A SYSTEM FOR MANAGING CRITICAL KNOWLEDGE FOR ACCELERATOR SUBSYSTEMS: *PANSOPHY**

C. Reece, V. Bookwalter, and B. Madre[†]
Thomas Jefferson National Accelerator Facility, Newport News, Virginia, USA

Abstract

 Accelerator development and construction projects often intentionally push the envelope of well-established technical performance and manageable complexity. In addition, the desire for efficient retention and exploitation of accumulated experience across the multi-decade life cycles of major installations calls for a robust yet user-friendly knowledge management system. To meet these needs, we are deploying a new web-based system at Jefferson Lab: *Pansophy*. This system is a custom integration of several commercial software utilities, DocuShare™, ColdFusion™, Matlab™, Ingres™, and common desktop programs. Users of the system range from process managers, shop-floor technicians, and test engineers to after-the-fact data miners and operations staff. The system integrates important quality assurance elements of procedural control, automated data accumulation into a secured central database, prompt and reliable data query and retrieval, and online analysis tools, all accessed by users via their platform-independent web browsers. A system overview, a completed pilot project, and implementation experience to date will be presented.

1 MOTIVATION

 Accelerator projects often intentionally push the envelope of technical performance and manageable complexity. In addition, the desire for efficient retention and exploitation of accumulated technical experience across the multi-decade life cycles of major installations calls for a robust yet user-friendly knowledge management system.

Past experience highlights the need for good systems. The construction and commissioning of cryomodules for CEBAF, for example, was accomplished using paper "travelers". These served to establish controlled procedures for each step in the process: fabrication, assembly, processing, calibration, testing, and commissioning. They also prescribed the required quality control data for each step and accumulated the data via hand-written or stapled-in entries. The data from subassemblies accreted, producing a rather large, intimidating notebook for each completed cryomodule.

Only one complete copy of the resultant traveler was

produced for each of the 42 cryomodules. Systematic analysis of the wealth of data generated by the traveler system was and remains impeded by the physical form of the stored data. In effect, the data collected served almost exclusively a pass/fail acceptance function. Little learning was derived from it during or subsequent to construction.

The Institute for SRF Science and Technology at JLab carries responsibility for constructing the cryomodules for SNS and is also preparing to support the CEBAF 12 GeV upgrade and an array of possible future collaborative efforts.

With each of these large projects individually, and all of them collectively, there lie opportunities for significant learning from quality assurance during construction, and also from subsequent analysis of the accumulated database. To be viable in the lean laboratory context, any system aiming to be broadly employed must make convenience to the user and efficiency in implementation paramount. In addition, for portability and durability the interface coding systems and underlying databases must be structured in robust, industry-standard ways.

2 GENERAL DESCRIPTION

 To meet these needs, we are presently deploying a new web-based system at Jefferson Lab: *Pansophy*. This system is a custom integration of several commercial software utilities, DocuShare™, ColdFusion™, Matlab™, Ingres™, and common desktop programs. Intended users of the system range from process managers, shop-floor technicians, and test engineers to after-the-fact data miners and operations staff. See Figure 1.
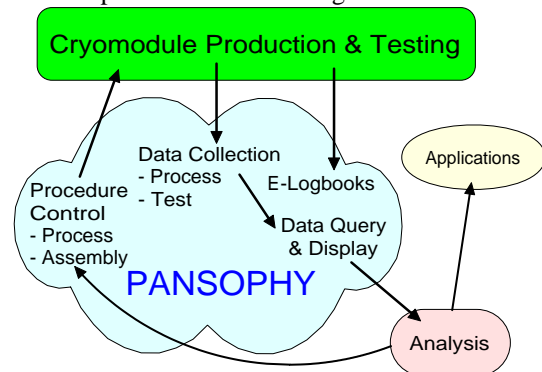

Figure 1. The Pansophy system

The system integrates important quality assurance elements of procedural control, automated data accumulation into a secured central database, prompt and reliable data query and retrieval, and online analysis tools,

all accessed by users via their platform-independent web browsers.

All files reside on secured servers, but are structured and presented for easy access on the local intranet. Process engineers and scientists establish procedural content and specify necessary data to be collected. The system provides a common user interface across many activities, and, where possible, enables automated upload of locally acquired data from test equipment.

3 Implementation Choices

3.1 DocuShare

Since the project staff is small, a decision was made early in the analysis phase to build on local experience, and to search for compromises between capability and ease of management, especially in our document management component. DocuShare, from Xerox, was chosen for its performance, intuitive interface, and user-configurability. DocuShare allows users to share documents of all types, in a controlled, centralized and (full-text) searchable environment. It provides powerful and easy-to-use tools for adding content, modifying and moving documents, creating groups, and setting read/write security. To access it, the user needs only a browser. DocuShare can be configured to notify users when the content of an area changes. If the operating environment is MSWindows, there is also a Windows Client that allows the user to save documents directly to DocuShare from such programs as MSWord or PowerPoint. For adding large numbers of archived documents, the system provides a command line interface that can be used to perform batch uploads.

3.2 ColdFusion

To create an intuitive web interface, that enables database interaction, we chose Macromedia's ColdFusion. This product extends the HTML browser language with tags which access databases via Open Database Connectivity (ODBC). The tags, referred to as CFML, employ browser standards for database retrieval and storage. With a few modifications, a simple HTML page can provide a wealth of information with database access. This up-and-coming industry standard of browser-database access is quickly receiving national recognition as a viable solution to many business ventures. Our belief was that ColdFusion could be adapted to technical applications. Additionally, Jefferson Lab's Computer Center and electronic media group use and recommended ColdFusion, thereby broadening in-house expertise and support.

