academicJournals

Vol. 10(1), pp. 1-6, March 2018
DOI: 10.5897/JETR2018.0638
Article Number: 3E747E956778
ISSN 2006-9790
Copyright © 2018
Author(s) retain the copyright of this article
http://www.academicjournals.org/JETR

Journal of Engineering and Technology Research

Full Length Research

Online electronic laboratory notebook: A secured cloud storage system scripted in Hypertext Pre-processor (PHP) programming language

Shrikant Pawar[1, 2*] and Harshul Batra[2]

[1]Department of Computer Science, Georgia State University, 34 Peachtree Street, 30303, Atlanta, GA, USA.
[2]Department of Biology, Georgia State University, 34 Peachtree Street, 30303, Atlanta, GA, USA.

Received 10 February, 2018; Accepted 29 March, 2018

Notebooks in research are an important part in tracking information of both wet and dry analytical data. Paper books or excel sheets remain the most popular and conventional ways of record keeping. To meet these demands, an open source electronic laboratory notebook is developed that can track users research need. Based on the user specifications, this suite is written in Hypertext Pre-processor (PHP) and based on a MySQL relational database. It enables the set-up of single or a multi-users access controls. A Linux server hosts the application and database. The hardware requirements of suite on the server side are moderate. Bounds and ranges have also been considered and need to be used according to the user instructions. The sharing can be limited to single individual or between research groups. Adequate server and database security measures and daily backups have been provided making its availability long-lasting and protecting data from aging damage. Important notable advantages of this system are that it runs entirely in the web browser with no client software need, industry standard server supporting major operating systems (Windows, Linux and OSX), and ability to upload and store external files. After testing and validating this suite on beta-users for 48 days, it has proven to bring higher quality documentation, data recording and retrieving information with un-parallel higher security. This suite is scalable to the size of laboratory by increasing server space with limited resources. Availability: The electronic notebook is hosted on personal linux server and can be accessed at: http://131.96.32.229/login-system/index.php

Key words: Internet/Web-based learning, pedagogy, database, hypertext pre-processor (PHP), electronic laboratory notebook.

INTRODUCTION

An efficient laboratory notebook should be easy and intuitive in usage. It should efficiently create, store, retrieve and share records with minimum training in a secure environment. The recording experimental data should be flexible on a common platform with privacy (LIMS, 2018). Ideally, the system should improve the

*Corresponding author. E-mail: spawar2@gsu.edu.

Author(s) agree that this article remain permanently open access under the terms of the Creative Commons Attribution License 4.0 International License

Table 1. Summarization of the conventional ways of research tracking and their features.

Feature	Paper notebook	Spreadsheets	Developed product
Integrated environment	No	No	Yes
Data sharing	No	No	Yes
Easy interface	No	Yes	Yes
Flexibility	No	Yes	Yes
File storage	No	Yes	Yes
Workspaces	No	Yes	Yes
Permission system	No	Yes	Yes
History track	No	Yes	Yes

work productivity by facilitating sharing and be affordable and user friendly. An electronic notebook is developed to target researchers in non-commercial institutions. A primary beta testing with the current version of tool on 25 users for a time period of 48 days who are currently working with wet laboratory data is performed. Following up of usage provided a general positive response towards this notebook. The data is stored in a secured server in MySQL database which only user with verified account can have access to. Such notebooks can also improve effective collaboration amongst users. Quality control and assurance are important features of any product in several research intensive procedures, which attempt was made to maintain in this system (Elliot, 2018). Furthermore, quality is usually compensated with the cost, which stretches the research budget of majority of non-profit institutions. This is a major issue with some of the currently available online electronic record keeping suites (LIMS, 2018). Some suites are tailored for specific research purposes which makes them complex, inflexible and require specific IT for support. Further, since some of these software suites were developed in early 2000's since they lack scalability, have poor interface with no client-server model (LIMS, 2018). Local file storage software suites are limited to certain number of users and information sharing is also limited or none. On the other hand, web applications can be accessed from a web browser, and data can be shared on the location anywhere. The web applications can be development intensive in start, but once marketed on complete completion they may only require server-side maintenance, thereby reducing support and the software maintenance costs (Table 1).

