Combining Structured and Unstructured Data in a Configurable Web-based Logbook

J. M. Nogiec, K. Trombly-Freytag

Fermi National Accelerator Laboratory, Batavia, Illinois

2009 ICALEPCS

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 must 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.

THE NEED FOR STRUCTURED DATA IN LOGBOOKS

One can try to augment the traditional functionality of logbooks as repositories of unstructured information by extending the set of attributes associated with log entries and, therefore increasing the amount of structured data stored in the system. Increasing the number of entry attributes assists by associating the described event or condition with the activity or process, the equipment and procedures used, 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.

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

STRUCTURED DATA IN WebLog

The focal point of the WebLog data model is an entry, 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.

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 sub-categories 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.

WebLog FEATURES

- Database-configurable
- Domain/application –specific attributes added to standard attributes
- Multi-log system
- Configurable viewing filters
- Locking and unlocking entries
- APIs from Java, C, and LabVIEW
- Printer-friendly views
- Calendar-based navigation
- Search on attribute values and entry contents
- Attachments
- E-mail notification mechanism
- Entry priority model (alert, non-alert)
- Selectable view order: chronological, reverse, threaded, by author, by importance
- Short (title, timestamp) and long entry view
- Entries attachable to other entries as comments
- Forms for standardized entries
- Entry preview and spelling checker
- Entry may contain hyperlinks and links to other entries
- User customizable entry text style and formatting

SUMMARY

The authors advocate the enhancement of the traditional electronic logbook paradigm to include configurable application domain-specific structured data. The resultant Content Management System 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 natural 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.

