COMBINING STRUCTURED AND UNSTRUCTURED DATA IN A CONFIGURABLE WEB-BASED LOGBOOK*

J. Nogiec[#], K. Trombly-Freytag, Fermilab, Batavia, IL 60510, U.S.A.

Abstract

A typical electronic logbook is designed as a general-purpose system for recording time-ordered events and actions and, therefore, allows for a great flexibility in recording information, but the data is unstructured. To better position it in a specific context (e.g., a, test facility, a group activity log) it needs to support both structured data (keyword, authors, etc) and unstructured data (text, title, attachments) in that context. To do this, a logbook system can define a set of attributes, possibly built as a hierarchy. These application-specific attributes will be associated with each entry. To be flexible, such a system has to be configurable to allow for tailoring it to each specific environment. The paper describes a design, functionality, and experiences with WebLog, a database-configurable electronic logbook developed with the J2EE Web technology. Various functional and technical properties of the system are discussed, including views, searches, threads of entries, an automated alerting system as well as integration with other applications.

STRUCTURED DATA IN LOGBOOKS

Electronic logbooks can be regarded as a category of Content Management Systems (CMS). They are used as a replacement of more traditional paper-based logbooks to keep track of various activities. As with their paper predecessors, they focus on maintaining a chronological order of entries, which are sometimes grouped into days or shifts. Entries can also be associated with categories and/or keywords and contain attachments. Basically, these systems keep unstructured data (text, pictures) with a small standard set of associated attributes. This characteristic is common to a number of electronic logbook implementations, which differ in their technologies, features and targeted application area, such as an experiment shift log, facility operation log, group activity log and many others (see [1] for a review of logbooks used in high energy physics facilities).

One can try to augment the traditional functionality of logbooks as repositories of unstructured information by extending the number of attributes associated with log entries and, therefore increasing the amount of structured data stored in the system. This approach allows for emphasizing the logbook application-specific context and for adding domain-specific data, and provides better integration with operational processes, and better correlation with domain-specific data. It also greatly improves the searchability of entries.

For such a domain-enhanced logbook to be universal enough to be deployed in varying contexts, it must allow for easy tailoring to a new deployment domain, possibly without any programming efforts, using the common code base. Therefore, such a system should be configurable to match the specifics and requirements of the application domain.

WEBLOG SYSTEM

WebLog is an example of a configurable electronic logbook system with extended structured data, which has been developed by the authors and deployed in the Technical Division of Fermilab. WebLog is a Web-platform application developed using JSP, servlets, JavaScript, DHTML and CSS.

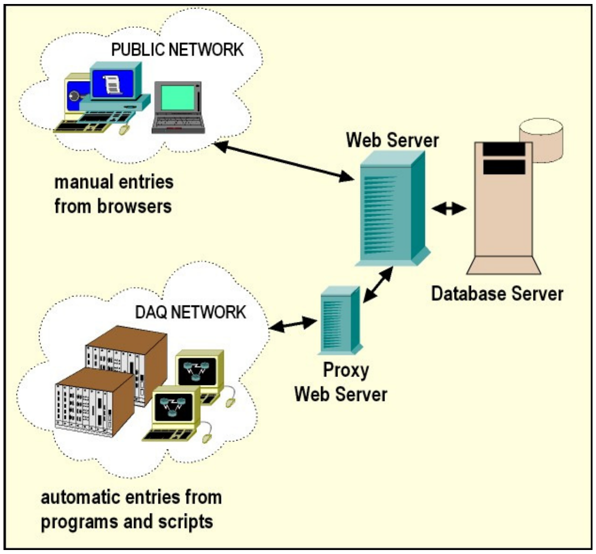

PUBLIC NETWORK
manual entries from browsers
Web Server
Database Server
DAQ NETWORK
Proxy Web Server
automatic entries from programs and scripts

Figure 1: 3-Tier architecture of WebLog.

Architecture

WebLog's architecture (see Fig. 1) is based on a 3-tier model, where logbook users, application logic and data storage are all distributed between three separate tiers: user workstations, an application server and a database/file system server. This classical architecture for Web applications is augmented by real-time control, monitoring and DAQ applications accessing the application server via proxies due to network firewall protection.

The WebLog has been successfully deployed and used with either of the popular open source database systems MySQL and PostgreSQL.

