Electronic laboratory notebooks in a public–private partnership

Lea A.I. Vaas[1,*], Gesa Witt[1,*], Björn Windshügel[1], Andrea Bosin[2], Giovanni Serra[2], Adrian Bruengger[3], Mathias Winterhalter[4], Philip Gribbon[1], Cindy J. Levy-Petelinkar[5] and Manfred Kohler[1]

[1] ScreeningPort, Fraunhofer Institute for Molecular Biology and Applied Ecology (IME), Hamburg, Germany
[2] Department of Physics, University of Cagliari, Monserrato, Italy
[3] Research, Basilea Pharmaceutica International AG, Basel, Switzerland
[4] Department of Life Sciences & Chemistry, Jacobs University Bremen, Bremen, Germany
[5] Innovation, Performance and Technology, GlaxoSmithKline, Collegeville, PA, United States of America
[*] These authors contributed equally to this work.

ABSTRACT

This report shares the experience during selection, implementation and maintenance phases of an electronic laboratory notebook (ELN) in a public–private partnership project and comments on user's feedback. In particular, we address which time constraints for roll-out of an ELN exist in granted projects and which benefits and/or restrictions come with out-of-the-box solutions. We discuss several options for the implementation of support functions and potential advantages of open access solutions. Connected to that, we identified willingness and a vivid culture of data sharing as the major item leading to success or failure of collaborative research activities. The feedback from users turned out to be the only angle for driving technical improvements, but also exhibited high efficiency. Based on these experiences, we describe best practices for future projects on implementation and support of an ELN supporting a diverse, multidisciplinary user group based in academia, NGOs, and/or for-profit corporations located in multiple time zones.

INTRODUCTION

Laboratory notebooks (LNs) are vital documents of laboratory work in all fields of experimental research. The LN is used to document experimental plans, procedures, results and considerations based on these outcomes. The proper documentation establishes the precedence of results and in particularly for inventions of intellectual property (IP). The LN provides the main evidence in the event of disputes relating to scientific publications or patent application. A well-established routine for documentation discourages data falsification by ensuring the integrity of the entries in terms of time, authorship, and content (*Myers, 2014*). LNs must be complete, clear, unambiguous and secure. A remarkable example is Alexander Fleming's documentation, leading to the discovery of penicillin (*Bennett & Chung, 2001*).

Table 1 Long-term benefits of an electronic laboratory notebook (ELN) compared to a paper based LN.

| Definitely	Potentially
• Create (standard) protocols for experiments • Create and share templates for experimental documentation • Share results within working groups • Amend/extend individual protocols • Full complement of data/information from one experiment is stored in one place (in an ideal world) • Storage of data from all experiments in a dedicated location • Search functionality (keywords, full text) • Protect intellectual property (IP) by timely updating of experimental data with date/time stamps	• Exchange protocols/standard operating procedures (SOPs) • Remote access of results/data from other working groups • Ensure transparency within projects • Discuss results online • Control of overall activity by timely planning of new experiments based on former results • Search for chemical (sub)structures within all chemical drawings in experiments

The recent development of many novel technologies brought up new platforms in life sciences requiring specialized knowledge. As an example, next-generation sequencing and protein structure determination are generating datasets, which are becoming increasingly prevalent especially in molecular life sciences (*Du & Kofman, 2007*). The combination and interpretation of these data requires experts from different research areas (*Ioannidis et al., 2014*), leading to large research consortia.

In consortia involving multidisciplinary research, the classical paper-based version of a LN is an impediment to efficient data sharing and information exchange. Most of the data from these large-scale collaborative research efforts will never exist in a hard copy format, but will be generated in a digitized version. An analysis of this data can be performed by specialized software and dedicated hardware. The classical application of a LN fails in these environments. It is commonly replaced by digital reporting procedures, which can be standardized (*Handbook: Quality Practices in Basic Biomedical Research, 2006*; *Bos et al., 2007*; *Schnell, 2015*). Besides the advantages for daily operational activities, an electronic laboratory notebook (ELN) yields long-term benefits regarding data maintenance. These include, but are not limited to, items listed in Table 1 (*Nussbeck et al., 2014*). The order of mentioned points is not expressing any ranking. Beside general tasks, especially in the field of drug discovery some specific tasks have to be facilitated. One of that is functionality allowing searches for chemical structures and substructures in a virtual library of chemical structures and compounds (see Table 1, last item in column "Potentially"). Such a function in an ELN hosting reports about wet-lab work dealing with known drugs and/or compounds to be evaluated, would allow dedicated information retrieval for the chemical compounds or (sub-) structures of interest.

Interestingly, although essential for the success of research activities in collaborative settings, the above mentioned advantages are rarely realized by users during daily documentation activities and institutional awareness in academic environment is often lacking.

Since funding agencies and stakeholders are becoming aware of the importance of transparency and reproducibility in both experimental and computational research (*Sandve et al., 2013*; *Bechhofer et al., 2013*; *White et al., 2015*), the use of digitalized documentation, reproducible analyses and archiving will be a common requirement for funding
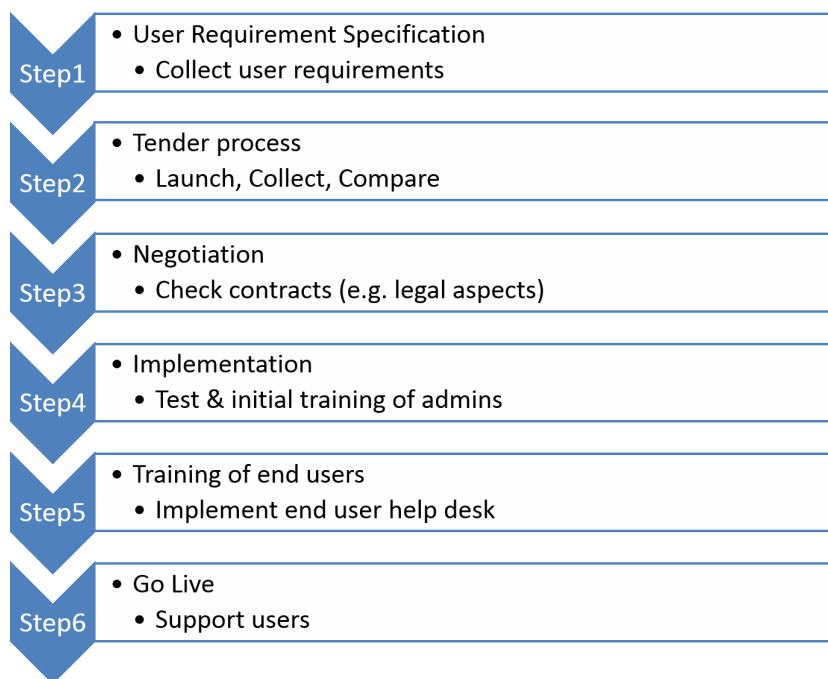
New Drugs for Bad Bugs (ND4BB) Structure
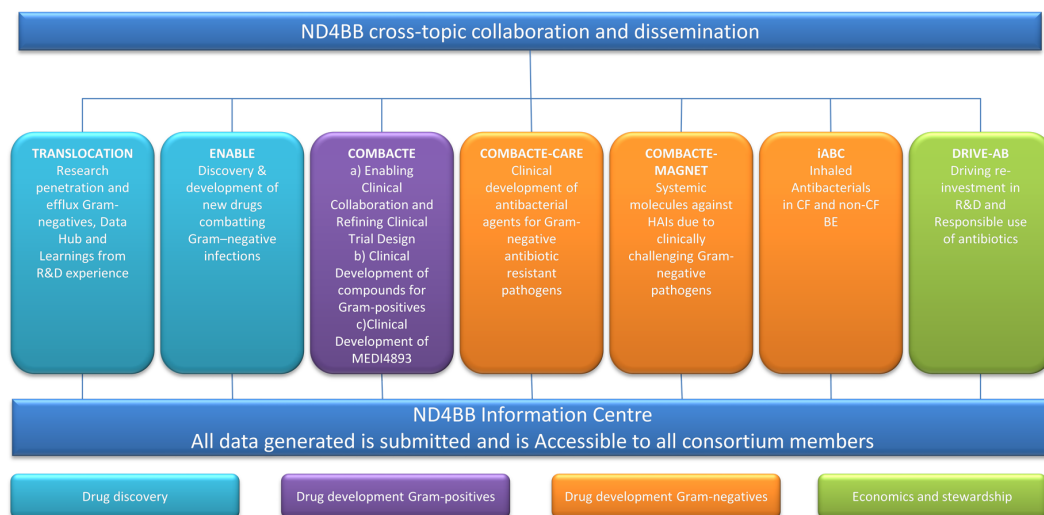
Figure 1 Structural outline of the New Drugs for Bad Bugs (ND4BB) framework.

