UNIVERSITY OF SOUTHAMPTON

FACULTY OF NATURAL AND ENVIRONMENTAL SCIENCES

Chemistry

Asking the right questions: Designing digital tools to support researchers recording scientific experiments

by

Cerys Willoughby

Thesis for the degree of Doctor of Philosophy

April 2015

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF NATURAL AND ENVIRONMENTAL SCIENCES

Chemistry

Thesis for the degree of Doctor of Philosophy

ASKING THE RIGHT QUESTIONS: DESIGNING DIGITAL TOOLS TO SUPPORT RESEARCHERS RECORDING SCIENTIFIC EXPERIMENTS

Cerys Willoughby

The paper notebook is the trusty companion of the majority of those who engage in scientific research. In this digital era though, there is pressure to abandon paper in favour of the benefits of digital tools. For many researchers, particularly in academic environments, there are a number of barriers to adopting Electronic Laboratory Notebooks, not least of which are anxieties about their ease of use. Another concern is the impact using these tools may have on the quality of experiment records. In order to design digital tools that support researchers recording scientific experiments it is necessary to understand how researchers work with paper and digital tools, and the impacts different designs may have upon what they record. In this thesis, research is presented on the note-taking and experiment capture behaviour of a range of scientific researchers together with requirements for digital tools obtained from this research. A number of original studies are presented that explore how different techniques for eliciting information from the memories of researchers can be used to capture information of value to the experiment record, with both positive and negative consequences. This thesis also includes an overview of the development of a mobile ELN, designed for use by synthetic chemists, to explore the potential to bridge the gap between desktop ELNs and paper-notebooks. The conclusions of this thesis put forward suggestions as to how digital tools could be enhanced in the future to both support researchers recording scientific experiments and to facilitate the management of research projects.

Table of Contents

List of tables

List of figures

List of accompanying materials

Appendices are provided on the accompanying CD for this version of this thesis.

DECLARATION OF AUTHORSHIP

I, CERYS WILLOUGHBY

declare that this thesis and the work presented in it are my own and has been generated by me as the result of my own original research.

ASKING THE RIGHT QUESTIONS: DESIGNING DIGITAL TOOLS TO SUPPORT RESEARCHERS RECORDING SCIENTIFIC EXPERIMENTS

I confirm that:

1. This work was done wholly or mainly while in candidature for a research degree at this University;
2. Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
3. Where I have consulted the published work of others, this is always clearly attributed;
4. Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
5. I have acknowledged all main sources of help;
6. Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;
7. Parts of this work have been published as:

Willoughby, C., Bird, C. L., Coles, S. J., & Frey, J. G. (2014). Creating Context for the Experiment Record. User-Defined Metadata: Investigations into Metadata Usage in the LabTrove ELN. *Journal of Chemical Information and Modeling*, *54*(12), 3268–83. doi:10.1021/ci500469f

Willoughby, C., Bird, C.L., & Frey, J.G. User-Defined Metadata: Using Cues and Changing Perspectives. *International Journal of Digital Curation. 10*(1), 18-47.

Signed: ....

Date: ....

Acknowledgements

I would like to express my deepest thanks to my supervisor Jeremy Frey for both giving me the opportunity to do this research, and for his guidance, patience, and ever-present enthusiasm. I also express my deepest thanks to Colin Bird for introducing me to Jeremy and helping to arrange the opportunity, but also for his endless supply of ideas, encouragement, and friendship.

I must also thank the various people who have contributed to my studies and user research, especially students from the Chemistry Department at Southampton University. In particular I would like to thank Thomas Logothetis for his considerable assistance with running and grading the two studies as part of the Organic Chemistry Summer School; and Colin Flowers for including the viscosity experiment and metadata questionnaires as part of his first year practical course. I would also like to thank Simon Coles, for helping to set up the Bath user studies, providing me with access to student experiment records produced as part of the Advanced Crystallography experiment, and for the opportunity to develop Notelus. I would also like to thank Emma Tonkin for helping to devise and run the Bath user studies, and the students that took part in those studies. Thank you also to Susanne Coles for providing guidance and support whilst working on Notelus.

I have been fortunate enough to visit a number of locations around the world to engage in user research, and wherever I have gone have received a great welcome and hospitality. Thanks in particular go to Brynn Hibbert and his team, Adrian Plummer and Danmar Gloria at UNSW, Sydney; Mat Todd and his team, Paul Ylioja, Swapnil Anam, and Murray Robertson at the University of Sydney; Karl Mueller and his team, Kelly Murphy, Nikki LaFemina, and Eric Poweleit at Penn State University; and Mark Rutland from KTH and his family. Apologies to anyone I have forgotten, my memory is more fallible than most!

Thanks to Laura Cowen for being my remote teammate and providing me with guidance, especially in the beginning. Thank you to Phil for enabling me and for being the nagging voice in my ear saying “Not until you’ve finished”! Finally, thank you Eclipse for your companionship. I shall miss you, but you will always be with me in my memory and my heart.

Definitions and Abbreviations

Accessibility	In the context of this thesis, accessibility refers to the ease of access to information.
Affordance	An attribute of an object that allows people to know how to use it (Rogers, Sharp, & Preece, 2011).
Archiving	A curation activity which ensures that data is properly selected, stored, can be accessed and that its logical and physical integrity is maintained over time, including security and authenticity.[1]
Autobiographical memory	Another term for episodic memory.
Blank canvas effect	A situation when users want to perform an action or use a tool, but do not know where to start, typically because the tool lacks the appropriate affordances or assistance.
Cognitive interview	A well-established technique for interviewing witnesses using a series of memory retrieval and communication techniques designed to increase the amount of information that can be obtained from the interviewee (Memon et al., 2010).
Collaboratory	A system to support scientific collaboration across geographical boundaries by sharing access to instrumentation, data and computational resources (Bly, 1998).
Concept	A representation in memory of a specific entity or class of entities together with their properties, or the defining or prototypical features of the entity or class (Coleman, 2009).
Confirmation bias	The tendency to test one's beliefs or conjectures by seeking evidence to confirm them and ignoring