METHODOLOGY

Implementation

Based on the user specifications, this suite is written in Hypertext Pre-processor (PHP) and based on a Structured Query Language (MySQL) database. It enables the set-up of single or a multi-users access controls. A Linux server hosts the application and database. It has been developed on a Mac OS X operating system using MySQL as the relational database management system and PHP as the scripting language. The hardware requirements of suite on the server side are moderate. The server utilized is Georgia State University (GSU) Orion with CentOS 6.7 64-bit, 6x IBM System x3850 x5, Intel Xeon Processor E7-4850, 4 CPUs (10 cores per Core Processing Unit), 2.0 gigahertz (GHz) processors, 512 gigabytes (GB) Random Access Memory and 2 terabytes (TB) of scratch storage for jobs. The database has a column text-field table which is updated interactively with the user. Further, the uploading of information is standardized as certain parameters like date has to be inputted in specific format only which helps in retrieval process. Checking of data type (float, integer, text, Boolean) and dates is important before putting into database. Errors with dates or invalid parameters or wrong data type cause a halt of workflow. Bounds and ranges have also been considered and need to be used according to the user instructions.

Security

Implementation of strong security measures for authentication of SQL server is provided. Kerberos protocol uses a number of encrypted messages to authenticate SQL server and the passwords are not passed across the network. Authentication is more reliable and managing it can be reduced by leveraging active directory groups for role-based access to SQL server. The sysadmin (sa) account is vulnerable when it exits unchanged so the sa account has been disabled on the SQL server instance. We chose to give options of complex passwords for sa and all other SQL-server-specific logins on SQL server and checked in the 'Enforce password expiration' and 'Enforce password policy' options for sa. Explicit grant control server permission has been blocked because logins with this permission get full administrative privileges. For permissions to users, a built-in fixed server roles and database roles or creating own custom server roles and database roles can be achieved. Guest user exists in every user and system database, which is a potential security risk in a lock down environment because it allows database access to logins those who do not have associated users in the database. This access and also accesses to user and system stored procedures were restricted. Furthermore, usage of a common specific TCP port (excluding 1433 and 1434) instead of dynamic ports has been implemented. SQL server browser service is only running on SQL servers and secure SQL server error logs and registry keys using NTFS permissions are utilized as they can provide great deal of information about the SQL server instance and installation.