applications on national and international levels (*Woelfle, Olliaro & Todd, 2011*; *DFG, 2013*; *European Commission Directorate-General for Research & Innovation, 2013*).

A typical example for a large private-public partnership is the Innovative Medicines Initiative (IMI) *New Drugs for Bad Bugs* (ND4BB) program (*Payne et al., 2015*; *Kostyanev et al., 2016*) (see Fig. 1 for details). The program's objective is to combat antibiotic resistance in Europe by tackling the scientific, regulatory, and business challenges that are hampering the development of new antibiotics.

The TRANSLOCATION consortium focus on (i) improving the understanding of the overall permeability of Gram-negative bacteria, and (ii) enhancing the efficiency of antibiotic research and development through knowledge sharing, data sharing and integrated analysis. To meet such complex needs, the TRANSLOCATION consortium was established as a multinational and multisite public–private partnership (PPP) with 15 academic partners, five pharmaceutical companies and seven small and medium sized enterprises (SMEs) (*Rex, 2014*; *Stavenger & Winterhalter, 2014*; *ND4BB - TRANSLOCATION, 2015*).

In this article we describe the process of selecting and implementing an ELN in the context of the multisite PPP project TRANSLOCATION, comprising about 90 bench scientists in total. Furthermore we present the results from a survey evaluating the users' experiences and the benefit for the project two years post implementation. Based on our experiences, the specific needs in a PPP setting are summarized and lessons learned will be reviewed. As a result, we propose recommendations to assist future users avoiding pitfalls when selecting and implementing ELN software.

Step1
- User Requirement Specification
  - Collect user requirements

Step2
- Tender process
  - Launch, Collect, Compare

Step3
- Negotiation
  - Check contracts (e.g. legal aspects)

Step4
- Implementation
  - Test & initial training of admins

Step5
- Training of end users
  - Implement end user help desk

Step6
- Go Live
  - Support users

Figure 2 Schematic outline of stepwise procedure for the implementation of a new system (*Nehme & Scoffin, 2006*).

METHODS

Selection and implementation of an ELN solution

The IMI project call requested a high level of transparency enabling the sharing of data to serve as an example for future projects. The selected consortium TRANSLOCATION had a special demand for an ELN due to its structure–various labs and partners spread widely across Europe needed to report into one common repository–and due to the final goal–data was required to be stored and integrated into one central information hub, the ND4BB Information Centre. Fortunately, no legacy data had to be migrated into the ELN.

The standard process for the introduction of new software follows a highly structured multiphase procedure (*Nehme & Scoffin, 2006*; *Milsted et al., 2013*) with its details outlined in Fig. 2.

For the first step, we had to manage a large and highly heterogeneous user group (Fig. 3) that would be using the ELN scheduled for roll out within 6 months after project launch. All personnel of the academic partners were requested to enter data into the same ELN potentially leading to unmet individual user requirements, especially for novices and inexperienced users.

As a compromise for step 1 (Fig. 2), we assembled a collection of user requirements (URS) based on the experiences of one laboratory that had already implemented an ELN. We further selected a small group of super users based on their expertise in documentation processes, representing different wet laboratories and in silico environments. The resulting
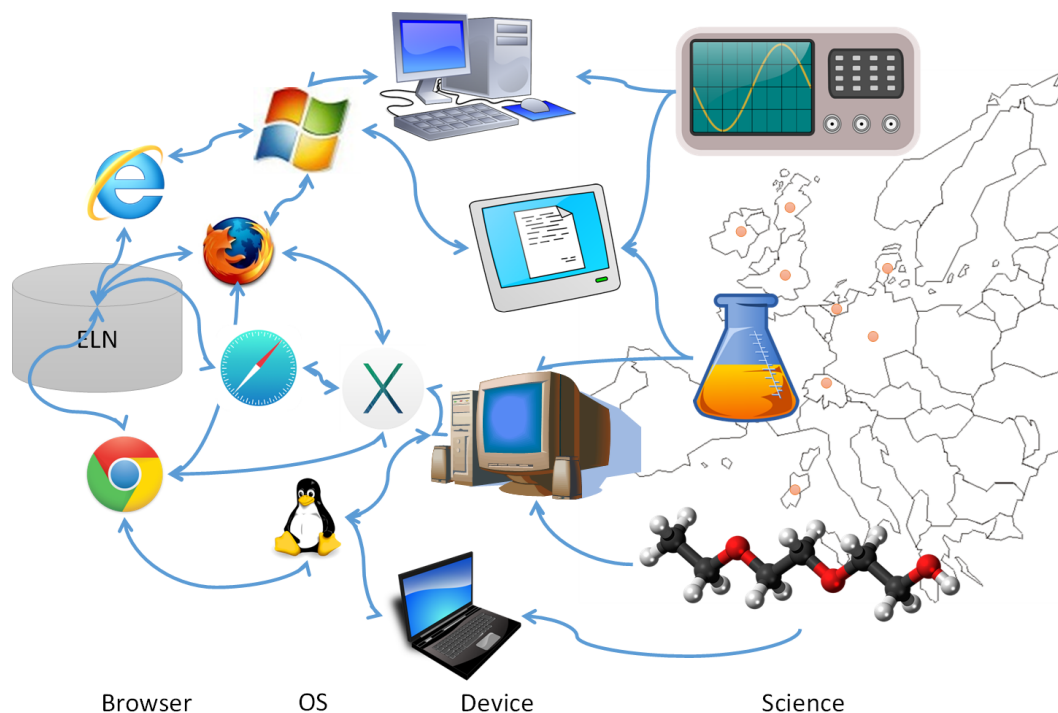
Figure 3 Technical and organizational challenge: schematic overview of paths for sharing research activity results within a public–private partnership on antimicrobial research.