Entry View
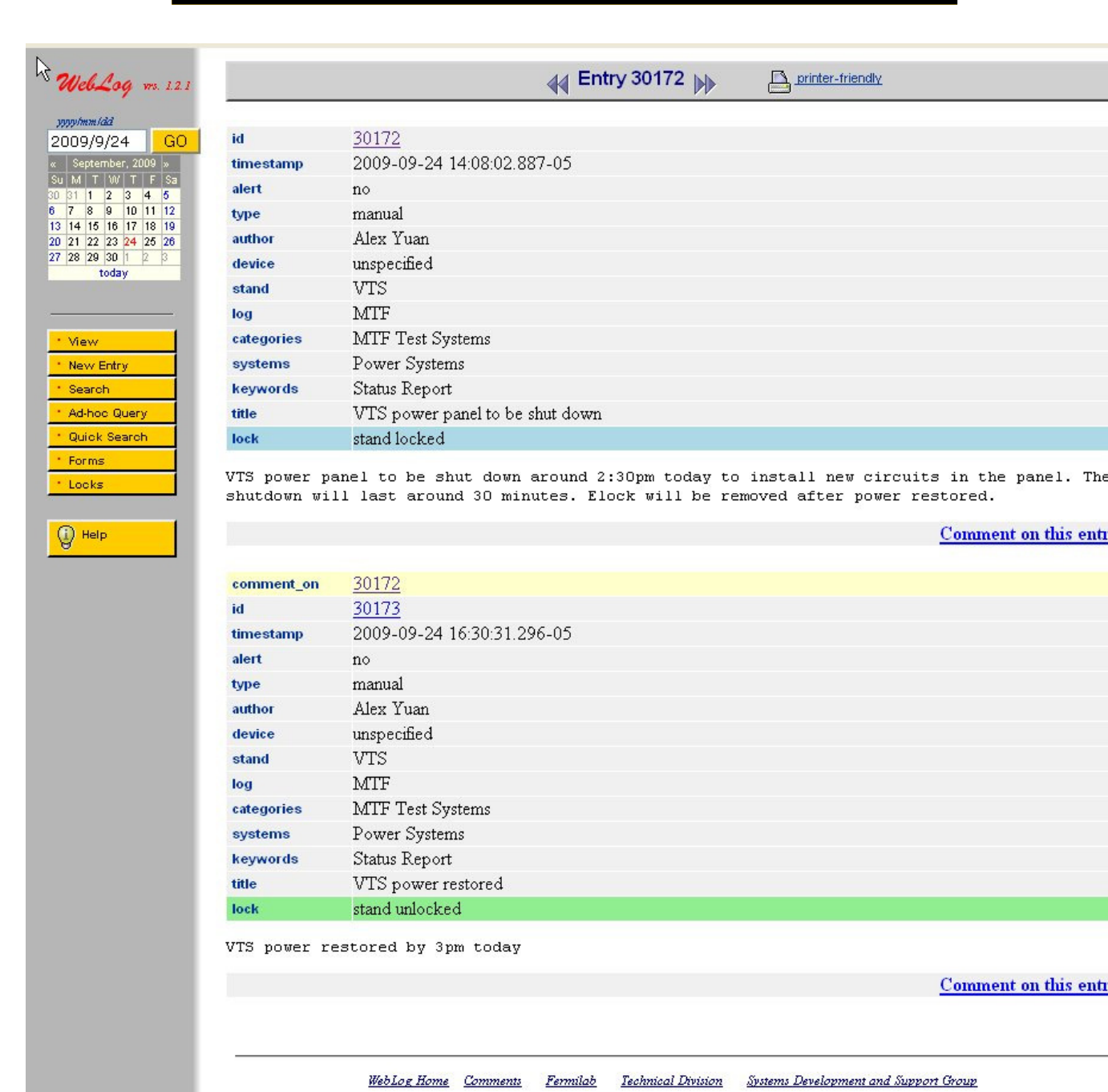
This view of a single entry with its connected comment entry displays the values of all attributes of the entry together with the long description and highlighting of the locked/unlocked status. Note that it is possible to create another comment entry for either of these entries from this view, and to navigate to the previous or next entries.
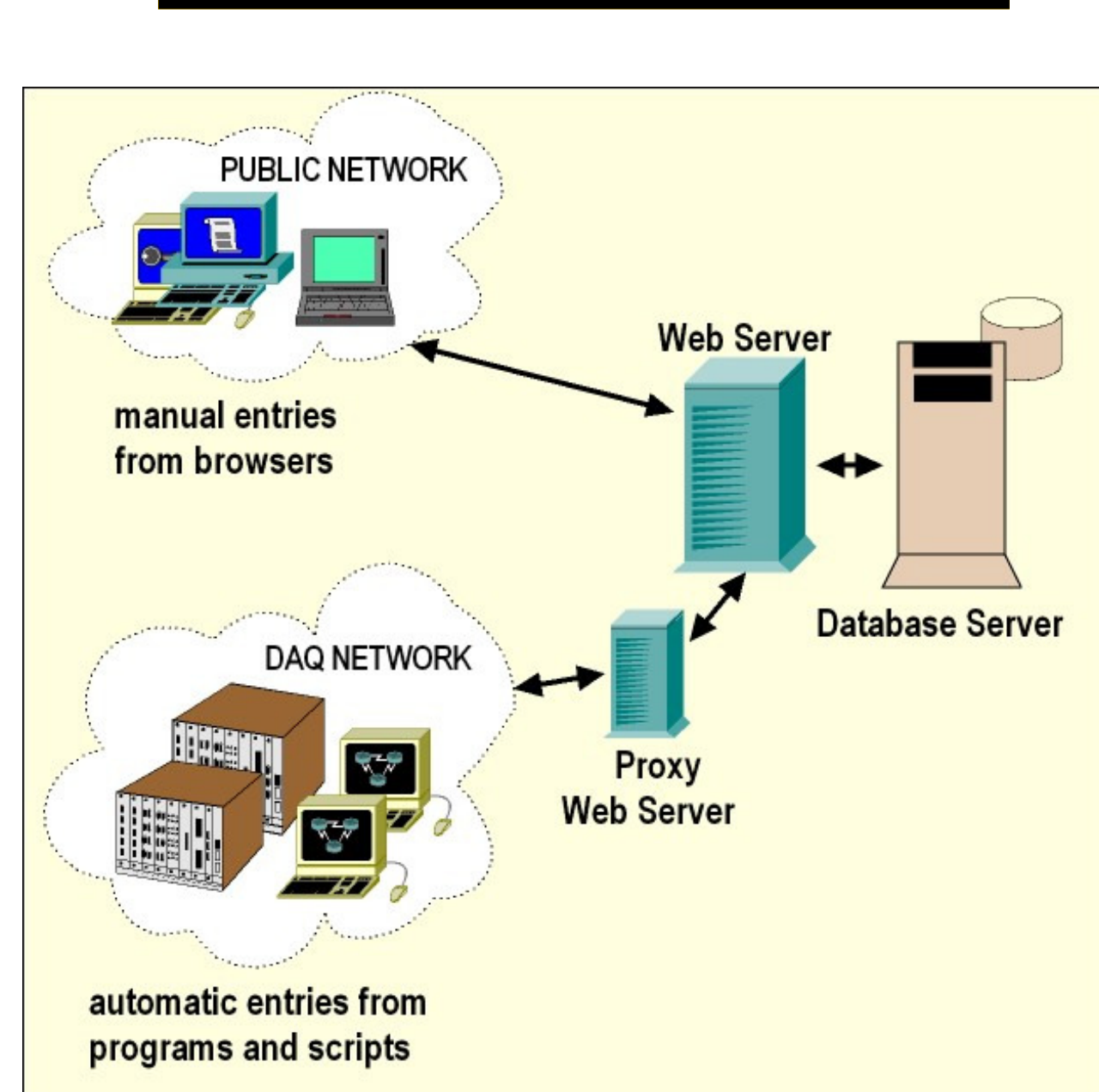
Day View

The screenshot above shows the day view with the short format in chronologically reversed order. By default, entries are grouped by day, with normal entries showing as gray, comment entries as yellow and alert entries as red. The highlighted timestamp is a locked entry.

