OpenLabNotes - An Electronic Laboratory Notebook Extension for OpenLabFramework

Markus List[1,2,3,*], Michael Franz[4], Qihua Tan[3,5], Jan Mollenhauer[1,2†], Jan Baumbach[4,6†]

[1]Lundbeckfonden Center of Excellence in Nanomedicine (NanoCAN), University of Southern Denmark, 5000 Odense, Denmark

[2]Institute of Molecular Medicine, University of Southern Denmark, 5000 Odense, Denmark

[3]Clinical Institute, University of Southern Denmark, 5000 Odense, Denmark

[4]Computational Biology group, Department of Mathematics and Computer Science, University of Southern Denmark, 5230 Odense, Denmark

[5]Epidemiology, Biostatistics and Biodemography, Institute of Public Health, University of Southern Denmark, 5000 Odense, Denmark

[6]Max Planck Institute for Informatics, 66123 Saarbrücken, Germany

Summary

Electronic laboratory notebooks (ELNs) are more accessible and reliable than their paper based alternatives and thus find widespread adoption. While a large number of commercial products is available, small- to mid-sized laboratories can often not afford the costs or are concerned about the longevity of the providers. Turning towards free alternatives, however, raises questions about data protection, which are not sufficiently addressed by available solutions. To serve as legal documents, ELNs must prevent scientific fraud through technical means such as digital signatures. It would also be advantageous if an ELN was integrated with a laboratory information management system to allow for a comprehensive documentation of experimental work including the location of samples that were used in a particular experiment. Here, we present OpenLabNotes, which adds state-of-the-art ELN capabilities to OpenLabFramework, a powerful and flexible laboratory information management system. In contrast to comparable solutions, it allows to protect the intellectual property of its users by offering data protection with digital signatures. OpenLabNotes effectively closes the gap between research documentation and sample management, thus making OpenLabFramework more attractive for laboratories that seek to increase productivity through electronic data management.

[*]To whom correspondence should be addressed. Email: mlist@health.sdu.dk
[†]joint last author

1 Introduction

In laboratory research, experimental work is typically documented in paper laboratory note-books (PLNs). This does not only help researchers to keep track of the details of their experiments, but also protects their interests, and those of their institutions, in legal disputes. For a long time, electronic data storage was not considered a legally acceptable alternative, since information could be falsified more easily than in PLNs. Therefore, electronic laboratory note-books (ELNs) were only considered as a supplement, but not as a substitute for PLNs. However, ELNs have clear advantages compared to their paper-based alternatives. Laboratories are open environments and thus documents can get accidentally destroyed, lost, or, in the worst case, stolen. ELNs, on the other hand, are typically paired with an automated and reliable backup system. Time stamps and additional meta information can further contribute to the reliable documentation of costly experimental efforts. Moreover, there are immediate advantages in daily use, such as full text search, which allows researchers to retrieve information quickly and to share documents with their colleagues. In this way, ELNs can significantly increase productivity and make it easier to keep track of experimental details despite of frequent turn-over in laboratory staff [1]. Thus, when the U.S. Federal Drug Administration introduced a guideline for electronic information storage (CFR21 part 11[1]), ELNs slowly began to replace PLNs in many laboratories.

In order to follow the regulations, an ELN needs to guarantee that documents are safely stored. The history of a document, such as the creation date and the author, must be traceable. To have credibility in a law case, documents must be protected from changes. This implies that a document in the ELN should be signed by its author. Ideally, it is possible for a colleague or a supervisor to countersign a document to further increase its credibility. The latter is particularly important when it comes to patenting, since a countersigned document that includes a timestamp can be considered strong legal proof for an invention. In a typical software system, a user proves his identity by securely logging into a system with a unique username and a secret password. However, instead of merely marking a document as signed by a logged in user, a so-called digital signature should be used to prevent scientific fraud not only on the application, but also on the database level. Digital signatures rely on a pair of asymmetric keys, a private key, to sign the document, and a public key, to validate the document [2].

ELNs focus on the thorough documentation of experimental procedures, including applied methods, consumables, as well as experimental results. Another type of electronic system that can increase productivity in laboratories is the laboratory information management system (LIMS), which is mostly concerned with sample management. Integrating an ELN with a LIMS enables researchers to link experimental documentation directly to the samples that were part of the experiment. This is particularly useful, if an experiment is to be repeated or continued by another researcher not familiar with the experiment.

