Guidelines for introducing an ELN/LIMS in academic research laboratories

This document provides practical guidelines on how to introduce an Electronic Laboratory Notebook (ELN) and Laboratory Information Management System (LIMS) in an academic research laboratory. Whilst it is based mainly on our experiences with life sciences laboratories, these guidelines could be easily extended across other research domains. This document is complementary to the **lists of ELNs and LIMS available on the market**, which can be used as guidance to select the appropriate tool for a research laboratory.

This document includes the following sections:

1. Introduction to ELNs and LIMS
2. Why should academic laboratories adopt an ELN/LIMS?
3. Key roles for the ELN/LIMS deployment
4. Stages of ELN/LIMS deployments in an academic lab
5. Key success factors

DLCM | swissuniversities

1. Introduction to ELNs and LIMS

Research scientists are required to document their work in laboratory notebooks. Traditionally, these were paper notebooks compiled following specific guidelines set by each company and university. The advent of automation in the last century led to a revolution in laboratory workflows and practices. The number of measurements that can be taken in a single lab increased exponentially over time, with an associated increase in the number of samples handled, procedures to follow and generated data. This required a change in the procedures used for documenting all experimental (and computational) steps and raised issues of data storage and tracking. ELNs were developed to provide a single platform where researchers could document their work and connect their experimental data. At the same time, Laboratory Information Management Systems (LIMS) were developed for managing samples and laboratory operations. These systems were first introduced in companies, to fully automate all processes and they are now state of the art. Whilst ELNs and LIMS started as separate platforms to address different laboratory needs, nowadays many systems have combined ELN/LIMS capabilities. The number of ELNs and LIMS solutions available on the market is constantly increasing, with more and more specific solutions for academic labs.

1.1 Experimental workflow in academic laboratories

Accurate documentation of lab work is necessary to ensure data reproducibility and transmission of knowledge. This process involves four main steps:

1. **Management of materials** All biological, chemical, electronic devices, and any other types of samples or materials bought or produced in a lab need to be tracked.
2. **Management of protocols.** All standard experimental and theoretical procedures need to be recorded.
3. **Description of experimental and computational procedures**. All details of experimental, processing, analysis steps need to be documented. These include instruments used, set up parameters, experimental conditions. All samples used and protocols followed need to be specified in each experimental procedure. All processing and analysis steps, scripts, software (with version), operating system need to be documented to ensure data reproducibility.
4. **Data management**. Raw data from measuring instruments, processed data, analyzed data and codes need to be safely stored, backed up, and annotated with enough metadata to be interpretable and codes should be versioned.

These four steps are usually separate in academic research labs. Materials are often stored in excel files or dedicated databases; protocols are written in Word files or wikis. Experimental and analysis procedures are still often documented in paper laboratory notebooks. The data is usually stored in different locations, such as on local hard disks of measuring instruments, on researcher's personal computers, on group shares.
A combined ELN with data management capabilities and a LIMS enables researches to store all these different kinds of information in a single platform, easily searchable and retrievable by anyone with access to the system.

2. Why should academic laboratories adopt an ELN/LIMS?

2.1 Reduce data loss

The academic world is currently struggling with the "data paradox": on one hand a lot of time and effort is spent on generating and analysing ever-increasing amounts of data; on the other hand, often very little is done to safeguard this data. It is, unfortunately, still very common to have data spread across different storages and resources, such as local hard disks of measuring instruments, researchers' personal computers, USB sticks, external hard drives, group shares. However, data without context is meaningless: if there is no information about the data stored in a given location, this data, which in principle could be very valuable, becomes useless and can be considered lost. It is also of utmost importance to have a backup mechanism in place for the data: it is not uncommon to have data stored only in one location and lose it because of hard-disk failures or thefts (see http://science.sciencemag.org/content/354/6317/1242.1?rss=1 as an example).
ELN/LIMS systems ensure that data is stored in a known location, data is contextualized by metadata and backup mechanisms are always in place.

2.2 Meet data management guidelines from funding bodies

More and more funding agencies are now requesting data management plans (DMPs) in their application proposals. For example, Horizon 2020 and the Swiss National Science Foundation (SNSF) made the DMP mandatory for all disciplines in 2017.
The usage of a specific ELN/LIMS system can be included in the DMP.

2.3 Facilitate data sharing

Nowadays, collaborations between research groups in different geographical locations across the globe are the norm. There is therefore a need to share data and information between collaborating research groups. If each research group had an ELN/LIMS to describe and store their data, they could provide external access to their collaborators, thereby putting an end to data exchange by email, which is still the most common practice.

2.4 Improve data traceability and reproducibility of results

The FAIR principle of scientific data management state that data should be Findable, Accessible, Interoperable and Reproducible (*Scientific Data* **3**, Article number: 160018 (2016) DOI: doi:10.1038/sdata.2016.18). This applies to all data generated in a lab: raw data from measuring instruments, processed data, and analysed data. The data needs to be safely stored, backed-up, and annotated with enough metadata to be interpretable and reproducible by anyone with sufficient knowledge in the field. The usage of an ELN/LIMS helps in ensuring data FAIRness.