*Work supported by the U.S. Department of Energy under contract no. DE-AC02-07CH11359

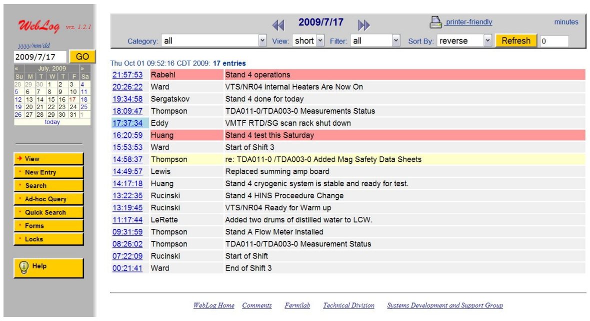
[#] nogiec@fnal.gov

Features

WebLog, by default, shows entries grouped by days. One can navigate to the previous or next day by following the supplied links, or, alternatively, use an embedded calendar to select a date. One can also navigate to any known entry directly using its unique identifier or navigate from one entry to the next.

WebLog offers a selection of views consisting of a series of entries for a given day. Entries can be shown in the short view consisting of a timestamp, author and entry title, and the long view showing the text of the entry and attachment icons in addition to the short view items. Each entry can be selected for full viewing with all its attachments. All the entries associated as comments with the entry are also visible in the full entry view. The day view can present entries in any of the following orders: chronological, reversely chronological, by author, by importance and threaded. A threaded view shows all associated entries in groups ordered chronologically. Each user can define separately how often his view is automatically updated. Any view can be also presented in a printer-friendly way, with no navigation menus or links.

Among useful features available in WebLog is the ability to configure email alerts. Entries marked as alerts will result in sending emails to those people listed in the configuration for a given combination of metadata, such as subsystems, test stands, categories, and subcategories.



2009/7/17

Category: all View: short Filter: all Sort By: reverse 0

Thu Oct 01 09:52:16 CDT 2009: **17 entries**

21:57:53	Rabehl	Stand 4 operations
20:26:22	Ward	VTS/NR04 internal Heaters Are Now On
19:34:58	Sergatskov	Stand 4 done for today
18:09:47	Thompson	TDA011-0/TDA003-0 Measurements Status
17:37:34	Eddy	VMTF RTD/SG scan rack shut down
16:20:59	Huang	Stand 4 test this Saturday
15:53:53	Ward	Start of Shift 3
14:58:37	Thompson	re: TDA011-0 /TDA003-0 Added Mag Safety Data Sheets
14:49:57	Lewis	Replaced summing amp board
14:17:18	Huang	Stand 4 cryogenic system is stable and ready for test.
13:22:35	Rucinski	Stand 4 HINS Proceedure Change
13:19:45	Rucinski	VTS/NR04 Ready for Warm up
11:17:44	LeRette	Added two drums of distilled water to LCW.
09:31:59	Thompson	Stand A Flow Meter Installed
08:26:02	Thompson	TDA011-0/TDA003-0 Measurement Status
07:22:09	Rucinski	Start of Shift
00:21:41	Ward	End of Shift 3

Figure 2: WebLog with the short view.

WebLog, unlike many other electronic logbooks, allows for commenting on any entry, with each comment also being an entry on its own. In this way, one can associate any given entry with any of the previous comment or non-comment entries. Original entries and comment entries are distinguished in the various views via different title colours (Fig. 2 shows normal entries as gray, comment entries as yellow and alert entries as red. The highlighted timestamp is a locked entry, see Advanced WebLog Features).

The WebLog system offers an extensive search capability, utilizing not only all available entry attributes, such as time spans, author, category, or subsystem, but also by looking for a phrase in the title, text or both. Canned, quick search options are offered to search for a number of recent entries or entries made in the last week.

When authoring an entry, the user is guaranteed that his original text formatting will be preserved when needed, can change the fonts and style, and incorporate hyperlinks and links to other logbook entries. The system also offers a preview feature and a spelling checker with the capability of adding domain specific vocabulary. An entry has various text and select boxes that are either mandatory or optional, allowing the user to enter only the information that he/she knows and considers to be correct. Lastly, the settings of these properties can be preserved to be used for the next entry, or cleared by the user.

A set of standard forms has been added to the system to standardize the view and contents of entries for certain repeatable events.

UNSTRUCTURED AND STRUCTURED DATA

WebLog depends heavily on the database for storing its configuration, entry structured and unstructured data, and metadata.

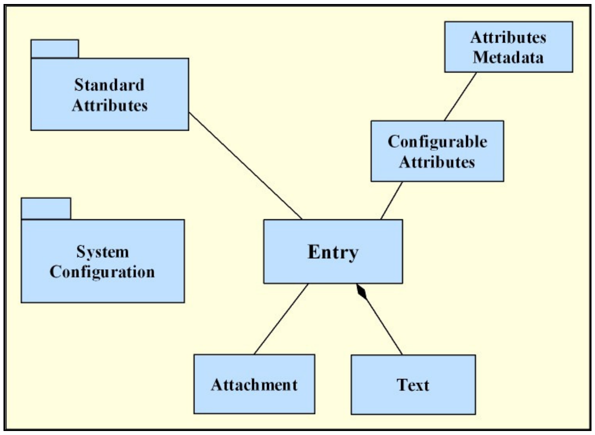


Figure 3: WebLog data model.

