Ouvrir la science !

Report of the Working Group on Electronic Lab Notebooks

French Committee for Open Science

September 2021

Report of the Working Group on Electronic Lab Notebooks

French Committee for Open Science

Tovo Rabemanantsoa - Pilote
INRAE

Dominique Pigeon - Pilote
Inserm

Gilles Mathieu – Pilote
Inserm

Christophe CHIPEAUX
ICOS/INRAE

Simon DUVILLARD
Réseau CURIE

Célya GRUSON-DANIEL
Université Lyon I

Marie HERBET
Université Lyon I

Arnaud LEGRAND
INRIA/CNRS

Nathalie LEON
CNRS

Domenico LIBRI
CNRS/INSB

Jean-Baptiste LILY
Université Lyon I

Jean-François PEYRAT
Université Paris Saclay

Agnès PINET
CEA

François SABOT
IRD

Moussa SEYDI
IRD

Véronique THEISEN
CEA

This document is the translation of a report first published in French :

Gilles Mathieu, Dominique Pigeon, Tovo Rabemanantsoa, Christophe Chipeaux, Simon Duvillard, et al.. Rapport du groupe de travail sur les cahiers de laboratoires électroniques. [Rapport de recherche] Comité pour la science ouverte. 2021, 68 p., DOI : 10.52949/3

September 2021

Conception graphique : opixido

Table of contents

1 Document background and objectives

1.1 Background

The laboratory notebook is a logbook used to record the day-to-day activities of research projects. It serves to track experiment descriptions and employed protocols, as well as log individual contributions. Using it ensures that requirements in terms of quality and scientific integrity are met by guaranteeing the traceability of the scientific methods employed and the reproducibility of research data and results. It therefore provides patent offices with proof of an invention, of its inventors, and thus of its rightholders.

The electronic lab notebook is the dematerialised digital version of the lab notebook. The term "electronic", often replaced by "digital" in other fields, is the most frequently used adjective for lab notebooks.

Due to its crucial place in research activities and its role in the management and protection of scientific knowledge, the electronic lab notebook is an essential strategic tool that is fully in line with open science. Moreover, it must meet the same scientific objectives and challenges as the physical notebook, which is why the choice of tool, its configuration, and how it is used are so important.

In order to inform scientific teams in choosing between the various solutions currently available, a working group was set up as part of the Research Data College of the Committee for Open Science within the French Ministry of Research and Higher Education.

Terminology: In the remainder of this document and in a ll the work produced by the working group, the Electronic Lab Notebook is referred to by the acronym ELN.

1.2 The working group's objective

The group's mission statement (in the appendices) specifies the objective of the working group, which is to propose an analysis method to scientific teams that have to choose an electronic lab notebook to meet their specific needs. To carry out this study, the working group defined the following objectives:

- Establish a shared vision of the definition, framing, uses and functional scope of the "electronic lab notebook" (see section 2).
- Produce a set of recommendations for interoperability between different electronic lab notebook tools and other tools or information systems already in use. Laboratory notebooks must be able to seamlessly merge with existing computer environments (including data

repositories), especially in a French academic multi-funder[1] context (see section 3).
- Produce a set of recommendations on the criteria for choosing a tool, based on functional needs, disciplines, and fields of research, as well as institutional requirements (see section 4).
- Draw up a comparative list of some existing tools based on the recommended criteria in order to measure their effectiveness in and/or suitability to different usage contexts (see section 5).
- Produce a set of recommendations to successfully implement an ELN (see section 6).

Note that the group's objective is to propose decision-making support criteria and not to provide definitive methods of comparison between different tools, nor to advocate the choice of a single tool for the entire higher education and research community.

1.3 Group composition

The working group is composed of 16 people from different institutions, representing the scientific diversity of higher education and research. Each member, according to their discipline and experience, was able to contribute their expertise, from the following points of view:

- **Scientific/experimental**: with a hands-on view, as a user
- **Legal and strategic:** consideration of aspects related to scientific integrity, intellectual property, and content promotion.
- **Technical:** both in terms of infrastructure and tool hosting, and of issues relating to data management, interoperability and semantics.

[1] especially in France where academic research is driven and funded by different bodies and organisations.

1.3.1 Leaders

The following people have been commissioned by the Committee for Open Science to lead the group's work:

- Gilles Mathieu, Research engineer in computer science at the Inserm Information System Department (ISD)
- Dominique Pigeon, Research engineer in computer science at the Inserm Information System Department (ISD)
- Tovo Rabemanantsoa, Member of the INRAE Directorate for Open Science (DipSO)

1.3.2 Members

The following people have been members of the working group:

- Christophe Chipeaux (ICOS/INRAE)
- Simon Duvillard (Réseau CURIE)
- Célya Gruson-Daniel (Inno^3/UTC)
- Marie-Emilia Herbet (Université Lyon I)
- Arnaud Legrand (INRIA/CNRS)
- Nathalie Leon (CNRS)
- Domenico Libri (CNRS/INSB)
- Jean-Baptiste Lily (ICOS/INRAE)
- Jean-François Peyrat (Université Paris Saclay)
- Agnès Pinet (CEA)
- François Sabot (IRD)
- Moussa Seydi (IRD)
- Véronique Theisen (CEA)

1.4 How the WG worked

The working group carried out its work between November 2020 and July 2021. Discussions took place by videoconference only, in the following format:

- plenary meetings held monthly for a duration of one hour.
- workshops that were offered to WG members, to work on particular topics.
- Exceptionally, external participants took part in sessions or were asked to discuss certain topics or contribute their thoughts on them. The list of persons concerned is provided in the appendices.

In addition, the work was coordinated with that of the "Successfully appropriating open science”[2] working group led by Anne Vanet at the Committee for Open Science and with the “Electronic Lab Notebooks” project which is underway at the CNRS, led by Nathalie Léon and Domenico Libri. To this end, the working group included a member appointed from the Data College, a member of the “Successfully appropriating open science” working group, as well as the two CNRS project co-leaders.

1.5 Purpose of the report

The working group has drawn up recommendations for electronic laboratory notebooks in two forms :

- Firstly, this report, which sets out the working group's approach, presents the tools and examples of their use.
- Secondly, a methodological tool: a grid of criteria to help choose an electronic lab notebook (editable table, with calculation formulas included); these criteria are accompanied by proposals for weighting according to use.

Two files make up this tool: “ELN WG - Tools - Criteria Grid.odt” provides instructions for the use of the different tabs in “ELN WG - Tools - Criteria Grid.ods”.

.

[2] https://www.ouvrirlascience.fr/reussir-lappropriation-de-la-science-ouverte/

2 Definition and scope: a shared vision of the electronic lab notebook

2.1 Definition

An *Electronic Lab Notebook*, or ELN, is a software program that collects and pools all the essential information for the reproduction of a scientific experiment. Its archival value and its ability to provide proof for patent applications echo those of a physical notebook, but the ELN goes much further, creating an "upgraded" and "amplified" notebook comparable to a complete laboratory knowledge management tool. The most significant differences are found in its collaborative capabilities (online collaboration), its improved traceability (use of standard protocols or data models) and time savings (using a search engine to look up experiments). In addition to combining operating procedures, the electronic notebook provides an integrated approach to scientific work. It establishes continuity between an experiment and its raw data, without being designed to store the latter. It facilitates data processing thanks to the possibility of interfacing with specialist software, laboratory instruments and databases. Finally, when the ELN is paired with an LIMS-style solution (see section 2.2.3 "related tools" below), its scientific role is complemented by logistical features likely to improve the organisation of the laboratory in terms of day-to-day stock management (substances, samples, animals, etc.) or equipment management (booking, maintenance).

2.2 Scope

2.2.1 Persons involved

An ELN is used by the following people:

- **The primary research participants,** for whom the ELN is primarily intended. These are the main users of the ELN, as far as tracking and logging their work is concerned;
- **The team leader,** acting as a specialist or administrator. The term principal investigator (PI) may also be used;
- **The signatory**, who will ultimately be responsible for validating the research recorded in the ELN ;
- **The Legal and Promotion departments**, who will be able to rely on the ELN to establish or confirm the date and authorship of the research work recorded in the ELN, particularly for patent filing processes;
- **The different partners** of the principal research participants (communities, industrial partners, hospitals, etc.), who will be able to use the ELN as a collaborative and information sharing tool;
- **Archivists**, who will be responsible for ensuring the long-term preservation of the information stored in the ELN.

2.2.2 Uses

The uses of an ELN differ enormously according to scientific discipline, but also—and especially—according to research practices. The names field notebook, handling book, logbook, archealogy field notebook are therefore preferred to "lab notebook", manifesting the various forms (digital in particular) and purposes according to the requirements of the respective communities. Outside the "laboratory", field surveys involve the use of a number of materials and tools whose functions differ depending on research methods. These can serve as:

- an aid for gathering a variety of field observations or analytical notes throughout "qualitative" research, whether individually or collectively as part of ethnographic, sociological, or anthropological approaches
- a tool for the systematic collection and organisation of data in archaeology and environmental sciences, resulting in databases and information systems

Whether it be note-taking using a text editor or the use of geographic information systems, the nature of the digital tools employed varies according to practice.

In the field of humanities and social sciences (sociology, anthropology, etc.) or life sciences (ecology, etc.), software tools have begun to replace paper in the field for taking observational or analytical notes, etc. While not having the same needs as experimental research, these disciplines can benefit from the use of ELNs as regards data security, indexing, integration with other digital tools, etc. A well-documented and non-exhaustive overview of these topics is provided in the appendices.

In terms of research practices, across all scientific fields, the working group identified three principal approaches that characterise the various uses of ELNs:

- Predominantly **experimental** research. The research is primarily based on controlled and deliberate testing, whether under laboratory or real-world conditions.
- Predominantly **observational** research, where the main focus of the research is observing objects, phenomena and behaviours outside of a controlled setting.
- Predominantly **analytical** research, which mainly focuses on computer analysis and processing of data, whether of an observational or experimental origin.

Note that most scientific disciplines have areas and topics of research relating to these three approaches and that there is no strict boundary between them.

2.2.3 Related tools

With advances in digital technology, the various tools available are seen less and less as separate objects: at different scales, digital work environments are being developed that emphasise the interoperability of different software. The lab notebook in a digital environment is therefore far from being isolated like a physical notebook, and it works with various tools and research infrastructures, allowing for fluid and secure data management.

The following tools are among those most likely to interact with an ELN:

- The **LIMS** = (***Lab Information Management System*****),** These are integrated management software packages that manage, for example, field data entry and sample traceability, users, instruments (mainly robots), stocks and supplies, and product and equipment tracking. For research units such as laboratories, working with a primarily experimental approach, it is particularly useful to have features dedicated to the management of this information. It should be noted that many ELN solutions integrate LIMS-like features.
- ***Notebooks***. These are interactive programming interfaces that make it possible to combine natural language text content with sections in computer language. These tools are often used by research communities working in a predominantly analytical manner.
- **Software forges**. These are systems for the collaborative management and maintenance of code, text and documents often linked to an IT project or software product. These tools are mainly used in analytical research.
- **Discipline-specific computer systems**, such as GIS (Geographic Information Systems), used in archaeology.
- **Data repositories**, allowing the link between the research process as recorded in the ELN and the final and potentially published data resulting from this work.

Linking the recorded progress of a project's research work with the data handled, the equipment used, or the code used, is often necessary to meet traceability and reproducibility requirements. Depending on the context, interoperability with these different tools may be necessary for the correct use of an ELN.

3 Recommendations for interoperability between different tools

3.1 Interoperability

Interoperability, as widely understood[3], is a product or system's ability to work with other existing or future products or systems without restrictions on access or implementation. Interoperability concerns only the external behaviour of products and does not influence their internal design. This capability is ensured by known standards and formats. In France, these standards can be found in the general interoperability framework (RGI – référentiel général de l'interopérabilité)[4] the current version of which was approved by the Ministerial Order of 20th April 2016[5]. The RGI is a document that outlines a set of standards and good practices used by French administrative bodies in the field of information technology.