[1] www.jisc.ac.uk/publications/ publications/pub_escience.aspx

	evidence that contradicts them (Coleman, 2009).
Cue	In memory research, a stimulus that aids the recall of information not recalled spontaneously (Coleman, 2009).
Cued recall	The retrieval of information from memory with the help of cues (Coleman, 2009).
Curation	The activity of managing and promoting the use of data from its point of creation, to ensure it is fit for contemporary purpose, and available for discovery and re-use. For dynamic datasets this may mean continuous enrichment or updating to keep it fit for purpose.[2]
Curation @ Source	Selecting, preparing, and curating data items at the point the data is created (Frey et al., 2015).
Data management	The process of managing, selecting, retaining and storing information generated during a research project.
Digital Research Notebook	A generic digital notebook that can be used by multiple disciplines with data management, access control, and provenance inspection capabilities (Badiola et al., 2014).
DRN	Digital Research Notebook
Electronic Laboratory Notebook	An electronic replacement for paper notebooks that enables the storing and organization of data while maintaining the data entry flexibility and legal recording functions of paper notebooks; additional functions often include searchable indices, data sharing, and ease of duplication, but can also include extended functionality such as inventory control (Wright. 2009).
ELN	Electronic Laboratory Notebook

[2] www.jisc.ac.uk/publications/ publications/pub_escience.aspx

Enactment	Carrying out an activity, in particular carrying out the physical processes involved in an experiment.
Encoding	Converting information into a different form or representation, particularly the transformation of physical sensory information into a representation suitable for storage and retrieval from memory.
Episodic memory	A type of long-term memory for personal experiences and events (Coleman, 2009).
Explicit memory	Memory of previous experiences and learning that can be consciously recalled when needed.
Frame	A knowledge structure representing an everyday aspect of the world containing fixed structural information and slots for holding variable information (Coleman, 2009).
Free recall	Retrieval of information from memory without the help of cues (Coleman, 2009).
GIS	Geographical Information Systems
Implicit memory	Memory of previous experiences and learning that aid performance of a task without conscious remembering of the previous experiences. Sometimes known as procedural memory.
In silico	Experiments performed on a computer using methods such as simulations and modelling (Bird et al., 2013a).
Institutional memory	Historical documents, data, tacit and explicit knowledge, and experiences shared within a group (Elliot, 2004).
IP	Intellectual Property
Long-term memory	A type of memory containing information for periods ranging from 30 seconds to many decades.
Mental models	An internal representation of some aspect of the external world.

Metadata	Descriptive information and classification labels that group related items, provide context, and facilitate the reuse of data and other research outputs (Bird et al., 2013b)
Open Science	The open sharing of scientific knowledge supported by collaboration and interoperability (Bird & Frey, 2013).
OCSS	Organic Chemistry Summer School
Persona	Rich descriptions of typical users of a product under development (Rogers, Sharp, & Preece, 2011).
PI	Principal Investigator
Preservation	An archiving activity in which specific items of data are maintained over time so that they can continue to be accessed and understood through successive change and obsolescence of technologies.[3]
Protocol	A written scientific procedure for an experiment. The protocol is typically written during the plan stage of an experiment and then used as a reference during the experiment. A protocol may also be used and shared to standardise a laboratory method.
Provenance	Metadata about experiment processes, sources of data, processing and transformations applied, and details of other experimental components, for example, the scientists themselves and related literature. It is typically used to identify the originator of data, changes that have been made, and to track versions of experiment runs and associated data (Zhao et al., 2004).
Recall	The act or process of retrieving information from memory spontaneously with or without the use of cues (Coleman, 2009).

[3] www.jisc.ac.uk/publications/ publications/pub_escience.aspx

Scenario	An informal narrative description providing a story of user activities or tasks for exploring contexts, needs, and requirements (Rogers, Sharp, & Preece, 2011).
Schema	A mental representation based on prior experience and knowledge, structured in a way that facilitates perception, cognition, and inferring or interpreting new information (Coleman, 2009).
Scripts	A schema of an event sequence (Coleman, 2009).
Semantic memory	Memory for factual information excluding personal experiences.
Spreading-activation Theory	A theory of memory, where memory is structured as a network of nodes organised by semantic similarity; more connections exist between concepts with more properties in common (Collins and Loftus, 1975).
Storyboard	A sequence of illustrations or images that are used to help identify specifications for software or visualise important steps in the user experience.
Taxonomy	A taxonomy or taxonomic scheme is a classification of things or concepts, typically with a hierarchical structure.
Template	An electronic or paper document with a standard structure typically used to generate new pages with the same structure, style, or pattern.
User-defined metadata	Metadata that is consciously created by humans, as opposed to being automatically generated by a computer.

Chapter 1: Introduction

"The like frailties are to be found in the Memory; we often let many things slip away from us, which deserve to be retain'd, and of those which we treasure up, a great part is either frivolous or false; and if good, and substantial, either in tract of time obliterated, or at best so overwhelmed and buried under more frothy notions, that when there is need of them, they are in vain sought for."

Robert Hooke, Micrographia, 1665.

1.1 The scientific record

In science, communication is of the utmost importance. Ideas need to be shared, evidence disseminated, plans discussed, findings recorded, and errors corrected. An individual researcher may work alone, but their research is of little value to the scientific community if it remains inside their heads. Historically the printed page has been the centre of this scientific communication, but the centrality of scientific journals is being eroded with the arrival of digital technologies and the Internet in particular (Boulton et al., 2012).

Although the first recording of scientific endeavour is a matter for debate, scientists have been recording and sharing their observations and ideas for millennia (Bird et al 2013a). Through most of history, the primary medium for recording and sharing scientific discoveries has been on paper through the two principle information artefacts of paper notebook and scientific paper (Coppin & Hockema, 2009). Scientific papers, since their emergence in specialized journals in the seventeenth century, have been tools for the shared communication of scientific investigations, ideas, and discoveries, and also the means by which scientists gain recognition (Holmes, 1987). Paper notebooks serve as a personal diary for the recording of the daily research activities of researchers and other scientists (Shankar, 2006, 2009). The artefacts of scientific paper and paper notebook represent two ends of a process of 'inscription' (Holmes, 1987; Latour & Woolgar, 1986). Both artefacts can be considered to be part of the scientific record, together with other information artefacts produced during the process of inscription, as shown in Figure 1-1. These information artefacts include data from instruments and 'intermediate' documents produced during research activities, such as research reports, presentations, drafts of papers, and meeting minutes.