Search Form
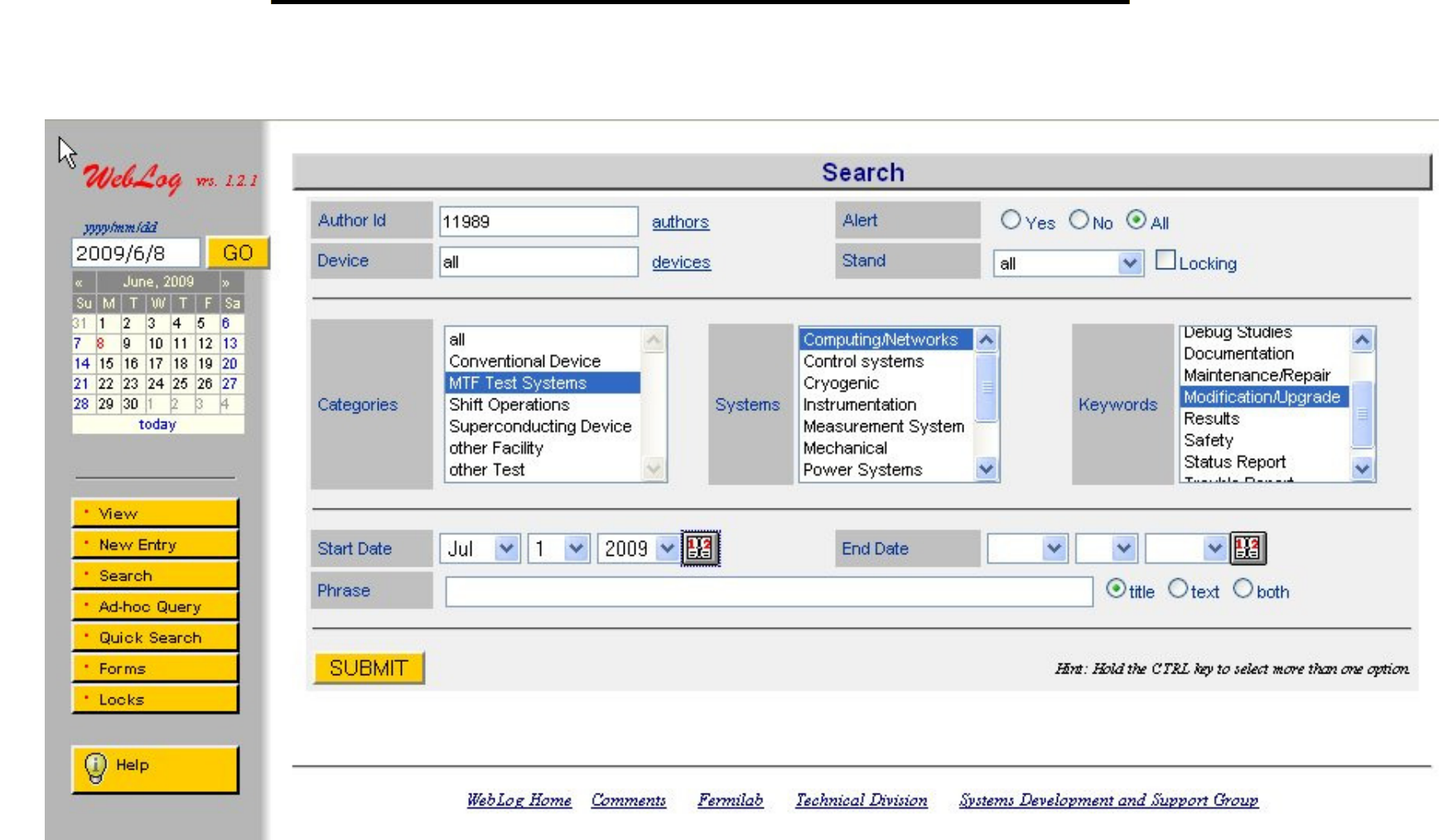
Searching on any or all attributes of an entry is provided, as well as on date ranges and long description and/or title text phrases. Multiple selections from drop down boxes are also searchable.

Architecture

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

WebLog is a Web-platform application developed using JSP, servlets, JavaScript, DHTML and CSS. WebLog's architecture is based on a 3-tier model, where logbook users, application logic and data storage are all distributed between three separate tiers.