The challenges of interoperability lie in the formats of the data being exchanged between different products as well as in the communication interfaces and protocols.

A research laboratory generally has a rich digital ecosystem, often referred to as an information system (IS). It can be composed of data management infrastructures (databases, document management, repository, etc.), computing infrastructures (HPC, *Notebook*, processing chains, etc.), virtual research environments (VRE) or simply office tools. Research staff also have access to resources outside of the laboratory such as public data repositories, catalogues and storage and/or computing space. In order for the adoption of an ELN to add value, it needs to seamlessly interface with existing IS components. The ELN's capacities in terms of interoperability are key to its integration into the IS.

3.2 Recommendations

3.2.1 Joint supervision

When each supervisor already recommends or uses an ELN, it is important that each ELN be able to export its content in a standard, open and documented format

This will allow different scientists to distribute and share content, and thus decompartmentalise research work.

As there is currently no standard format for ELN content, the accessibility and usability of these export formats is paramount in order to create bridges between existing ELNs.

[3] https://en.wikipedia.org/wiki/Interoperability

[4] https://www.numerique.gouv.fr/publications/interoperabilite/

[5] https://www.legifrance.gouv.fr/jorf/id/JORFTEXT000032438896

It is essential that the ELN can at least assign a Persistent IDentifier (PID) and a timestamp to each of its inputs. Ideally, the ELN should be able to address these identifiers directly through repositories (DOI, ORCID, etc.)

3.2.2 Integration

One should Prioritise ELNs that integrate seamlessly with tools already in use in the institution.

This integration can be done natively without the need for programming, but can also be achieved through the use of programming interfaces (APIs). In this case, favour an ELN with well-documented APIs.

The ELN must be able to interface with existing systems such as identity management, EDMs, VREs, repositories (Dataverse, software forge, etc.) or computational environments (*notebooks* such as Jupyter and Scilab, R-studio, etc).

3.2.3 Access to data

Consider the ability of the ELN to access and read data stored outside the system - especially large datasets that cannot be imported into it.

It is important that the ELN can access data and metadata stored in repositories (Dataverse, CKAN, etc.) to enable reuse of research data.

3.2.4 Backup and archiving

If the ELN includes a backup mechanism, it must be configurable and operable without limitation for the user institution's managers. Backed-up items must be accessible and restorable at will.

Otherwise, the ELN’s structure must make it possible. Backup of ELN data must be possible externally through images of the ELN file system tree and the ELN’s databases. These images must be restorable, at least, on the same version of the ELN and, ideally, with backward and forward compatibility.

In any case, it is in the interest of the user institution to be able to access a history of the contents of the ELN it uses. In the event of corruption or loss of data, it must be able to cope through a restoration of this history.

The ELN must be able to export all or part of its entries in an open format in order to establish a persistent archive. These exports must include all necessary metadata, associated media and references so that they can be used independently of the ELN.

3.3 Summary

Ensure that the ELN can store and handle data in open input AND output formats.

Ensure that the ELN provides a well-documented API to facilitate its integration into the IT ecosystem if it does not do so natively.

Ensure that the content of the ELN can be backed up and archived.

4 Recommendations for the tool selection procedure

4.1 Needs

4.1.1 Determine the wishes of stakeholders

Since the objective is to meet the needs of stakeholders who want an electronic solution suited to specific research practices, one of the first phases is to broadly identify functional requirements. This process involves a study to gather user needs and define the technical specifications derived from the contributions of the different stakeholders (management, scientists, legal, information systems, security, etc.) In order to ensure the traceability and scientific integrity of research results, careful consideration must be given to providing all the security, confidentiality, and traceability necessary to protect the results and scientific heritage of the institution and of its research partners. Also, collaboration between partner institutions and most often unit co-supervisors is essential in order to issue shared recommendations facilitating access to the tools for the primary target, i.e. the users.

The managing body must provide guidelines for the expected objective(s), the scope of the study or need, and the overall organisational structure necessary to respond to the achievement of this or these objectives according to indicators which will be set in advance. These will define the arbitration and validation steps allowing for consistent progress of the study with regard to preliminary results and allowing for a certain agility in adjusting milestones over time if necessary.

4.1.2 Take into account the research discipline and its practices

Usually used for bench or field experiments, the electronic version of the lab notebook has wider applicability, as shown in the "uses" section (2.2.2) of this report. Most tools have a shared core set of features (creation of experiments, advanced search, experiment models, etc.), supplemented by specialised functionalities that are sometimes very specific to a given discipline.

Selecting a tool can thus be difficult due to specialised requirements in the different fields and sub-fields of research. Within a single discipline,

expectations can be very different: a synthetic chemist will not use the ELN in the same way as a chemist who does analytical work with code. However, these two profiles may very well be found within the same laboratory. Consequently, the choice should hinge on research practices rather than the discipline itself.

Staying with the example of chemistry and more specifically synthetic chemistry, the added value of an ELN lies in features such as a stoichiometric table performing mass and quantity calculations automatically, an efficient search by chemical structure and substructure, a molecule editor and a solution for managing inventories of reagents and products used during tests. It is this comprehensive set of discipline-specific tools that can make the ELN an asset in a researcher's daily work.

The inventory management mentioned above is a functionality that relates more to LIMS-style (*Laboratory Information Management System*) tools. However, most electronic lab notebook solutions currently available include stock management features, with varying degrees of sophistication. In situations where the laboratory is required to manage products, equipment or apparatus on a daily basis, as may be the case in chemical or biology laboratories, the choice of ELN may be partly determined by this feature.

For researchers from analytical and software development fields, tools that strictly correspond to the definition of an electronic notebook are not very suitable: advanced calculation, editing and execution of code, version control, etc.; these are indispensable features that do not exist at time of writing–. Moreover, many very powerful tools performing most of these functions exist and are already in use by various communities. Examples include Git for version control, Gitlab for collaborative development, Emacs for code editing and Jupyter Notebook for code formatting and execution. The ELN cannot replace these tools, but can complement them. Some ELN publishers provide an API that allows the tool to be interconnected with other software. It may be of interest to analytical researchers to see if they are able to “piggyback” their tools onto the ELN using the API.

Therefore, the importance to be placed on a particular feature will vary according to the discipline and the research methods favoured by the target audience.

4.1.3 Consider the level of security required

One of the most important things to consider when choosing is the expected level of security for the storage, processing and retrieval of data. Depending on the technical, scientific, and human circumstances specific to each study, it will be necessary to establish the extent of the security requirements. To this end, it is recommended to use a risk analysis process such as that from the French National Agency for the Security of Information Systems (ANSSI – Agence Nationale de la Sécurité des Systèmes d'Information), also known as the EBIOS[6] method, while referring to standardised criteria such as those found in the ISO/CEI 27001[7] standard. Whether or not a laboratory has restricted access policies should also be taken into account when choosing. The scope of an ELN's functions tends to imply high security requirements, particularly according to the four commonly applied assessment criteria:

- The confidentiality of processed data must be ensured through an authorisation system : only authorised persons should be able to access the information intended for them (through access rights or permissions) and any unwanted access must be prevented.

[6] https://www.ssi.gouv.fr/administration/management-du-risque/la-methode-ebios-risk-manager/

[7] https://www.iso.org/isoiec-27001-information-security.html

- **Authenticity** must be ensured via an **authentication** system: users must prove their identity through the use of an access code. This allows the management of access rights to the resources in question, and maintains a level of confidence when sharing.
- **Integrity**: the data must be as expected, and must not be altered in an accidental, illicit or malicious way. This means that the data in question must be accurate and complete. Usually *checksums* or hashes are used for this purpose.
- **Availability**: the ability to access information system resources must be maintained continuously and without interruption during planned periods of use. Services and resources should be accessible quickly and regularly.

Security requirements may also be increased according to the nature of the data being processed, for example data subject to professional secrecy, industrial and commercial secrets, personal data and data requiring protection to safeguard scientific potential. The country and conditions in which the data in the notebook is hosted will require careful scrutiny, as will the conditions for its storage and access by third parties.

In addition to these four main criteria, there are other aspects to be considered in the selection of an ELN, including:

- **Traceability** (or "proof"): a guarantee that access and attempts to access the data in question are traceable and that these traces are stored and usable.
    - **Non-repudiation** and **imputation**: no user should be able to disavow tasks that they carried out as part of the activities they were authorised to perform, and no third party should be able to attribute the tasks of another user to themselves.

4.1.4 Consider the extent of the tool's deployment

A distinction must be made between the level of deployment of the ELN and that of its hosting:

- An ELN can be created for each contributor to a research project or, more commonly, for each research team; it is less common for a multi-team research unit or even an entire research centre to have a single ELN.
- ELNs can be set up on a researcher's workstation, in the building hosting the research unit, in a regional or national data centre, or even in the cloud.

Having the ELN set up on an individual workstation can provide its owner with autonomy when it comes to backing up data and ensuring data quality (e.g. timestamping activities). However, this makes it difficult to share data.-While one of the advantages of an ELN over its physical counterpart is the ability to share information within a research team. An advantage of having the ELN on an individual workstation is the ability to work offline or during network outages. Setting up ELNs on a central server makes sharing easier but also makes them dependent on the quality of networking between the workstations and the server. As various combinations of the above are possible, they will need to be assessed with the following in mind:

- Sharing information

- Data security
- Potential interconnectivity with other digital tools
- Efficiency, reliability

4.1.5 Identify the systems that will interact with the ELN

Depending on the context, methods, and fields of the research in question, the chosen ELN should be able to interact with different third party systems. This will be the case for the related tools described in section 2.2.3. Recommendations covering the interoperability of the chosen solution, as described in section 3, would therefore need to be closely followed.

4.2 Establish the selection criteria

4.2.1 Technical and functional characteristics

The process of choosing an ELN starts with drawing up a list of desired features and any essential requirements and technical specifications. The following list of features was drawn up by the working group:

Table 1: shortlisted features

Feature
Description of experiments and results
Ability to duplicate experiments (without using research output), such as a model/template for a new experiment
Description of protocols
Ability to duplicate protocols as a model/template for a new protocol
Attachment management: office (PDF, audio, video, etc.) or scientific and technical (git, chromatograms, spectra, images, etc.)
Ability to export materials from the ELN to a printable format
Access to external data (i.e. not stored in the ELN), raw or not, for reading and for reference
Data structuring with internal search engine and filter; use of keywords (tags)
Accessible content (for mining)
Possibility to associate metadata with the data sets being handled, whether they are generic (dates, authors, etc.) or specific (ontologies, taxonomies, etc.)
Use and maintenance of equipment (including notifications)
Stock management (reagents, cell lines, etc.)
Tools for simple calculations (basic mathematical functions, unit conversion, etc.)
Tools for complex calculations (statistics or analyses requiring a programming language)
Molecule building/design tool
Notes on experiments or protocol (during experimentation), reports, etc.
Ability to cite content from bibliographies (such as labtex.bib Text)
Access to workstation office tools
Management of software tools and codes (continuous integration, continuous deployment, version control, etc.)
Management of sharing by users
Shared agenda
Collaborative document editing (online and simultaneously by several users)
Annotation of content and comment insertion

As for technical specifications, the group came up with the following list:

Table 2: shortlisted technical specifications

Category	Specification
Data protection	Compliance with the GDPR for personal user data
	Compliance with the GDPR for research data: personal data of research participants if stored in the ELN
	Have a managed, redundant and backed-up storage (remote sites)
Integrity	Compliance with regulatory or legislative requirements governing the hosting location of research data
	Enable document signing (hashes)
	Enable activity and document timestamping
	Enable unique input identification (protocol, experiment, note, etc.) – UUID
	Enable versioning/change tracking
	Ena ble the implementation of predefined workflows (e.g. a set of required actions in the description of an experiment or result)
Authentication, security	Enable electronic signature by participants
	Use a secure authentication system
	Enable the creation of user management rules: access permissions by role, project, and life cycle (arrival/departure of participant, project launch/completion, etc.)
Data backup	Encrypt data transport (TLS, RGS-compliant algorithms)
	Encrypt data
	Include regular backups and restores
	Define data retention periods and archiving or destruction terms
Confidentiality	Ensure the confidentiality of exported data for backup or archiving.
	Safeguard the confidentiality of information that may represent valuable know-how or confidential business information (safeguard the innovation in the event of a patentable invention).
	Safeguard the confidentiality of information exchanged with a partner in the context of a collaboration

Compatibility, interoperability	Compatibility with different operating systems (Windows, Linux, etc.), for servers
	Allow data to be managed (*Import/Export*) in a structured format (e.g. in case of a software or administrative modification)
	Allow integration with a local information system (databases, directories, etc.)
	Enable links with data repositories
	When available, use open standards for input or output data to allow information to be permanently re-used.
	Enable data sharing through data integration and exposition services, via *WebServices* or REST APIs if possible
Use, usage	Ability to easily install updates and upgrades, while guaranteeing accessibility, if using a web browser
	Ergonomic and intuitive
	Enable modularity/customisation/flexibility
	Have the necessary documentation (guide, tutorials, instruction manuals, etc.)
	Access to source code in order to manage, if necessary, the development of the tool, its maintenance, and most importantly the data it contains
	Supported by a community of users
	Multilingual support (at least English in addition to the local language)
	Allow offline use with subsequent data synchronisation
	Be usable on the move (on tablets and smartphones)
	Take environmental footprint into account : electricity consumption of servers, etc.