We have previously developed OpenLabFramework (OLF), an open-source LIMS for sample management in mid-sized laboratories tailored towards the needs of the functional genomics community [3]. OLF has a modular structure and can be easily extended. Modules imple-

---

[1] http://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfcfr/CFRSearch.cfm?CFRPart=11

mented so far include OpenLabGeneTracker, for management of functional genomics samples, OpenLabStorage, for physical sample tracking across customizable storage hierarchies, Open-LabBarcode for printing and scanning of QR-barcode labels, and OpenLabAttachments for uploading arbitrary documents. OLF can be accessed through a web browser and offers a flexible and intuitive user interface. Managing experimental documentation in OLF was so far limited to uploading additional files or to use the comment fields of samples. While this allows for some level of documentation, it is not a convenient solution and falls short of the legal requirements an ELN is expected to fulfill.

This motivated us to develop OpenLabNotes, a new module that adds ELN features to OLF and in this way, combines the advantages of an ELN with that of a LIMS. The focus of this module is the tight integration with existing functionality, as well as a robust digital signature mechanism.

2　Related Work

A large number of commercial ELNs are already on the market [4]. However the costs for these systems can typically not be covered by small to mid-sized laboratories. Free alternatives range from general note-keeping tools to dedicated ELNs. A comprehensive list of commercial and free ELNs can be found online[2]. The majority of the free systems, however, do not sufficiently protect the documents to prevent scientific fraud. Moreover, many of these are web applications that are only accessible online and cannot be deployed in a local network. Thus, researchers using these systems might be in violation of strict data protection rules found in many academic institutions. Voegele *et al.* found that none of the existing ELNs could sufficiently fulfill the requirements found in a typical laboratory and, thus, developed a new open-source ELN based on WordPress [5]. Their IARC ELN prevents scientific fraud on the application level by user authentication and on the database level by securing and monitoring database access. The lack of digital signatures in this system, however, may pose a problem in case of patenting, where it is advantageous if a document describing an experiment that led to an invention is actively signed and protected by an officially recognized digital signature algorithm. Finally, none of the available open-source ELNs, including the IARC ELN, feature the integration with a LIMS. We thus sought to extend OLF to close this gap. A detailed comparison between the features of the IARC ELN and OpenLabNotes can be found in Table 1.

3　Implementation

3.1　Integration with OLF Functionality

OLF is implemented using the Java and Groovy based Grails web application framework[3]. Grails utilizes hibernate[4] to encapsulate database interactions with domain classes. In OLF,

---

[2]http://www.atriumresearch.com/html/eln.htm
[3]http://www.grails.org
[4]http://www.hibernate.org

Feature	IARC ELN	OpenLabNotes
Workspaces / user groups	✓	✗
Notebooks	✓	✓
Share with specific user	✓	✓
Projects	✗	✓
Comments / annotations	✓	✗
Navigation interface	✓	✓
Rich text editor	✓	✓
Editor templates	✓	✗
File attachments	✓	✓
Sort by title or date	✓	✓
Userstamp and timestamp	✓	✓
Search by key words	✓	✓
Full text search	✗	✓
E-mail notifications	✓	✗
Export to PDF	✓	✓
Export to docx	✗	✓
Notes stored as HTML	✓	✓
Integration with LIMS	✗	✓
Digital signing (author)	✗	✓
Digital cosigning (supervisor)	✗	✓

Table 1: Feature comparison between the IARC ELN and OpenLabNotes.

domain classes typically extend one of several base classes, which allows OLF modules to provide functionality irrespective of the domain class that will ultimately rely on it. In this way, existing modules such as OpenLabFileAttachment can accommodate the additional domain class NoteItem that is introduced by OpenLabNotes without any changes. Vice versa, OpenLabNotes allows samples to be linked to DataObjects that are generated by modules such as OpenLabGeneTracker (Figure 1). Moreover, the NoteItem class is recognized by the Compass search engine[5], enabling full text search on all database fields. Finally, any OLF module can serve as a content provider for other modules or the front-end. Templates for building additional tabs, menus, or add-ins can be added. This mechanism is utilized, for instance, to automatically list linked notes in views specific to arbitrary domain classes without the need for additional customization.