RESULTS

Ease of usage

The web application is easy to use; the data can be

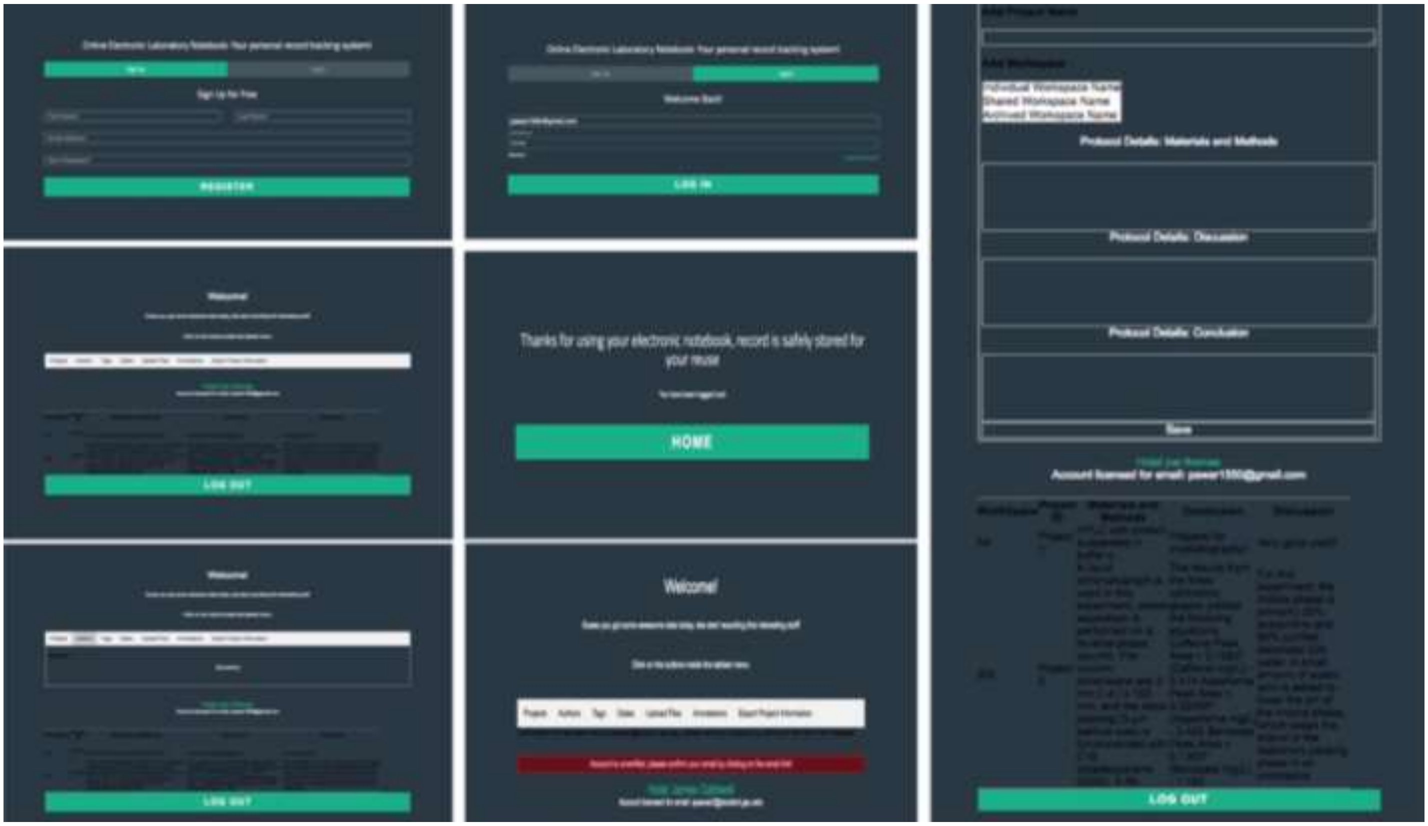

Figure 1. The snapshots of the interface windows (Login, Registration, Account Verification, Home Page).

quickly entered, retrieved, and shared. The sharing can be limited to single individual or between research groups. The very specific usages of this tool can be towards tracking project developments and updating manuals. The snapshots of the interface windows are as shown in Figure 1.

Access

It is a web based application and therefore accessible through any web browser connected to the firewall-protected network. Upon signing in, the user will get a person email, wherein a key encrypted verification is needed to get access with a controlled personal login id and password. Since it is written in PHP, it is compatible with any operating system and on a flexible hardware including and not limited to PC, Mac, tablets and smartphones. On logging, the user gets an option to choose his workspace, which can be an individual workspace dedicated to personal notebooks which is non-sharable and accessed only by the user, shared workspace containing shared projects which can be shared, read and written by all the members in a group and archived workspace, detailing closed notebooks from past.

Features

Some important features of this tool are easy to navigate interface and menus with minimal or no informatics skills, a clean and research targeted text editor with standard information on experiment details like name, objectives, material and methods, results and conclusion, a dedicated space to handle storage of files such as pictures, word documents, excel sheets, pdfs, hyperlinks, annotations/comments space, and writing protocols to pdf files.

Security

An adequate server and database security measures as explained in methods have been provided, which can ensure prevention from scientific frauds. Appropriate daily backups make it long-lasting and protecting data from aging damage. It also prevents users from non-authorized login accesses with tracking of network and access attempts.

Advantages

Important notable advantages of this system are that it runs entirely in the web browser with no client software needed, industry standard server supporting major operating systems (Windows, Linux and OSX), ability to upload and store external files, exporting protocols in pdf formats, permissions system to control data access, a group, individual and archived workspaces, typed data support (strings, floats, integers) and a familiar user friendly interface similar to a traditional paper laboratory notebook.