Of course, the various features and technical specifications listed will not have the same weight in the final choice. They should mainly serve as a starting point for the creation of an assessment grid, which will be outlined below. Further details on some of these features and specifications can be found in the appendices.

4.2.2 Other criteria

Features and technical requirements are not the only criteria to consider. For a comprehensive assessment, it is also important to consider:

- The **overall cost of the solution**, whether financial costs or resource requirements, such as human resources. This cost is to be estimated for the acquisition, deployment, and long-term use of the chosen solution. In addition to the potential licence costs for a proprietary solution, the study must also include maintenance costs related to hosting, setting up the required technical environment, support, and change management.

- **Ease of adoption and available support.** This can mean the quality of support provided by the publisher, if any, the availability of documentation, but also the size and vitality of the user community.
- **Subjective suitability criteria specific to the target users** While difficult to quantify, it is nevertheless a major factor in successfully choosing a solution. For some communities this can mean the ability of the ELN in adapting to the typical user workflow, and this can lead to a more effortless interplay between the thought processes of the researcher and the actual input of information into the ELN. It is difficult to objectively assess this factor and it is often necessary to carry out a use case analysis with test users followed by feedback and opinion surveys.
-

4.3 Identify commercial options

The ELN market is booming. Between 2014 and 2020, revenue in this sector more than doubled in the US from $222 million to $576 million[8]. This dynamism was largely driven by the pharmaceutical industry's digital transformation due to its early recognition of the potential of ELNs as a driver for improving quality processes, innovation, and competitiveness. Some industry giants such as BMS, Astra Zeneca, Johnson & Johnson, GSK, Eli Lilly and Company, and Roche, therefore opted for electronic lab notebooks in the early 2000s.

This drive for transformation resulted in a boom in available solutions. In 2006, a US report counted under 30 suppliers of electronic lab notebooks[9]. In 2017, a study conducted by the University of Southampton counted 103 different solutions[10].

The ELN software market is so intricate as to complicate the selection of a solution, and this is reflected in the many current attempts at comparative analysis (i.e. *benchmarking*). Between 2018 and 2021, Harvard Medical School and the Boston University Medical Campus conducted a benchmarking study with 33 individual products to “help researchers (…) effectively identify ELN tools that meet the specific needs of their research.”[11] The resulting table is still available online [12]and was created by sending questionnaires to the relevant publishers. Other such efforts are also freely available, such as the more succinct survey by Cambridge University[13], which includes

[8] https://www.grandviewresearch.com/industry-analysis/electronic-lab-notebook-eln-market

[9] https://journals.sagepub.com/doi/10.1016/j.jala.2009.01.002

[10] https://jcheminf.biomedcentral.com/articles/10.1186/s13321-017-0221-3

[11] https://zenodo.org/record/4723753#.Y9Ks562ZPIV

[12] The working group behind the document decided to stop updating it after April 2021: https://docs.google.com/spreadsheets/d/1ar8fgwagOh30E31EAPL-Gorwn_g6XNf81g3VDQnQ_I8/

[13] https://www.data.cam.ac.uk/data-management-guide/electronic-research-notebooks/electronic-research-notebook-products

short case studies focusing on four products (Onenote, Hivebench, Benchling, and LabArchives)[14]. The latter software, which is used by Monash University in Melbourne, Australia, is extensively featured on the university library's website[15]. Similarly, the university libraries of Lyon 1 and Grenoble Alpes have drawn up a comprehensive overview of eight electronic lab notebook solutions, some of which excel at providing features specifically suited to chemistry[16].

An initial assessment of the different products available can involve separating ELNs according to usage, business model, and host type:

Usage	General-purpose
	Discipline-specific
	Integrated LIMS features
Business model	Open Source
	Proprietary
Host type	Cloud provided by publisher only
	In-house hosting only
	Cloud or in-house hosting

This basic grid enables products to be filtered out based on the features of currently available products and the priorities of individual establishments. The template provided below includes four examples of ELN that are representative of the variety of products on offer:

Criterion / Solution		**Chemotion**	**elabFTW**	**Mbook**	**Labstep**
Usage	General-purpose		X		X
	Discipline-specific	X		X	
	Integrated LIMS features			X	X
Business model	Open Source	X	X		
	Proprietary			X	X
Business Host type	Cloud provided by publisher only				X
	In-house only	X			

[14] https://www.data.cam.ac.uk/data-management-guide/electronic-research-notebooks/electronic-notebook-case-studies

[15] https://www.monash.edu/library/researchers/data-collection-management/eln

[16] https://www.datacc.org/bonnes-pratiques/utiliser-un-cahier-de-laboratoire-numerique/

	Cloud or in-house hosting		X	X	

Given the diversity of available solutions, it is not possible to identify a general trend favouring one particular tool in recent years in the field of research and higher education. Some institutions have favoured *open source* solutions (elabFTW at Inria, the Curie Institute, the Institut Jacques Monod, the University of Düsseldorf, Graz University of Technology, etc.), while others have gone for proprietary solutions (Labguru at Inserm, eLabJournal at the Pasteur Institute and the University of Geneva [where they also use Rspace][17], etc.). Among the software with purposefully discipline-specific designs, we see products primarily catering to the needs of biologists and chemists.

The ELN market is also experiencing an incursion by large publishers, some of whom are buying up products initially developed by start-ups. Examples include the acquisition of Hivebench by Elsevier in 2016 and Digital Science's purchase of shares in Biodata (publisher of Labguru), whose parent company is Holtzbrinck. ELNs are also part of the business strategies of global groups such as Dassault (Biovia), PerkinElmer (Signals), and Agilent (Slims).

These solutions stand in contrast to strategies pursued by the academic sector, which have developed in recent years: elabFTW (Curie Institute), Chemotion (Karlsruhe Institute of Technology) and OpenBIS (University of Zurich).

4.4 Assess options based on stated criteria

The comparison of tools in technical terms (i.e. features and specifications) should ideally rely on a list of objective criteria that allows solutions to be classified according to the weight given to a particular criterion in the context of the study.

This requires a two-tiered analysis: firstly with regard to the desired scale and granularity of criteria (scoring), and secondly with regard to the relative importance of criteria (weighting).

4.4.1 Scoring

The minimum degree to which a tool can be rated according to the established criteria is binary: the assessed tool either has a feature or does not; it either meets a requirement or doesn't. Solutions can therefore be assessed by feature as part of an objective assessment that does not imply any kind of value judgement.

However, depending on the study's context, different scoring scales can be employed according to the profile of the assessors or the level of detail required. The template grid provided by the working group presents a scoring system with four tiers:

[17] https://www.unige.ch/researchdata/fr/collecter-organiser/eln/

- **0** = does not meet requirement or information unavailable
- **1** = partially meets requirement
- **2** = meets requirement
- **3** = meets requirement with merit, or has specific qualities

If the study involves numerous assessors, a continuous scale (e.g. from 0 to 10) could be used to get an average of the scores given by the different assessors for each feature.

In any case, the type of scoring scale should be explained to assessors in order to avoid confusion and misuse.

4.4.2 Weighting

While scoring allows different tools to be compared based on a specific criterion, when making the final choice a distinction must be made between essential criteria and those of a more auxiliary nature. For example, it makes sense to place more emphasis on the “allows description of experiments” feature (without which an ELN would not be an ELN) than on the ability to export text in a printable format.

Consequently, to get a rounded impression of a solution's quality these individual scores should be weighted in a way that treats the results of the assessment as a whole.

Two weighting scales have been put forward by the working group. The first has four tiers:

- **1** = useless
- **2** = secondary
- **3** = important
- **4** = essential

The second is a non-linear scale, put forward in order to allow for a more detailed analysis. It works by summing the weighting values to 100 and then dividing the values according to their corresponding importance. This scale is used in the grid shown in the following paragraph.

4.4.3 Presentation of results

Proper use of the grid should allow for the results of the technical comparative study to be displayed. Depending on needs and context, the results can be displayed in several ways:

- A binary “does/doesn't” table: this format can prove effective and useful as a first step in filtering a large list of tools (see section 5)
- An overall score for each tool, calculated by multiplying the score for each criterion by its weighting value and then summing the results.
- A graphical representation of each tool’s strengths and weaknesses. This could take the form of a radar chart.

Figures 1 and 2 show examples of a scoring grid and are based on some of the tools discussed in section 5.

A radar chart can be created by comparing the subtotals for all features and the subtotals for each category of specifications with the maximum score of 2 per criterion (3 meaning that based on stated requirements, the assessed solution surpasses 100% suitability), and then scaling them to 5. The eight values obtained are represented on the radar chart in a way that can be particularly useful for comparing different solutions. However, it is worth noting the bias inherent to the fact that the relative importance of the different groups of specifications has not been taken into account. Examples of a radar chart and the table of values used to generate one are shown in Figures 3 and 4.

4.4.4 Comments:

- In the example given above (Figure 1), we can see that the features of the tool being assessed are potentially better suited to an experimental approach, while its technical specifications are not significantly different.
- We can see that there are a few items with a score of 0 (Figure 2). This is actually because the ELN in question is only available as a hosted solution in the United States (i.e. as SaaS - Software as a Service), since its publisher is American. Although the terms of use state that the publisher is compliant with the GDPR with respect to users' personal data, they make no mention of the protection of research data stored in the ELN. It is therefore reasonable to assume that personal research data is not protected. However, neither the documentation nor the publisher's website indicate whether content is encrypted or not. This information is therefore missing.
- As for the items with a score of 1: with the “use and maintenance of equipment” feature in mind, this ELN does indeed allow access to the laboratory's equipment inventory and also allows us to see whether equipment is free or in use, but we are not able to see whether a piece of equipment is overdue for inspection.
- For items with a score of 3 (meaning this specific feature or specification of the ELN exceeds expectations): if we look at the “Molecule building/design tool” feature as an example, we see that this ELN exceeds the typical functionality of such features as it also enables the verification/validation of atomic bonds, molar mass calculation, etc.

