Introducing an Electronic Laboratory Notebook to PhD Students Undertaking Chemistry Research at A Research Intensive University

Rosanne QUINNELL
School of Biological Sciences, The University of Sydney,
Sydney, NSW, 2006, Australia

and

D. Brynn HIBBERT
School of Chemistry, The University of New South Wales,
Sydney, NSW, 2052, Australia

ABSTRACT

The School of Chemistry at the University of New South Wales, Australia, is trialing an electronic laboratory notebook (ELN) to determine its suitability as a means to capture the research conducted by PhD students. ELN implementation required a partnership to be established between the Library (server space and digital curation), ICT services (technical expertise and IT management), researchers in School of Chemistry (end-users) and the creators of the ELN at Southampton University in the UK. Students who opted to participate in the trial were in the first year of their three-year PhD program, with students in later stages of their PhD (year two and onwards) seeing little value in investing in learning a new work process. The students participating in the trial use an average of five different types of instruments when conducting their experiments and this represents more than five different types of file format. We are now 6 months into the trial and can report that all student participants have found this particular ELN sufficiently flexible to accommodate their experimental work. We are yet to realise the capacity of the ELN as a means to publish data directly from the kinds of instruments used in chemistry research (e.g. mass spectrometers, nuclear magnetic resonance). However, once this has been achieved, we expect to introduce the ELN into undergraduate chemistry curricula.

Keywords: electronic laboratory notebooks, ICT in postgraduate chemistry research, Web 2.0 technologies in science practice.

1. INTRODUCTION

Information and communication technologies (ICT) have changed the way we practice as educators and researchers in the higher education sector. For several decades, the online environment had enabled flexible access to learning and teaching resources via the internet and our reliance on email for sustaining asynchronous collaboration is profound. The Web 2.0 environment promises enhanced means of collaboration and given that the practice of science is collaborative, more so than other disciplines [1], one would predict that the use of Web 2.0 tools in science practice would be high; however, this is not the case. Although scientists are technically savvy there are some important aspects of current science practice that seem to constrain the adoption of *e*Research[†] across the science discipline. The most accessible *e*Research tool for scientists is the collaborative electronic laboratory notebook (ELN) (Figure 1).

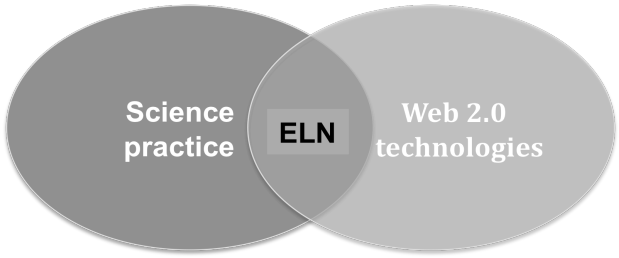

Figure 1. An electronic laboratory notebook is a blog that allows experimental data to be digitally annotated.

[†]The term *e*Research is used to describe research that exploits the Web 2.0 environment and advanced information and communication technologies.

Chemistry research practice

For laboratory-based science, such as chemistry, the details of each experiment are recorded in a paper-based laboratory notebook. Best practice requires a laboratory notebook to be date-stamped and signed at the end of each day and so, the laboratory notebook becomes a legal document that proves proof of when an experiment was conducted and by whom, i.e. the "provenance" of the experiment [2]. Keeping a laboratory notebook is a requirement of all science students undertaking research in a laboratory. As shown in Table 1, the laboratory notebook is the only part of our experimental practice that remains in hard-copy and is where details of how an hypothesis is being tested, i.e. the experimental design. Some data are recorded in the notebook as words and/or as annotated hand-drawings. The laboratory notebook is used when formally writing up the research as a thesis and/or publications. Replacing the laboratory notebook with an electronic version requires a major shift in both experimental practice and perceptions as to how well an ELN will be able to encapsulate an individual scientist's experimental design and associated datasets [2]. What a laboratory notebook cannot do is make electronic links from the experimental design to the data sets generated from scientific instruments. This is the key advantage afforded by ELN.

Drivers for change in the Higher Education sector

The move to a digital commons: In Australia, as in the UK and USA, there is a push to digitally archive (a data commons) all publicly funded research, including the original experimental records [3]. For this to be realised, it means a shift in the way research in science is currently practiced where the primary means of record keeping is paper-based laboratory notebook to keeping all records in a digital format.