3. Key roles for the ELN/LIMS deployment

The introduction of an ELN/LIMS in academic laboratories is a complex process that involves several people with different expertise. It is important to identify key roles in this process.

3.1 Principal Investigator (PI)

In order to ensure successful ELN/LIMS deployment, the PI should fulfil the following tasks:

a. Initiate the project
b. Assume overall responsibility for the project and the objectives
c. Ensure that the objectives of the project match the strategies and requirements of the laboratory organization
d. Allocate resources and ensure their efficient use
e. Make timely decisions on proposals
f. Designate the people in charge of ELN/LIMS in the laboratory (lab. referent) and ensure adequate cooperation

3.2 ELN/LIMS laboratory referent

One person in the research lab should be in charge of coordinating the implementation of the ELN/LIMS. This role should be given to senior scientists or technicians, who have in-depth knowledge of the laboratory workflows.

3.3 ELN/LIMS application expert

The application expert is responsible for configuring the system according to the laboratory referent's requests, and to provide advice on how to best implement the system in the laboratory environment. In most cases, this will be someone from the Research IT group of the Institution or from partner institutions.

3.4 End-users: researchers and other laboratory members

ELN/LIMS end-users use the tool in their everyday work. Active participation of end-users in the deployment is highly encouraged, as it fosters the quality of the process. Their remarks and comments are crucial to deploy a tool that fits with their habits.

4. Stages of ELN/LIMS deployments in an academic lab

4.1 Initialisation

This pre-deployment phase starts with informal discussions and exchanges between the ELN/LIMS application expert and the PI or the main laboratory referent. It is essential to organize at least one demonstration of the software based on the expressed needs.
The initialization phase aims to confirm that the laboratory needs this tool. If so, a precise list of objectives is defined by the laboratory and prioritized.

4.2 Analysis of laboratory needs

In this phase, the ELN/LIMS application specialist analyses the requirements of the laboratory more in depth. He/she needs to:

- Understand current laboratory notebook practices (documentation of experiments, management and location of samples, instrument management, centralized procedures)
- Understand laboratory workflows
- Validate and, if needed, update the prioritized requests list
- Identify the risks that could compromise the smooth progress of the deployment
- Find the right working methodology and adapt it in accordance with the laboratory referent agenda and laboratory imperatives

4.3 Test phase

After identification of the needs of the laboratory, and commitment of the PI on the necessity of the introduction of a specific ELN/LIMS in the lab, there should be a limited period of time in which the laboratory can test the system, with the support of the ELN/LIMS application specialist. During this period, members of the laboratory should be able to test the application and provide their feedback on whether it fits their needs. The test phase should include in-depth training for the ELN/LIMS laboratory referent and user training for all laboratory members.

4.4 Production phase

At the end of the test phase, the PI should provide his/her commitment to using the tested ELN/LIMS. The production phase can then start and should include:

a) Migration of existing databases for tracking samples already in use. Often this information is stored in excel files, which should be imported in the new system.
b) Migration of standard protocols to the new system.
c) Customization of the platform with the support of the ELN/LIMS application specialist.
d) Integration with measuring instruments for direct raw data upload. Please note that not all laboratories will require this feature.
e) Integration with analysis pipelines and workflows.

5. Key success factors

To ensure a successful implementation of an ELN/LIMS in an academic laboratory, we identify three important requirements:

1. **Sponsorship & commitment.** Strong sponsorship and commitment at the institutional level (research board, the PI and laboratory manager) is crucial. A sustainable business model must be created.
2. **Training.** The data management problem should be widely acknowledged by the researchers. Training plays a fundamental role in raising awareness about this issue. Very often, researchers think that spending time in setting up an ELN/LIMS deprives them of valuable time that could be spent doing research. Researchers need to become aware that this time is an investment for the future, which will lead to time saved whenever the need arises to look for old data and information.
3. **Communication.** Strong communication between stakeholders, through regular meetings, and permanent reassessment of the requirements is essential.

5.1 Understanding the new paradigm in specialized IT business

In general, IT professionals focus on technology. However, it is important to distinguish between two types of IT groups within academic centres:

- **Central or administrative IT:** support the administrative computing needs of large organizations, and is generally organized in one specific IT department and project office.
- **Research IT:** support the computing needs of researchers by providing highly specialized tools such as ELN/LIMS, high performance computing, big data or other scientific IT tools.

Working in today's life sciences research environment requires IT support that goes beyond the installation of computers on researchers' desktops or even digitizing all of the data. This task requires various skills in project management, IT engineering, the ability and the background knowledge to understand trends in scientific research and laboratories organization.

5.2 Understanding the new paradigm in life sciences research

The amount of data generated in research labs is constantly increasing. Most researchers are aware of the limitations of general-purpose applications such as spreadsheets. However, the majority of them continue to rely on those due to their ease of use, low cost, and habit. Today, working in a life sciences research environment requires more than a perfect knowledge of the core business. Managing the growing amount of information should systematically become part of a research project. In order to perform this task efficiently, it is essential to master data management tools, such as ELN/LIMS.