Figure 1 – Example of feature scoring

Feature	Score assigned	Exp. Weighted score	Analyt. Weighted score	Obs. Weighted score	Weighting per approach		
					experimental	analytical	observational
Description of experiments and results	3	30	30	30	*10*	*10*	*10*
Ability to duplicate experiments (without the research output) as a model/template for a new experiment	3	15	21	12	*5*	*7*	*4*
Description of protocols	1	10	7	10	*10*	*7*	*10*
Ability to duplicate protocols as a model/template for a new protocol	2	10	14	10	*5*	*7*	*5*
Attachment management: office (PDF, audio, video, etc.) or scientific and technical (git, chromatograms, spectra, images, etc.)	3	21	15	21	*7*	*5*	*7*
Ability to export materials from the ELN to a printable format	2	4	4	4	*2*	*2*	*2*
Access to external data (i.e. not stored in the ELN), raw or not, for reading and for reference	2	20	20	20	*10*	*10*	*10*
Data structuring with internal search engine and filter; use of keywords (tags)	2	14	12	16	*7*	*6*	*8*
Accessible content (for mining)	0	0	0	0	*2*	*2*	*3*
Ability to associate metadata with processed datasets, whether generic (dates, authors, etc.) or specific (ontologies, taxonomies, etc.)	2	16	12	16	*8*	*6*	*8*
Use and maintenance of equipment (including notifications)	1	3	1	1	*3*	*1*	*1*
Stock management (reagents, cell lines, etc.)	2	6	0	4	*3*	*0*	*2*
Tools for simple calculations (basic mathematical functions, unit conversion, etc.)	3	9	9	9	*3*	*3*	*3*
Tools for complex calculations (statistics or analyses requiring a programming language)	0	0	0	0	*0*	*10*	*5*
Molecule building/design tool	3	12	3	0	*4*	*1*	*0*
Notes on experiments or protocol (during experimentation), reports, etc.	1	2	1	3	*2*	*1*	*3*
Ability to cite content from bibliographies (such as labtex.bib Text)	1	3	2	4	*3*	*2*	*4*
Access to workstation office tools	0	0	0	0	*3*	*2*	*4*
Management of software tools and code (continuous integration, continuous deployment, version control, etc.)	0	0	0	0	*2*	*10*	*0*
Management of sharing by users	2	10	10	10	*5*	*5*	*5*
Shared agenda	0	0	0	0	*2*	*1*	*2*
Collaborative document editing (online and simultaneously by several users)	0	0	0	0	*2*	*1*	*2*
Annotation of content and comment insertion	1	2	1	2	*2*	*1*	*2*
Features subtotal	**34**	**187**	**162**	**172**	**100**	**100**	**100**

Figure 2 – Example of technical specifications scoring

Technical specifications						Weighting per approach		
Category	Specification	Score assigned	Exp. Weighted score	Analyt. Weighted score	Obs. Weighted score	experimental	analytical	observational
Data protection	Compliance with the GDPR for personal user data	2	8	8	8	4	4	4
	Compliance with the GDPR for research data: personal data of research participants if stored in the ELN	0	0	0	0	4	4	4
	Have managed, redundant, and backed-up storage (remote sites)	2	8	8	8	4	4	4
Integrity	Compliance with regulatory or legislative requirements governing the hosting location of research data	2	8	8	8	4	4	4
	Enable document signing (hashing)	0	0	0	0	4	4	4
	Enable timestamping of activities and timestamping of documents	1	4	4	4	4	4	4
	Enable unique input identification (protocol, experiment, score, etc.) – UUID	0	0	0	0	4	4	4
	Ensure versioning/monitoring of changes	1	3	3	3	3	3	3
	Enable the implementation of predefined workflows (e.g. a set of required actions in the description of an experiment or result)	1	3	3	3	3	3	3
Authentication/security	Enable electronic signature by participants	1	4	4	4	4	4	4
	Use a secure authentication system	2	6	6	6	3	3	3
	Enable the creation of user management rules: access permissions by role, project, and life cycle (participant arrival/departure, project launch/completion, etc.)	1	4	4	4	4	4	4
Data backup	Encrypt data transport (TLS, RGS-compliant algorithms)	2	8	8	8	4	4	4
	Encrypt data	0	0	0	0	3	3	3
	Include regular backups and restores	2	8	8	8	4	4	4
	Define data retention periods and conditions for archiving or destruction	2	4	4	4	2	2	2
Confidentiality	Ensure the confidentiality of exported data for backup or archiving.	0	0	0	0	4	4	4
	Safeguard the confidentiality of information that may represent valuable know-how or confidential business information (safeguard innovation in the event of a patentable invention).	0	0	0	0	4	4	4
	Safeguard the confidentiality of information exchanged with a partner in the context of a collaboration	1	2	2	2	2	2	2
Compatibility/interop	Compatibility with different operating systems (Windows, Linux, etc.), for servers	1	3	3	3	3	3	3
	Allow data to be managed (Import/Export) in a structured format (e.g. in case of software or administrative modifications)	1	4	4	4	4	4	4
	Allow integration with a local information system (databases, directories, etc.)	2	8	8	8	4	4	4
	Enable links with data repositories	0	0	0	0	3	3	3
	When available, use open standards for input or output data to allow information to be permanently re-used.	1	4	4	4	4	4	4
	Enable data sharing through data integration and exposition services, via WebServices or REST APIs if possible	1	4	4	4	4	4	4
Use/Usage	Ability to easily install updates and upgrades, while guaranteeing accessibility, if used from a web browser	2	6	6	6	3	3	3
	Ergonomic and intuitive	2	8	8	8	4	4	4
	Enable modularity/customisation/flexibility	1	2	2	2	2	2	2
	Have the necessary documentation (guide, tutorials, instruction manuals, etc.)	3	12	12	12	4	4	4
	Access to source code in order to manage, if necessary, the development of the tool, its maintenance, and most importantly, the data it contains	0	0	0	0	3	3	3
	Supported by a user community	1	3	3	3	3	3	3
	Multilingual support (at least English in addition to the local language)	1	4	4	4	4	4	4
	Allow offline use with subsequent data synchronisation	0	0	0	0	4	4	4
	Be usable on the move (on tablets and smartphones)	2	8	6	8	4	3	4
	Take environmental footprint into account: consumption of electricity by servers, etc.	0	0	0	0	2	2	2
	Specifications subtotal	**38**	**136**	**134**	**136**	**123**	**122**	**123**
		Score assigned	experimental	analytical	observational			
	Grand Total	72	323	296	308			
	Score out of 100		**72**	**67**	**69**			

Figure 3 – Example of calculations for graphical representation

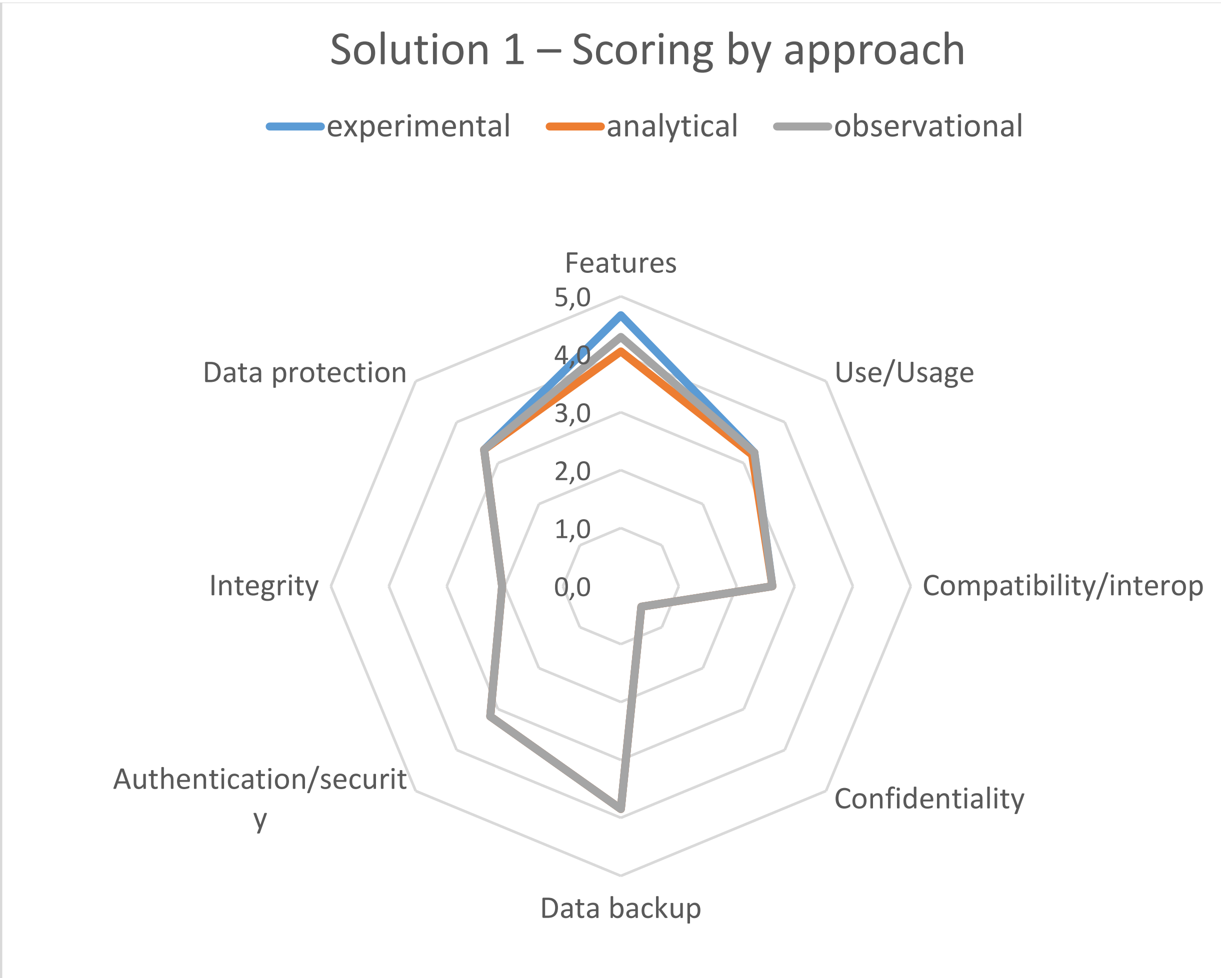
experimental approach			
Category	score per cat.	max score	score out of 5
Features	187	200	4,7
Data protection	2	24	0,4
Integrity	1	44	0,1
Authentication/security	1	22	0,2
Data backup	2	26	0,4
Confidentiality	1	20	0,3
Compatibility/interop	1	44	0,1
Use/Usage	0	66	0,0

analytical approach			
Category	score per cat.	max score	score out of 5
Features	162	200	4,1
Data protection	16	24	3,3
Integrity	18	44	2,0
Authentication/security	14	22	3,2
Data backup	20	26	3,8
Confidentiality	2	20	0,5
Compatibility/interop	23	44	2,6
Use/Usage	41	64	3,2

observational approach			
Category	score per cat.	max score	score out of 5
Features	172	200	4,3
Data protection	16	24	3,3
Integrity	18	44	2,0
Authentication/security	14	22	3,2
Data backup	20	26	3,8
Confidentiality	2	20	0,5
Compatibility/interop	23	44	2,6
Use/Usage	43	66	3,3

	Score out of 5 per approach		
Category	experimental	analytical	observational
Features	4,7	4,1	4,3
Data protection	0,4	3,3	3,3
Integrity	0,1	2,0	2,0
Authentication/security	0,2	3,2	3,2
Data backup	0,4	3,8	3,8
Confidentiality	0,3	0,5	0,5
Compatibility/interop	0,1	2,6	2,6
Use/Usage	0,0	3,2	3,3

Figure 4 – Example of graphical representation using a radar chart



4.4.5 Relative importance of different criteria

While an assessment grid makes it possible to thoroughly and objectively assess functional and technical criteria, it is important to bear in mind that it should not be used as the sole basis for ELN selection; it should be compared with other relevant criteria (cost, ease of adoption, etc.). The relative importance of these different factors can only be established with a clear understanding of the economic and policy context of the ELN implementation project.

5 Comparative list of tools

The working group compared a selection of available solutions based on the methods described in section 4.