3.2 Rich Text Editor

In order to motivate researchers to implement and to use an ELN, it is crucial to provide the text editing experience known from popular office tools such as Microsoft Word®. We chose TinyMCE[6] as an appropriate javascript rich text editor, due to its excellent documentation and

---

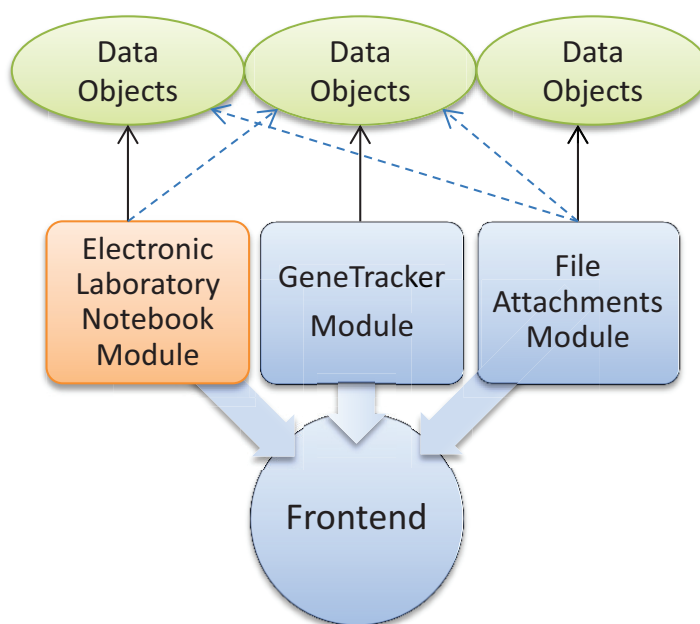[5]http://www.compass-project.org/
[6]http://www.tinymce.com/

Figure 1: Integration of OpenLabNotes in OLF. To maintain its flexiblity, OLF keeps track of arbitrary domain classes that extend a common class DataObject. This allows for complex interactions between modules. OpenLabNotes, for instance, generates DataObjects representing experimental notes and links them to other DataObjects, such as samples generated by OpenLabGeneTracker. The FileAttachment module can generate DataObjects encapsulating uploaded documents and links them to either DataObjects irrespective of whether they represent samples or ELN notes.

its rich plug-in ecosystem. TinyMCE relies on HTML for formatting and can be configured to allow images to be pasted in binary format.

3.3 Digital Signatures

ELNs are expected to comply with regulations for effective data protection. This implies that access is generally limited to authorized users. In OLF, this is facilitated through the industry-prove Spring Security project[7]. In addition, documents need to be digitally signed to guarantee origin, authenticity and integrity. Hash functions, such as SHA-1, can be used to generate unique identifiers of an input string representing a document. Hash functions are one-way functions that produce a different output if the input string is changed. In this way, scientific fraud can in principle be detected. This, however, does not establish a sufficient level of protection, since the hash string in the database can simply be recomputed based on the modified document. Thus, the hash string is protected through asymmetric public-key cryptography (Figure 2). First, a public/private key pair needs to be generated for each user. When a user signs a document, the private key, together with a password, is used to generate a digital signature. Subsequently, a public key can be used by any user to decrypt the digital signature and to validate the integrity of the signed document. Digital signatures are typically embedded in a document. However, both the author and a supervisor can sign a document, which would

---

[7]http://projects.spring.io/spring-security/

thus be stored redundantly in the database. Hence, we decided to generate detached signatures that are kept in addition to the original document. We apply the Digital Signature Algorithm (DSA), which is an officially recognized standard based on a variant of the ElGamal digital signature [2], which in turn is based on the Diffie-Hellman key exchange method [6]. To add DSA capabilities to OpenLabNotes, the Bouncy Castle Java library[8] was added. Apart from the document itself, it is imperative to protect timestamps, such as the creation date and the date of signing. Therefore, this information is included when creating the digital signature and for the validation.
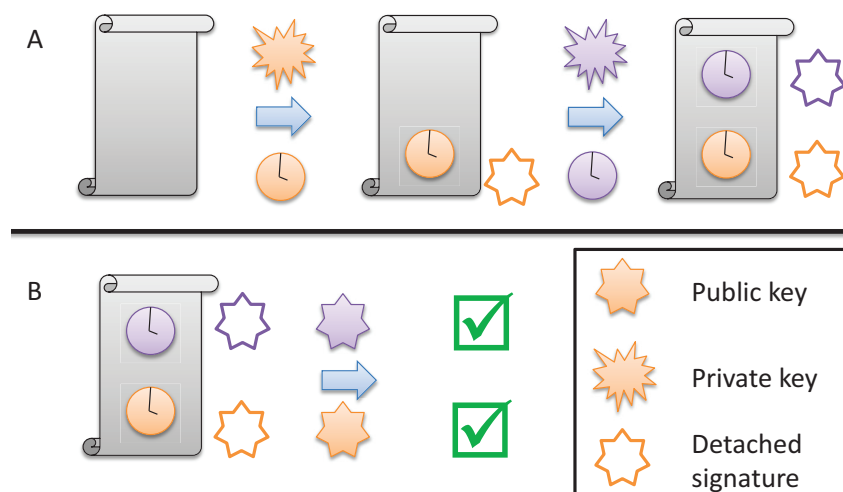