URS was reviewed by IT and business experts from academic as well as private organisations of the consortium. The final version of the URS is available as Article S1.

In parallel, based on literature (*Rubacha, Rattan & Hossel, 2011*) and Internet searches, presentations of widely used ELNs were evaluated to gain insight into state-of-the-art ELNs. This revealed a wide variety of functional and graphical user interface (GUI) implementations differing in complexity and costs. The continuum between simple out-of-the-box solutions and highly sophisticated and configurable ELNs with interfaces to state-of-the-art analytical tools were covered by the presentations. Notably, the requirements specified by super users also ranged from "easy to use" to "highly individually configurable." Based on this information it was clear that the ELN selected for this consortium would never ideally fit all user expectations. Furthermore, the exact number of users and configuration of user groups were unknown at the onset of the project. The most frequently or highest prioritized items of the collected user requirements are listed in Table 2. We divided the gathered requirements into 'core' meaning essential and 'non-core' standing for 'nice to have, but not indispensable.' Further, we list here only the items, which were mentioned by more than two super users from different groups. The full list of URS is available Article S1.

Based on the user URS, a tender process (Step 2, Fig. 2) was initiated in which vendors were invited to respond via a Request for Proposal (RFP) process. The requirement of the proposed ELN to support both chemical and biological research combined with the need to access the ELN by different operating systems (Windows, Linux, Mac OS) (see

Table 2 Overview of user requirements organized as 'core user requirements' for essential items, and 'non-core user requirements' representing desirable features.

Core user requirements	Non-core user requirements
• System set-up and implementation should be fast and simple • Access from different platforms should be possible: Windows, Linux, Mac OS • Low training requirements (for high level of acceptance) • Hosted system with state-of-the-art security settings • Simple user management (only limited support by project members possible) • Suitable for both chemical (including e.g., drawings of molecules) and biological (including e.g., capture fluorescent images) experiments • Low costs, especially for long-term usage in the academic area	• Conform with Good Laboratory Practise • User management with dedicated access permissions (expectation: all users working on the same project, but in different work packages) • Workflow management • Order management • Chemical structure handling • Dedicated tree structure for storing experiments • Legally-binding procedures (signatures) • Modular expandability • Appropriate integrated analytical features • Social networking and collaborative (chat) features • Storage for large sets of "raw" data for re-analysis

Fig. 3) reduced the number of appropriate vendors. Their response provided their offerings aligned to the proposed specifications.

Key highlights and drawbacks of the proposed solutions were collected as well as approximations for the number of required licenses and maintenance costs. The cost estimates for licenses were not comparable because some systems require individual licenses whereas others used bulk licenses. At the time of selection, the exact number of users was not available.

Interestingly, the number of user specifications available out of the box differed by less than 10% between systems with the lowest (67) and highest (73) number of proper features. Thus, highlights and drawbacks became a more prominent issue in the selection process.

For the third step, (Fig. 2), the two vendors meeting all the core user requirements and the highest number of non-core user requirements were selected to provide a more detailed online demonstration of their ELN solution to a representative group of users from different project partners. In addition, a quote for 50 academic and 10 commercial licenses was requested. The direct comparison did not result in a clear winner–both systems included features that were instantly required, but each lacked some of the essential functionalities.

For the final decision only features which were different in both systems and supported the PPP were ranked between 1 = low (e.g., academic user licenses are cheap) and 5 = high (e.g., cloud hosting) as important for the project. The decision between the two tested systems was then based on the higher number of positively ranked features, which revealed most important after presentation and internal discussions of super users. The main drivers for the final decision are listed in Table 3. In total, the chosen system got 36 positive votes on listed features meeting all high ranked demands listed in Table 3, while the runner-up had 24 positive votes on features. However, if the system had to be set up in the envisaged consortium it turned out to be too expensive and complex in maintenance.

The complete process, from the initial collection of URS data until the final selection of the preferred solution, took less than five months.

Table 3   List of drivers for final decision about which ELN-solution to be set up and run in the project consortium.

Main drivers for the final decision

- Many end users were unfamiliar with the use of ELNs, therefore the selected solution should be intuitive
- Easy to install and maintain
- Minimal user training required
- Basic functionalities available out-of-the-box (import of text, spreadsheet, pdf, images and drawing of chemical structures), with as few configurations as possible
- ELN does not apply highly sophisticated checking procedures, which would require a high level of configuration, restricting users to apply their preferred data format (users should take responsibility for the correct data and the correct format of the data stored in the ELN instead)
- Web interface available to support all operating systems to avoid deploying and managing multiple site instances
- Vendor track record—experience of the vendor with a hosted solution as an international provider
- Sustainability:
  - affordable for academic partners also after the five year funding duration of the project (based on the maturity of the vendor, number of installations/users, the state-of-the-art user interface and finally also the costs)
  - Easy to maintain (minimal administrative tasks, mainly user management)
  - Support of different, independent user groups
  - Configurable private and public templates for experiments
- Proposed installation timeframe
- Per user per year costs for academics and commercial users

Following selection, the product was tested before specific training was offered to the user community (*Iyer & Kudrle, 2012*). Parallel support frameworks were rolled out at this time, including a help desk as a single point of contact (SPoC) for end users.

The fourth step (Fig. 2) was the implementation of the selected platform. This deployment was simple and straightforward because it was available as Software as a Service (SaaS) hosted as a cloud solution. Less than one week after signing the contract, the administrative account for the software was created and the online training of key administrators commenced. The duration of training was typically less than 2 h including tasks such as user and project administration.

To accomplish step 5 (Fig. 2), internal training material was produced based on the experiences made during the initial introduction of the ELN to the administrative group. This guaranteed that all users would receive applicable training. During this initial learning period, the system was also tested for the requested user functionalities. Workarounds were defined for missing features detected during the testing period and integrated into the training material.

One central feature of an ELN in its final stage is the standardization of minimally required information. These standardizations include, but are not limited to:

- Define required metadata fields per experiment (e.g., name of the author, date of creation, experiment type, keywords, aim of experiment, executive summary, introduction)
- Define agreed naming conventions for projects

- Define agreed naming conventions for titles of experiments
- Prepare (super user agreed) list of values for selection lists
- Define type of data which should be placed in the ELN (e.g., raw data, curated data).

We did not define specific data formats since we could not predict all the different types of data sets that would be utilized during the lifetime of the ELN. Instead, we gave some best practise advice on arranging data (Tables S1 and S2) facilitating its reuse by other researchers. The initially predefined templates, however, were only rarely adopted, also some groups are using nearly the same structure to document their experiments. More support especially for creating templates may help users to document their results more easily.

For the final phase, the go live, high user acceptance was the major objective. A detailed plan was created to support users during their daily work with the ELN. This comprised the setup of a support team (the project specific ELN-Helpdesk) as a SPoC and detailed project-specific online trainings including documentation for self-training. As part of the governance process, we created working instructions describing all necessary administrative processes provided by the ELN-Helpdesk.