5.1 Selection of tools studied

The tools selected are listed in the table below. The analyses were carried out on the software versions available in May-June 2021, unless stated otherwise:

Table 3: list of assessed tools

Name	Publisher	Version	Type
BIOVIA	Dassault Systèmes		Commercial
Chemotion	Karlsruhe Institute of Technology	0.5.0	Open Source
ElabFTW_dev	Deltablot	4.0	Open Source
ElabFTW_QeR	Deltablot	3.6.6	Open Source
Elog			Open Source
FindMolecule	FindMolecule, inc.		Commercial
Jogl	Just One Giant Lab MVAC		Open Source
Jupyter notebook			Open Source
LabArchives	Labarchives LLC		Commercial
LabBook	eNovalys		Commercial
LabCollector	Agilebio	6.032 with ELN add-on 4.13	Commercial
Labforward	LabForward gmbh		Commercial
LabGuru	Biodata		Commercial
Mbook	mestrelab	2.1.1	Commercial
OSF.io	Center for Open Science		non-profit or ganisation nisation
RSpace	ResearchSpace		Commercial

The tools that we studied were chosen with an even distribution of several criteria in mind:

- Tools under open licence and commercial licence
- Tools specialising in certain scientific fields (biology, chemistry, etc.) and general-purpose tools
- Tools only available online (Software as a Service – SaaS) and tools that can be installed on-site (on-premise software).

We also chose to focus on tools that are already in use or deployed at our own institutions in order to ensure the highest assessment quality possible.

Additionally, we also included a few outsiders that were not originally ELNs, but ended up being used for this purpose.

Appendix 5 contains links to the publishers of the assessed solutions and a list of other tools that were considered but could not be analysed due to time limitations.

5.2 Results

An overview of the results is provided in Figures 5 and 6. We chose the simplest scoring method—a colour code:

- green = the solution meets the requirement with varying degrees of success
- red = the solution does not meet the requirement or the information is unavailable

Note: We did not have the means to fully assess eLog, so its technical specifications column is blank.

5.3 Comments

This type of table does not take into account the relative weight of the criteria or the requirements specific to each research approach. It can therefore only be used to eliminate solutions that do not meet a criterion or a set of criteria that would be considered essential. The working group was not expected to give out of context recommendations for one tool or another.

The list of solutions is therefore non exhaustive and scoring remains incomplete. This is partly because it is divorced from real-world conditions, and partly because assessments were conducted in a limited period of time by a diverse group of people.

Nevertheless, these assessments still serve to emphasise the importance of correctly establishing selection criteria and their weight in a thorough assessment process.

Figure 5 – Feature scores for the 16 assessed solutions

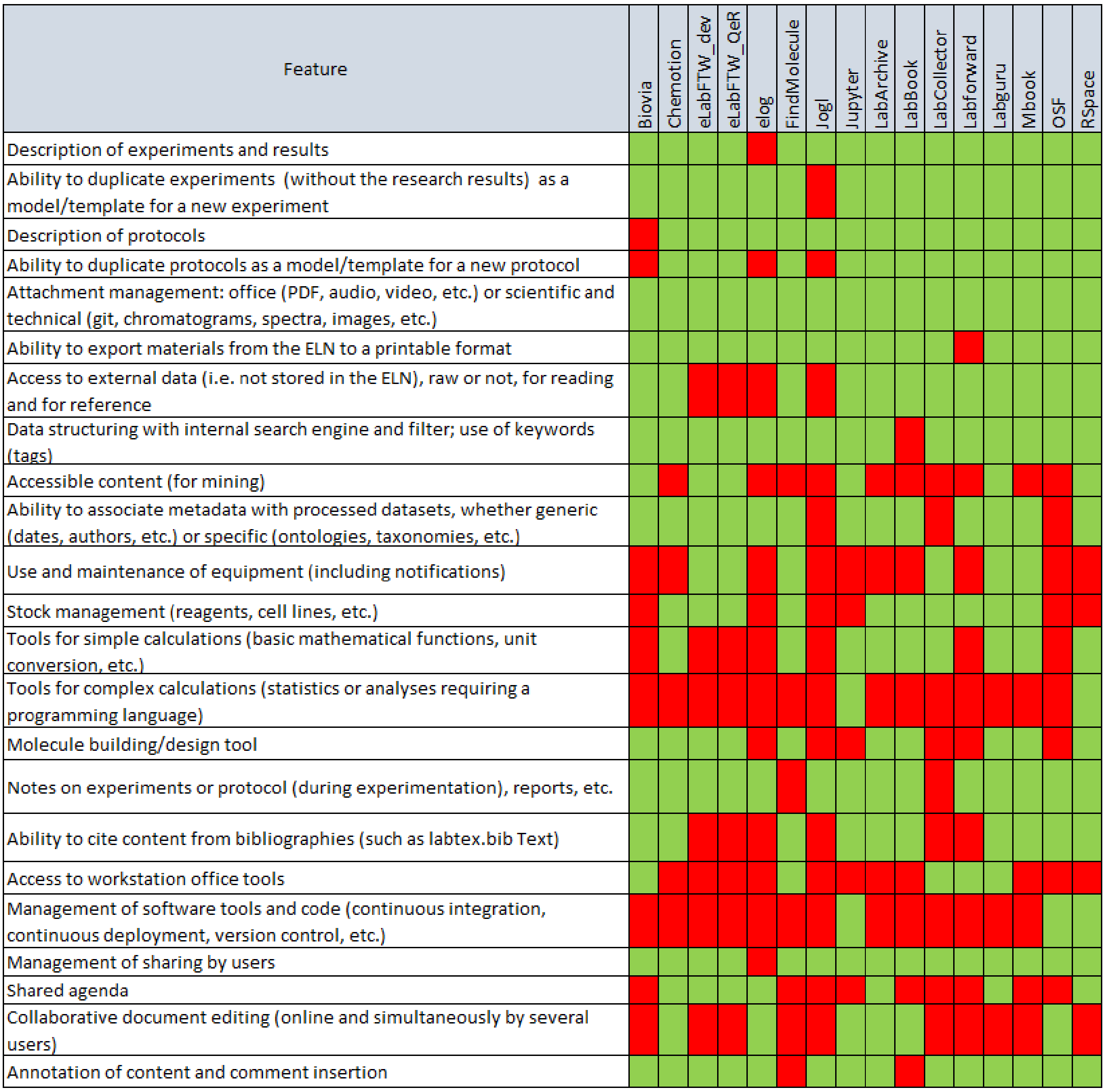


Figure 6 – Technical specifications scores for the 16 assessed solutions

6 Successfully conducting an ELN implementation project

6.1 Introduction

The aim of this chapter is not to dictate how a project should be executed in general, but rather to underline the particularities and challenges inherent to a project of this nature.

6.2 Project stakeholders and leadership

In addition to the stakeholders typically involved in a software implementation project (i.e. the purchasing department, legal department, and information systems department), it is also important to involve very early on the departments that stand to gain the most from the following benefits of an ELN's use:

- Quality
- Open science
- Promotion
- Scientific integrity
- Archiving
- Data security
- Data management

Stakeholders and project leaders need the input of people that represent future users. Support from the management team of either a research unit or a national research institution is a guarantee of the project's success.

It can be difficult to reconcile the duties of legal experts (who need to ensure GDPR compliance) with the principles of Open Science.

6.3 Pre-project

In the pre-project phase, it is necessary to:

- Find out about any recommendations from the relevant supervisory ministry(ies) or institutions: do they recommend or require a particular solution or selection method?
- Find out what the situation is at other institutions of potential co-supervisors: do they recommend or require a particular solution?
- Properly identify target users: e.g. based on discipline and also on research approach (experimental, analytical, or observational); see paragraph 2.2.2
- Consider the appropriate number of solutions: the simplest option would be to choose not one, but several solutions that

could better address the needs of different communities. This could also be achieved by choosing a simple solution that has additional modules.
- Consider whether the use of the selected tool(s) is mandatory or not.
- Check data security constraints (e.g. hosting).
- Think about how ELNs will be allocated: to a research project, a team, a sub-group, a unit, or to a particular device.

6.4 Pilot phase

ELN deployment is easier if the solution was chosen by a panel of users during a pilot phase. This can start with the use of a small number (2 to 4) of selected tools by a group of future users during a pre-study phase (see Chapter 4). This group should be as representative as possible of the range of future applications of the tool (team structure, disciplines, research approaches, location, etc.). This phase can last between a quarter and a year with periodic reviews and user satisfaction surveys.

6.5 Call for proposals

Institutions or laboratories can use the features and technical specifications grid in Chapter 4 as a guide for their requirements, so that they can analyse tools that may meet their needs and ensure that factors are not omitted. If a tool hosted by a third party is being sought, the requirements should stipulate that details regarding the location, hosting conditions, access conditions, etc. are to be provided.

It can be beneficial to request:

- that the response to the call for proposals include a planned development roadmap for the submitted solution. This will allow an assessment, for example, of whether the solution could be supplemented with components that could make it similar to a LIMS, or whether its insertion into the researcher's digital working environment is likely to be easy.
- that the tool's use can be maintained in read-only mode once the contract has ended.

6.6 Deployment

Deployment of the chosen solution is best done gradually so that the quality of the technical architecture and available support can be assessed prior to heavy usage.

6.7 Change of tools

One major consideration to keep in mind is that the limitations of public tenders can mean that it may be necessary to switch to a different tool once the conditions for renewal of the licence are exhausted[18] . This can be very laborious if data transfer cannot be automated and/or if there is a large number of notebooks. The integrity of all data must be guaranteed. Training all users to work with a new ELN can also have significant financial costs and require considerable effort.

In short, research would be slower down during the transition from one ELN to another

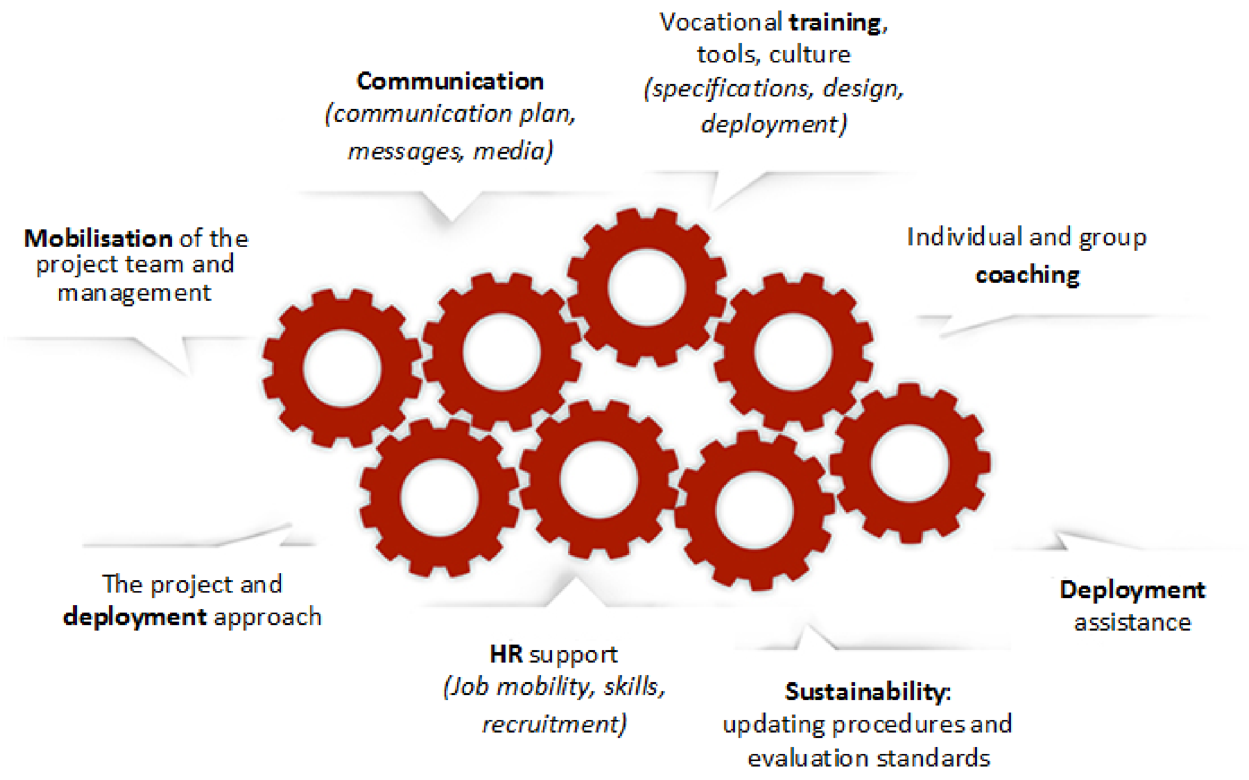
6.8 Change management

This is an important part of a project's success. For many researchers, the standard laboratory notebook is an everyday tool. This is why it is important for them to quickly master the electronic lab notebook and to be able to quickly find answers to their questions. Additionally, certain ELN features can lead to significant changes in the way work is organised: collaborative work (notebooks are usually shared by a team and are constantly accessible by the Principal Investigator (PI)), remote access (enabling telecommuting), etc. Therefore, it is necessary to provide appropriate support for both the end user and the notebook administrator (team leader, PI).