Figure 2: Digital signatures require a pair of asymmetric keys that is generated by the system for each user. A: The author uses his/her private key, his/her password and a timestamp to sign a note. A second signature is added by the supervisor or a colleague based on his/her private key, the current timestamp and the timestamp of the author. The result is a detached signature for the author and the supervisor, respectively. B: The public keys of the author and the supervisor, together with the original note and the original timestamps can be used to validate the signatures.

3.4　Document Export

To be able to export content generated with TinyMCE to other formats, we first parse HTML to valid XHTML using JTidy[9]. Subsequently, we utilize docxj4[10] to obtain docx or PDF documents for export.

4　Results

We developed OpenLabNotes to extends OLF with ELN capabilities (Figure 3). A rich text editor allows documents, which are called notes in OpenLabNotes, to be created with the various formatting options that a user typically expects from a modern text editor. Images can be placed

---

[8] http://www.bouncycastle.org
[9] http://jtidy.sourceforge.net/
[10] http://www.docx4java.org/trac/docx4j

within the text and will be stored as part of the note. Each note can be edited until it is finalized by the user (draft status). In this case, the note, as well as a timestamp, are digitally signed with the private key and password of the user (final status). When signing a note, the user needs to appoint a supervisor, which can then countersign it (signed status). These signatures are subsequently validated each time a note is shown to a user. In order to document the origin and current location of material used in an experiment, notes can be linked to arbitrary samples in OLF. Moreover, notes can be attached to specific projects or organized in notebooks, which are user-defined collections of notes. The documentation of an experiment can be further complemented by uploading additional files. Frequently, notes need to be accessible to researchers other than the author or the assigned supervisor. To this end, each user with access to a note can also grant access to other users. To retrieve notes, several types of lists are available, including a list of authored notes, a list of accessible notes, a list of notes that were signed in the role of a supervisor and most importantly, a list of notes a supervisor still has to sign. Lists can be filtered by author, supervisor, or project, as well as by several types of dates such as the creation date or signature dates. Each list includes a preview of the notes and can also be exported to various formats, including xlsx, odt, doc, and PDF. When browsing samples, a list of linked notes is shown in a dedicated tab. Most importantly, notes can be retrieved through the search functionality, which extends to entire text of the note. Finally, notes can be printed or exported to either PDF or docx format.

5 Discussion

ELNs revolutionize the way in which experimental notes are kept by making them more secure, accessible and shareable. The requirements to data safety make ELNs more reliable than their paper-based alternatives and it can be expected that researchers will be required to use ELNs in the future [1]. The choice of an appropriate solution, however, is difficult. The market offers a wide variety of products resulting in a large choice of attractive and reliable products that are mostly superior to free alternatives. As a disadvantage, these solutions are often not affordable to laboratories. Moreover, it is unclear if a company will endure in the competition and still be able to provide support in a long term perspective. It might therefore become necessary to migrate from one ELN to another, which is complicated by the lack of a common data format for ELN data storage and the fact that some products do not offer a backup mechanism in the first place. We therefore consider open source ELNs an attractive alternative to commercial products.

OpenLabNotes is an open source ELN extension for OLF that comes with advantages typically found in state-of-the-art note keeping tools, such as sharability, full text search, a reliable backup system, and a modern text editing experience. Using HTML for storing notes simplifies data export, long term storage, and the potential migration to another ELN. To our knowledge, OpenLabNotes is the first open source solution to provide a digital signature mechanism to prevent scientific fraud. Being open source, it can also be easily adapted to use alternative digital signature methods should the legal requirements change. The intended work-flow suggests that each note is signed by two authors, which serves to protect the researchers in a potential law case.

