

Managing Your SAS® Data Warehouse Project from Design to Delivery

Linda S. Sullivan, University of Central Florida, Orlando, FL

M. Paige Borden, University of Central Florida, Orlando, FL

ABSTRACT

A data warehouse project is a significant resource investment for an institution. Effective management of the project is critical to ensure its growth and continuing administrative support. This paper presents how the Office of Institutional Research developed and implemented a university data warehouse which serves as the foundation for the development of a wide variety of reporting applications using SAS Business Intelligence software. It provides a full discussion of the project management from design to delivery, including both functional and technical perspectives, the development and participation of a university stakeholder group, project plan challenges, lessons learned, and tips on what to include and avoid in a project of this type and scope.

WHO WE ARE

Mission

The mission of the Office of Institutional Research is to provide information of the highest quality, and which is both timely and easily accessible, and to facilitate and enhance decision-making, strategic planning, and assessment at the University of Central Florida.

Purpose

Institutional Research (IR) provides electronic and web-based dissemination of official information to the University community (including the Board of Trustees, the various colleges, departments and other academic and administrative units), external agencies, and the Florida Board of Governors (BOG). IR generates, supervises or develops all official University data reports and state-required reports, as well as providing end-user data solutions and training to facilitate decision-making processes. The director and staff serve on numerous university-wide committees and workgroups and assist with the collection and interpretation of institutional data, assist in planning academic programs, and participate in the implementation of evaluative procedures. The functions of the office support the entire university enterprise.

Background

Institutional Research has, as one of its primary objectives, the task of developing and implementing a university data warehouse. This enterprise-wide data system will facilitate the creation, access, and dissemination, by internal and external stakeholders, of institutional knowledge pertinent to the university. In collaboration with the data warehouse project stakeholder group, which represents a cross-section of key data users, the IR staff has established a data source of ten years' worth of student reporting data. This warehouse of information serves as the foundation for the development of a wide variety of reporting applications using SAS Business Intelligence software.

INTRODUCTION

Today, many higher education institutions are struggling with increased competition and limited resources. Administrators are looking for accurate and timely information that will help guide their strategic and operational decision making: admissions offices are attempting to analyze where applicants are coming from and what percentage actually enroll so that marketing dollars can be spent wisely and efficiently; academic affairs departments are monitoring graduation rates and time to graduation so that the right mix and frequency of class offerings can be scheduled to enable students to complete their degrees on time; student development offices need to identify at-risk students so that the appropriate intervention and support programs can be made available; data on cost efficiencies of program offerings and faculty salaries are necessary for resource allocation and future growth planning.

Data warehouse projects can provide a robust and flexible solution to these information needs. However, as with other enterprise-wide technology solutions, a data warehouse solution is a significant resource investment for an institution. Effective management of the project is critical to ensure its growth and continuing support.

This paper chronicles the development and implementation of the SAS® data warehouse project for the University of Central Florida (UCF). The warehouse serves as the foundation for the development of a wide variety of reporting applications, utilizing SAS Business Intelligence software.

PROJECT HISTORY

Initial Efforts

In 2000, a campus-wide survey effort was undertaken to review student and non-student reporting and data needs, and to identify those needs that were unmet. The findings of this effort led to the incorporation, as an objective action item in the UCF 2002-2007 *Strategic Action Plan*, the development “of an institutional level data warehouse with appropriate data marts that will permit analysis of student, employee, financial, and infrastructure data for planning and management purposes.” The outcome of this university-level endorsement for a data warehouse project was a more detailed data needs survey which was compiled into a University Data Administrators Committee (UDAC) document, *UDAC Vision Survey 2002*. This survey document provided the impetus for administrative support and funding to investigate a reporting information system solution. A university “special” funding grant was approved for UDAC to start the initial investigation into reporting solutions. The Office of Institutional Research (IR) was designated to lead the effort because the prevailing opinion was that IR best understood reporting output needs.

IR and the Office of Information Technologies and Resources (ITR) began collaboration to forward the development of a university information system plan, which would provide a business intelligence and reporting solution to support and enhance strategic and operational decision making. The resulting plan outlined a) the driving forces behind the need for a comprehensive information system, b) benefits to the university community, c) desired system architecture, d) project deployment (including scope, iterations, and tasks), e) data and information quality management, f) the implementation team make-up, g) estimated costs, h) identification of critical success factors (CSF's), i) and caveats to the plan that included assumptions, constraints, issues and risks.

Following administrative approval of the plan, two or three vendors were invited to come to campus to demonstrate their solutions. Members of the UDAC committee as well as other representatives from vested constituencies across campus observed their demonstrations and provided feedback to IR and ITR. The momentum behind the project, however, stalled at this point due to a number of factors that occurred almost simultaneously: lack of funding to purchase a solution; solution options that didn't, at that time, quite fit the university's needs; loss of IR project personnel; and shifting of university focus and resources to prepare for re-accreditation.

Solution Purchased

In 2004, IR again applied for university special funding to purchase a data warehouse solution as outlined in the earlier information system plan. This time it was approved and allocated. The remainder of the year was spent evaluating vendors and working out the details to enter into a purchase agreement. The SAS data warehouse product with business intelligence tools was purchased in January 2005.