New Entry Form
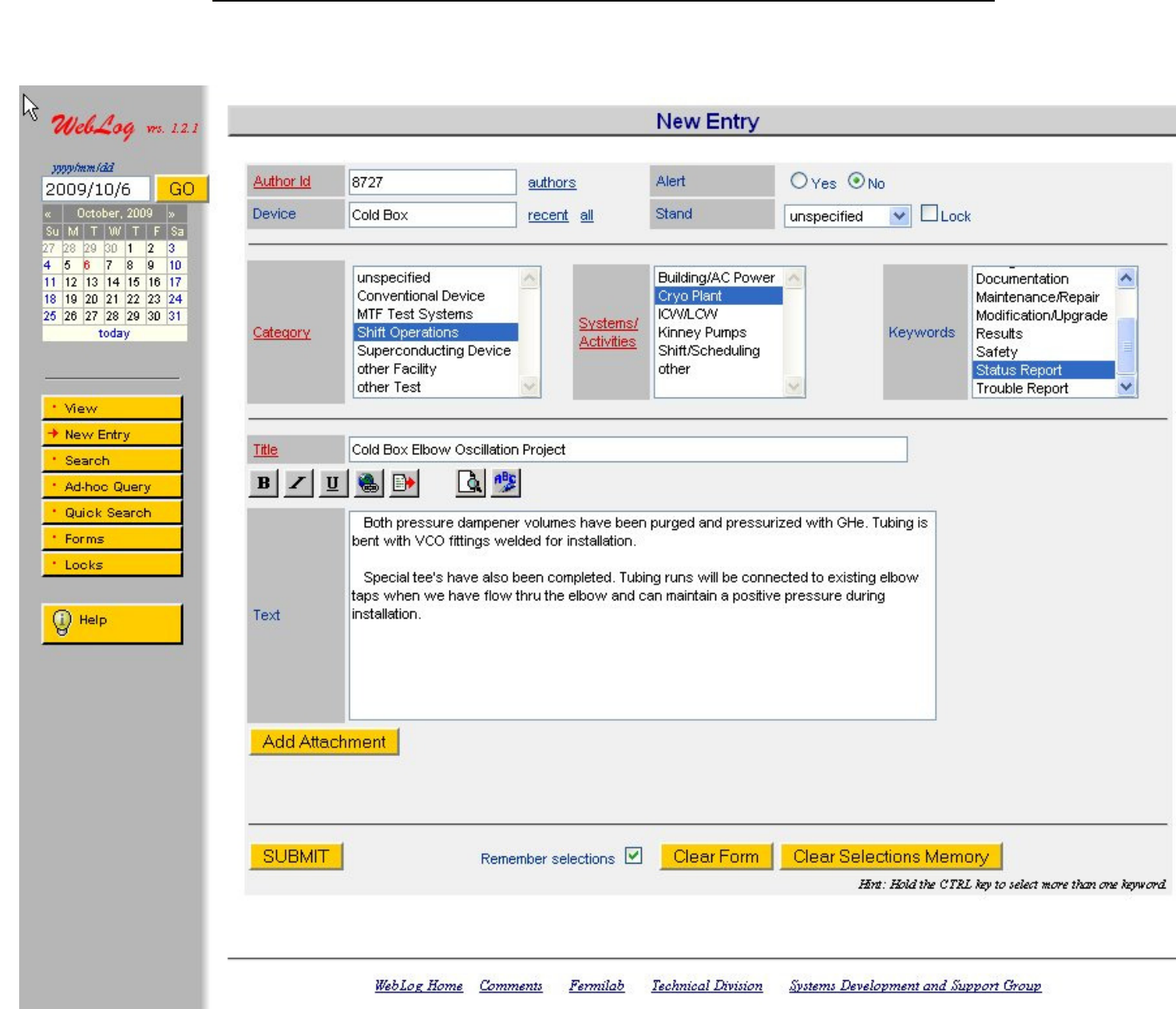
The New Entry Form contains mandatory (red text in field name) and non-mandatory attributes. The Text (description) entry is free-form, and all author specified formatting is kept. Preview and spellchecking are also provided. A toggle switch allows certain fields to retain their values between entries for this instance of the form. An entry may have an unlimited number of attachments.