We should be mindful of our current data management practice as, according to a recent survey of academics in Australia, 81% of survey respondents said that they store their research data on internal hard drives [4]. It would seem that that there is still some considerable distance to cover before we see a research data commons.

The research-teaching nexus: In the higher education sector there is a strong push to improve the linkages between our tertiary level teaching programs and research by adopting research-led, or research-informed, teaching methods and asking students to complete authentic assessments. Brew and Boud [5] assert that teaching and research share "a concern for learning" and critical to this is being able to introduce students the science discipline territory as termed by Becher and Trowler [6]. In Australia, PhD students undertake an intensive research program under supervision and are required to produce a thesis. In order to compete for employment on the international market, ideally students need to produce scholarly publications. For PhD students, the challenge is to complete their research program in a three-year timeframe and the ELN offers a way to manage all records related to this program in the one place.

Table 1. Alignment of the Scientific method with how is practiced and communicated.

Scientific method	Science practice	Science communication
Examine the literature posit an hypothesis	**Access literature** (*digital* .pdf) in the office/home.	**Introduction**
Test hypothesis experimentally	**Record** experimental design in *hard-copy* lab notebook in the lab, requires date stamp and sign-off.	**Materials and methods** (reproducibility**)**
Critique experimental findings (evidence)	**Prepare** samples, **measure** key characteristics of samples using instruments (*digital* data files) in the lab. **Data** 'crunched' to **generate** evidence in the office/home.	**Results**
Generate new ways of understanding the world (physical/biological /chemical)	**Discuss** findings with research team (students, supervisors, colleagues) and write paper in the office/home.	**Discussion**

There is an obligation on our part to align our research students' practices to the how scientific research will be done in the years to come. With the impending government requirement to have the data generated from publicly funded research housed in digital commons, it is judicious to trial ways of achieving a data repository that are sympathetic to discipline practices.

Introducing the electron laboratory notebook

The primary impetus for this trial was to offer a means by which students could keep their experimental records in a digital format rather than a paper-based laboratory notebook. The ELN offers a

way to archive data sets directly from their source, i.e. the instruments that generate the data, on an institutional server. The ELN offers other advantages over a paper laboratory notebook in that it provides:

- a way for data files to be annotated and linked to the context of the experiment (experimental design)
- a portal for sharing data and collaboration, i.e. between student and supervisor who may not be co-located
- a means of archiving experiments so that experiments can be easily revisited
- a place where the data files generated by scientific instruments used in experiments can reside.
- the experimental record templates in the Southampton ELN are created by each researcher lending the required level of adaptability for research purposes.

Here we seek to characterize both the benefits and sticking points of introducing *e*Research into a PhD program, focusing on an ELN in chemistry research. Do the advantages that ELN proffers translate in practice for research in chemistry?

2. APPROACH

In August 2009 we undertook a trial of an electronic laboratory notebook (ELN) in the School of Chemistry at UNSW. The ELN selected was developed at Southampton University, UK, (see [2, 7]) for use in a research environment and is "open-access" software grounded in the philosophy "open science" [8].

Project implementation
Prior to implementation of the ELN we identified stakeholders that included:

- ELN creators at Southampton, UK
- The Library: to provide server space and associated curation
- IT support: to manage the ELN at UNSW
- Researchers: PhD supervisors and their supervisors as participants in the trial.

The ELN trial
Staff and students in the School of Chemistry were surveyed at the start of the ELN trial as to their:

- main research area within chemistry
- the instruments they rely on to generate experimental data
- how they retrieve their data from these instruments and the software they use for data-handling
- student computer and internet access, including home access
- how they share experimental findings with their collaborators (e.g. for students this includes their supervisor)

3. RESULTS

We are 6 months into the pilot and generally speaking, those students who opted to participate in the trial were in the first year of their three year PhD program. Students in later stages of their PhD (year two and onwards) saw little value in investing in learning a new work process. From surveys and focus groups with our student participants (n = 9) we report that the users of the ELN have written templates tailored to the experiments they are performing. This tells us that the ELN format has been able to accommodate the kind of chemistry research being undertaken at UNSW and that the experimental design (the context) is linked to the datasets generated from those experiments.