To start with, it seems sensible to (re)acquaint users with good practices regarding the use of a lab notebook, whether in physical or electronic form. One specific obstacle involves teams that have never used physical laboratory notebooks because the tool was considered to be totally unsuitable for their work, and they did not recognise its value for the protection and promotion of their results.

[18] This is the case when the purpose of the contract is to acquire an ELN solution and not to implement a given solution

Figure 7 – Factors for successful software deployment (taken from the book "Pro en Conduite du changement", Juliette Ricou and Valérie Moissonnier, 2018, Vuibert)



The primary fears regarding the use of such a tool are to do with aspects related to computing in general: ergonomics, difficulties in implementation (during field assignments in particular), costs (licences, equipment), human resources (technical skills, assistance, support), as well as a lack of confidence in the reliability of computing tools and the continuity of research results.

The change management process must be structured and planned, and involves two major project phases:

- assistance and troubleshooting during the setup phase and in the event of technical problems;
- assistance and user support during the usage phase if difficulties arise when using the tool.

In the first case, several options are possible depending on the implementation of theselected tool: support can be provided externally (by a dedicated service provider or by a solution provided as part of the tool's acquisition) or internally by local or national information systems departments, etc.

In the second case, support can be provided internally through training, education, and dialogue with the user community. The publisher of the selected tool may also provide support in the event of technical problems (chat, online help, hotline, etc.)

Change management should also be provided in various forms:

- learning initiatives, presentations of the study and choices made, individual and group training in the use of the tool,
- the creation of user communities,
- availability of documentation and video tutorials. In the context of Higher Education, Research, and Innovation, the creation of MOOCs (Massive Open Online Courses - remote learning that can accommodate a large number of participants) could be explored as a means of validating students' knowledge, for example.

Additionally, it is necessary to set up support services to address users' difficulties and troubleshooting needs during the installation phase or when using the tool.

6.9 HR profiles and scaling

The relatively rapid turnover of doctoral and post-doctoral students in laboratories means that support services are essential.

The institution must have one (or more) designated point(s) of contact whose responsibility is to handle questions regarding research data traceability and management (a specific department, committee, technical advisor, etc.), data security, etc.

A *Lab Data Manager* role (which covers more than just the ELN) can be recommended for laboratories in order to manage data, configure the ELN, and ensure that newcomers are familiar with the system.

6.10 Joint supervision

How does one deal with situations where a laboratory's co-supervisors recommend different solutions? In such circumstances a list should be made of the policies and strategies employed by the unit's co-supervisors, if they exist. The next step would be to check whether a co-supervisor recommends an electronic solution and whether its use would be worthwhile to the unit in terms of technical functionality and deployment (access, storage, and security conditions, etc.) If necessary, the solutions can be reviewed taking into account the unit's stated needs and expectations and in accordance with its rules (e.g. competition rules or procurement rules).

7 Conclusion

The fact that so many ELN options have recently been developed and that research organisations are keen to adopt such solutions clearly demonstrates the growing importance of ELNs in the development of digital tools for research teams, in data quality assurance, and in research practices. Given the proliferation of available options, the diversity of research approaches, and the need to ensure interoperability both with other tools and other ELNs, making a well-informed choice is of paramount importance.

With this in mind, the working group produced:

- Lists of the features and technical specifications expected of an ELN
- An example of how to weight these criteria according to the relevant approach: experimental, analytical, or observational.
- An example of how to score ELNs and graphically represent these scores to allow for comparison.
- A comparison of around fifteen proprietary and open source tools, depending on whether or not they met the required criteria.
- Recommendations on how to go about the implementation of an ELN after its selection.

These materials are freely available and reusable, and represent the main deliverables of the group's assignment, which ended in July 2021.

This report represents all of the work that was carried out, and it was delivered in its present state in September 2021.

A Appendices

A.1 Non-WG contributors

In addition to the members of the working group, its co-leads would like to thank the following people for their contribution:

- Jérémie Leonard, Data Processing Officer, University Lyon 1
- Nicolas Argento, Applications Team Leader, École polytechnique fédérale de Lausanne
- Philippe Rizand, ISD Director, Curie Institute
- Nicolas Carpi, designer of eLabFTW, UMR144, CNRS, and Curie Institute
- Nicolas Renard, UMR6521, CNRS
- Alain Rivet, UPR530 – Henri Valeins, UMR5536 – Yaël Hersant – UMR6296, members of the Qualité en Recherche network, CNRS
- Ambra Mari, team leader, CNRS and University Lyon 1
- Frédéric Leroux, director, UMR7042, CNRS
- Auriane Denis Méyère, Céline Charavay, Eric Faudry – ELN administration team, UMR 5075 – IRIG, CEA
- Fanny Brizzi, ELN project leader, DSI, Inserm

The WG co-leads would also like to thank the CNRS electronic lab notebook WG, which shared its results before they were made final.

A.2 WG mission statement

The mission statement sent to the project leaders, as well as the working group's roadmap, are included in full on the following 4 pages.

MINISTÈRE
DE L'ENSEIGNEMENT
SUPÉRIEUR,
DE LA RECHERCHE
ET DE L'INNOVATION
Liberté
Égalité
Fraternité

Direction générale
de la recherche et de l'innovation

Paris, le 17 novembre 2020

Affaire suivie par :
Claire LEYMONERIE
Tél : 01 55 55 80 51
Mél : claire.leymonerie@recherche.gouv.fr

1 rue Descartes
75231 Paris SP 05

Messieurs,

Les cahiers de laboratoire sont des outils scientifiques et juridiques qui permettent de garantir la traçabilité des expériences de laboratoire, de maintenir la continuité du cheminement intellectuel aboutissant aux résultats scientifiques, d'identifier les contributions de chacun et d'établir les dates d'acquisition des résultats. Ils contribuent à la reproductibilité et au caractère cumulatif des connaissances scientifiques et s'inscrivent donc pleinement dans une démarche de science ouverte.

Il importe donc d'éclairer les choix qui sont faits par les équipes scientifiques parmi les différentes solutions actuellement disponibles.

Nous vous confions à cet effet le pilotage d'un groupe de travail qui aura pour objectifs d'établir une vision partagée sur la définition, le périmètre fonctionnel et les usages du cahier de laboratoire électronique, de produire des recommandations sur les critères de choix d'un outil et son interopérabilité avec d'autres outils, dont notamment les entrepôts de données. Vous établirez ainsi une analyse comparative d'outils existants (sans viser à l'exhaustivité) et en vous appuyant sur les études déjà réalisées par certains établissements ou le réseau CURIE.

Vos travaux s'inscriront dans le cadre du Plan national pour la science ouverte et seront menés en lien avec la Direction générale de la recherche et de l'innovation du ministère de l'Enseignement supérieur, de la Recherche et de l'Innovation, et le collège des données de la recherche du Comité national pour la science ouverte. Vous présenterez les résultats des travaux menés devant le Secrétariat permanent de la science ouverte.

Gilles Mathieu
INSERM, Département Systèmes d'Information

Dominique Pigeon
INSERM, Département Systèmes d'information

Tovo Rabemanantsoa,
INRAE, Direction pour la science ouverte

Copies :
Véronique Stoll et Pierre-Yves Arnould, pilotes du collège données de la recherche du Comité pour la science ouverte
Anne Vanet, responsable du groupe de travail « réussir l'appropriation de la science ouverte »

Pièce jointe : Feuille de route du groupe de travail « Cahiers de laboratoire »

Vous veillerez à coordonner vos travaux avec ceux du groupe de travail « réussir l'appropriation de la science ouverte » piloté par Anne Vanet au sein du Comité pour la science ouverte et avec le projet « cahiers de laboratoire électronique » en cours au CNRS. Pour cela, votre groupe de travail comptera dans sa composition un membre nommé au sein du collège de données et du groupe de travail « réussir l'appropriation de la science ouverte ». Plus globalement, la composition du groupe devra autant que possible représenter la diversité disciplinaire et structurelle de l'ESR.

Vous trouverez joint à ce courrier une feuille de route précisant les objectifs et les modalités de fonctionnement du groupe de travail.

Nous vous remercions pour votre engagement dans cette mission et vous prions de recevoir, Messieurs, nos cordiales salutations.

Isabelle Blanc
P/O le bureau de la Science Ouverte

Committee for Open Science
“Laboratory Notebooks” Working
Group Roadmap

1. Background

The laboratory notebook is a logbook used to record the day-to-day activities of research projects, experiments, and employed protocols. For patent offices in particular, it provides proof of an invention, of its inventors, and therefore of its rightholders. The notebook simplifies quality processes, contributes to the reproducibility of data and research results, and meets legal and contractual requirements.

Due to its crucial place in research activities, the electronic lab notebook represents a strategic tool for the team and unit, as well for institutional supervisors. It is an essential tool for research projects committed to the open science movement. The choice of tool, its configuration, and above all how it is used are of capital importance

2. Objectives and deliverables

The primary objectives of the “electronic lab notebooks” working group will be to:

- **Establish a shared vision of the definition, functional scope, and purposes of the “electronic lab notebook”.** It will be particularly important to establish what such a tool does and does not do, what it is suitable for, and—according to the context—which tools can be used to supplement it.
- **Produce a set of recommendations on the criteria for choosing a tool.** The criteria will be established in an objective manner according to functional needs, disciplines, and fields of research, as well as institutional requirements. These criteria will also be used for comparative analyses of different tools and to measure their effectiveness in and/or suitability to different usage contexts.
- **Produce a set of recommendations for interoperability between different tools.** In the French academic multi-supervisor context, the use of different tools should not result in research work being compartmentalised and isolated. Laboratory notebooks must be able to seamlessly (meaning without the need for specific licences, reverse engineering, etc.) merge with existing computer environments and allow, in particular, for data movement (i.e. data exposition) to data repositories.
- Draw up a (non-exhaustive) comparative list of tools based on the recommended criteria.
  This list can serve as an example of how to conduct a selection study.

3. Methodology

3.1. Group composition

It is recommended that a group of between 8 and 12 people be assembled, with each person representing one or more of the following:

- **Fieldwork**: users or potential users of notebooks, preferably from different scientific communities

- Scientific and legal integrity, and intellectual property: lawyers, data specialists, quality managers, etc.
- **Technical and infrastructural expertise**: service operators who are experienced in efficiency, backups, security, etc.
- **Technical and data expertise**: data managers, platform managers who are experienced in formatting, interoperability, archiving, etc.
- **Strategic and content promotion expertise**: institutional or scientific representatives, content promotion managers, etc.

The group's composition should represent the disciplinary and structural diversity of the higher education and research community as best as possible. Distribution by establishment of origin and by scientific discipline will therefore serve as selection criteria for the group's final composition. The group should also include a member of the Data College.

3.2. How it works

The group will carry out its work between October 2020 and June 2021.
Due to the dispersed nature of the group, the majority of meetings and workshops are expected to be held by videoconference, according to a schedule that will be jointly determined once the group is established.
In addition to these regular meetings there will be quarterly progress and feedback meetings, in January 2021 and April 2021, with the MESRI open science team and the Data College. The group's findings and final deliverable will be presented at the end of their work in June 2021.