Parallel to the implementation of the ELN-Helpdesk, a quarterly electronic newsletter was rolled out. This was used advantageously to remind potential users that the ELN is an obligatory central repository for the project. The newsletter also provided a forum for users to access information and news, messaging to remind them the value of this collaborative project.

Documents containing training slide sets, frequently asked questions (FAQs) and best practice spreadsheet templates (Table S2) have been made available directly within the ELN to give users rapid access to these documents while working with the ELN. In addition, the newsletter informed all users about project-specific updates or news about the ELN.

RESULTS

Operation of the ELN
Software operation can be generally split into technical component/cost issues and end-user experiences. The technical component considers stability, performance and maintenance whereas the end-user experience is based on the capability and usability of the software.

Technical solution
The selected ELN, hosted as a SaaS solution on a cloud-based service centre, provided a stable environment with acceptable performance, e.g., login < 15 sec, opening an experiment with 5 pages < 20 sec (for further technical details, please see Article S1). During the evaluation period of two years, two major issues emerged. The first involved denial of access to the ELN for more than three hours, due to an external server-problem, which was quickly and professionally solved after contacting the technical support, and the other was related to the introduction of a new user interface (see below).

The administration was simple and straightforward, comprising mainly minor configurations at the project start and user management during the runtime. One issue was the gap in communication regarding the number of active users causing a steady increase in number of licences.

A particular disadvantage using the selected SaaS solution concerned system upgrades. There was little notice of upcoming changes and user warnings were hidden in a weekly mailing. To keep users updated, weekly or biweekly mails about the ELN were sent to the user community by the vendor. Although these messages were read by users initially, interest diminished over time. Consequently, users were confused when they accessed the system after an upgrade and the functionality or appearance of the ELN had changed. On the other side, system upgrades were performed over weekends to minimize system downtime.

The costs per user were reasonable, especially for the academic partners for whom the long-term availability of the system, even after project completion, could be assured. This seemed to be an effect of the competitive market that caused a substantial drop in price during the last years.

User experience

In total, more than 100 users were registered during the first two years runtime, whereas the maximum number of parallel user accounts was 87, i.e., 13 users left the project for different reasons. The number of 87 users is composed of admins ($n = 3$), external reviewers ($n = 4$), project owners ($n = 26$) which are reviewing and countersigning as Principal Investigators (PIs), and normal users ($n = 41$). Depending on the period, the number of newly entered experiments per month ranged between 20 and 200 (Fig. 4 blue bars). The size of the uploaded or entered data was heterogeneous and comprised experiments with less than 1 MB, i.e., data from small experimental assays, but also contained data objects much larger than 100 MB, e.g., raw data from mass spectroscopy. Interestingly, users structured similar experiments in different ways. Some users created single experiments for each test set, while other combined data from different test sets into one experiment.

In an initial analysis, we evaluated user experience by the number of help desk tickets created during the live time of the system (Fig. 4 orange line).

During the initial phase (2013.06–2013.10) most of the tickets were associated with user access to the ELN. However, after six months (2013.12–2014.01), many requests were related to functionality, especially those from infrequent ELN users. In Feb 2015, the vendor released an updated graphical user interface (GUI) resulting in a higher number of tickets referring to modified or missing functions and the slow response of the system. The higher number of tickets in Aug/Sept 2015 were related to a call for refreshing assigned user accounts. However, overall the number of tickets was within the expected range (<10 per month).

There was a clear decline in frequent usage during the project runtime. The ELN was officially introduced in June 2013 and the number of experiments increased during the following month as expected (Fig. 4 blue bars). The small decrease in Nov 2013 was anticipated as a reaction to a new release, which caused some issues on specific operating system/browser combinations and language settings. Support for Windows XP also ended at this time. Some of the issues were resolved with the new release (Dec 2013). The increase in new experiments in Oct 2014 and subsequent decline in Nov 2014 is correlated to a reminder sent out to the members of the project to record all activities into the ELN. The
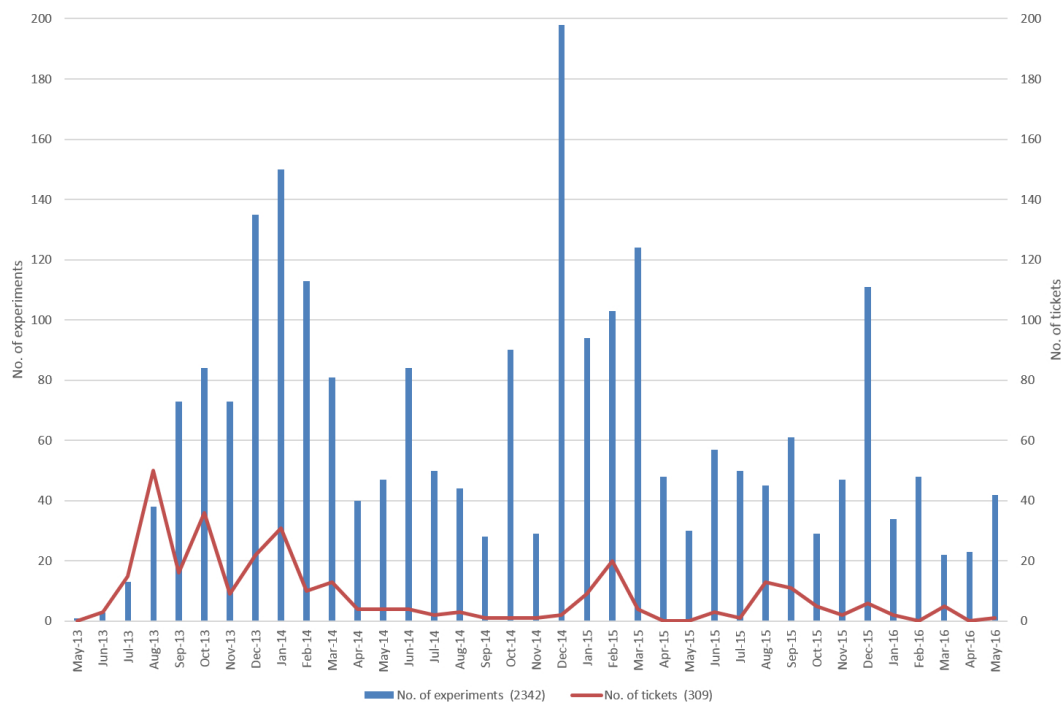
Figure 4 Overview of usage of the ELN and workload for the helpdesk over time. *Y*-axis on the left shows number of experiments created per month (blue bars) overlaid by number of help desk tickets created per month (orange line) scaled on the *y*-axis on the right.

same is true for Sept/Oct 2015. The last quarter of each year illustrates year-end activities, including the conclusion of projects and the completion of corresponding paperwork, as reflected in the chart (Fig. 4 orange line).