Data Model
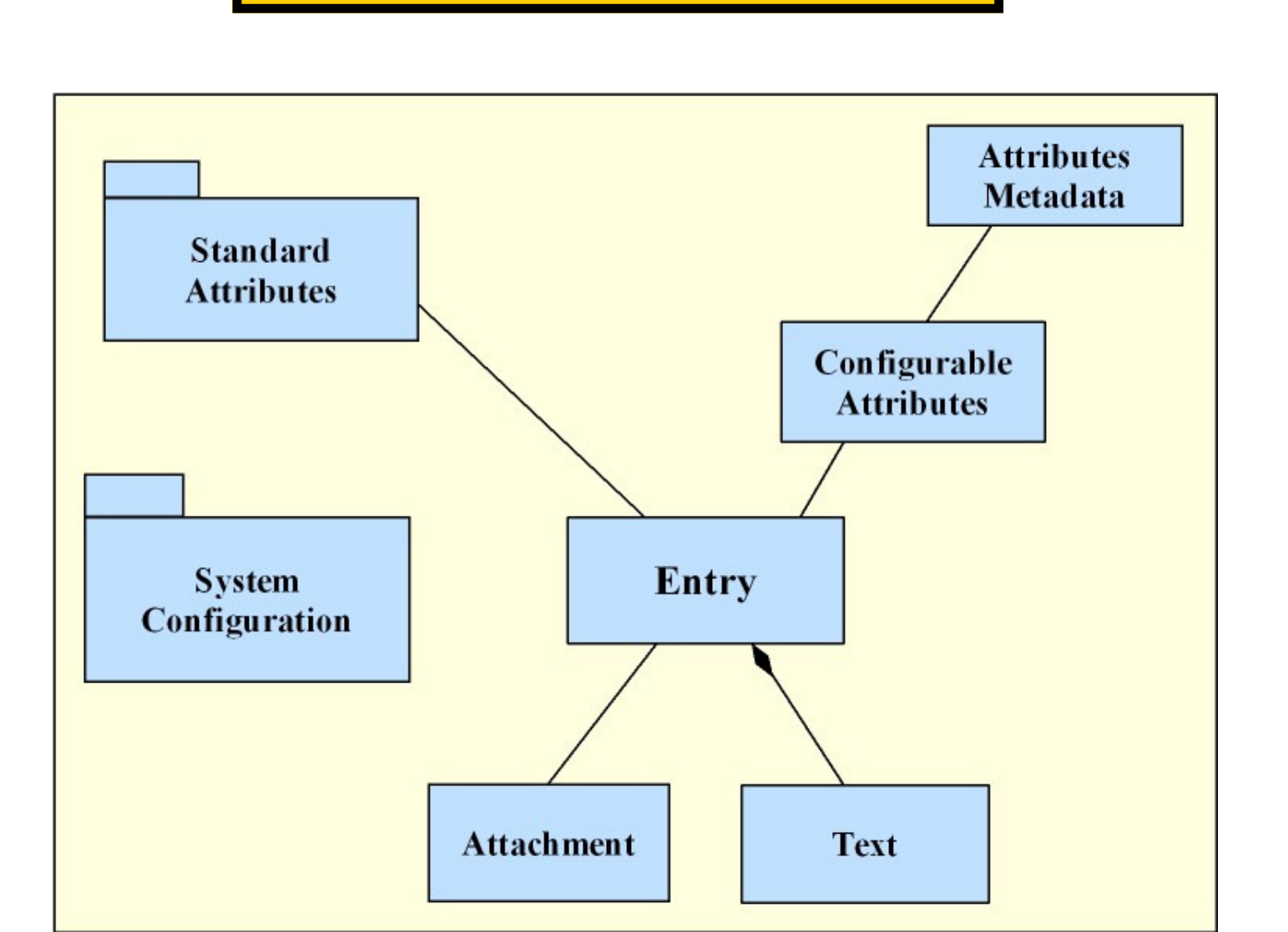
The focal point of the data model is a time-stamped entry describing a single logbook event. It is associated with unstructured text and attachments (images, documents or data files) and with structured data stored in various database tables. The model also contains system configuration information, used to specify log-generic attributes and other general data.