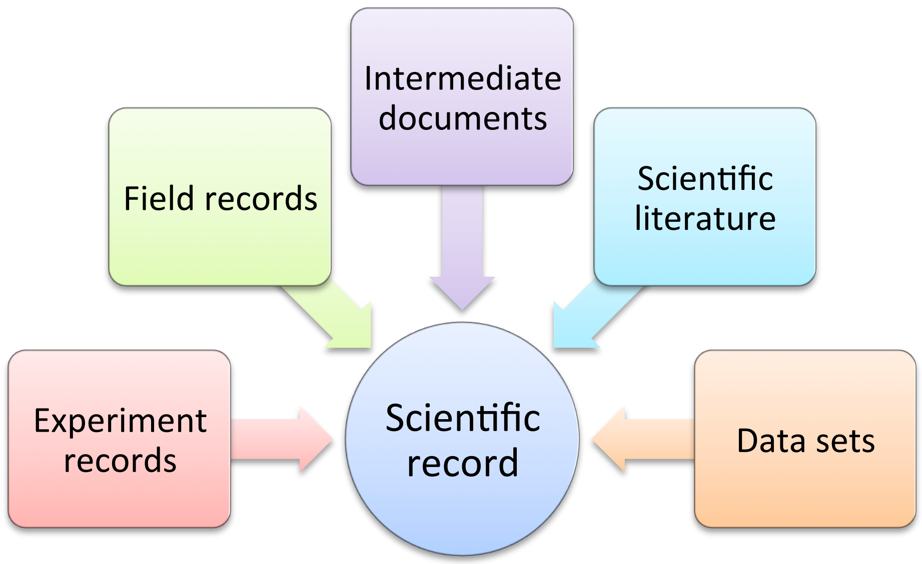


Figure 1-1: Elements contributing to the scientific record

Notebooks are written to be read by others in order to construct an 'institutional memory' for the immediate research group, university, or other organization, fulfilling functions as organizational documents that must fill expectations beyond those of the individual scientist who created them (Elliott, 2004; Shankar, 2009). In reality, however, most notebooks are shared with very few people and are usually viewed as a profoundly personal document representing a personal history of learning, expertise, and meaning (Shankar, 2009). This confused perception and role of the notebook can lead to both surprise and anger when researchers find out they are expected to surrender their notebooks to the institution when they move on (Shankar, 2009).

Notebooks that are not in current use can typically be found residing unused on shelves 'just in case', or even locked away in an archive (Tabard et al., 2008; Elliott, 2004). This is not always the case, and some notebooks may be passed over to a new researcher to inform their research. Even if notebooks are readily available and shared there are still significant problems involved in extracting information of value from these records. Content in the notebook may be illegible or difficult to read and records may be incomplete. Notebooks are difficult to search and may require considerable specialist knowledge to interpret, often they are impossible to understand without significant extra context. Often the only person that can decipher the content of the notebook is the person who wrote them (Elliot, 2004). The frustration of trying to understand the content of someone else's notebook is neatly summed up by the following paragraph from Christoph Hoffman's investigation of writing in scientific research (2013: p280):

> Yet almost everyone who examines research notebooks shares the experience of how incommensurable, fragmentary, crude, puzzling, and thus hard to conceive such material is. Apparently research notebooks do not at all permit a direct look into the researcher's mind or – through his or her eyes – at the course of events at the research site. Only when contextualized by additional sources and supplementary assumptions does the content of research notebooks become intelligible.

As a consequence knowledge gained from previous experiments can be effectively lost to the current generation of researchers and lead to unnecessary duplication of unproductive effort (Elliott, 2004). This is particularly true for failed experiments; although the lessons learnt from such activities are valuable for learning and protocol improvement, they are rarely published and therefore remain trapped within the notebooks recording them.

The content of notebooks and their associated data can be brought into sharp public attention during intellectual property disputes or when fraud is suspected (e.g. Levelt, 2012; Ince, 2011; Harding, 2005; McAllister, 2012). When the notebook does come to the forefront, it is vitally important that the content of the notebooks is 'good', for the protection of intellectual property and 'defending the research', but also for providing a record of the thoughts of the experimenter and detailed procedures of the experiments (Eisenberg 1982). The record must be usable for many years hence, and must not only allow the experiments and conditions within them to be replicated by others, but must also be sufficient for the scientist producing it to be able to understand it throughout their duration of their research (Shankar, 2009).

Paper notebooks are by nature blank and provide no guidance on what information should be recorded within them. Consistency in recording comes from standard guidelines for recording experiments used in both academia and formalised by regulatory organisations, for example the Good Laboratory Practice Handbook (GLP, 2009; Eisenberg, 1982). Many academic researchers are still keeping their paper notebooks in exactly the same way they were when they were taught decades ago (Carpenter, 2012).

1.2 The adoption and impact of technology

Technology is undoubtedly having a major impact on scientific research and teaching - shaping the way that experiments are performed, results captured, and findings disseminated. Computers enable a myriad of functions that benefit scientists, not least interactive computation, electronic communication, multimedia, and digital information management (Yeh et al., 2006). Within the lab most instruments are now computer controlled; computers are the main tools for capturing, analysing, and annotating data. Some experiments are even performed 'in silico' using computational methods such as simulations and modelling (Bird et al., 2013a).

The increasing dominance of technology, digitization of data, and the drive towards more openness in science brings with it a need to ensure that scientific records and their associated data are effectively managed and preserved; and where appropriate accessible to others. Two essential elements of adapting to this changing environment are the effective curation of digital records and the expanding use of Electronic Laboratory Notebooks (ELNs) (Bird et al., 2013a, 2013b). ELNs provide a range of features that can help to improve data management, data retrieval, and collaboration - providing positive benefits for teaching, learning, and research (Badiola et al., 2015).

Despite rapid technological change and a history for ELNs spanning more than three decades, there are still barriers to their adoption (Bird et al 2013a). ELNs are still relatively rare in academic environments with a variety of challenges to overcome (Rudolphi & Goossen, 2012; Goddard et al., 2009). The majority of lab and field scientists still use some kind of paper notebook or journal to record their activities. The many benefits and affordances of paper are given as a primary reason for a reluctance to change, even when the scientists are already heavy computer users (Mackay et al., 2002). ELNs in comparison are often perceived to be difficult to use, lacking in functionality, and unreliable as described in more detail in **Chapter 2: Digital tools for capturing the scientific record**.