Overall, the regular documentation of experiments in the ELN appeared to be unappealing to researchers. This infrequent usage prompted us to carry out a survey of user acceptance in May/June 2015 (detailed description of analysis methods including KNIME workflow (*Berthold et al., 2007*) and raw data are provided in Article S2). The primary aim was to evaluate user experiences compared to expectations in more detail. In addition, the administrative team wanted to get feedback to determine what could be done to the existing support structure to increase usage or simplify the routine documentation of laboratory work. Overall, 77 users (see Table 4) were invited to participate in the survey. Two users left the project during the runtime of the survey. We received feedback from 60 (= 80%) out of the remaining 75 users. Two questionnaires were rejected due to less than 20% answered questions. The number of evaluated questionnaires is 58 (see Dataset S1).

There are also some limitations of the survey which should be discussed. The number of invited active ELN users was low ($n = 77$), thus we refused to collect detailed demographic data in order to ensure full anonymity of the participants, so we expected a higher participation and especially more detailed answers to the free text questions. In addition, some interesting analysis could not be answered by of the questionnaire due to the low number of returned forms ($n = 58$). For example only six users had some experience with ELNs. Three out of the six users found the ELN is changing the way of personal

Table 4 Number of users of the ELN and participants in the survey.

Description	n	Remarks
Maximum number of parallel users of the ELN	87	
Number of parallel users at the time point of the survey	84	
Administrators	3	Not invited to participate in survey
External Reviewers	4	Not invited to participate in survey
Project Owners = Principal Investigators	26	Invited to participate in survey
Normal User	51	Invited to participate in survey
Deactivated users during the runtime of the questionnaire	2	
Questionnaires returned	60	
Questionnaires rejected due to insufficient (less than 20%) number of answered questions	2	
Evaluated questionnaires	58	

Table 5 Overview about most important results from survey, which was sent out to 77 users from 18 academic and SME organisations. A set of 58 comprehensively answered questionnaires was considered for evaluation,

Major results from survey

- Most user never used an ELN before (51 out of 58 users replying to the survey stated "I never used an ELN before this project"; no info from remaining 19 invited users)
- Most users (76%) are using a paper notebook in addition to the ELN
- Many users ($n = 23$) would not recommend using an ELN again
- No of Operating systems: Linux = 7, Mac OS = 14, Windows = 37
- ELN typically used
    - Rarely or sometimes with < 1 h per session (53%)
    - Frequently < 1 h per session (16%)
    - Sometimes or frequently 1–2 h per session (9%)
- Frequent users ($n = 13$) realized an increase in quality of documentation (46%) and 38% would recommend this software to colleagues while even three out the 13 frequent users would **not** use an ELN again if they could decide
- Rarely users ($n = 19$) are skeptical about ELN functionality (42%)
- While 52% of the Mac and Linux users are satisfied about the performance 35% of the Windows users are unhappy compared to 27% which are satisfied about the performance of the system
- Helpdesk support in general is O.K. (36%), but some users (10%) seem to be not satisfied, especially with training (47%)
- Most users demand higher speed ($n = 15$) and/or better user interface ($n = 14$)

documentation positively, while the others didn't answered the question or gave a neutral answer. Thus, we did report these results as not representative.

It should also be mentioned that this survey reflects the situation of this specific PPP project. The results cannot be easily transferred to other projects. It would be of interest if the same survey would give different or the same results (i) either in other projects (ii) during the time course of this project.

A summary of the most important results of the survey is presented in Table 5.

Despite the perceived advantages of an ELN compared to traditional paper-based LNs as described above, users encountered several drawbacks during their usage of this ELN.

Table 6 Summary of acceptance of ELN and user experience given the mainly used OS plus an overview of the main result from survey for the different user-groups.

Frequently used platform to access the ELN	No. of users	Percentage of all users	No. of wet-lab researchers[a]	No. of in-silico researchers[a]
Mac	14	24%	5	8
Windows	37	64%	30	12
Linux	7	12%	1	5

Summarized result based on the answers by OS:
- Windows users mainly conduct wet lab work while Linux and Mac users perform in silico work
- Windows users find the software too slow and too labor-intensive
- Windows users know the functionality of the ELN better, as they are using the system more frequently
- Mac OS and Linux users are more comfortable with the speed of the software, but they would not use or recommend it again. This may be related to the specific in silico work which might not be supported by the ELN sufficiently

Notes.
[a] Those columns do not add up to 58 because some participants stated that they are both wet-lab and in-silico researchers.

Users criticized the provision of templates and cloned experiments, which were considered to impede the accurate documentation of procedural deviations. The standardized documentation of experimental procedures made it difficult to detect deviations or variations because they are not highlighted. Careful review of the complete documentation was required in order to check for missing or falsified information.

For many users (44 out of 58 = 76%), a paper-based LN was still the primary documentation system. They established a habit of copying the documentation into the ELN only after the completion of successful experiments rather than using the ELN online in real time. This extra work is also a major source of dissatisfaction and could create difficult situations in case of discrepancies between the paper and the electronic version when intellectual property needs to be demonstrated. In these cases, failed experiments were not documented in the ELN, although comprehensive documentation is available offline. For failed experiments, the effort to document the information in a digital form was not considered worthwhile by the users. In other cases, usage of the ELN for documentation of experimental work was hindered due to performance issues by technically outdated lab equipment.

The survey helped to understand what factors where contributing to the low usage of the ELN. Additional results of the survey grouped by operating system or usage are shown in Tables 6 and 7. For a more detailed analysis, see Dataset S2.

Conclusions from the survey

About 40% of the users did not find the selected solution appropriate for their specific requirements. Either the solution did not support specific data sets or experiment types, or the solution did not respond fast enough to be used adequately. This indicates that the solution was not fit for purpose. More individual user demands would have to be considered to improve the outcome. This would require more resources in time and manpower than can be accommodated in a publicly funded project. Time would be a key factor as

Vaas et al. (2016), *PeerJ Comput. Sci.*, DOI 10.7717/peerj-cs.83

12/22

Table 7 Overview of self-assessed frequency and usage-time of the ELN given the used OS plus summarized results of the survey for the different user-groups.

Self-assessed frequency of the ELN	Self-assessed length of usage-time of the ELN	Frequently used platform to access the ELN	No. of users	Percentage[a]	Percentage/OS
Rarely	<1 h	Linux	2	3%	29%
Sometimes	<1 h	Linux	5	9%	71%
Rarely	<1 h	Mac	3	5%	21%
Rarely	1–2 h	Mac	4	7%	29%
Sometimes	<1 h	Mac	5	9%	36%
Frequently	<1 h	Mac	2	3%	14%
Rarely	<1 h	Windows	9	16%	24%
Rarely	1–2 h	Windows	1	2%	3%
Sometimes	<1 h	Windows	7	12%	19%
Sometimes	1–2 h	Windows	8	14%	22%
Sometimes	>2 h	Windows	1	2%	3%
Frequently	<1 h	Windows	7	12%	19%
Frequently	1–2 h	Windows	4	7%	11%

Summarized results based on frequency of usage:

- In silico users enter into the ELN less frequently, which is not unexpected as computational experiments generally run for a longer period of time than wet lab experiments
- More frequent users operate the ELN online during their lab work
- Frequent users would like to have higher performance (this might be related to Windows)
- Better quality documentation was associated with more frequent use
- Frequent users are not disrupted by documenting their work in the ELN, they like the software and would use an ELN in future
- Frequent users of an ELN obtain a positive effect on the way documentation is prepared
- More frequent users like the software and feel comfortable about using the software while Infrequent users find the ELN complex and are frustrated about functionality
- Infrequent users are disappointed about the quality of search results

Notes.
[a] This column does not add up to 100% due to rounding errors.

experimental work begins within 6 months after project kick-off and the documentation process of experiments starts parallel to the experimental initiation. Keeping in mind that every user needs training time to get acquainted to a new system and there are always initial 'pitfalls' to any newly introduced system, an electronic laboratory notebook must be available within 4–5 months after project start. About one month should be allocated for the vendor negotiation process. Another month or two are required for writing and launching the tender process. This reduces the time frame for a systematic user requirement evaluation process to less than one month after kick-off meeting. It should be acknowledged that not all types of experiments will be fully defined nor will all users be identified. Thus the selection process will always be based on assumptions as described above.