On average, our PhD students are generating data from five instruments (e.g. mass spectrometers, nuclear magnetic resonance, UV-Vis spectrophotometers). Because of the nature of instrument software, the number of different file formats that each student is handling is in excess of this number. Most of the data generated from instruments used in experiments are in file formats that can only be accessed via proprietary software residing only on the computer attached to the instrument that generated the data. Effective workflow and archiving of experimental data files means that both the original data file (generated from the instrument) and the derived data file (in a more usable file format e.g. Excel or Origin) need to be archived. Currently students are sending their data files from the instrument to their computer using email.

All students participating in the trial are aware of the requirement to document their experimental work by way of a laboratory notebook. The rate for completing a standard paper-based laboratory notebook ranges from less than one book to over two laboratory books per year of study. This suggests that there are differences in how students currently use their laboratory notebook to record their data and/or that there are differences between

the types of research conducted.

Students are using the ELN as both a place to store and to annotate their dataset. Most of our participants are linking their records to the relevant published literature. We expect that the practice of linking experiments to the published literature will, in theory, make the process of writing up their work as a thesis and papers much easier for students.

The use of the ELN for collaboration is low. Currently we are devising systems to directly link instruments to the ELN so that data files can be available to users over the internet. We have identified deterrents to using the ELN for research in chemistry, which are:

- the ELN does not interface directly with the proprietary software such as instrument software generally and ChemDraw specifically
- having to login to the ELN to make minor entries

The survey indicated high-use instruments and we are discussing how to proceed with putting instruments on the intranet/internet so that data files can be accessed via the ELN and linked directly to the experimental protocol and the inferences from the data.

We are continuing to work with the creators of the ELN to improve functionality for research in chemistry. We will continue to monitor how students are using the ELN particularly with regard to the level and quality of collaboration and interaction with their supervisor(s).

4. CONCLUSIONS AND FURTHER WORK

We are working towards having high-use instruments 'pushing' data to the ELN. This is no trivial task; it requires 'good will' and resourcing at all levels of the institution to a sustainable system of having instruments online. Once instruments are blogging data the ELN we will be in a position to introduce the ELN into the undergraduate chemistry curriculum. Using an ELN as a data portal will allow undergraduate students, wherever they are located, access to a much more diverse range of experimental equipment. The potential of the ELN to address access issues in the science higher education sector is enormous, particularly as not all universities have invested in the same range of equipment.

5. REFERENCES

[1] A. Biglan, "Relationship between subject matter characteristics and the structure and output of university departments", **Journal of Applied Psychology** Vol. 57, 1973 pp. 204-213.

[2] m.c. schraefel, G.V. Hughes, H.R. Mills, G Smith, T.R. Payne, J. Frey "Breaking the book: translating the chemistry lab book into a pervasive computing lab environment", E Dykstra-Erickson, M. Tscheligi (Eds.): **Proceedings of the 2004 Conference on Human Factors in Computing Systems, CHI** 2004, Vienna, Austria, April 24 - 29, 2004 pp. 25-32. [http://eprints.ecs.soton.ac.uk/8780/1/schraefel_smarttea Chi04.pdf; accessed March 3, 2010]

[3] The Australian National Data Service Technical Working Group "Towards an Australian Data Commons", **Australian Government Dept Science, Education and Training,** October 2007, [https://www.pfc.org.au/pub/Main/Data/TowardstheAustr alianDataCommons.pdf; accessed April 20, 2010]

[4] L. Markauskaite, A. Aditomo, L. Hellmers, "Co-developing eResearch infrastructure: Technology-enhanced research practices, attitudes and requirements", **The University of Sydney and INTERSECT Australia Ltd**. [http://www.intersect.org.au/sites/default/files/eResearch Summary_01.pdf; accessed April 20, 2010]

[5] A. Brew and D. Boud, "Teaching and research: Establishing the vital link with learning", **Higher Education** Vol. 29, 1995 pp. 261 – 273.

[6] T. Becher and P. Trowler, **Academic Tribes and Territories: intellectual enquiry and the cultures of disciplines** (2nd edition), 2001, Buckingham: Open University Press/SRHE.

[7] Frey, Jeremy (2008) "Curation of laboratory experimental data as part of the overall data lifecycle", **International Journal of Digital Curation,** 3, (1) [http://eprints.soton.ac.uk/64612/; accessed March 3, 2010]

[8] C. Neylon, "Opennetware blogs." [http://blogs.openwetware.org/ accessed November 9, 2009].