Figure 3: The show note view in OpenLabNotes. A: The menu allows to select different types of lists or to create a new note. B: The history of the note, including checkmarks for valid signatures. C: Notes can be linked to samples, projects or notebooks. Moreover, files can be uploaded and attached and notes can be shared with other users. D: Here, the note is shown in a read only viewer. E: The note can be exported to PDF or docx.

To optimally protect the digital signatures in OpenLabNotes, it is imperative to include an external identity proof. To this end, a variety of plug-ins are available for the spring security library used here. These offer support for established protocols, such as OAuth2, OpenID, CAS, LDAP, or Kerberos. Exemplary, we included the spring-security-cas plug-in, allowing OLF to be configured to use a CAS server commonly found in many academic institutions. A disadvantage of this approach, however, is that the public keys generated for the user will still be managed by OLF. A more secure approach would thus allow public keys to be managed by an external certificate authority. Support for this will be implemented in the future.

A common issue when introducing electronic data storage to a laboratory is that often not only an ELN, but also an efficient sample management solution is required. Tightly integrating an ELN with a LIMS allows experimental information to be linked efficiently, thus increasing experimental documentation overall. While any open source ELN can be integrated with some developmental effort into a LIMS or vice versa, a single solution offers the advantage of re-

duced maintenance effort. Furthermore, laboratories and IT services at academic institutions are typically concerned with the longevity of data management systems. Open source projects such as this one are at risk of being abandoned if not adopted by a larger community. While OpenLabNotes is currently used actively by laboratory personal within the NanoCAN center, broad support for the project is still lacking. This might change in the foreseeable future, since several groups have contacted us with plans to adopt OLF for their research. Nevertheless, it is imperative to provide optimal support for data export to address these concerns. Notes stored in OpenLabNotes can thus be exported to a variety of formats, such as CSV or XML, allowing content to be migrated to an alternative system should the need arise. In a long term perspective, however, a standard format for ELN data storage needs to be defined to allow research notes to be migrated more easily between systems.

In conclusion, OpenLabNotes is a powerful and user-friendly alternative to costly commercial ELNs. The digital signature mechanism implemented here follows accepted standards and should thus be suitable to fulfill legal requirements for data protection. Finally, the addition of OpenLabNotes to OLF effectively closes the gap between experimental documentation and sample management. This is expected to make OLF significantly more attractive as a comprehensive solution for laboratory information management.

Acknowledgements

Conflict of Interest: none declared.

Availability and Requirements

Source Code:	`https://github.com/NanoCAN/OpenLabNotes`
Manual:	`https://github.com/NanoCAN/OpenLabFramework/wiki`
Demo:	`http://www.nanocan.dk/openlabframework/demo`
	(user: admin, password: demo0815)
Operating system(s):	Platform independent
Programming language:	Java, Groovy, Java-script
Other requirements:	Java 1.6-1.8 with the Java Cryptography Extension
	Tomcat 6.0 or higher for server deployment
License:	GNU GPL v3.

References

[1] S. Y. Nussbeck, P. Weil, J. Menzel, B. Marzec, K. Lorberg and B. Schwappach. The laboratory notebook in the 21st century: The electronic laboratory notebook would enhance good scientific practice and increase research productivity. *EMBO reports*, 2014.

[2] T. Elgamal. A public key cryptosystem and a signature scheme based on discrete logarithms, 1985.

[3] M. List, S. Schmidt, J. Trojnar, J. Thomas, M. Thomassen, T. A. Kruse, Q. Tan, J. Baumbach and J. Mollenhauer. Efficient sample tracking with openlabframework. *Scientific Reports*, 4, 2014.

[4] M. Rubacha, A. K. Rattan and S. C. Hosselet. A review of electronic laboratory notebooks available in the market today, 2011.

[5] C. Voegele, B. Bouchereau, N. Robinot, J. McKay, P. Damiecki and L. Alteyrac. A universal open-source electronic laboratory notebook. *Bioinformatics (Oxford, England)*, 29(13):1710–1712, 2013.

[6] W. Diffie and M. Hellman. New directions in cryptography. *IEEE Transactions on Information Theory*, 22(6), 1976.