Locking
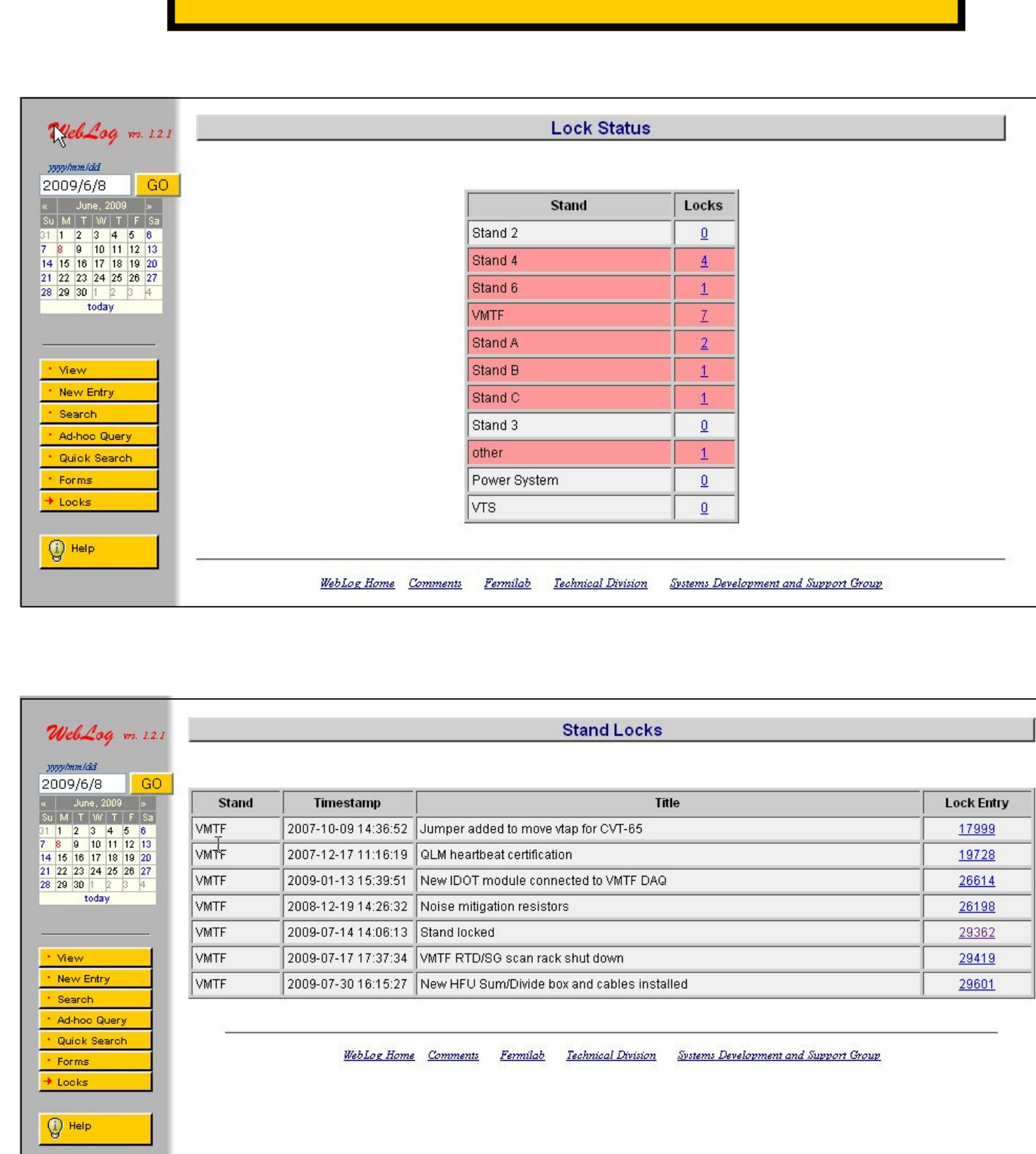
WebLog entries can indicate that they are associated with the locking or unlocking of a particular real-world system. This is frequently used to note that work is being done on a subsystem and to alert other logbook users that they may not perform certain activities on that system until it is unlocked. Multiple locks