Even if researchers are reluctant to adopt ELNs for their work, the proliferation of instruments producing data in a digital format means that researchers that are exclusively using paper notebooks have had to change their behaviour by printing out electronic data and pasting it into their notebooks or by including a reference to stored files (Carpenter, 2012). Other researchers are choosing to make use of tools other than ELNs for recording their notes electronically, such as using Word and Excel, or generic digital notebook such as Microsoft OneNote[4] and Evernote[5] (Hayes, 2012). These tools are typically more familiar or easier to use than ELNs, but they do not have common ELN functions that are important for capture and preservation, for example automatic timestamps, storing all changes made to a file, and capturing data provenance to enable authenticity (Carpenter, 2012).

1.3 Motivations and research aims

The primary motivations for this research sit within the wider aims of the Computational Systems Chemistry group at the University, which aims to make high quality chemistry data more available to the community for use and re-use. In order for chemistry data to be effectively used and re-used by the community, it must be both discoverable and trustworthy. This requires high quality context for the data ideally added by the researcher creating the data - but getting researchers to do this is often a very difficult task. To this end we seek to develop tools to facilitate capture of the research record, but also to understand how such tools might influence the quality of the research record.

The relatively low rates of adoption of ELNs within academic environments highlight that it is not sufficient just to provide tools such as ELNs and expect researchers to use them. In order for any software to be used and to be useful, it must be usable and meet the needs of the researchers that it is trying to support. With this in mind, the research presented in this thesis strives to put the user at the centre of the discussion about ELNs and other digital tools. If we wish to design an ELN that will be adopted by researchers we need to consider two primary questions:

---

4 http://www.onenote.com
5 http://evernote.com

Introduction

- How do we design tools that researchers will want to use?
- How do we ensure that the tools accurately capture the research record?

Researchers will only be willing to use tools that make their lives easier, and therefore those tools must enable the researchers to at the very least do what they can do now, without additional effort. Ideally any tool will also be able to provide researchers with additional functionality that is perceived to be of value. In the case of laboratory scientists, the tools we develop must enable them to achieve what they can already do with paper notebooks – and must also replace or integrate with the digital tools that they do already use.

Any tools we develop must also not have a negative impact on the primary purpose of the tool – to facilitate the capture of the scientific record. The records captured must be at least effective as the current methods used, but if we can encourage researchers to adopt our tools for capturing their research, we can not only provide them with the added value that will make their lives easier and their research better, but we can also ensure the capture of a high quality scientific record for the benefit of the researchers themselves, their colleagues and collaborators, and also for the future.

In order to be able to create such tools we need to understand how researchers work – what information do they currently capture, how do they capture this information, and how might our tools influence what they record? Additional information that may not normally be recorded using non-digital methods may also be useful for providing functionality that would add additional value for both the researchers and the community, for example, capturing metadata about the research would enable more effective curation of the record, enhance discoverability, and provide additional descriptive context. We therefore would benefit from an understanding of what metadata researchers currently use and how our tools might influence the capture of metadata.

In this thesis, research is presented on the note-taking and experiment capture behaviour of a range of scientific researchers together with a number of original studies that explore how different techniques for eliciting information from the memories of researchers can be used to inform the designs of interfaces to capture the experiment record. The primary aims of this research are:

- To understand the way that researchers currently keep and organise their work, and how they currently use and interact with ELNs and other tools, in order to identify requirements to contribute to the design of future ELNs and other digital tools.
- To investigate how different interface designs might impact on the information that is captured by an ELN system, in order to maximise the quality, consistency, and usefulness of experiment records and associated materials.
- To produce a proof of concept ELN on an iPad that highlights the potential of mobile platforms for providing a bridge between paper notebooks and desktop ELNs, in order to demonstrate added value of ELNs for the researcher by utilising the extended capabilities of mobile devices.

1.4 Contributions

The objective of this research is to contribute knowledge to the design of effective digital tools to support researchers recording their scientific experiments. Electronic Laboratory Notebooks and Digital Research Notebooks provide opportunities to not only capture research activities for future reuse and preservation, but also as tools to support researchers as they work, and to facilitate the capture of a quality scientific record. In order to serve these functions the tools need to be easy to use and meet the needs of both the researchers and future audiences of the records.

The contributions of this work are:

- Identifying current recordkeeping and curation behaviour by students and researchers, with and without the aid of digital tools.
- Identifying the requirements of digital tools for recordkeeping and curation, and the limitations of current ELN designs.
- Assess user behaviour and patterns of metadata usage within ELNs and on a variety of platforms that support the creation of user-defined metadata.
- Assess the impact of using different templates for the capture of experiment information.

- Assess the impact of different strategies for the generation of metadata across different information types and their applicability to experiment records.
- Design and implement a mobile ELN application to support synthetic chemists in the laboratory.
- Proposing future functions and designs for digital tools such as ELNs and DRNs.

1.5 Thesis structure

This thesis is divided into ten chapters (including this chapter) as follows:

Chapter 2: Digital tools for capturing the scientific record: An overview of the role of ELNs, barriers to their adoption, and the potential benefits of mobile ELNs. An overview is provided of the role of recordkeeping as memory and the usefulness of cognitive psychology in designing interfaces for capturing records. A brief overview is also provided of considerations in curation in the management of records and the need for tools for effective sharing and collaboration.

Chapter 3: Keeping and managing records: A review of how researchers keep and manage their records based on interviews and user studies with researchers in our communities. These findings are compared with similar activities from the literature to build up a view of the common recordkeeping practices and requirements in the scientific community, including the use of ELNs and other technologies.

Chapter 4: Interaction design and tools for keeping and managing records: This chapter extends the findings about the capture of experiment records from **Chapter 3: Keeping and managing records** to identify possible user roles, scenarios, and personas for helping to design tools to support researchers. The generated scenarios and personas are used to create a series of example storyboards that could be used to develop requirements for a social-networking style ELN or Digital Research Management platform, as further described in **Chapter 10: Human and computer interfaces for digital tools: The bigger picture**.

