2010-11 Budget 'CALL'
Data Warehousing Initiative

April 8, 2010

FY 2009-10 Update

The budget for the Data Warehousing initiative in FY 2009-10 was $30,000 (although to date only $18,000 actually has been allocated). All DW personnel reside in both IT and IPA. The DW Team includes all staff in IPA: Director, two Principal Research Analysts, one Systems Manager (Programmer), and one Data Administrator (Programmer), all of whom have full-time responsibilities in IPA. Also on the DW Team are four IT staff members: CIO and SIS Team Manager (both of whom with full-time responsibilities in IT), DW Engineer and ODS Programmer/DBA. The DW Engineer is assigned 100% to the DW project. Until recently, the ODS Programmer (who joined the Team in Spring 2009) was 50% but, as of February 2010, is now 100% assigned to the project. (Temporarily, however, he continues to support SIS projects unrelated to the ODS .)

Despite the resource challenges, we have made some significant progress in the past nine months:

- Cognos report training (provided by a consultant over five consecutive days in July 2009). In addition to the DW Team, staff members in 4 functional offices were trained: Admissions, Financial Aid, Registrar's Office, Student Business Services, as well as IT staff supporting COEUS and other applications.
- Following good practices for any applications, two ODS/DW environments were set up: development and production.
- Initial security was established both within the database and the Cognos tools.
- Data feeds, with nightly refreshes, were established for Banner SIS and the payroll/personnel and financial/budget system, QDB.
- In addition to the delivered ODS data views, ODS/DW programmers designed and launched other views to support known reporting needs. These views imbed sometimes very complex logic into the database tables to make reporting much easier and more consistent. For example, the logic for defining variables such as employees' primary appointments, total annual salary and fte (across all appointments), and flags to indicate whether employees are included in official snapshots will greatly simplify reporting for many campus users.
- ODS/DW was upgraded from 3.1 to 8.1. This was a prerequisite to the Banner SIS upgrade, which currently is underway.
- Crystal Report programmers are implementing a plan to migrate to Cognos. The projected completion date is Summer 2011. This will allow UC Merced to phase out Crystal Reports, as the campus terminated Crystal support and upgrades over 2 years ago.
- The first (prototypical) dimensional model (Admissions) has been designed and is being developed. Target deployment of the first iteration is planned for May 2010. This model, integrated with an Enrollment dimensional model, will be the basis of an

Enrollment Management & Planning data mart. Dimensional models transform the raw data (from various sources) into logical data groups that make it easier for reporting and allow analysts to drill down from summary to detailed data. They optimize decision-support query performance.

- In addition to the report writers who were trained last July, other users have been granted access to the ODS. Two staff members in the Dean's Office in the School of Natural Sciences have access to a view created by the DW Team. This has enabled the Dean's staff members to run simple reports, much like the Crystal reports that had been provided by SIS programmers, except that now they can modify the reports as needed. Access to OD also was given to an analyst in Academic Personnel to load course data into Digital Measures. Sponsored Projects uses personnel data and specialized functions from the ODS in their new COEUS application.
- The DW Team is in process of getting all the programs, tables, and scripts from UC Davis to implement our own version of Desii, an application that supports faculty workload, instructional activity, and classroom/space utilization analyses. We have nicknamed this new application MERii (Merced instructional information).
- Training (framework manager) planned for DW programmers (May).
- Used QDB data from the ODS for the Applied Math Academic Program Review.

DW Role in Assessment and WASC-Related Areas

The Data Warehouse represents the campus' access to assessment data and analyses. It is designed, within a secure environment, to provide the *windows and doors* to integrated campus data to be used in assessing and evaluating programs and strategies, helping campus leaders to make decisions about resource allocations and the efficiency and effectiveness of programs. The DW transforms data into information that is easily understood, consistent, comparable across time, and transparent to the campus and public constituents. Ever evolving, the DW will expand as the need for data and information grows.

Examples of the types of data marts or decision-support systems that will be developed include: Enrollment Management/Planning (including Admissions, Enrollments, Financial Aid), Faculty Workload & Academic Resource Management, Graduate Support and Enrollment Planning, Contracts and Grants Management, Space Utilization and Planning, Human Resources Management, Budget & Staffing, Donor Tracking, Student Achievement/Success.

DW Staffing Needs

Since much of the Data Warehousing work depends, first of all, on bringing the source data into the ODS staging area, a high priority is to add ***one full-time (permanent or contractual) ODS Programmer Analyst (III-IV)*** to support the ODS. This infrastructure work is critical to the pace at which the team can deliver usable reports and data marts and to the ultimate success of the DW. In addition to setting up the appropriate pathways to bringing source data into the ODS, this position would help support the Cognos server setup, security, upgrades, and ETL functions such as data cleansing options, as well as provide backup to the existing ODS programmer analyst.

Request: IPA requests one permanent fte Programmer III- IV to help develop and manage the ODS data-load processes (ETL, etc.) for various source data systems, work with business users to understand their information/reporting needs from the ODS and create views and frameworks to satisfy those needs.

$71,000-$96,000 + Benefits

In order to make better progress in creating dimensional models and integrated data marts, as well as dashboards and scorecards, it is critical to add at least ***one permanent, full-time DW programmer*** to the Team. This would be a high-level programmer who has experience building dimensional models and data marts, has worked on a comprehensive DW project, and has demonstrated success in working with business users, translating their reporting and information needs into appropriate deliverables (data views, data marts, etc.).

Request: IPA requests one permanent fte Programmer III- IV to work with business users to translate their information, reporting and decision-making needs into reports, data marts, dashboards, and/or scorecards that successfully meet those needs.

$71,000-$96,000 + Benefits

DW Application/Hardware Needs

When purchasing the Cognos BI tools, the DW Team assumed that the Oracle Data Warehouse Builder (OWB) ETL tool that the campus already had licensed for other purposes would be adequate for the DW initiative. An ETL tool is extremely important in any DW project, as it is the utility that moves the data from various source systems into the ODS, then transforms and moves the data into the DW (dimensional models & data marts). A good ETL tool also performs some data cleansing activities by identifying data problems (missing data, bad codes, etc.) while moving the data and then notifying the appropriate source system users who need to fix the problems. This data quality function is one of the important added-values of any DW project. The ETL also performs over 30 other tasks that meet the demanding requirements placed on the data warehouse. Within a few months, it was obvious that OWB would not be an adequate ETL tool for UC Merced's DW environment. OWB does not integrate well with Cognos, as it primarily was meant to integrate with Oracle's BI tool: Discoverer.

In the absence of an adequate ETL tool, the programmers resorted to using SQL to develop their own processes for moving and transforming the data, as well as for some quality control functions. Continuing down this path will greatly impact the timeline for deliverables and ultimately will increase the overall cost of the project. The longer we postpone the purchase of a ***suitable ETL tool***, the more SQL code that will need to be reworked.

Request: Funding for ETL application. (One-time cost; in subsequent years there will be annual maintenance costs, estimated @ 15% of purchase price.)

$90,000 - $150,000

Training and Consulting Needs

Because of the complexity of developing a data warehouse, adequate and targeted training to team members is critical to their success in transitioning from the relational model to the dimensional model. The team will be more productive and efficient if provided training on best practices in developing data warehouse data marts. Timely design reviews by professionals would also be cost-effective, accelerate the learning curve, and ensure adherence to standards.

Request: Funding for training and consulting. (One-time cost; with annual review to determine subsequent needs.) **$40,000 - $60,000**