The slow response of the selected system might have occurred due to many potential issues. It could be related to the bandwidth available at the location, but more frequently

we believe this is based on the hardware available. We tested the ELN on modern hardware with low and high bandwidth (Windows: Core i7 CPU @ 2.6 GHz, 6 GB RAM tested on ADSL: 25 kbit/s download, 5 kbit/s upload and iMac Core i5 CPU @ 3.2 GHz, 4 MB RAM tested on ADSL: 2 kbit/s download and 400 bit/s upload) without a major impact on performance, but we did not test old hardware. During this project we learned that in certain labs computers run Windows XP, MS Office 2003 and Internet Explorer 8. Using outdated software and hardware can be a contributor to the slow response of the ELN. Another potential issue could arise from uploading huge data sets on slow Asymmetric Digital Subscriber Lines (ADSL). Users working on local file servers and downloading data from Internet face unexpected low performance when uploading data to a web resource using ADSL, which is due to low upload bandwidth, respectively high latency. This is true for all centralized server infrastructures accessed by Internet lines including SaaS and should be reflected when considering hosting in the cloud.

Finally, users demand similar functionalities as on their daily working platform. This is an unsolved challenge due to the heterogeneity of software used in life sciences; from interactive graphical user interface (GUI) based office packages to highly sophisticated batch processing systems. The evolution of new ELNs should provide more closely aligned capabilities to meet the users' requirements.

For the ongoing PPP project, a more individualized user support capability may have helped to overcome some of the issues mentioned above. Individual on-site training parallel to the experimental work could offer insight to users' issues and provide advice for solutions or workarounds. This activity would require either additional travel for a small group of super users or the creation of a larger, widely spread group of well-trained super users which is highly informed about ongoing issues and solutions.

The issues discussed above also constitute a social or a scientific community problem. Being the first, which often is considered as being the best, is the dictum scientists strive to achieve especially when performance is reduced to the number of publications and frequency of citation, not to the quality of documentation/reproducibility accounting for determination of quality of results. This culture needs to be replaced by ''presenting full sets of high quality results including all metadata'' for additional benefit to the scientific community (*Macarelly, 2014*). ELNs could contribute significantly to this goal.

DISCUSSION

We successfully selected and implemented an ELN solution within an ambitiously short timeframe for mostly first-time electronic lab book users. The implementation included the creation and arrangement of internal training material and the establishment of an ELN helpdesk as a SPoC.

Lessons learned from the selection and implementation of an ELN solution

Generally accepted strategies for software selection and implementation are recommended because they provide a structured and well-defined process for decision-making. Typically, any selection process involves balancing several contrasting features. In our case, the

functionality of the system was balanced against long-term maintenance options and costs. The choice was a compromise between these three aspects.

Performance, ease of use and functionality are the most important aspects for end users of newly launched software. The general expectation is that software should make work quicker, easier and more precise. Thus, users anticipate that software should simplify typical tasks and support their current workflow without adaptation. However, an optimal use of the potential capacity of an ELN requires a change in the working process. Replacing the paper-based LN by a fixed installed lab PC will not lower the entrance barrier. The benefit for the end-user must be communicated in data handling and flexibility in data reuse (Off-site use and rapid communication among partners).

For an ELN in a PPP, the documentation of daily work is the key issue. In particular, in a project with widespread activities ranging from fundamental chemical wet laboratory and in silico work to biological *in vitro* and *in vivo* studies, the communication between the sites requires certain data structures. The selected ELN must support many different types of documentation (e.g., flat text files, unstructured images and multidimensional data containers) and in parallel must be at least as flexible as a paper-based notebook (e.g., portable, accessible, ready for instant use and suitable for use when the researcher is wearing laboratory protective clothing). The concealed features of an ELN, such as comprehensive filing of all information for an experiment in one place, standardised structuring of experiments, and long-term global accessibility are not as important for the end user when they are working with the system. The end user requires fast access to his latest experiments and expects support of his specific workflow during documentation. For some users, documenting laboratory work in an electronic system also requires more time and attention than writing entries in a paper notebook.

No ELN can fully reproduce the flexibility of a blank piece of paper. The first draft for experimental documentation in a classical paper-based LN can be rough and incomplete because only the documenter needs to understand it. Later on, the user can improve the notes and add additional information. In contrast, the documentation process in an ELN is more structured and guided by mandatory fields, which appears to be more time consuming to the user. However, the overall process could be faster once the user gets familiar with the system. The opportunity to lose any necessary information would be lowered by using the system online during the practical work. There are also other advantages in the ELN, e.g., linking experiments to regular protocols.

The main value of an ELN is that it makes the data more sharable because it is

- constantly accessible
- more complete
- easier to follow.

Lessons learned from technical solution

Operating system or configuration dependent issues were also encountered frequently. Supporting different systems and browsers, and interfacing with other tools such as office software packages, is complex and requires testing in different environments. Only the most common combinations of systems and tools are recommended by vendors. One

possible source of difficulty was the laboratory computers in academia. For the purpose of data recording, older versions of software running on outdated hardware are often used with restrictions based on instrument software installed on them (Article S2). Many incompatibilities are based on atypical configurations. However, it is impossible to drive the computer upgrade path for laboratory computers from the requirements of a single system, particularly for academic partners. Furthermore, in large academic institutions, many systems are not updated due to frequently changing personnel structures.

Data sharing

Science, by definition, should be a discipline sharing "knowledge" with the scientific community and the general public (DFG, 2013), particularly if funded by public organizations. Creating knowledge only for self-interests makes little sense, because knowledge can be verified and extended only by disputation with other researchers. Why do we struggle to share and discuss our data with colleagues?

Scientists often display a strong unwillingness to share their data. They often believe they are the data owner, i.e., the entity that can authorize or deny access to the data. Nevertheless, they are responsible for data accuracy, integrity and completeness as the representative of the data owner. The data generator should be granted primary use (i.e., publication) of the data (DFG, 2013), but the true owner is the organization that financially supports the project.