The focal point of the data model is an entry (see Fig. 3), a time-stamped data item describing a single logbook event. An entry is associated with unstructured text and attachments, which could be images, documents or data files. It is also associated with structured data stored in various database tables. All these data specify various properties and create the context of an entry.

Increasing the number of entry attributes assists in associating the described event or condition with the activity or process, the equipment and procedures used, and the processed objects, conditions, causes, and specific details, all of which are used to both filter and search for a specific event. Some of the attributes also correspond to information in other systems, and can help retrieve data and/or information from them, such as log traces or data results.

Non-configurable attributes are the same for all WebLog installations and include such properties as author, timestamp, entry alert status or title.

Configurable attributes can be tailored to a specific installation, and include lists or categories and their subcategories or associated activities, and other domain/application specific entry attributes, such as keywords. For instance, a test facility logbook may have test stands, test objects or devices, infrastructure systems, measurement systems, and other similarly specified properties, whereas a logbook for group activities may have such configurable entry attributes as meeting type, assignment type, report type, sub-group functions, etc. Configurable properties are set up with the help of metadata, also stored in the database, that define attribute names, allowed values, user interface labels and tooltips.

WebLog also has several database tables used to keep data that is not associated directly with entries. These data include definitions of email alert lists, log definitions for multiple log installations, and view filter definitions.

ADVANCED WEBLOG FEATURES

The WebLog system offers number of advanced features that are novel or rarely present in typical logbook implementations.

One completely novel feature is locking. WebLog entries can indicate that they are associated with locking or unlocking of a particular real-world system, specified by the structured data accompanying the entry. This feature is frequently used to mark that work is being done on a subsystem and to alert other logbook users that they may not perform certain activities on that system until it is unlocked. Multiple locks can be taken on a given system and it is possible to check any of the currently held locks for the specific information contained in the associated logbook entries.

Important for industrial-type applications is the ability to integrate the electronic logbook system with their data acquisition, monitoring and control applications. WebLog offers C, Java and LabVIEW Application Programming Interfaces (API) for seamless integration to such systems, and allows real-time programs to make logbook entries, noted as automatic entries. In this way decisions and events observed and recorded by human operators can be correlated with events, exceptions and alarms reported by software systems into one coherent chronologically ordered set.

Depending on the size of the facility and/or intensity of operations, the number of automated entries made in even one hour can be significant enough to make it difficult to be processed effectively by a human operator and to guarantee him reacting on them appropriately. WebLog provides two mechanisms to help alleviate this problem. One is the capability to have separate logs, while retaining the ability to view joined entries from a number of selected logs. Another is the ability to customize a view for a specific user by applying filters. A user can filter entries reported to him using configurable filters, which limit the visible entries by specifying any condition based on the combination of available properties, e.g., subset of originators, specific importance/urgency level, subset of controlled systems or categories, etc.

FUTURE WORK AND EXTENSIONS

The electronic logbooks provided by WebLog are used by technicians, engineers and scientists with varying computer skills and somewhat different expectations. Therefore, the authors have tried to strike the right balance between simplicity and convenience of use and the number of included features which are, inevitably, connected with increased complexity.

WebLog could be extended by adding sketching via a drawing tablet, drag and drop of attachments, user-customizable profiles, statistics and administration pages, and entry templates mechanism. None of the above features has been deemed necessary or have yet been requested by our users. However, it is likely that modifications stemming from the need for different authentication methods and authorization mechanisms or from the expectations to modernize the presentation layer with more Ajax-based features will be necessary.

SUMMARY

The authors advocate the enhancement of the traditional electronic logbook paradigm to include configurable application domain-specific structured data. The resultant CMS is characterized by better integration with domain-specific applications and an improved context for storing otherwise unstructured data.

The authors presented WebLog, a system built on these premises and verified in a production-like environment over several years. WebLog has proven to be a completely indispensable tool for running operations of a facility including such systems as a cryogenic plant, a number of test stands for testing and performing measurements of superconducting and conventional accelerator components and several high-current power supply systems.

ACKNOWLEDGMENTS

The authors would like to express their gratitude to John Tompkins, Camille Ginsburg, and Mike Tartaglia for their help in understanding the needs of logbook users and many useful comments that helped in refining the WebLog application.

REFERENCES

[1] C\S. Gysin, E Harms, W. Kissel, J. Patrick, M. Tartaglia, C. Saunders, J. Nelson, “Evaluation of Electronic Logbooks for Use at FNAL ILC Test Areas”, Batavia, IL, Jun 2006, ILC-doc-305-v4, http://docdb.fnal.gov/ILC/DocDB/0003/000305/004/Reportv4.doc.