can be taken on a system and it is possible to check any of the currently held locks for the specific information contained in the associated logbook entries.
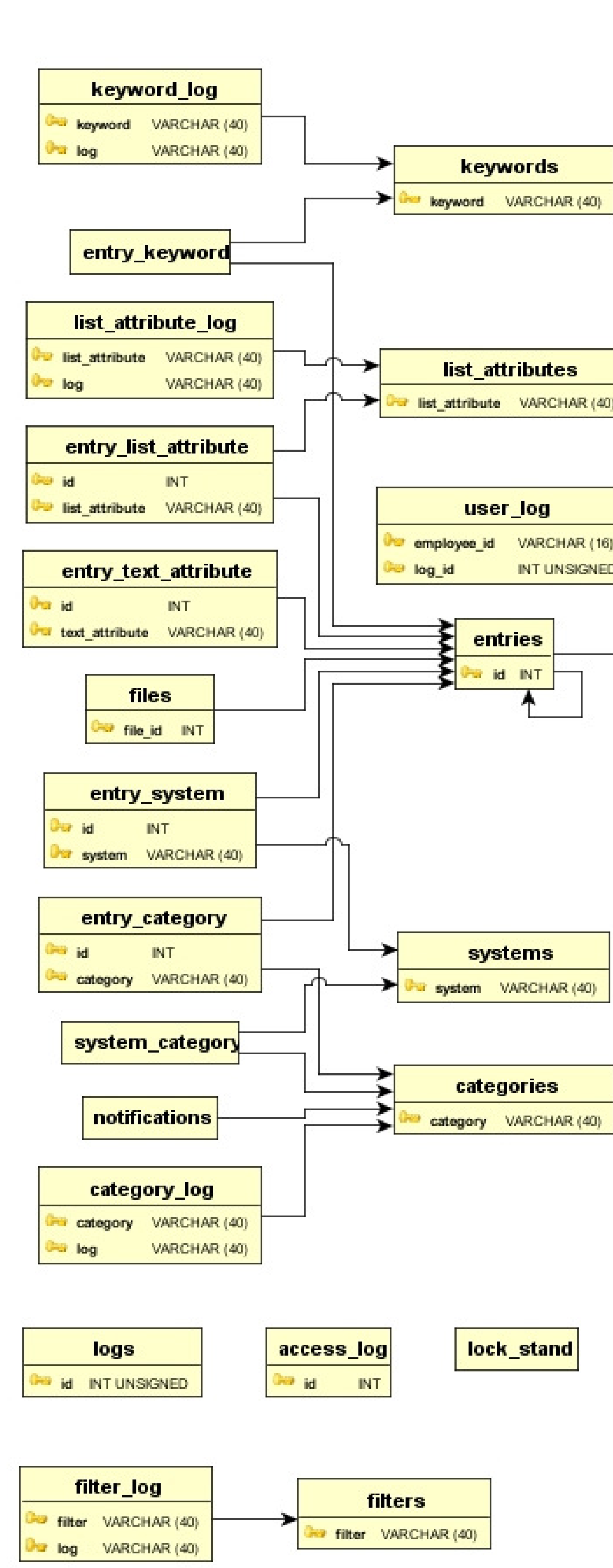
Special Entry Forms
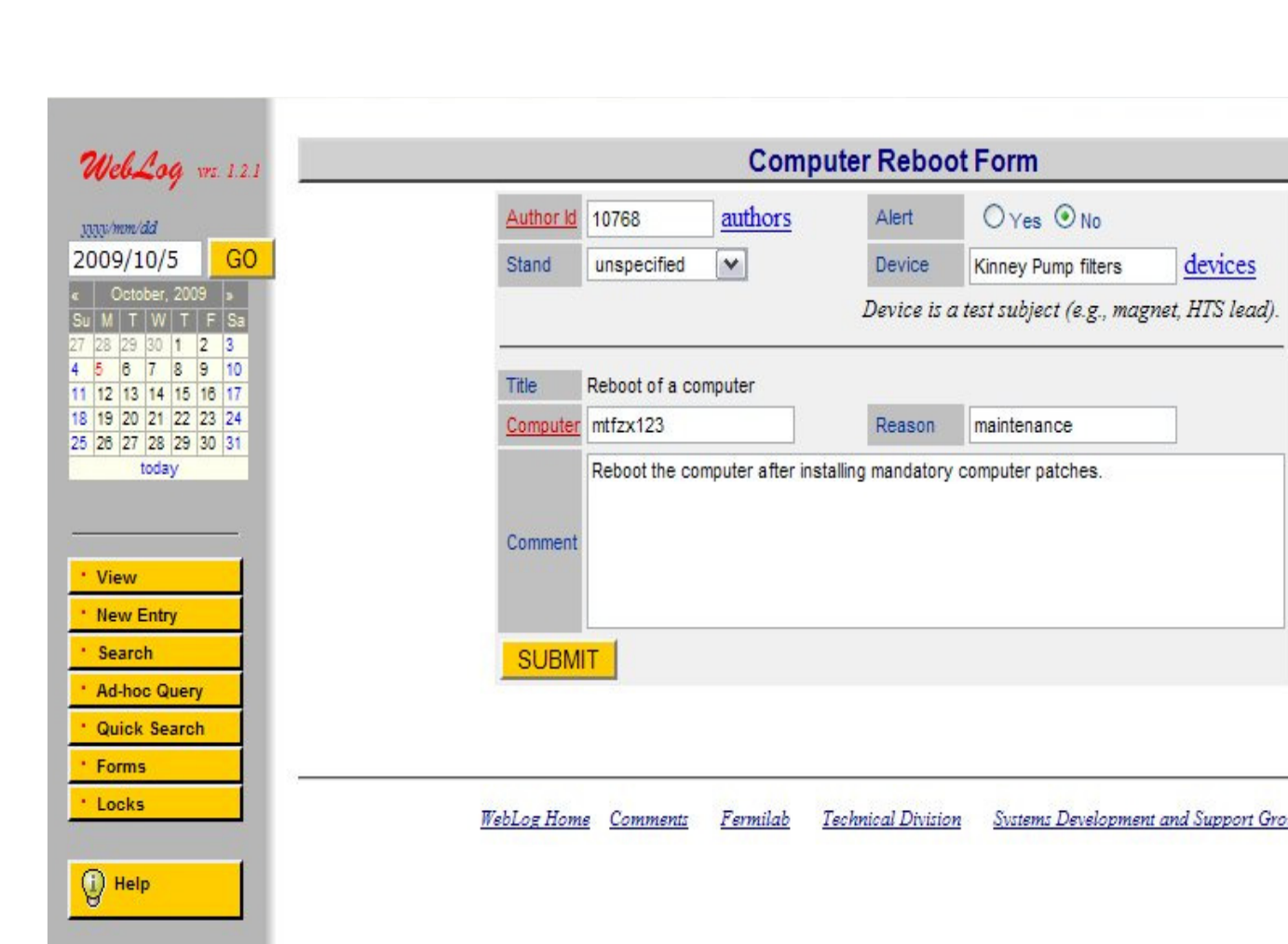
WebLog allows for the creation of special entry forms for standardized entries describing reoccurring events.

