Data Warehouse-General

Administrative and Enterprise Applications
Office of Data Administration
Data Warehouse

Introduction ..................................................................................................................................... 3
DW Plan/Methodology ................................................................................................................... 3
  Concept ........................................................................................................................................ 3
  Principles ..................................................................................................................................... 3
  Methodology for Building the UM Data Warehouse ................................................................. 4
Architecture .................................................................................................................................... 5
  Relational tables holding atomic level data + metadata .............................................................. 5
  Snapshot of DW data becomes “frozen” data for institutional reporting .................................... 5
  Data Marts built from atomic level data ..................................................................................... 5
Infrastructure ................................................................................................................................... 5
  Hardware & Configuration .......................................................................................................... 5
Personnel Support and Partners ..................................................................................................... 6
  OIT- Office of Data Administration: .......................................................................................... 6
  OIT - AEA Administrative Enterprise Application Support Services ....................................... 6
  OIT - Technical Services and Support ........................................................................................ 6
  Institutional Research, Planning & Assessment ........................................................................ 6
  Data Stewards/Data Managers .................................................................................................. 6
  Campus Users of Data ................................................................................................................ 6
Security/Access ................................................................................................................................. 6
  UM Data Warehouse Account Request Form ............................................................................ 7
Users ............................................................................................................................................... 7
  Client Users ............................................................................................................................... 7
  Web Users ................................................................................................................................. 7
Training ........................................................................................................................................... 7
  Brio Intelligence Explorer Tool Training .................................................................................... 7
  Data Subset Training .................................................................................................................. 8
User Support .................................................................................................................................. 8
  Help Desk ................................................................................................................................... 8
  DW User Group .......................................................................................................................... 8
  DW Web Site .............................................................................................................................. 8
Data Warehouse Uses ..................................................................................................................... 8
Benefits ........................................................................................................................................... 10
**Introduction**

This document outlines general information about the Data Warehouse. This document includes DW: concepts; principles; methodology; architecture; infrastructure; personnel support and partners; security access; account request form; Users; training; DW uses and benefits.

There also exists a companion document that outlines details that concern the Data Warehouse. These details include: the current status of the construction by DW data subset and the administration of the data warehouse. Also included is: a list of current, officially released tables; a list of current web server groups used to publish reports using the online data warehouse tool (Hyperion); and measures of accountability. It can be found in dataadmin/all/dwstats/dwsummary_details.doc.

**DW Plan/Methodology**

**Concept**

In 1996, as an outgrowth of the recommendations of the PARADO CQI Team, the Office of Data Administration, in partnership with other offices on campus, began to design and implement a data warehouse. The goals of the project were to provide an integrated data infrastructure which would serve as the institutional database for the Institutional Research, Planning & Assessment, provide data access to campus service offices for management of their data subsets, and provide data access to departmental and college end users as appropriate for management and strategic planning purposes. The strategy was to extract data from legacy transactional systems, restructure it by data topic, de-normalize it, and place it on a platform(s) accessible to end users. The resulting data warehouse originated as a joint project with the Office of Data Administration, Administrative and Enterprise Applications (AEA), the Institutional Research, Planning & Assessment, campus data stewards and managers, and end users of institutional data.

The UM Data Warehouse contains a collection of data subsets which have been extracted from various campus transactional databases, structured for query and analysis, and located on platforms which can be accessed using standard SQL query tools. The data warehouse consists of several different types of data: transactional, frozen (moment in time) and year-end. In order to achieve consistency of data, a core of defined data elements are identified for each data topic to be included in the warehouse, data standards and principles are applied to the data, derived data elements are developed, and data from transactional systems are loaded into a transactional data warehouse. Frozen data, for institutional reporting, are then taken as a snapshot (moment in time) of the transactional data and placed in a frozen data warehouse. Some data subsets, such as financial data, may be loaded with cumulative totals at the end of a fiscal year.

