Databases and ontologies

A universal open-source Electronic Laboratory Notebook

C.Voegele[1], B.Bouchereau[2], N.Robinot[1], J.McKay[1], P.Damiecki[2], L.Alteyrac[2*]

[1] Genetic Cancer Susceptibility group, International Agency for Research on Cancer (IARC)

[2] Information Technology Services group, International Agency for Research on Cancer (IARC)

ABSTRACT

Motivation: Laboratory notebooks remain crucial to the activities of research communities. With the increase in generation of electronic data within both wet and dry analytical laboratories and new technologies providing more efficient means of communication, Electronic Laboratory Notebooks (ELN) offer equivalent record keeping to paper based laboratory notebooks (PLN). They additionally allow more efficient mechanisms for data sharing and retrieval, which explains the growing number of commercial ELNs available varying in size and scope but all are increasingly accepted and used by the scientific community[1]. The International Agency for Research on Cancer (IARC) having already a LIMS[2] and a Biobank Management System[3] for respectively laboratory workflows and sample management, we have developed a free multidisciplinary ELN specifically dedicated to work notes that will be flexible enough to accommodate different types of data.

Availability and implementation: information for installation of our freeware ELN with source codes customizations are detailed in supplementary data.

1 INTRODUCTION

The requirements for the development of IARC ELN consisted of integrating all the features existing in the PLN together with additional functionalities to enable easy and intuitive navigation and information search capacities upon the notebooks [4].; following the definition of the Collaborative Electronic Notebook Systems Association: "System to create, store, retrieve and share fully electronic records in ways that meet all legal, regulatory, technical and scientific requirements; records being collections of information or data associated with an experiment to enable a suitably skilled person to repeat it".

The specifications included (i) recording of experimental data in a flexible way through a common platform suitable for the various research activities of the institute, and at the same time maintaining the rigorous scientific practises required for day-to-day notekeeping (traceability of each single modification), all within a highly secure environment; (ii) improving lab efficiency and work productivity by facilitating protocol and data sharing, record keeping, particularly in terms of limiting repetitive tasks; and (iii), providing an efficient means by which to search archived material,

finally (iv) being free, open-source and web-based requiring no installation of software on user's computer; (v) enabling long-term and secure storage in basic electronic format (text).

2 METHODS

Based on these specifications, we have tested several tools (cf. supplementary data) and finally chose WordPress 3.4.1 with Atahualpa 3.7.6. theme. WordPress is a modern, well adopted Content Management System (CMS), easy to design and to customize, written in PHP and based on a MySQL database. This user friendly open-source publishing software was found to fit the best the Agency's needs mainly because it enables the set-up of multi-users with tight access controls and provides a lot of options for customization.

A simple Linux server hosts both the web application and the database. A large number of free plugins offering different functionalities have been tested to find the most adapted ones (cf. suppl. data). The system is currently being developed, maintained and administrated by an IT developer and a bioinformatician.

The correspondence between the CMS terminology and the laboratory notebook has been set up as following: categories are equivalent to notebooks or groups of notebooks, posts are equivalent to pages, and comments to annotations.

3 RESULTS

ELNs provide significant advantages over PLNs. Data can be quickly entered, retrieved, located, shared and are easy to archive and backup; data including text, images, attached documents, links. An ELN can belong to one single person or be shared within and/or between research groups when being specific to one common project or one instrument or equipment (for example for maintenance follow up and troubleshooting issues). It can be used for proposal developments, meetings notes as well as centrally updated manuals or SOPs.

We will detail in the following chapters and also in the supplementary data how we managed to fulfill all the requirements and rules of a PLN (*) and the additional features making the ELN a far better alternative to PLN.

3.1 Access

The tool developed is entirely web-based and therefore accessible through any web browser connected to the internal firewall-

---

*To whom correspondence should be addressed.

protected network making it accessible from both laboratories and offices. The access is controlled by personal login and password (notably thanks to the "Login configurator" plugin). The IARC ELN is compatible with PC, Mac, tablets and smartphones. The access to data in reading and editing is defined very precisely and controlled by specific user permissions thanks to the 'Role Scoper' plugin (from completely private to public). The system offers the option for some notebooks to have more than one author or to be read only.

By default we put in place a 3-workspaces design for each research group in the Agency:

- "group workspace" dedicated to personal notebooks and containing one sub-workspace for each member of the group, each sub-workspace containing one or more notebooks (each member of the group has write access to his own notebook and read access to the other members notebooks).
- "share workspace" containing notebooks for instruments, robots or shared projects (with read and write access for the whole group).
- "archives workspace" containing all closed and archived notebooks of staff who have left the Agency (set to read-only access).