By securing the funding for the purchase, IR's role as the project leader for the university data warehouse initiative was solidified. During the remainder of the 2005 fiscal year, 14 IR and other key division staff members prepared for the data warehouse implementation by completing 31 live and webinar SAS training classes. To jump start the project, consultants were hired for the initial setup of the warehouse environment. During the spring and summer months of 2005, the consultants worked with the IR SAS programmer and an additional database administrator to complete the setup, import the initial set of data files, and to develop an enrollment pilot project utilizing OLAP cube functionality.

Although the project was off to a good start, it significantly slowed with the departure of the consultants. This time the delay was due to insufficient technical and project staffing. The single IR SAS programmer and the database administrator were unable to continue the project development and at the same time keep up with current IR operational reporting demands.

THE PROJECT TODAY

New Project Management

FY 2006 brought renewed activity to the data warehouse project. Another university funding special for the fiscal year enabled the IR director to hire three additional staff members with primary responsibilities for the data warehouse: two full-time programmers (one in a permanent position and the other in a one-year time limited position) and an associate director who was assigned as the functional project manager. The new staff provided the IR office with the resources to refocus data warehouse development efforts.

The first step in regenerating the project was SAS training for the new staff members. This was immediately followed by the establishment of a project tracking database so that individual developer assignments could be assigned, documented and tracked by the job status. This simple database proved to be a very effective management tool that was useful to the entire project team.

A new three-year project plan was developed. The plan identified the immediate tasks to be completed in the current fiscal year. It also incorporated future project tasks (beyond the scope of the initial implementation goals) in the second and third years. Personnel training requirements for staff and campus user groups, campus communication plans, and system interface impacts were also included in the plan.

Technical Development and Challenges

The first priority for the data warehouse developers was to initiate data cleansing and validation of the data files loaded during the initial setup. The file load merged official state-required historical census data with additional elements from the student information system. Data from 27 academic terms were pushed through this process. Because UCF had converted from an in-house legacy system to an ERP system four years ago, the data cleansing and validation efforts also required mapping of values between the legacy and ERP data so that it could be used effectively for trend analysis. Over 7 million rows of data had to be reviewed and validated.

The data cleansing and validation task has proven to be one of the biggest project management challenges. Due to the sheer volume and complexity of the data, this task has taken approximately three times as long as originally planned, with three developers working on it. Although the project plan has had to be revised to accommodate the lengthy process, and several tasks which were to have been completed in FY 2006 have now been moved to FY 2007, the importance of completing this task thoroughly and accurately can't be over emphasized.

Data Warehouse Stakeholder Group

Because the project management and technical support was housed within the IR office, it would have been easy for its visibility to be lost and its value as a university-wide resource to be forgotten. It was therefore critical for the life of the project to have a dedicated and informed stakeholder group. Individuals and departments across the campus community with a vested interest in university data and information for decision-making were invited to join the stakeholder group. The group currently has approximately 30 members who represent a cross-section of key data users in the UCF community. They include several high-level administrators who act as project champions in sharing warehouse information and soliciting support for project funding.

The stakeholders provide the IR development team with guidance, feedback, and support on project initiatives. Members of the group recently provided valuable feedback during testing of the SAS[®] Information Delivery Portal by which the users will access data warehouse reporting functionality. The group also voted on the project name.

Project Website

As part of the project communications plan, a project website was developed and recently added as a menu link on the IR office home page. The website provides project status updates, general information on the mission of the warehouse, a Resources & Information page, a page with Training & Support information, and an FAQ page. Minutes and information from the stakeholder's group meeting are included on the Stakeholders Group page. (See the following examples).



UCF
Pegasus Mine
Data Warehouse

Pegasus
info
Mine

The Pegasus Mine is supported
by Institutional Research.



Institutional Research	UCF Facts	UCF News & Info	UCF Search	UCF Home
Mission Project Information What's New? Training & Support Resources & Information Stakeholders' Group FAQs	Project Information Please click on the 'Data Warehouse Project Information' link to access the current version of the project plan, the initial project scope document and other information related to the overall project."  Training & Support  IR staff members are available to provide you with the training & support you will need to navigate the data mine. Use the 'Training & Support' link for your connection. Frequently Asked Questions  If you don't find the question or answer you're looking for in the FAQs, email your question to dwarehouse@mail.ucf.edu	Resources & Information Information/demo source on SAS Business Intelligence is available from this site. Use the 'Resources & Information' link for your connection. Stakeholder's Group  The Data Warehouse Stakeholder Group represents a cross-section of key data users in the UCF community who are helping to guide the development of the project. Please click on the 'Stakeholder's Group' link to access meeting minutes, handouts and further information. What's New?  With each new program development there is a new file made available with new data. Use the 'What's New?' link for your connection to updates as they become available.		



UCF
Pegasus Mine
Data Warehouse

Pegasus
info
Mine

The Pegasus Mine is supported
by Institutional Research.