**Principles**

- The data warehouse will be designed to meet the needs of multiple constituencies, including service offices, campus administrators, end users.
• The data warehouse activity involves a **partnership** with the Office of Data Administration, Institutional Research, Planning & Assessment, Administrative and Enterprise Applications, campus data stewards and managers, and end users to build a data repository for the campus.

• The warehouse data will be **de-normalized for ease of use** by the end user. If it is determined that a normalized table should be maintained for analytical purposes (may be desirable for IRPA analysis), it must be structured so that it can be rolled into a de-normalized version for the end users.

• Each data subset (transactional, frozen, year-end) will be based upon the same standard set of tables and field element names so that there is consistency across data subset types.

• All tables and elements in the UM Data Warehouse will be reviewed by the Office of Data Administration before database structures are built to ensure consistency and adherence to campus data standards.

• Derived fields will be used where possible to provide consistent data definition for pulling complex data sets and to provide more effective user-oriented access to data.

**Methodology for Building the UM Data Warehouse**

The strategy has been to create a solid, supportable infrastructure for the UM Data Warehouse. To date, resources have been supplied by the Office of Data Administration, Administrative and Enterprise Applications, and the Institutional Research, Planning & Assessment. With this infrastructure, the following iterative process has been established for the development of the warehouse:

• Target the data subsets that have targeted business requirements as determined by the Institutional Research, Planning & Assessment, service offices and end users for inclusion into the warehouse.

• Partner with Office of Data Administration, Institutional Research, Planning & Assessment, data stewards and managers, and the campus community to model the data table structures.

• Write the transaction system extracts of the data, build the warehouse structure, and load the data into the warehouse.

• Research and catalog the corresponding metadata to thoroughly describe the data and make it available to the campus community.

• Determine security and access needed for each data subset as determined by the corresponding Data Steward and campus policy.

• Thoroughly test data before release.

• Build repository and/or web queries.

• Train the campus community on supported query tool (optional) and data subsets before access to data is granted.

• Provide help desk and user group support.
Architecture

Relational tables holding atomic level data + metadata

The UM Data Warehouse is a collection of atomic data that are arranged around subject areas - personnel, financial accounting, budget, student records, courses, etc. It represents the lowest level of granularity of university activity over time.

Snapshot of DW data becomes “frozen” data for institutional reporting

Frozen data, for institutional reporting, are then taken as a snapshot (moment in time) of the transactional data and placed in a frozen data warehouse.

Data Marts built from atomic level data

Using atomic data, Data Marts have been developed for decision support based on the needs of a given department or targeted campus-wide business requirements. UM Data Marts are dependent and their source data come from the UM Data Warehouse. This strategy enables a consistency in standard data naming conventions, data definitions, data quality, manageable interfaces and efficient data maintenance.

Infrastructure

Hardware & Configuration

DW Server
   Database: Oracle 10.2
   Server/OS: Sun Fire T2000/Solaris 10
   Processor: 32
   Memory: 16 GB
   Size: Oracle tablespace total = ~400 GB
   Server Name: Datum (aka Utopia)

On-Demand Server/WOW Server
   Operating System: Win 2000
   Server: Dell PowerEdge 6450
   Processor: 4- 700MHz processors
   Memory: 4 GB of memory
   Size: Hard Drives 1-7.8 GB, 2- 58.9 GB
   Server Name: Hypprod

Frozen DW Server
   Database: Oracle 10.2
   Server/OS: Sun Fire T2000/Solaris 10
   Processor: 32
   Memory: 16 GB
Personnel Support and Partners

OIT- Office of Data Administration:
- Directs DW project (determine scope, analyze requirements, evaluate source systems, design table structure, develop data access, research and catalog metadata, test data, administer repository and web server, administer data access, develop training courses, train users, user support).