A.3 Detailed look at uses – observational approach

A.3.1 Ethnographic approach and "qualitative" analysis (sociology, anthropology, design, participatory research, etc.)

In ethnographic research (e.g. in anthropology or so-called "qualitative" sociology), the "logbook" or "field notebook" provides a framework for both recording the data collected and explaining the rationale behind decisions made during the research. These documents play an important role in structuring thought processes throughout research. Elucidating this data collection process and its subsequent analysis represents a benchmark in terms of quality and reliability in the research communities employing these approaches (Vanlint, 2019).

During research, ethnographers may use several physical notebooks or loosely formatted documents (Weber, 1991) that serve different functions (observational or analytical notes, etc.). In a digital medium, these often consist of a variety of text files containing analysis reports/memos, notes from interviews, comments, etc. These text files (from a text editor) may also include sketches, graphics, and audio or video recordings. There are currently few digital ready-to-use tools that assemble these sometimes disparate notes and files. General-purpose online note-taking tools (such as Evernote) can be used. In the field of design, some platforms are being developed that facilitate collective note-taking in various formats (audio, text, video) and provide a "shared history" of events for participatory research projects.

During the analysis of observations obtained and data gathered, Qualitative Data Analysis (QDA) tools are used to elucidate interpretations of the coding of information from a body of text (verbatim accounts, etc.).

7.1.1 A few examples:

- An example of a physical sociological research field notebook used in participatory research:

https://reflexivites.hypotheses.org/11932

- The "Lévi-Strauss and his Nambikwara fieldnotes – NAMBIKWARA" project.

An ANR (French National Research Agency) project to decipher, digitise, and produce a scholarly edition of Claude Lévi-Strauss' field notebooks written during his second visit to the Nambikwara people (May 1938 – January 1939):

https://anr.fr/Projet-ANR-18-CE27-0017

- Various digitised and anonymised surveys undertaken as part of the BeQuali project (a qualitative social science survey bank): https://bequali.fr/en/

For example, the qualitative survey on "The working conditions of pregnant women" by Anne Marie Devreux (with anonymisation assistance from the BeQuali team):

https://cdsp.sciences-po.fr/en/ressources-en-ligne/ressource/cdsp_bq_s12/

- L'Atelier des Chercheurs ("Researchers' Workshop") provides a number of open source prototypes employed in the context of a variety of standard and collaborative observational practices in design projects:

https://latelier-des-chercheurs.fr/infos

- Summary of tools aimed at assisting content analysis, interpretation of real-life accounts, or interviews in qualitative sociology:

https://www.squash.uliege.be/logiciels/

A.3.2 Data collection and recording in challenging environments (archaeology, environmental sciences)

In archaeology, excavation notebooks used to exist in a format that allowed observations to be shared as research progressed. In the 1980s, these observational practices became more systematic and standardised, leading to the introduction of physical files that allowed for an analytical overview of the situation in the field and the recording of information under defined headings (Desachy, 2016). This information is now directly recorded digitally using field apps/tablets, resulting in the creation of relational databases and the development of field-based Geographic Information Systems (GIS).

Field data collection practices in challenging environments are also a central issue in ecology and the environmental sciences. What we now call "electronic field notebooks" actually represent various media and tools associated with these approaches. For example, these include mobile devices (phones, tablets, etc.) that are designed to withstand extreme environments. These are used with data collection and organisation software. The latter often include features like an "offline" mode and postponed data sync (for situations in which a network connection is not available), as well as collaborative features allowing the shared input of data into the research project's databases and GIS.

7.1.2 A few examples:

- OpenData Kit is an *open source* application for the collection, management, and use of data in challenging environments.
- Pi4X4 is a rugged tablet project built using a Raspberry Pi.

https://odk-x.org/

- Presentations of various projects in ecology and the environmental sciences during an "electronic field notebook" workshop in March 2018. https://rbdd.cnrs.fr/spip.php?article270
- An example of data collection in a public health project in Africa using mobile phones and open data kit. https://www.youtube.com/watch?v=vB0jHQgVkEs

The tools most closely related to observational and analytical research work, as well as to field and investigative research, take a variety of forms depending on knowledge production modes and their validity. The idea of reproducing an experiment (which underlies the notion of a "laboratory notebook") is antithetical to qualitative approaches that seek to observe and interpret social phenomena, either through the study of specific communities or people, or through the critical analysis of a body of text.

Nevertheless, there are a number of needs that digital tools are able to meet depending on specific usage methods and practices, e.g. use on mobile devices with "offline" mode, collected data backup and protection, internal search engine, linking to data repositories, collaborative features, and access rights management.

7.1.3 Further information:

- Baribeau, Colette. 2005. *L'instrumentation dans la collecte de données - Le journal de bord du chercheur* (Instrumentation in data collection - The researcher's logbook) Hors-Série (2): 17.
- Desachy Bruno. 2016. *Du Carnet de Fouilles Aux Systèmes d'information* Archéologiques de Terrain: (From the Excavation notebook to Field Archeology Information Systems:) Quelques Remarques Sur l'évolution de l'en- *registrement et l'impact de l'informatisation* (Some remarks on the evolution of information registration and the impact of computerisation) Médiathèque de La MSH Mondes · Médiathèque MSH Mondes.

http://mediatheque.mae.cnrs.fr/s/fr/item/883

- Gruson-Daniel, Célya. 2019. *Les outils numériques d'enquête* (Digital research tools). In Guide décolonisé et pluriversel de formation à la recherche en sciences sociales et *humaines* (decolonised and pluriversal research training in the social and human sciences). *Éditions science et bien commun* (Science and common good editions).

https://scienceetbiencommun.pressbooks.pub/projetthese/chapter/les-outils-numeriques-denquete/

- Olivier de Sardan, Jean-Pierre. 1995. "La politique du terrain Sur la production des données en anthropologie" ("Fieldwork policy for the generation of data in anthropology"). Survey. Archives de la revue Enquête, issue 1 (October): 71–109.

https://doi.org/10.4000/enquete.263

- Pecqueux, Anthony. n.d. L'écoute comme politique de l'enquête (Listening as a policy of inquiry): 5/ Écoute acousmatique, ou Le nez dans le carnet ou dans le guidon ? (5/Acousmatic listening: absorbed with the notebook or absorbed in the task) Bulletin. Espaces *réflexifs (blog)*. Visited on 27 May 2020.

https://reflexivites.hypotheses.org/11932

- Vanlint, Alice. 2019. "Le journal de bord ou de terrain." dans Guide décolonisé et pluriversel de formation à la recherche en sciences sociales et humaines. ("The log or field log." in "A decolonised and pluriversal guide to research training in the social sciences and humanities"). Éditions science et bien commun (Science and common good editions). https://scienceetbiencommun.pressbooks.pub/projetthese/chapter/le-journal- de-bord-ou-de-terrain/

- Weber, Florence. 1991. *L'enquête, la recherche et l'intime ou : pourquoi censurer son journal de terrain ?* ("Investigation, Research, and Intimacy, or Why You Should Censor Your Field-Diary?") Espace Temps 47 (1): 71–81. https://doi.org/10.3406/espat.1991.3788

A.4 Description of the criteria

The features and specifications listed in section 4 are further described below when deemed necessary:

Feature	Description
Description of experiments and results	
Ability to duplicate experiments (from search results) as a model/template for a new experiment Description of protocols	
Ability to duplicate protocols as a model/template for a new protocol new protocol	
Attachment management: office (PDF, audio, video, etc.) or scientific and technical (git, chromategrams, spectra, images, etc.)	
Ability to export materials from the ELN to a printable format	The software must allow the export of all its content for subsequent printing. This means formatting all kinds of materials (images, diagrams, texts, etc.).
Access to external data (i.e. not stored in the ELN), raw or not, for reading and for reference	The ability to save links to data sets, ideally with a DOI (digital object identifier)
Data structuring with internal search engine and filter; use of keywords (tags)	The ability to execute simple searches of material stored in the notebooks, primarily via the user interface
Accessible content (for mining)	A TDM (Text and Data Mining) tool can access the data in the ELN in order to run searches.

Ability to associate metadata with processed datasets,whether generic (dates, authors, etc.) or specific (ontologies, taxonomies, etc.)	
Use and maintenance of equipment (including notifications)	A CMMS component (Computerised Maintenance Management System) must be proposed to the user. This will allow for the management of equipment reservation and use schedules, as well as the management of equipment servicing and maintenance. A notification feature (equipment not returned, preventive maintenance or calibration necessary, equipment not returned from supplier, etc.) must also be included.
Stock management (reagents, cell lines, etc.)	A stock management component must be available in the electronic lab notebook. It provides users with an overview of supplies and allows them to plan the procurement of consumables (reagents, cell lines, chemical products, etc.).
Tools for simple calculations (basic mathematical functions, unit conversion, etc.)	
Tools for complex calculations (statistics or analyses requiring a programming language) programming language)	
Molecule building/design tool	
Notes on experiments or protocol (during experimentation), reports, etc.	Ability to create simple documents, linked to others.
Ability to cite content from bibliographies (such as labtex.bib Text)	Ideally, it should be possible to connect to a bibliographic database, (e.g. PubMed or HAL).
Access to workstation office tools	This allows files to be viewed or modified without leaving the ELN. Clicking on the file in the ELN opens the corresponding application, if it is installed on the workstation.
Management of software tools and code (continuous integration, continuous deployment, version control, etc.)	A feature rarely found in ELNs, the alternative being a connection to a specialised server such as GitLab or or SourceSup.

Management of sharing by users	Each user must be able to allow other users access to information related to a project or experiment. This allows for quick but fully informed sharing of information with rightholders.
Shared agenda	If an electronic lab notebook includes a project, for example, and this project includes milestones that need to be met or experiments that need to be carried out (i.e. involving several people), the availability of a shared agenda for the rightholders is essential.
Collaborative document editing (online and simultaneously by several users)	The electronic lab notebook must allow several users to edit documents online and simultaneously. This ensures that all changes are taken into account when editing simultaneously and prevents a document from being locked because someone is already editing it.
Annotation of content and comment insertion	This involves adding to descriptions of experiments or protocols without creating a new document.

The list of technical specifications cited in section 4 is detailed below:

Specification	**Description**
Compliance with the GDPR for personal user data	
Compliance with the GDPR for research data: personal data of research participants if stored in the ELN	
Have managed, redundant, and backed-up storage (remote sites)	
Compliance with regulatory or legislative requirements governing the hosting location of research data	

Enable document signing (hashes)	This is not an individual's signature (see "Authentication/security/Enable electronic signature by participants" criterion) but a signature by a program that generates a sequence of characters specific to the file (its "hash"). If the file is made available to a third party, the third party must be able to retrieve the same sequence of characters by using the same signing mechanism (even if the file was encrypted and decrypted), thus proving that the authenticity of the file has been protected. See https://www.cnil.fr/fr/securite-chiffrer-garantir-lintegrite-ou-signe
Enable activity and document timestamping	
Enable unique input identification (protocol, experiment, score, etc.) – UUID	
Enable versioning/change tracking	Each action carried out must be indicated, for example with incremental version numbers and/or a "Last Modified" date. Versioning ensures that no action can be carried out without being recorded.
Enable the implementation of predefined workflows (e.g. a set of required actions in the description of an experiment or result)	
Enable electronic signature by participants	Each person using the electronic lab notebook must be able to sign it electronically. Typically, a witness (a unit head, a colleague not working on the project, or someone from another team, etc.) must regularly sign the notebook in order to certify what is recorded in it, thereby guaranteeing its legal value.
Use a secure authentication system	
Enable the creation of user management rules: access permissions by role, project, and life cycle (participant arrival/departure, project launch/completion, etc.)	