Future scope

In the second release, we plan to refine the interface and add a set of new features. The performance can be improved and the application can be scaled up to support millions of records. A user feedback support will be incorporated in the next version.

DISCUSSION

The laboratory notebook is the last remaining paper component of an experiment this century. Several electronic laboratory notebooks have been developed to digitize record keeping processes, but none has been successful in gaining popularity because of many adoption barriers. There are several issues identified by current electronic laboratory notebooks, some of them being cost, ease of use and their accessibility across different operating platforms, making scientists to make use of generic notebook softwares, paper laboratory notebook and spreadsheets. These methods of record tracking are being slowly replaced by the arrival of digital technologies, where the data can be easily searched, shared, easily backed up, and is readily accessed (Boulton et al., 2012). Electronic laboratory notebook can benefit researchers by facilitating long-term storage, reproducibility, and enhanced availability of experiment records (Bird, 2013; Hice, 2017). Recently, semantic laboratory notebooks (SLNs) have been utilized to explore these functions as a formalized metadata (Coles, 2013). There are multiple forms of electronic laboratory notebooks, some easy to use are Microsoft Word, Office 365, Google Docs, Evernote and OneNote (Evernote, 2016; Microsoft, 2016). These forms provide accessibility from any online computer, search ability, it can be simple and practical for some laboratories, but may not be compatible for others (TechCrunch, 2016). Furthermore, the presently available electronic laboratory notebooks have a major concern of cost. Table 2 states some of the barriers of adopting electronic laboratory notebooks in academic and non-academic setting.

Any developed software can be licensed as: paid for which is a proprietary piece of software that can be purchased, paid (with free version) is a proprietary piece of software that can be purchased, but which also has a version of this software which can be used for some time as free, open source which is a product where the code behind the actual software has been made openly available so that anyone can redistribute it, and finally a free product which is free to use. Cost is a significant barrier to electronic laboratory notebook adoption, further stay hours, troubleshooting, and maintenance with support add up to the cost (Bird, 2013; Goddard, 2009; Rudolphi, 2011). Free softwares would clearly have the advantage of being cheaper to run and test but may have disadvantage in terms of features. Another challenge to the electronic laboratory notebooks is ease of use when compared with the paper notebooks. Paper notebooks are easier to use, input data, read, transport, and inexpensive, with no training and minimal IT support (Bird, 2013; Brandl, 2010; Cooke, 2017). Web based tools weigh more as compared to platform dependent softwares, as web based are platform independent, while non-web based rely on certain operating systems with platform compatibility. In one of the academia survey, 74 researchers using currently available electronic laboratory notebooks expressed concerns about needing to enter data in both the laboratory and write-up area, leading to copying and pasting printouts into paper notebooks and manually transcribing data between notebooks and computers leading to data loss, transcription errors (Bruce, 2017; Myers et al., 2017). They also raised concerns about the difficulties of instrument integration, leading to difficulties in data compatibility and portability. Table 3 explains how the proposed system could mitigate the barriers of the current electronic laboratory notebooks.

Conclusion

After testing and validating this suite on beta-users for 48 days, it has proven to bring higher quality documentation, data recording and retrieving information with un-parallel higher security. It should immensely help researchers to strengthen collaborations between researchers. This suite is scalable to the size of laboratory by increasing server space with limited resources. While most of the beta-users were inclined to its usage, some stated their preference for a more free form environment with a desire to be able to make all data public and so were not inclined in incorporating its usage. We plan to introduce this suite as an open source program and can commercialize depending on its future demand. A video will be supplied with the package overviewing the main activities of suite including what you can do with this suite, who are the target audience, its features and supported image formats. Availability and requirements: Project name: Online Electronic Laboratory Notebook. Operating system: platform-independent. Programming language: PHP, HTML. Other requirements: server: Apache 1.3 or above, MySQL 4.23 or above; client: OS:

Table 2. States some of the barriers of adopting electronic laboratory notebooks (ELN) in academic and non-academic setting.