3.3 Matlab

Matlab was selected as a means of enabling modular data processing between the database query and the web browser. Matlab provides a robust mechanism for data analysis and allows us to reuse analytical routines constructed by users. This well-known package permits sharing of data and analysis between scientists and between labs. Matlab has also made a venture into the web browser world by providing a web server. This server allows users to submit data and script files and receive back HTML formatted pages containing their analyses.

3.4 Ingres

Since we use the database only as a raw engine, our choice of a DBMS was completely arbitrary. We therefore based our selection on the availability of internal support. Our Computer Center has numerous applications in Ingres from Computer Associates and many years of experience in developing Ingres databases. From the beginning of the project, it was understood that we would not have a full-time database administrator (DBA), and would have to rely heavily on assistance from the Computer Center. Relative ease of use was another factor. Once the Computer Center installed Ingres we were able to get a prototype running and ready for data entry within a month.

3.5 Visible Analyst Database Engineer

Visible Analyst Database Engineer from Visible Systems Corporation is our database design tool. Database Engineer provides entity-relationship diagram capabilities for designing and documenting changes to a database. It also produces SQL scripts for creating databases, which we customize for our environment using a perl script.

3.6 Common desktop programs

Since the writers of the procedures and travelers are not likely to be web programmers, a simplified approach was needed to get from their current paper forms to web-based forms. Most authors are familiar with MSWord, which now has an HTML interface, so MSWord was selected as the base tool for authoring of travelers.

Adobe Acrobat is used for posting static versions of documents for reference from the web-based "travelers".

4 TECHNICAL FEATURES

4.1 Pilot project

Our first project, a prototype, was to incorporate historical data from CEBAF production cavity testing in our Vertical Test Area. Electronic files existed in several forms, but were not generally accessible. We were able to simultaneously reformat and load the data using Ingres' embedded Structured Query Language (SQL) for C. Some data was only available on paper, but by using an online ColdFusion form, data was conveniently entered by hand. A Matlab analysis/plotting module was written to process queried data into graphics that are directly

served to the user by ColdFusion. The pilot project demonstrated that the combination of Ingres, ColdFusion, and Matlab would work for a technical application, gave us a number of significant challenges, and prepared us for the next phase. An additional benefit was universal access to original cavity performance data for CEBAF.

4.2 The DocuShare Connection

DocuShare facilitates organizational separation of the travelers and the common procedures referenced by the travelers. The hierarchical system files travelers and procedures by development or production, then by project, and still further by approval status. When the MSWord version of the traveler has gone through the approval process, a copy is uploaded to the appropriate controlled area in DocuShare. From there, it feeds into the automated conversion process. Procedures are saved into DocuShare in Adobe Acrobat format. In the database the reference to a procedure is saved as a URL. URLs display in the user interface as live links to the referenced procedures.

4.3 Automated Processes

An MSWord template containing formatting rules and web interface boxes for data input was established for the authors to use. Once a traveler is written and approved it is saved in HTML format. From that point an automated process converts MSWord's HTML format into ColdFusion CFML browser pages. This system allows the authors to write within a familiar environment and results in an HTML document with Pansophy-specific information.

The same script that performs the CFML conversion also creates a list of form pages, and the fields on each page. A C program takes the information and creates a database table for that traveler. One of the challenges we encountered was a limitation in the ODBC/Ingres "middleware" that would only allow 2000 bytes in a database row. Theoretically, we wanted each traveler to be a table, but the limit made that impossible. The solution was for the program to examine the traveler, determine if the record is more than 2000 bytes, and if so, split the traveler up so that each page is a table. If the traveler is under 2000 bytes, one table is created. This problem has apparently been addressed by Computer Associates, and an upcoming release will eliminate the limitation.

Where possible, the user is relieved of the responsibility of data transfer to the database through automated collection processes. Data generated by LabView, test equipment, and production equipment is stored in native formats, and background processes read, reformat and upload the data to the database.

4.4 Connectivity

All user access is through the web, which is the "glue" that pulls the system together. DocuShare facilitates file administration. ColdFusion provides connectivity between the user and the data for input/collection, querying, viewing, and administration. The travelers provide live links to reference drawings and specifications, stored in DocuShare. Matlab Web Server provides the capability for modular data analysis. The combination of HTML, Java Script, perl, and the Internet Information Server (IIS) permit format conversion (non-web to web format), access and configuration controls, and data security.

5 FUNCTIONAL RESULTS

At this point, the database and the automatic traveler table generation are in place. While enhancements to these areas will continue, they are functional and in use. The MSWord template to create traveler documents and the perl converter from HTML to CFML are in place and in use. Standard naming conventions are in place and the tools are available for managers to establish project-specific traveler identifiers and control the availability of traveler headers. Interactive input and viewing of data are in place.

Facility-specific electronic logbooks are implemented for the Cryomodule Test Facility, the electron beam welder, the Vertical Test Area and the processing furnaces.

6 PLANS

Still to be implemented:
- Automatic production data uploads to the database.
- Automation of the SQL-ColdFusion link in traveler generation is not completed.
- Traveler revision and user access control.
- Additional data manipulation functions: update, query, analysis (including the Matlab web connection), and export.
- Facilities Test Requests/Test Plans scheduling, and travelers.

About 500 pages of web-based procedures with integrated data entry are expected for each of several major projects to be implemented over the next several years by the JLab Institute for SRF Science and Technology. Pansophy will enable efficient capture and utilization of data generated during these projects as well as any parallel R&D efforts that the institute undertakes.