Encrypt data transport (TLS, RGS-compliant algorithms)	
Encrypt data	Each data backup must be protected by encryption to ensure confidentiality. It is inconceivable that a backup be used without any restrictions, just as applies to the electronic lab notebook.
Include regular backups and restores	In order to safeguard the contents of the electronic lab notebook, the user should be able to set up regular backups (with a specified regularity and possibly location) that are then automatically executed. It must also be possible to restore content in the event of errors.
Define data retention periods and data archiving or destruction data	
Ensure the confidentiality of exported data for backup or archiving.	
Safeguard the confidentiality of information that may represent valuable know-how or confidential business information (safeguard innovation in the event of a patentable invention).	
Safeguard the confidentiality of information exchanged with a partner in the context of a collaboration.	
Compatibility with different operating systems (Windows, Linux, etc.), for servers	
Allow data management (Import/Export) in a structured format (e.g. in case of a software or administrative modification)	
Allow integration with a local information system (databases, directories, etc.) Enable links with data repositories When available, use open standards for input or output data to allow information to be permanently re-used.	

Enable data sharing through data integration and exposition services, via WebServices or REST APIs if possible.	
Ability to easily install updates and upgrades, while guaranteeing accessibility, if used from a web browser	
Ergonomic and intuitive	
Enable modularity/customisation/flexibility	
Have the necessary documentation (guide, tutorials, instruction manuals, etc.)	
Access to source code in order to manage, if necessary, the development of the tool, its maintenance, and most importantly the data it contains	The proposed solution is open source, meaning that the software's source code is made available under a free, open licence. Ideally the source code is available on a software forge and is documented. The open source aspect can also be combined with a community development process. As far as the data itself is concerned, the Terms of Use and "data policy" specify the conditions of data use and retention. If there are any APIs, these are also documented and open.
Support by a community of users	It is crucial for the user to be able to contact other users (e.g. via an online forum) to find a solution to any questions or problems they may have. The user should also be able to post any solutions they find themselves.

Multilingual support (at least English)	As with most scientific communities, internationalism represents a fundamental aspect. Not everyone speaks (or is sufficiently fluent in) the home language of the laboratory they are working in, so it is essential for users to be able to select the language they use. This setting obviously only affects the UI of the electronic lab notebook and not its content. So it is fair to say that at least English should be an option.
Allow offline use with subsequent data synchronisation	When network access is not available (e.g. when conducting experiments outdoors or in a building with no network connection), the user must be able to use the electronic lab notebook in offline mode. Once a network connection is established, any new data and/or changes must then be synchronised.
Be usable on the move (on tablets and smartphones)	The electronic lab notebook must be compatible with mobile devices such as tablets and smartphones. This implies that the software's ergonomics are compatible with touchscreens and varying screen resolutions (icons not too small, numpad, etc.). It is also essential that the software can run in offline mode and can sync data afterwards.
Taking the environmental footprint into account: electricity consumption of servers, etc.	

A.5 List of tools analysed and not analysed

The list of tools assessed in section 5 is given below, along with notes for each tool:

Name	Publisher	Type	Score
BIOVIA[19]	Dassault Systems	Commercial	USA storage by default.
Chemotion[20]	Karlsruhe Institute of Technology	Open Source	Originally developed for organic chemistry, but can also be used for other fields. The development of new discipline-specific solutions is in progress.
ElabFTW_dev[21]	Deltablot	Open Source	Standard version. Examples of usages are available[22]
ElabFTW_QeR	Deltablot	Open Source	Version of the CNRS Qualité en Recherche network.
Elog[23]		Open Source	Open source tool developed and maintained by the particle and nuclear physics community as a whole.
FindMolecule[24]	FindMolecule, inc.	Commercial	ELN and chemical substances inventory management. Under Quebec law, free for the academic community
Jogl[25]	Just One Giant Lab MVAC	Open Source	SaaS project management solution; cannot replace an ELN but can complement one.
Jupyter notebook[26]		Open Source	Allows for development, documenting, and execution of code, as well as sharing of results.
LabArchives[27]	Labarchives LLC	Commercial	
LabBook[28]	eNovalys	Commercial	Tool used by units in the DR10 of the CNRS and by universities (Polytechnic University of Hauts-de-France, Paris Sud, Strasbourg).
LabCollector[29]	Agilebio	Commercial	Used at the CEA (French Alternative Energies and Atomic Energy Commission), mainly in biology. https ://lab-collector.com/
Labforward[30]	LabForward gmbh	commercial	
LabGuru[31]	Biodata	Commercial	Deployed at Inserm. ELN with some LIMS features.

Mbook[32]	mestrelab	Commercial	for synthetic chemists.
OSF.io[33]	Center for Open Science	non-profit organ isation nisation	
RSpace[34]	ResearchSpace	Commercial	Interfacing with repositories already implemented - Under British legislation.

The following list contains tools that were selected but not assessed by the group:

Name	**Publisher**	**Type**	**Score**
Scinote[35]	scinote llc	Commercial	only available in cloud mode
Benchling[36]	benchling	Commercial	mainly used in the private sector (including Sanofi, Syngenta, BP)
Labstep[37]	Labstep Ltd.	Commercial	Only available in cloud mode, free for the academic community
ISIA			AnaEE France consortium

19 https://www.3ds.com/products-services/biovia/products/laboratory-informatics/electronic-lab-notebooks/biovia-notebook/

20 https://www.chemotion.net/chemotionsaurus/

21 https://www.elabftw.net/

22 https://archive.fosdem.org/2021/schedule/event/open_research_using_elabftw/

23 https://elog.psi.ch/elog/

24 https://findmolecule.com/

25 https://jogl.io/fr

26 https://jupyter-notebook.readthedocs.io/en/stable/

27 https://www.labarchives.com/eln-for-research/

28 https://www.lab-book.org/

29 https://labcollector.com/

30 https://labforward.io/

31 https://www.labguru.com/

32 https://mestrelab.com/software/mbook/

33 https://osf.io/

34 https://www.researchspace.com/

35 https://www.scinote.net/

36 https://www.benchling.com/notebook

37 https://www.labstep.com/

Cassandre[38], NVivo[39]	University de Liège	Open Source	Cassandre and NVivo are Open Source programs for analysing qualitative data, used in HSS.
Hivebench	Elsevier	Commercial	
Bookitlab[40]	Prog4biz Software Ltd	Commercial	
OpenDataKit[41]		Open Source	
Elabjournal[42]	Bio-itech-Ependorf	Commercial	Used by the Pasteur Institute
Laby[43]	KYLI SAS	commercial	Has some LIMS features. Developed in France.
Evernote[44]	Evernote corporation	Commercial	

[38] https://github.com/Hypertopic/Cassandre

[39] https://www.qsrinternational.com/nvivo-qualitative-data-analysis-software/home

[40] https://bookit-lab.com/

[41] https://getodk.org/#features

[42] https://www.elabnext.com/products/elabjournal/

[43] https://laby.io/

[44] https://evernote.com/intl/fr

B Glossary and abbreviations

Term or Concept	Translation	Meaning
ANSSI – Agence Nationale de la Sécurité des Systèmes d'Information	French National Agency for the Security of Information Systems	French national authority for the security and defence of information systems.
API	Application Programming Inter-face	Standardised set of classes, methods, features, and constants that serve as a front end through which software provides services to other software
Checksum, hashing, hash	n.a.	A process that uses input data to generate a digital signature that can be used to quickly identify the initial data, similar to the way a signature identifies a person.
CKAN	Comprehensive Knowledge Archive Network	Web application enabling the storage and sharing of data such as spreadsheets or database content.
CoSO - Comité pour la Science Ouverte	French Committee for Open Science	Attached to the Directorate-General for Research and Innovation of the French Ministry of Higher Education, Research, and Innovation.
Dataverse	Comprehensive Knowledge Archive Network	Open source web application for storing, sharing, citing, searching, and analysing research data.
DMP	Data management *plan*	A living document that helps the researcher or research project leader to establish a plan for the management of data used and generated during their research activity or project.

DOI	Digital Object Identifier	System for identifying resources, which can be digital resources, such as a film, a report, or scientific articles, but can also be people or any other type of object.
EBIOS (*Expression des Besoins et Identification des Objectifs de Sécurité*)	Expression of Needs and Identification of Security Objectives	Comprehensive information systems security risk management tool, compliant with the RGS and the latest ISO standards. It enables such risks to be assessed and dealt with.
ELN	Electronic Lab Notebook	See chapter 1 of the report.
Emacs	n.a.	A series of text editors with an expandable set of features. It is popular among programmers and more generally among people with more advanced computer skills.
ESRI (Enseignement Supérieur, Recherche et Innovation)	Higher Education, Research, and Innovation	
EDM	Electronic Document Management	A computerised process for identifying, creating, producing, organising, distributing, and managing information and physical or electronic documents within an organisation.
Git	When asked why he called his software "git", which means "unpleasant person" in British English slang, Linus Torvalds quipped: "I'm an egotistical bastard, and I therefore name all my projects after myself. First 'Linux', now 'Git'."	Free distributed version control software.

GitLab	n.a.	Free forge software based on Git providing wiki functionality, a bug tracking system, continuous integration, and continuous delivery.
CMMS	Computerised Maintenance Management System	A software-assisted management method for a company's maintenance departments to help them in their work.
HAL (*Hyper Articles en Ligne*)	n.a.	Multidisciplinary open archive intended for the upload and dissemination of research-related scientific articles (published or not) and theses from French or foreign academic and research institution, as well as from public or private laboratories.
HPC	High-performance computing	Supercomputing.
ICOS	Integrated Carbon Observation System	European infrastructure dedicated to observing and quantifying greenhouse gas fluxes.
INSB (*Institut National des Sciences Biologiques*)	French National Institute of Biological Sciences	One of the CNRS' ten specialised institutes.
LIMS	Laboratory Information Management System	See paragraph 2.2.3 of the report.
ORCID	Open Researcher and Contributor ID	Non-proprietary alphanumeric code to uniquely and perpetually identify researchers and authors of scholarly and scientific contributions.

Benchmarking		The practice of finding companies that perform a given process or task most efficiently, studying them, and then tailoring this process to one's own company.
PI	Principal Investigator	In North America, refers to the holder of an independent grant and the lead researcher for the grant project, usually in the sciences; often synonymous with "head of the laboratory" or "research group leader".
PID	Persistent IDentifier	A string of alphanumeric characters that persistently identifies a resource regardless of its location. This resource can be any real or conceptual object (a person, a structure, research results, etc.).
PubMed	n.a.	Leading search engine for bibliographic data from all fields of biology and medicine. PubMed is free and provides access to the MEDLINE bibliographic database, which contains citations and abstracts of biomedical research articles. Developed and hosted by the National Institutes of Health (NIH, USA).
REST	Representational state transfer	A software architectural style that defines a set of constraints for creating web services. Web services that conform to the REST architectural style (also called RESTful web services) allow for interoperability between computers on the Internet.

RGI (*Référentiel Général de l'Interopérabilité*)	General Interoperability Framework	A set of recommendations listing norms and standards that bolster interoperability within the French government's information systems.
GDPR	General Data Protection Regulation	EU regulation on the protection of natural persons with regard to the processing of personal data and on the free movement of such data.
RGS (Référentiel Général de Sécurité)	General security framework	Regulatory framework for establishing trust in electronic exchanges between the French government and citizens.
RStudio	n.a.	Free, open, and multi-platform development environment for R, a programming language used for data processing and statistical analysis.
HSS	Humanities and social sciences	A group of disciplines studying various aspects of human experience from both individual and collective perspectives.
SourceSup	n.a.	Software management platform for research and higher education institutions. It is managed by the RENATER public interest group and allows software with open source licences to be made public, and projects to be hosted without making them public.
TDM	Text and Data Mining	The process of obtaining information from material read by a computer. It works by copying large quantities of textual material, extracting the data, and combining it to identify patterns.

TLS	Transport Layer Security	A protocol that allows communications security over computer networks, particularly over the Internet.
UUID	Universally Unique IDentifier	A system that allows distributed systems to uniquely identify a piece of information without depending on central coordination.
VRE	Virtual Research Environment	An online system that helps researchers collaborate.