Chapter 5: Asking the right questions: Effects of Using Structured Templates for Recalling Chemistry Experiments An introduction to the research contribution investigating how the design of different templates impacts the information that students and researchers record after completing an experiment. This is a summary of the research contribution that will be submitted for publication as:

> Willoughby, C., & Frey, J.G. Effects of Using Structured Templates for Recalling Chemistry Experiments (In preparation, April 2015)

Chapter 6: Use of user-defined metadata in ELNs and other platforms: An introduction to the research contribution investigating patterns of metadata use within the LabTrove ELN[6] (Milsted et al. 2013) and other platforms that enables users to create their own user-defined metadata. This is a summary of the research contribution that was published as:

> Willoughby, C., Bird, C. L., Coles, S. J., & Frey, J. G. (2014). Creating Context for the Experiment Record. User-Defined Metadata: Investigations into Metadata Usage in the LabTrove ELN. *Journal of Chemical Information and Modeling*, *54*(12), 3268–83. doi:10.1021/ci500469f

Chapter 7: Generating metadata by using cues and changing perspectives: An introduction to the research contribution investigating the impacts of different strategies for generating user-defined metadata for different information types including photographs and structured information. This is a summary of the research contribution that was published as:

> Willoughby, C., Bird, C.L., & Frey, J.G. User-Defined Metadata: Using Cues and Changing Perspectives. *International Journal of Digital Curation. 10*(1), 18-47.

Some new research and analysis not published in the above paper is also included in this chapter.

---

[6] http://www.labtrove.org

Chapter 8: Mobile Electronic Notebooks: An introduction to mobile electronic notebooks and other solutions for capturing and using experiment information within the laboratory and for other scientific research.

Chapter 9: Notelus: An Exemplar Mobile ELN: A discussion of the design and development of a mobile ELN to help support and capture the experiment records of synthetic chemists in the laboratory.

Chapter 10: Human and computer interfaces for digital tools: The bigger picture: This thesis concludes with a summary of the research presented in this thesis and a discussion of the implications of the findings of the studies for designing the interfaces of tools for supporting researchers and the capture of the scientific record.

The papers related to Chapters 5, 6, and 7 can be found in the **Published papers** section at the end of this thesis, together with a list of contributions made to other papers during the PhD project discussed in this thesis.

Chapter 2: Digital tools for capturing the scientific record

2.1 The role of electronic laboratory notebooks

Electronic Laboratory Notebooks (ELNs) facilitate the digital capture of experiments. Early ELNs were initially seen by some as a revolutionary replacement of paper notebooks in the laboratory (Borman, 1994). Over time they have developed from tools primarily for the protection of IP and adherence to regulatory requirements, ensuring standard operating procedure compliance, and providing interfaces to instrumentation (Hice, 2009); to tools that provide extended functionalities such as interdisciplinary and multidisciplinary collaboration, improving data quality, and facilitating open science (Bird et al. 2013a; Taylor, 2011; Wilson, 2011; De Roure & Frey, 2007). An ELN can support the whole experiment lifecycle from planning to the execution of the experiment (Fakas et al., 2005; Hughes et al., 2004), providing long term storage and preservation of experiment records (Myers 2003), as well as storing raw data and results. ELNs can be considered to be an organizational ‘knowledge repository’ for supporting decision-making and providing the ability to learn from others, and to create a more efficient and collaborative atmosphere (Elliot, 2004).

Computers are now ubiquitous, and as research activities move on-line they are transitioning from personal to more collaborative (Tabard et al., 2008). Scientists are already using computers with standard office and analysis software as part of their research activities, often leading to a complex juggling act of working with and keeping track of both computer and paper-based resources. As a result scientists are under increasing pressure to use fully electronic notebooks (Tabard et al., 2008).

2.1.1 Barriers to adoption

As mentioned in **Chapter 1: Introduction**, there are a variety of challenges and barriers to the adoption of ELNs, particularly in academic environments, where their use is still relatively scarce (Rudolphi & Goossen, 2012; Goddard et al., 2009).

Capturing the scientific record

The primary barrier to ELN adoption appears to be the costs associated with commercially available systems such as licenses, database infrastructure, installation, and maintenance, leaving them out of reach of non-profit institutions and researchers that tend to be decentralized and independent (Bird et al., 2013a; Rudolphi & Goossen, 2012; Goddard et al., 2009).

Another concern of researchers is the lifespan of systems and their associated data formats. Although a paper-notebook can be read centuries after it was created, the data from a computer system may be unusable five years later - when the software and data formats have moved on (Mackay et al., 2002). ELNs have a reputation for using proprietary formats leading to potential loss of data in the future and a dependence on the supplier for customisation, fixes, updates, and data conversions (Coles et al., 2013; Rudolphi & Goossen, 2012). Other concerns about ELN adoption are issues of Intellectual Property protection, data security, and confidentiality with researchers expressing anxieties about who owns the IP of data entered into a system, together with a reluctance amongst many researchers to share information captured within their notebooks (Bird et al., 2013a). Although not an issue for all, legal and regulatory requirements are an important concern for some (Bird et al., 2013a; Coles et al., 2013; Carpenter, 2012).

Many scientists use computers extensively, yet the majority still choose to use paper notebooks, highlighting an appreciation for paper's simplicity and flexibility, and also highlighting that computer illiteracy or an unwillingness to use technology cannot fully explain resistance to ELNs (Tabard, 2008; Klokmose & Zander, 2010; Mackay et al., 2002). Perhaps the largest challenge for ELN providers to overcome is the perceived 'ease of use' of paper notebooks compared to electronic systems. Paper notebooks are easier to use, easier to read, easier to transport, inexpensive, and readily available. They 'turn on' instantly, have infinite battery life, are socially acceptable during meetings, require no training, and require minimal IT support (Bird, Willoughby & Frey 2013a; Mackay et al., 2002; Goddard et al., 2009; Yeh et al., 2006; Brandl et al., 2010). ELN software for taking notes on the other-hand is more difficult to use, takes more time, and is less flexible. Researchers also have anxieties about the stability and accessibility of ELNs in terms of availability both when and where the researcher needs it (Drake, 2007).