Institutional Research	UCF Facts	UCF News & Info	UCF Search	UCF Home
Mission Project Information What's New? Training & Support Resources & Information Stakeholders' Group FAQs Pegasus Mine Home	Stakeholders' Group The Data Warehouse Stakeholders' Group is a project advisory group with individuals representing a cross-section of key data users in the UCF community. The stakeholders' preview warehouse development projects and provide development recommendations; they are also project decision-makers who are helping guide the data warehouse project into an excellent enterprise-wide data system that will facilitate the creation, access, and dissemination of institutional knowledge pertinent to UCF. Stakeholder minutes for 3/31/06 meeting and handouts attached below. UCF Data Warehouse Minutes 3/31/2006 • UCF Data Warehouse Name Suggestions • Data Warehouse Structure Draft-Modified Stakeholder minutes for 5/4/06 meeting and handouts attached below. UCF Data Warehouse Minutes 5/4/2006 • Data Warehouse SDCF WVR DIM Contents • Action Items 3/31/2006 • Data Warehouse ADMISSIONS DIM Contents • Data Warehouse SDCF COURSE DIM Contents • Data Warehouse SDCF DATA DIM Contents • Data Warehouse SDCF FLWSHP DIM Contents • Data Warehouse SDCF STPND DIM Contents • EG - Graphic OutPut Options			

UCF Pegasus Mine Data Warehouse

The Pegasus Mine is supported by Institutional Research.

Navigation: Institutional Research | UCF Facts | UCF News & Info | UCF Search | UCF Home

Left Sidebar:

- Mission
- Project Information
- What's New?
- Training & Support
- Resources & Information
- Stakeholders' Group
- FAQs
- Pegasus Mine Home

Main Content - Resources & Information:

SAS® Business Intelligence

An excellent information/demo source on SAS® Business Intelligence, particularly Enterprise Guide and Add-In for Microsoft Office, is available at the SAS e-Learning Training web site:
[http://www.sas.com/apps/elearning/elearning_courses.jsp?cat=Business+Intelligence.](http://www.sas.com/apps/elearning/elearning_courses.jsp?cat=Business+Intelligence)

Click on: **Overview: SAS® Business Intelligence.**

For a quick start, we recommend that after Starting the Overview, click on **BI CLIENTS**, then select **INDIVIDUAL CLIENTS**. From there you'll find Enterprise Guide and Add-In for Microsoft Office demos. The **BI FOUNDATION** is also very good for those interested in seeing the system structure in a graphical displayEnterprise Guide

- [EG - Graphic OutPut Options](#)
- [Formats & Informats](#)

Additional Resources

- [SAS Customer Support Center](#)

Use of Consultants for Specific Project Tasks

Due to the additional project time required for the data cleansing and validation, it was necessary to determine how to best keep the project moving forward for other key pieces that would need to be in place and functioning before user groups could help test functionality. It was decided to contract with SAS consultants to set up the security access control template and to develop single sign-on access to the data warehouse portal through the university portal page. Having highly skilled and knowledgeable consultants lead and develop these two key components was essential to continuing to progress with the project. This has definitely proven to be a worthwhile investment. Our developers were able to continue with the data cleansing and validation, while at the same time be involved in the consultant projects through knowledge transfer sessions.

Delivery of the Data Warehouse Project

Through the dedicated efforts of the development team, the data warehouse and Information Delivery Portal has been made available to a large university user group. Soon all faculty and staff will have access to the data warehouse environment through the portal and to reports based upon role security. Project demonstrations have been given to the stakeholder group, key university administrators, and academic groups. The project is now moving into its third phase, which will focus on completing the conversion of existing reports to the warehouse environment and to train users on how to extract data and generate reports using the Business Intelligence tools to support their decision making processes.

Conclusion

The UCF university data warehouse project has provided many management challenges. Key lessons that have been learned through this include the importance of keeping key administrators and other stakeholders informed of the project and current status; to utilize a stakeholder group to build and maintain momentum and support for the project; the need to be flexible for unforeseen technical or programming delays; be willing to invest in subject matter expert help when it is needed to support the in-house team and to keep the project from floundering.

REFERENCES

Office of Institutional Research website: <http://www.iroffice.ucf.edu/>

Pegasus Mine Data Warehouse Project page: http://www.iroffice.ucf.edu/info_mine/index.html

UCF Strategic Planning website: <http://www.spc.ucf.edu/>

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the authors at:

Linda S. Sullivan
Associate Director
Office of Institutional Research
University of Central Florida
P.O. Box 160021
Orlando, FL 32816-0021
Work Phone: (407) 823-6683
Fax: (407) 823-4769
Email: lindas@mail.ucf.edu
Web: <http://www.iroffice.ucf.edu>

M. Paige Borden
Director and University Data Administrator
Office of Institutional Research
University of Central Florida
P.O. Box 160021
Orlando, FL 32816-0021
Work Phone: (407) 823-5061
Fax: (407) 823-4769
Email: mborden@mail.ucf.edu
Web: <http://www.iroffice.ucf.edu>

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration.

Other brand and product names are trademarks of their respective companies.