Within a PPP project, it is necessary to establish a documentation policy that is suitable for all. An agreement must occur on standards for the responsibility, content and mechanisms of documentation, particularly in international collaborations where country-specific and cultural differences need to be addressed (Elliott, 2010). Furthermore, no official, widely accepted standards are pre-defined. As long as the justification for ELNs is for more control over performance than to foster willingness to cooperate and share data and knowledge in the early phase of experiments, user acceptance will remain low (Myers, 2014). Without user acceptance, the quality of the documented work will not improve (Asif, Ahsan & Aslam, 2011; Zeng, Hillman & Arnold, 2011).

In addition to these social and community-based challenges (Sarich, 2013; Sarich, 2014a; Sarich, 2014b), there are technical aspects and security concerns that remain to be addressed. A simple "copy and paste" functionality for any type of text and data, including special characters and symbols, was high on the list of user demands. Another issue is the speed and convenience of access to the ELN, especially for technicians in the laboratory. Although the first enhancement needs some improvements from vendors, the responsibility for the second feature lies with the policies of the research organizations and their IT departments. Typically, out-of-date hardware and slow Internet connections are installed in laboratories. This impedes adoption of ELNs because slow hardware and slow network connections both have a negative impact on the usability of software. Vendors should ensure that the minimal specifications for network access and hardware required to run their systems without excessive latency are clearly defined also on a long-term perspective. Before implementation, hardware/network configurations should also be

Table 8 Checklist for ELN implementation.

Phase	
Selection	Selection and implementation team should be sufficiently staffed
	Group leaders should be part of the 'super user' team
	All user groups (wet lab, in silico) should be represented during requirement evaluation
	Most important user requirements should be met with the selected solution
Implementation	Sufficient hardware must be available especially in the labs
	Dependencies in OS and installed software should not permit using the ELN efficiently
	Additional hardware (e.g., tablets) should be considered for using ELN efficiently
	Test copy/paste functionality for important software packages
	Change management for the documentation process should be suggested
Support	Adequate training time especially for first time ELN users should be planned
	Implement sufficient training throughout the whole life time
	Implement sufficiently staffed and trained helpdesk according to the phases of the project (more at beginning, less at end)
	Feedback from users must be obeyed carefully
	Establish a useful communication path (e.g., newsletter, webpage)

checked by responsible persons in the research organizations and replaced by adequate equipment.

Finally, a generalized standard export/import format for the migration of data between different ELN solutions would be beneficial, because this would provide independence from one selected product and would also support data archiving (*Elliott, 2009*).

CONCLUSIONS

The adoption and use of ELNs in PPP projects need a careful selection and implementation process with a change management activity in parallel. Without the willingness of users to document their experimental work in a constructive, cooperative way, it will be difficult to acquire all the necessary information in time. Mandating users to record all activities, which could easily be obliged in a company by creating a company policy, will not work in a PPP project with independent organisations. Forcing users to record all activities suggests control of users and will result in minimal, low quality documentation.

Although user buy-in is a key requirement, basic technical requirements must also be addressed. Up to date hardware and high-speed network connections for accessing the ELN by laboratory personnel are important for user acceptance. Finally, the selected ELN solution should support the daily work of each user by simplifying their documentation processes and adding value primarily to the user. Vendors need to complement the heterogeneous workflows that are common in life science research, particularly by adding drag and drop functionality to streamline the usage of an ELN. Another option to support users would be an easier transition of paper notebook pages e.g., by printed ELN templates and a dedicated scanning and importing procedure with optical character recognition

(OCR). However, this would require well prepared templates and accurate recordings by the user.

Within the current PPP project, the selection process focused on project centric considerations such as:

- Legal requirements/compliance (Open Access)
- Sharing data between different sites/locations
- Search functionality across work packages
- Optional long-term availability of the data
- The relevance, accuracy, authenticity, and trustworthiness of electronic records.

User demands were considered of secondary importance, which clearly had a negative impact on usage and acceptance:

- Instant availability
- Easy to use and straightforward GUI
- Satisfactory performance
- Support for various data formats
- Drag and drop between any type of software.

More testing of heterogeneous documentation activities should occur prior to the final decision. Lack of long-term planning and selection phases in collaborative projects can prevent an intensive testing phase. For these projects, another approach could be the provision of individual guidance by trained staff during the initial go-live phase to overcome any inconvenience, but this needs appropriate funding for trainers.

The implementation of support functions (e.g., trainings, ELN-Helpdesk with defined roles and responsibilities) mitigated some striking issues related to user obstacles, but could not solve local issues like outdated hardware or insufficient support by local IT departments on installing the required plugins.

The value of the quarterly newsletter cannot be measured directly, but it provided a platform to reach and convey to the participants of the pertinence of to maintaining the ELN as a central tool for the project.

A summary of our findings can be found in the checklist Table 8.

The survey comment "It requires a new way of documentation, this is unusual at Universities" summarizes best what we need: a new positive thinking and willingness to share data in a scientific community. An ELN can support this effort, but an ELN will not automatically create motivation for sharing data.

ACKNOWLEDGEMENTS

(BioSciConsulting) for fruitful discussions and consultancy on technical details. Further, RM Twyman is acknowledged for text editing and guidance during the writing process.

ADDITIONAL INFORMATION AND DECLARATIONS

Competing Interests

Lea A.I. Vaas, Gesa Witt, Björn Windshügel, Philip Gribbon and Manfred Kohler are employees of Fraunhofer Institute for Molecular Biology and Applied Ecology - ScreeningPort, (IME-SP).
Adrian Bruengger is an employee of Basilea Pharmaceutica International AG.
Cindy J. Levy-Petelinkar is an employee of GlaxoSmithKline.

Author Contributions

- Lea A.I. Vaas conceived and designed the experiments, performed the experiments, analyzed the data, contributed reagents/materials/analysis tools, wrote the paper, prepared figures and/or tables, performed the computation work.
- Gesa Witt conceived and designed the experiments, performed the experiments, analyzed the data, contributed reagents/materials/analysis tools, wrote the paper, prepared figures and/or tables.
- Björn Windshügel conceived and designed the experiments, performed the experiments, contributed reagents/materials/analysis tools, wrote the paper, prepared figures and/or tables, reviewed drafts of the paper.
- Andrea Bosin and Cindy J. Levy-Petelinkar wrote the paper, reviewed drafts of the paper.
- Giovanni Serra and Adrian Bruengger reviewed drafts of the paper.
- Mathias Winterhalter conceived and designed the experiments, wrote the paper, reviewed drafts of the paper.
- Philip Gribbon prepared figures and/or tables, reviewed drafts of the paper.

- Manfred Kohler conceived and designed the experiments, performed the experiments, analyzed the data, wrote the paper, prepared figures and/or tables, performed the computation work.

Data Availability

The following information was supplied regarding data availability:

The raw data has been provided as Supplemental Files.

Supplemental Information

Supplemental information for this article can be found online at http://dx.doi.org/10.7717/peerj-cs.83#supplemental-information.

REFERENCES

European Commission Directorate-General for Research & Innovation 2013. Guidelines on open access to scientific publications and research data in Horizon 2020. Brussels: European Commission. *Available at https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-pilot-guide_en.pdf* (accessed on 12 September 2016).