ELNs are also typically used on desktop or laptop computers. Desktop computers have the disadvantage that they may not be accessible in the locations or times that they are needed, and are rarely located at the bench where the researcher may want to capture notes or access information. Laptops have portability, but may still be difficult to use in the lab because of their size, a reliance on a keyboard for computer input, and they may not be permitted near chemicals or within cleanroom environments.

Despite ringing endorsements for paper, there are problems with paper; problems that can be solved through the use of ELNs, or at least using a digital tool of some kind. Although paper notebooks are inexpensive and portable, they are also a vulnerable storage medium, at risk of being damaged by chemicals and fire in the lab, or rain, mud, and wind in the field; they can also be misplaced or lost (Amorim et al., 2014; Bruce, 2002). Paper notebooks can pose a contamination risk in certain sensitive areas, either by transferring dangerous chemicals or radioactivity between locations or through problems caused by dust from the paper itself (Arnstein et al., 2002; Myers et al., 2001).

Paper notebooks are unsuitable for collaborative work making it difficult to share work with different teams at the same site, let alone when more and more research groups are working across different sites and countries. In contrast ELNs can facilitate access to research notes and materials at any time and in any location (Bruce, 2002; Elliott, 2004; Myers et al., 2001).

Paper is also time-consuming to write-up work in, difficult to create table of contents for, tedious to index, difficult to search, and expensive to archive (Bruce, 2002; Mackay et al., 2002; Myers et al., 2001). Paper materials can be retrospectively digitised and archived, but the process is extremely time-consuming, and simply scanning the notebooks does not enable you to make use of the material any more effectively (Bell & Gemmell, 2007). The coexistence of digital data and paper notebooks typically leads to physical cutting and pasting of printouts into paper notebooks or manual transcription of data - with at worse errors made in transcription and a loss of raw data - or at best records stored in different places with some kind of cross referencing system between them (Bruce, 2002; Myers et al., 2001). It is also difficult to incorporate other useful media types into a paper-notebook; for example, digital photographs of thin chromatography plates can provide an interpretation free version of the data that is more nuanced than a hand-drawn sketch (Badiola et al., 2015). These problems highlight the inefficiencies of managing the content of paper notebooks and incorporating and curating digital content; tasks that are easier, more efficient, and reliable within ELNs (Badiola et al., 2015). ELNs are also vastly superior to paper in terms of legal defensibility and provenance (Bird et al., 2013b; Myers et al., 2001).

A step on the way to ELNs is to use digital tools other than ELNs to replace the paper-notebook and to create the links between the record and the data. Some groups have seen benefits in using technologies such as Google docs, electronic notebooks on iPads, and Dropbox, in combination with specific lab software, for example for drawing molecules and analysing data. Using these tools can help to remove some of the perceived problems of ELNs such as costs and improved ease of use, but the fact that people are willing to engage with these technologies to solve their problems suggest that there is still an opportunity to create a good digital notebook with the flexibility and usability to meet the needs of researchers (Giles, 2012).

Some authors suggest that in order to encourage the adoption of ELNs it is necessary to develop tools that are not a revolutionary change from their current day-to-day experience and that maintain a close analogy with paper-notebooks (Lass, 2011; Myers, 2001). ELNs for mobile platforms perhaps have more potential to provide a closer experience to the paper notebook than desktop software can, and in turn may help to reduce some of the barriers of ELNs. One of the activities of this research project was to design and develop a prototype mobile ELN for the iPad that uses a ‘paper’ analogy for capturing notes and sketches, but also makes use of the extended capabilities of the tablet platform for experiment capture in the lab, as discussed in **Chapter 9: Notelus: An Exemplar Mobile ELN**. Examples of other mobile ELN projects from the literature and challenges for designing for mobile are discussed in **Chapter 8: Mobile Electronic Notebooks**.

2.2 Capturing the experiment record

The design of digital tools, both in terms of external usability and internal data structure, will have an impact on what information is captured and how effectively it can found and utilized by the content author and any intended audience. Any such tool needs the functionality to capture the experiment record, to management the experiment record, and to support sharing and collaboration. The consequences of an ineffective ELN design are likely to be a reluctance to adopt or use the tool and a failure to capture (or find) the required information for a quality experiment record. In order to create an effective design for ELNs is it necessary to understand the experiment process, the way that the scientists think about their experiments, and how different interfaces may affect their behaviour.

The following section provides an overview of the relationship between recordkeeping and memory, and suggests techniques that might be used within tool design to help capture the most appropriate information for the experiment record. The techniques described in this section have been used to develop studies to examine how information relevant to experiments can best be elicited from researchers so that interfaces can be designed to utilise the most effective approaches, as described in **Chapter 5: Asking the right questions: Effects of Using Structured Templates for Recalling Chemistry Experiments** and **Chapter 7: Generating metadata by using cues and changing perspectives**. **Chapter 3: Keeping and managing records:** provide a review of how researchers keep and manage their records based on interviews and user studies, and findings from similar activities in the literature. This chapter also gives insight into the functionality that a variety of researchers desire or expect from ELNs that can also contribute requirements to the design of tools to support researchers.

2.2.1 Recordkeeping as memory

Note-taking and recordkeeping can be thought of as acts of memory, allowing the author of the record to forget the specific details of an experiment or learning activity, but providing an 'aide-memoire' at a later time (Shankar, 2007, 2009; Yeo, 2014). Notebooks can be used to serve as prospective memory aids as a reminder to do something, for example, a guide to the steps in the procedure or protocol for an experiment, to keep an appointment, or make a note of an instrument setting. Notebooks can also be used as retrospective memory in order to help stimulate a memory of the original event (Hodges et al., 2006), for example to recall the circumstances of a particular experiment, action, or observation. In some situations the steps of an activity or important observations are only recorded after the event, because for example, a notebook cannot be taken into a lab, it is difficult to make notes at the time they are needed (e.g. working in the field at night), or because only significantly different details are recorded to save effort (Arnstein et el., 2002; Scholl & Van Laerhoven, 2014; Shankar, 2004). In these situations it is easy to lose important information because it is forgotten before it has had a chance to be recorded into the 'memory' of the notebook.