Reports/Quick Search
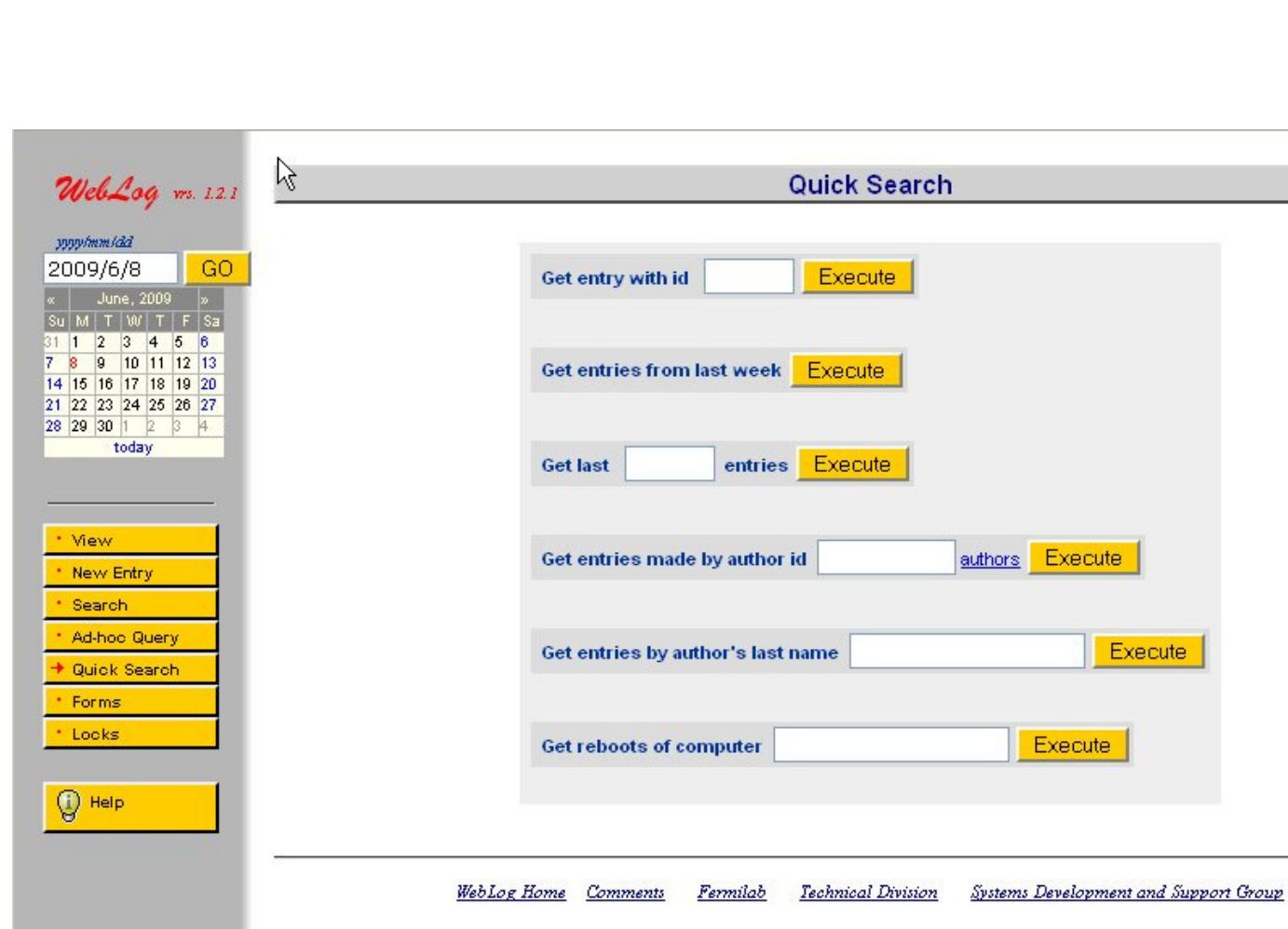
Canned, quick search options are offered to look for a number of recent entries, entries written by a given author, entries made in the last week, etc.

Database Schema

The WebLog database system is centered on the entry. Attributes are connected to specific entries and are customizable on a per logbook basis. Other tables provide log specific information and tailoring. WebLog has been successfully deployed and used with either the MySQL or PostgreSQL database systems.

Kobe, Japan