OIT - AEA Administrative Enterprise Application Support Services
- DBA staff provides DBA support, builds Erwin models, builds Oracle tables, attaches tables to roles, and builds data marts.
- Programmers & Analysts of administrative applications that have data in the warehouse write extracts, work with DBAs on table structure, and make changes as required.
- Programmers & Analysts of SIS and ARS applications that have data in the warehouse write extracts, work with DBAs on table structure, make changes as required.
- DBA staff installs & configures Brio web server.

OIT - Technical Services and Support
- Systems staff maintains current DW servers – back up, software installation & configuration and operational support.
- Systems staff maintains frozen warehouse server.

Institutional Research, Planning & Assessment
- DBA for frozen warehouse Oracle tables, builds additional IRPA tables as required, builds data marts/data views, builds and maintains Java-based metadata application/bridge application/security application, converts past 10 years of frozen data to match data warehouse structures, converts IRPA reports to Brio reports

Data Stewards/Data Managers
- Data Stewards - Officers of a department or unit, or their designees, having policy level responsibility for a specific subset of data; e.g. Associate Provost for Research.
- Data Manager - University officials and their staff who have operational level responsibility for a specific subset of data; e.g. Comptroller's Office.

Campus Users of Data
- Data User - UM faculty, staff member or student worker who has access to a specific subset of data; the data user may use specific subsets of data in the performance of his/her position (e.g. Department Chair) or may have access to information about himself/herself.

Security/Access
Data in the data warehouse is used by different groups on campus -- Institutional Research & Planning, service offices, and end users. Because of the unique and or sensitive nature of some data elements, data will be secured and access determined by the Data Steward responsible for each data subset included in the data warehouse.

Access is assigned to the data warehouse based on data subset access determined by the appropriate Data Steward and the person's functional role in his/her unit as authorized by the unit head. In addition, data are also secured by unit, college and division. Depending on a user's function and authorization, he/she may have access to some or all of the data for a defined data subset. In general, a departmental user may only have access to his/her own data (department, college, or division as appropriate), similar to viewing of transaction system data.

UM Data Warehouse Account Request Form

http://www.oit.umd.edu/DataAdmin/DataWarehouse/Checklist/DWacctform.html

Users

Client Users

Client users build their own ad-hoc queries utilizing query tool software. The query tool allows them to access the data warehouse tables/views, select their query criteria, retrieve data and format reports & charts. Any SQL-based Oracle compatible query tool can be used with the UM data warehouse. The query tool supported by the campus is Hyperion Intelligence Explorer which provides point-and-click access to the warehouse on multiple platforms. Purchase of the software is necessary and available via OIT Software Licensing.

Web Users

Web users access secured data and reports via the web for queries created by Hyperion client users and loaded to the Warehouse on the Web (WOW) server. WOW is accessible via current versions of web browsers such as Internet Explorer, Netscape, Firefox and Safari, No special plug-in or helper application software is needed for WOW.

Training

Brio Intelligence Explorer Tool Training

Training for Brio Intelligence Explorer, the query tool supported by campus, is provided by The Office of Data Administration. The course covers: familiarizing users with the Brio Intelligence Explorer client software, creating & refining (limiting) queries, formatting & exporting results, creating pivot & detail reports, using the repository, and printing & saving reports.
Data Subset Training

UM Data Warehouse policies and procedures are based upon the requirement that all data warehouse users must receive data training for the data subset(s) to which they are to be given access. Data training is developed and provided by the Data Manager responsible for the specific Data Subset. The Office of Data Administration, in an effort to assist the Data Managers, has developed and provided some of the Data Training courses. At some time in the future, the Data Managers will assume the maintenance and teaching of these classes.

User Support

Help Desk

http://www.helpdesk.umd.edu

DW User Group

http://www.oit.umd.edu/DataAdmin/DataWarehouse/Dwusergroup/

DW Web Site

http://www.oit.umd.edu/DataAdmin/DataWarehouse

Data Warehouse Uses

Offices from around the campus have increasingly utilized data from the University of Maryland’s Data Warehouse for analysis and decision-making. The information has been provided to campus information users in a variety of formats, ranging from: queries and results accessed through email; queries stored in the campus repository run at the users desktop; queries processed on line from a data view developed for a specific business application; or queries run by ODA staff and mailed to the user via paper.