An ELN can be used as 'an external memory' of experiments or other research activities, but the accuracy of the record is only as good as the information that is entered into to it. If the experiment record excludes information because the author did not remember it, did not think it was important, or it was obvious to them - but unknown to a later reader (so-called 'tacit knowledge'), then the record may cease to be usable in the future when the original context has been lost. The information that a researcher enters into such a system is affected by the knowledge in their own memories and by the interface that is provided for them. It may be possible to encourage the capture of some information by explicitly asking for it, but it is also possible to inadvertently cause a loss of information by structuring interfaces in a particular way. In order to design interfaces that have a positive impact on the users of these systems and that do not cause unexpected problems, it is important to have an understanding of the way that knowledge is stored in the mind, memory processes, and techniques for stimulating recall.

2.2.2 The role of memory and knowledge in capturing the experiment record

The knowledge that each of us has is vital to our ability to be able to make sense of the world around us and to function within it. This knowledge not only relates to our understanding of our environment, but also who we are, and how we achieve our goals and carry out plans. Knowledge is stored in our memory and we access this information whenever we perceive our environment, plan an action, remember the past, or imagine the future.

Traditionally memory and knowledge structures have been investigated with lab based psychology experiments. These have more recently been supplemented with studies of patients with brain damage and memory disorders, together with brain imaging techniques used on both normal and brain-damaged patients. There has also be a drive to increase the number of experiments that are 'ecologically valid', in that they apply to how memory is used and operates in everyday life, for example, eyewitness memory.

Information stored in long-term memory is traditionally divided into two main types, implicit and explicit memory. Explicit memory is conscious memory that can be retrieved and 'known', including memory for semantic or general knowledge information and 'episodic' memory or memory for events. Implicit memory on the other hand includes procedural and perceptual memory, information that influences our behaviour, but that is not consciously recalled, for example perceptual information about familiar words and objects; and learning of repetitive skills such as riding a bike and reading. However, the evidence that these types of memory are actually separate systems with different ways of representing knowledge is controversial. Although many studies apparently show evidence in favour of the traditional separation of implicit and explicit memory (e.g. Schott et al., 2005; Spiers et al., 2001), other groups argue that these can be explained with a single memory system (Reder et al., 2009; Berry et al., 2006). These disagreements are characteristic of the difficulties explaining the internal functions and systems of the brain from human behaviour and cognitive problems alone.

Despite these problems, findings from psychological research studies are useful for providing clues as to how we might be able to better design interfaces to support researchers and to identify how aspects of human behaviour might cause problems for the capture of experiment records or interactions with our tools.

2.2.2.1 Knowledge Representation

Studies have shown that the knowledge we have has an impact on how we understand, interpret, memorise and recall information. The way that the knowledge is represented, acquired and stored in memory is important for understanding why these effects might occur. Knowledge structures are also important in understanding how the different types of information that are stored in memory might be connected and related. Understanding these knowledge structures is also useful for designing information structures in computing (such as metadata, semantics, taxonomies, and schemas) for capturing and describing experiment data in such a way that it is meaningful for the humans that must make use of them. The knowledge structures in memory determine how effectively humans can both process and remember information. If we present information in a way that is incongruent with these knowledge structures or unexpected then we can create confusion and undesired affects. In the context of technology this can lead to interfaces that are confusing, create avoidance behaviour, or that fail to generate the expected response from the user. For an ELN this could mean a failure for the user to engage with the system, to provide inaccurate information, or to fail to record important information, such as missing parts of the experiment record or adding inappropriate metadata.

2.2.2.2 General knowledge in memory (Semantic memory)

Semantic memory is the memory for general knowledge about the world such as knowledge of objects, facts, word meanings, and people, which do not relate to specific instances of time or place,. Semantic memory is thought to consist of ‘concepts’ about things in the world. A concept is a representation in memory of a specific entity or class of entities together with their properties. For example the concept ‘canary’ may have the properties ‘is yellow’ and ‘sings’. At a higher level (or superset) is the concept ‘bird’ that has the properties ‘lays eggs’, ‘has feathers’, and ‘can fly’. The canary is a child concept (or subset) of bird, along with starling, ostrich, and so on. At a higher level still is the concept ‘animal’ with properties ‘moves’, ‘breathes’, ‘has skin’ and ‘eats’. Evidence for the hierarchical organisation of concepts in the brain comes from a condition called semantic dementia, where the meaning of words break down systematically from precise referential meanings, through semantically related meanings, and finally to a total loss of meaning. Less familiar meanings break down first, for example patients forget the names of less familiar animals before more common ones (Hodges & Patterson, 2007). What these symptoms show is that there is a continuum of meaning in semantic memory at different levels of abstractness. At the most extreme level, information is disconnected from context, and at the other it is embedded in specific contexts relating to personal experience. As higher-knowledge structures start to break down, the context free levels of knowledge are compromised, and the meaning of words and objects presented outside of a physical context can no longer be identified (Funnell, 2001).

A theory known as the spreading-activation theory suggest that semantic memory is a network of concept nodes organised by semantic similarity, such that the more properties two concepts have in common the more connections exist between them (Collins and Loftus, 1975). This is similar to the structure of an ontology used to represent relationships between concepts in computing. When a concept is processed because it is seen, read about, heard about or thought about the node is activated. The activation acts like a signal that spreads out from the node along the paths of the network activating the semantically most closely related nodes along the network making it quicker to retrieve information about more closely related concepts than distantly related ones. The spreading activation term is also used in computing to describe a mechanism for information retrieval that functions in the same way using semantic networks (Crestani, 1997), leading to suggestions that using spreading activation techniques on ontologies could be used to support users of personal and task based information management systems (Katifori et al., 2010).

The structure of memory as a semantic network can also help to explain why certain information is more likely to be retrieved from memory than others. For example if a semantic network contains nodes for vehicles with related nodes: bus, fire-engine, car, truck and so on, and there is also a node for red, fire-engine will have a close connection to red, along with other nodes such as: orange, apples, sunsets, cherries and so on. If 'vehicle' is use as a cue for information retrieval, then the spreading from the vehicle node will activate fire-engine, car, truck, bus and so on. Activation of each vehicle related node will retrieve others that are semantically related, such as 'have fuel', 'have wheels' and so on. If 'red' is used as a cue, the activation that spreads to fire-engine will be unlikely to retrieve apples, cherries and sunsets, because they share few semantic similarities.