Asif KH, Ahsan SM, Aslam M. 2011. Improving quality of experimentation procedure through electronic note book (ELN): a qualitative study. *Journal of Quality and Technology Management* **VII(II)**:83–90.

Bechhofer S, Buchan I, Deroure D, Missier P, Ainsworth J, Bhagat J, Goble C. 2013. Why linked data is not enough for scientists. *Future Generation Computer Systems* **29**:599–611 DOI 10.1016/j.future.2011.08.004.

Bennett JW, Chung K-T. 2001. Alexander Fleming and the discovery of penicillin. *Advances in Applied Microbiology* **49**:163–184 DOI 10.1016/S0065-2164(01)49013-7.

Berthold MR, Cebron N, Dill F, Gabriel TR, Kötter T, Meinl T, Wiswedel B. 2007. KNIME: the konstanz information miner. In: Bock H-H, Gaul W, Vichi M, Weihs C, eds. *Studies in classification, data analysis, and knowledge organization (GfKL 2007)*. Freiburg: Springer.

Bos N, Zimmerman A, Olson J, Yew J, Yerkie J, Dahl E, Olson G. 2007. From shared databases to communities of practice: a taxonomy of collaboratories. *Journal of Computer-Mediated Communication* **12**:652–672 DOI 10.1111/j.1083-6101.2007.00343.x.

DFG. 2013. *Proposals for safeguarding good scientific practice.* Weinheim: WILEY-VCH.

Du P, Kofman JA. 2007. Electronic laboratory notebooks in pharmaceutical R & D: on the road to maturity. *Journal of Laboratory Automation* **12**:157–165 DOI 10.1016/j.jala.2007.01.001.

Elliott MH. 2009. Thinking beyond ELN. *Scientific Computing* **26(6)**:6–10.

Elliott MH. 2010. What are the benefits of ELN? *Scientific Computing* **27(1)**:6–8.

Handbook: Quality Practices in Basic Biomedical Research. 2006. Geneva, Switzerland: special programme for research and training in tropical deseases (TDR). Von.

Available at http://www.who.int/tdr/publications/training-guideline-publications/ handbook-quality-practices-biomedical-research/en/.

Ioannidis JP, Greenland S, Hlatky MA, Khoury MJ, Macleod MR, Moher D, Schulz KF, Tibshirani R. 2014. Increasing value and reducing waste in research design, conduct, and analysis. *The Lancet* **383**(**9912**):166–175 DOI 10.1016/S0140-6736(13)62227-8.

Iyer R, Kudrle W. 2012. Implementation of an electronic lab notebook to integrate research and education in an undergraduate biotechnology program. *Technology Interface International Journal* **12**(**2**):5–12.

Kostyanev T, Bonton MJ, O'brien S, Steel H, Ross S, Francois B, Goosens H. 2016. The Innovative Medicines Initiative's New Drugs for Bad Bugs programme: European public–private partnerships for the development of new strategies to tackle antibiotic resistance. *Journal of Antimicrobial Chemotherapy* **71**:290–295 DOI 10.1093/jac/dkv339.

Macarelly A. 2014. Jump off the page. *Nature* **507**:523–525 DOI 10.1038/nj7493-523a.

Milsted AJ, Hale JR, Frey JG, Neylon C. 2013. LabTrove: a lightweight, web based, laboratory blog as a route towards a marked up record of work in a bioscience research laboratory. *PLoS ONE* **8**(**7**):e67460 DOI 10.1371/journal.pone.0067460.

Myers JD. 2014. Collaborative electronic notebooks as electronic records: design issues for the secure electronic laboratory notebook (ELN). *Available at* http://www.researchgate.net/publication/228705896_Collaborative_electronic_notebooks_as_electronic_records_Design_issues_for_the_secure_electronic_laboratory_notebook_%28eln%29 (accessed on 8 January 2015).

ND4BB. 2015. TRANSLOCATION. *Available at* http://translocation.eu/ (accessed on 22 September 2015).

Nehme A, Scoffin RA. 2006. Electronic laboratory notebooks. In: Ekins S, ed. *Computer applications in pharamceutical reserach and development (S. 209-227).* Hoboken: John Wiley & Sons, Inc.

Nussbeck SY, Weil P, Menzel J, Marzec B, Lorberg K, Schwappach B. 2014. The laboratory notebook in the 21st century: the electronic laboratory notebook would enhance good scientific practice and increase research productivity. *EMBO Reports* **15**(**6**):631–634 DOI 10.15252/embr.201338358.

Payne DJ, Miller LF, Findlay D, Anderson J, Marks L. 2015. Time for a change:addressing RandD and commercialization challenges for antibacterials. *Philosophical Transactions of the Royal Society B: Biological Sciences* **370**:20140086 DOI 10.1098/rstb.2014.0086.

Rex JH. 2014. ND4BB: addressing the antimicrobial resistance crisis. *Nature Reviews Microbiology* **12**:231–232 DOI 10.1038/nrmicro3245.

Rubacha M, Rattan AK, Hossel SC. 2011. A review of electronic laboratory notebooks available in the market today. *Journal of Laboratory Automation* **16**(**1**):90–98 DOI 10.1016/j.jala.2009.01.002.

Sandve GK, Nekrutenko A, Taylor J, Hovig E. 2013. Ten simple rules for reproducible computational research. *PLoS Computational Biology* **9**(**10**):e1003285 DOI 10.1371/journal.pcbi.1003285.

Sarich Z. 2013. ELN Incorporation into research. Digital Centers Internship Program. *Available at* https://blogs.cul.columbia.edu/dcip/2013/11/08/eln-incorporation-into-research/ (accessed on 3 August 2015).

Sarich Z. 2014a. Choosing an ELN. Digital Centers Internship Program. *Available at* https://blogs.cul.columbia.edu/dcip/2014/02/03/564/ (accessed on 3 August 2015).

Sarich Z. 2014b. ELN Presentation. Digital Centers Internship Program. *Available at* https://blogs.cul.columbia.edu/dcip/2014/05/19/eln-presentation/ (accessed on 3 August 2015).

Schnell S. 2015. Ten simple rules for a computaional biologist's laboratory notebook. *PLoS Computational Biology* **11**(9):e1004385 DOI 10.1371/journal.pcbi.1004385.

Stavenger RA, Winterhalter M. 2014. TRANSLOCATION project: how to get good drugs into bad bugs. *Science Translational Medicine* **6**:228ed7 DOI 10.1126/scitranslmed.3008605.

White TE, Dalrymple RL, Noble DW, O'hanlon JC, Zurek DB, Umbers KD. 2015. Reproducible research in the study of biological coloration. *Animal Behaviour* **106**:51–57 DOI 10.1016/j.anbehav.2015.05.007.

Woelfle M, Olliaro P, Todd MH. 2011. Open science is a research accelerator. *NATURE CHEMISTRY* **3**:745–648 DOI 10.1038/nchem.1149.

Zeng J, Hillman M, Arnold M. 2011. Impact of the implementation of a well-designed electronic laboratory notebook on bioanalytical laboratory function. *Bioanalysis* **3**(13):1501–1511 DOI 10.4155/bio.11.116.