

- Designing, architecting and implementing higher education BI/DW solutions
- Working with higher education stakeholders to plan the right data warehouse (either initial or advanced) for strategic decision analysis
- Taking inventory of disparate data sources, creating an approach to cleanse data
- Bringing disparate data sources together, performing the necessary ETL and creating the right data warehouse tailored to the stakeholder's needs
- Data modeling creating the proper logical and physical designs and prototyping the data warehouse
- Selecting the proper BI and analytical tools against the data warehouse
- Defining the associated processes that govern data usage, security, and reporting
- Project planning and training stakeholders