Category	Barriers States
Cost	Upfront costs and licensing fees, additional infrastructure costs (e.g. computers), future development and costs of applications
Ease of use	ELN was too difficult to use, does not capture the right or some kind of information for me
access	You'd need to enter data in both the lab and write-up area, no easy access to appropriate hardware in the lab
Data compatibility	Data will be tied into a commercial package

Table 3. Explains how the proposed system could mitigate the barriers of the current electronic laboratory notebooks (ELN).

Category	Desired features
Recording notes	Simple to install, personalisable, create default values, easy to write in as a paper notebook
Organizing notes	Highlightable, contents table, tag/classify notes and experiments, store metadata
Searching	Keyword search, data traceability
Linking data	Upload/Link files, images and data, links to notes between different notebooks
Writing reports	Integrate and store different types of documents, export functionality
Accessibility	Platform Independent, tablet/smartphone compliant
Archiving	Secure storage, backup and archives downloads/printing
Intellectual property	Secure access, different access levels for users
Collaboration	Shared notebooks
Project activities	Moderate comments
Cost	free of charge
Ease of use	Adding a semantic layer might improve attitude of scientists already using suite
Access	A cloud based ELN can be accessed anywhere with an internet connection
Data compatibility and portability	Using cloud software to store data would allow the user to export data in a variety of common data formats

Mac OS X. License: GNU General Public License (GPL). Any restrictions for use by non-academics: none. Known issues and test notes will be published with each release version.

CONFLICT OF INTERESTS

The authors have not declared any conflict of interests.

REFERENCES

Bird CL (2013). Laboratory notebooks in the digital era: the role of ELNs in record keeping for chemistry and other sciences. Chem. Soc. Rev. 42(20):8157-8175.

Boulton G, Campbell P, Collins B, Elias P, Hall W, Laurie G, O'Neill O, Rawlins M, Thornton J, Vallance P, Walport M (2012). Science as an open enterprise. The Royal Society London.

Brandl P (2010). Nicebook: supporting natural note taking. In: Proceedings of the SIGCHI conference on human factors in computing systems. ACM. pp. 599-608. https://dl.acm.org/citation.cfm?id=1753417

Bruce S (2017). A look at the state of electronic lab notebook technology. Sci. Comput. P 723. https://www.scientificcomputing.com/article/2002/12/look-state-electronic-lab-notebook-technology. (Retrieved Jan 2017).

Coles SJ (2013). First steps towards semantic descriptions of electronic laboratory notebook records. J. Cheminform. 5(1):1.

Cooke R (2017). Signature ip and clip: virtually clipping and dog earing pages in a Digital Lab Book. In: Proceedings of the user interface software (UIST), Fante Fe, USA, Mexico. http://eprints.soton.ac.uk/259251/.

Elliot MH (2018). Electronic Lab Notebooks: Market and

Regulatory Overview. http://www.atriumresearch.com/library/AtriumResearchSLA2008 (Retrieved Feb 2018).
Evernote (2016). Meet Evernote, your second brain. https://evernote.com/. (Retrieved Jan 2016).
Goddard NH (2009). eCAT: online electronic lab notebook for scienti c research. Automated Experiment.1(1):1.
Hice RC (2017). Roadmap to a clear de nition of ELN. Sci. Comput. https://www.scientificcomputing.com/blog/2009/05/roadmap-clear-definition-eln (Retrieved Jan 2017).
Laboratory Information Management System (LIMS) (2018). Laboratory Information Management System https://www.sunquestinfo.com/tag/laboratory-information-management-system-lims/ (Retrieved Feb 2018).
Microsoft (2016). Onenote. Evernote. https://www.onenote.com/. (Retrieved Jan 2016).
Myers JD, Mendoza ES, Hoopes B (2017). A collaborative electronic laboratory notebook. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.5.8267&rep=rep1&type=pdf (Retrieved Jan 2017).
Rudolphi F (2011). Electronic Laboratory Notebook: the academic point of view. J. Chem. Inform. Model. 52(2):293-301.
TechCrunch (2016). Evernote. Evernote's new privacy policy allows employees to read your notes. https://techcrunch.com/2016/12/14/evernotes-new-privacy-policy-allows-employees-to-read-your-notes (Retrieved Jan 2016).