The examples cited below reveal the many ways in which the warehouse is assisting the campus in becoming more productive data users. The list also shows the breadth of information currently available on the warehouse and hints at the possibilities for other analyses across information systems. Ultimately the Data Warehouse should allow the campus to more effectively serve students by providing information key to sound decision-making.

The following are examples of how information from the Data Warehouse (DW) is being used around campus:

- An integral part of the Payroll & Human Resources (PHR) system is the Warehouse on the Web (WOW) reporting mechanism. Setting up employees for HR and payroll involves inputting transactions into the web-based PHR application screens and managing those transactions via web-based reporting. The PHR WOW reporting
structure provides information on employee demographics, appointments, time/leave, pay and security. Reports fall into two categories: (1) reports for maintaining and processing PHR transactions and (2) reports for general information purposes only.

- Student Financial Aid is using the warehouse to produce the Scholarship Disbursement Account report on WOW so that authorized college business managers can produce the report on demand in their offices without having to request that the Student Financial Aid produce it manually for them.

- Office of Research Administration and Advancement is using the warehouse to produce WOW reports for internal use about the research awards they manage. Examples of this include reports that provide ORAA staff with detailed information on awards received from the State of Maryland; reports which detail projects that have been awarded $500,000 or more; and a report that compares award totals between two past fiscal years. ORAA is also using data in the warehouse to deliver reports to campus users which documents the university's proposal, award and expenditure activity detail, and make that data available for distribution via their website: http://www.umresearch.umd.edu/ORAA/data/

- Academic advisors in the departments and colleges are using the data warehouse as a management tool to retrieve data for students in their majors in conjunction with SIS and Advise on the Web.

- The Federal Work Study Office within the Office of Student Financial Aid has produced a query available on WOW that allows campus employers to track the status of the remaining award money allotted to their federal work-study students.

- ODA used the data warehouse report to develop a report on summer appointments and retirement pay for Academic Affairs. The report provided summary information on the decision of whether to include summer pay as part of the retirement calculation for 9.5/10 month faculty. The query listed the 3 types of summer appointments found in the Academic Resource System along with the retirement types for faculty in Human Resource System.

- IRPA uses the data warehouse extensively for all of their institutional freeze data (ex. degrees, student enrollment, FRS, scheduling, faculty workload, graduate applications, end of term, personnel) which is used to produce statistical data for the State of Maryland. They also use PHR data for the information reported in the Departmental Indicators Book that is used extensively by the College in their planning activities.

- IRPA used scheduling and personnel information from the data warehouse to support their faculty workload system.

- Payroll information from the warehouse was used by the College of Agriculture and Natural Resources to compile a list of employees who were being paid out of MCE accounts and the percent of their total funding coming from those accounts.
Advise on the Web uses the data warehouse as the source of its data.

OIT Business Services uses the data warehouse to access an OIT Disaster Recover report which contains OIT employee information such as business and permanent address and phone data, cell/phone, pager data and emergency contact information.

**Benefits**

- Ability to answer business questions that require data from multiple data subsets
- Ability to answer business questions on a more timely basis
- Ability to analyze data over time
- Ability to analyze impact of decisions before making them
- Provide focal point for IT systems to look at the institution as one enterprise, rather than disparate offices.
- Consistent data definitions, coding and naming across data subsets
- Consistent data definitions, coding and naming across current and frozen data sets
- Communication among diverse Services Offices
- Improvement in data quality by making data available then fixing errors in source application system
- Common source of data for official institutional reporting
- Allows for converting and merging data via crosswalk bridges
- Empowers users
- Reduction in time to build and run reports – eliminated batch report writing
- Elimination of batch report writing freed up programmers to work on new application development