The users are offered the possibility to leave comments on all the notebooks that they have read access to.
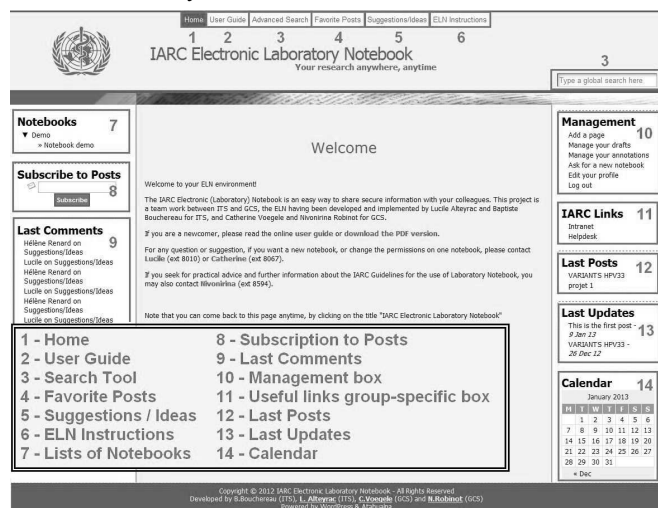

Fig. 1. ELN main page with navigation links and tool boxes.

3.2   Features

Our ELN system includes the following main functionalities:
- Easy to navigate interface and menus providing access to all users irrespective of their informatics skills (thanks to widgets, custom code and "Collapsing Categories" plugin)
- Advanced text editor similar to the well-known office ones and in which specific templates have been integrated for laboratory work including "Title of experiment, objectives, material and methods, results and conclusion"* and instrument troubleshooting with "Date, issue description, solution and conclusion" ("CKEditor" plugin with custom code).
- Insertion and attachment tool for different kind of files such as pictures, Word documents, Excel sheets, pdfs, screenshots, multimedia, hyperlinks (among which cross-referencing of notebook pages), plots, computer outputs (which allow incorporation of data from instruments avoiding printing and pasting into PLN)

- Automatic table of contents* of pages/records providing the possibility to sort by title or date (customization of Atahualpa code and "WP-PageNavi" plugin)
- Search tool by keywords ("Search by Category" plugin)
- Userstamp and timestamp for each update or annotations of a page* (recorded by default by WordPress and displayed by "Post Revision Display" plugin)
- Lists of favorite pages (Customized "WP Favorite Posts" plugin)
- Annotations or comments
- Post notification plugin that enables the user to be warned by email each time a new post is added into some selected notebooks ("Post Notification" plugin with custom code)
- Export to pdf ("WP-Print" and "Post2PDF Converter" plugins)
- Management of users permissions by ELN administrators ("RoleScoper" plugin)

3.3   Security

Security is ensured in terms of: (i) prevention from scientific fraud and protection of intellectual property: the fine-grained access controls make it more secure than any paper notebook; all amendments of records are stored in the database; a detailed documentation describes how to use the ELN which should prevent any deviation from the protocol; (ii) longevity: the html format of the raw data combined with appropriate daily backups makes it longlasting and prevented from damage or deterioration; the ELNs cannot be misplaced, lost or accidently destroyed; at departure of staff, the ELNs are closed and archived to read-only status; (iii) protection from external access (secured server and network, accesses and attempts of accesses are recorded).

Conclusion

After 6 months of testing and validation, the ELN accounts for around 50 regular users, 100 notebooks and over 1300 pages. This exploratory phase having been a complete success, the Agency has decided to switch from the PLN to the ELN since 1st of January 2013. The IARC ELN has been largely adopted and is now intensively used within the agency since it has proven to bring higher quality documentation, gain of time for data recording and retrieving as well as higher security. It should also strengthen collaborations and exchanges between researchers within the Agency. Last but not least it demands very low resources and is scalable whatever the size of the institution or laboratory is (cf. Supp. Data).

REFERENCES

Rubacha M, Rattan AK, Hosselet SC (2011) A review of electronic laboratory notebooks available in the market today. *J Lab Autom*. 2011 Feb;16(1):90-8

Voegele C and al. (2007) A Laboratory Information Management System for a high throughput genetic platform aimed at candidate gene mutation screening *Bioinformatics* 2007 Sep 15;23(18):2504-6

Voegele C, Alteyrac L, Caboux E, Smans M, Lesueur F, Le Calvez-Kelm F, Hainaut P (2010) A sample storage management system for biobanks *Bioinformatics* 2010 Nov 1 ; 26(21):2798-800

Wright JM (2009) Make it better but don't change anything. *Autom Exp* 2009 Nov 19; 1:5