There is evidence from some brain-damaged patients and neuroimaging studies that suggest that we store perceptual and functional information about objects in different parts of the brain. Living things are usually associated with perceptual information, such as what does it look like, taste like, smell like, or feel like? Man-made objects tend to be associated with functional information, what is it for, how do you use it? There are many brain damage cases where patients have a problem recognising living things, with a smaller number of patients show the opposite problem, and are unable to recognise non-living things (Marques et al., 2008).

Different authors use different names for these knowledge structures (such as schemas, categories, frames, and mental models) and have different theories to explain their workings, but they are all roughly talking about the same conceptual model of memory (Kohls & Scheiter, 2008). In order to make sense of the world and to interpret new information, it is necessary to be able to relate the new information to existing information. Our mental representation of previous experience and knowledge enable us to have expectations and predictions about the world. These knowledge structures also have a large impact on how we understand, process, and retrieve information. If we are able to associate information with what we already know through the presence of context then we are more likely to be able to make sense of it and to remember it. If information fails to be understood then it is less likely to be remembered and less likely to be remembered correctly (Bergman and Roediger, 1999; Bransford and Johnson,1972).

The need for context and understanding has potential impacts in a number of areas. For example, if a student is performing a task without the understanding the context of the task, they may not only struggle with the task itself, but then also struggle to remember what they did because of the difficulty of generating their own retrieval strategy based on conceptual understanding. If students do not understand particular concepts, they may rationalise these and make erroneous interpretations. These could in turn distort their memories for what they did in the experiment, and what they thought happened in the experiment.

2.2.2.3 Changing perspectives

Studies have demonstrated that our expectations and previous knowledge can interfere with what we remember. Information considered as ‘obvious’ in terms of our knowledge may be omitted from what we record about experiments, but ‘unobvious’ or unexpected information would be more likely to be highlighted. We also tailor the information we retrieve and how we present it based on who our audience is (Tversky & Marsh, 2000; Bransford & Johnson, 1972), or if we are asked to take a different perspective when recalling information (Marsh & Tversky, 2004; Dudukovic, Marsh & Tversky, 2004; Anderson & Pichert, 1978).

Social context also appears to shape the way that memories are recalled and retold, with exaggeration used in recall to entertain and simplification to inform (Dudukovic, Marsh & Tversky, 2004).

The findings of these studies suggest that deliberately changing the perspective of the students and researchers at recall time could elicit different information, and therefore result in different information being recorded about the experiment. The information recalled could be more or less accurate depending upon the perspective adopted, or the specific instructions given to the students. If we ask researchers to record information for a particular audience, a corresponding change would be seen in the information recorded based upon the perceived knowledge of the specified audience. For example, if a researcher was asked to write up an experiment for their supervisor the record may be less detailed than if the record is written for 'someone who needs to repeat the experiment'.

The researchers may chose to omit details because they are making assumptions about the shared knowledge with the audience, or because they assume that a 'gist' perspective is adequate instead of a fully accurate account. They may also make the assumption that an accurate record is not required because the experiment failed. In fact students may claim that an experiment worked because that is the expectation of their teachers; as a result they may fail to record pertinent observations or engage in reasoning about why the results were different than expected. Students may write up an experiment as though it worked even though it didn't because they do not view the failure as part of the learning experience.

That the audience of the communication affects the way memory is communicated is potentially very important in the context of electronic lab notebooks. The perception of who the audience is, what knowledge they are expected to have, and the perceived function of the record may be very important when the recording information viewable by others. Humans are not very good at judging what knowledge others have and tend to assume that others have the same knowledge as themselves (Nickerson, 1999; Bromme et al., 2001). This is another reason why ensuring the capture of as much detail and context as possible in the experiment record is important.

2.2.2.4 Scripts

The activities that we engage with usually have a recognisable sequence or procedure, where one thing has to happen before another. For example, when making a cup of tea, you have to heat the water and add the tea bag to the mug before you pour on the water, and remove the tea bag. A special kind of schema, known as a script, provides a knowledge structure for handling these kinds of ordered events and procedures. Scripts deal with events and the consequences of events, and contain the information about the order of events that are usually expected to happen in a given situation. For example when you go to a restaurant you know that reading the menu comes before ordering the food or paying the bill. Scripts may be important for a number of everyday activities such as planning and execution of activities. For example achievement of a particular goal may choose the script most likely to fulfil the goal (I'm hungry selects "make a sandwich" script). Scripts also enable understanding of what we observe or read about someone performing an instance of a given activity.

It is obvious to see how scripts might be important for science experiments where there is a step-by-step procedure that has to be followed. Even without the individual detailed steps, there are activities you would expect to complete such as coming to the lab, putting on a lab coat, fetching equipment, measuring chemicals, making observations, cleaning up, and leaving the lab, with many other steps along the way.

Studies have shown that there is a surprising amount of agreement between subjects on the 'basic' action language used to describe familiar activities, such as visiting a dentist and eating in a restaurant, with very few unique actions described; subjects also remember information in the stereotypical temporal order (Barsalou and Sewell, 1985; Bower et al., 1979). Subjects found it more difficult to recall the correct order if actions were presented out of sequence and also falsely remember actions that were not presented but fit with their expectations of what should happen. Subjects also tended to forget information that was not perceived as relevant to the script unless it was particularly interesting (for example the waiter was naked). There is evidence that scripts and semantic knowledge may be related to each other (Cosentino et al., 2006; Anderson & Conway, 1993).

It is possible that there is a script for experiments, and this may impact on the types of information that are remembered and subsequently recorded. The presence of an experiment script could mean that actions the students expect to take are recorded, even if they didn't actually do them; if an observation or occurrence is not perceived as being interesting or relevant to the script, it may also be forgotten and fail to be recorded.

2.2.2.5 Generating memories

Episodic memory is the memory for personal events and experiences including details such as what, when, where and with whom. Episodic memory is what allows us to ‘travel back in time’ and re-experience something that happened in the past (Tulving 2002). Our memories are not a literal account of events and what we remember. Although we may have very vivid memories for significant or personally meaningful experiences, more mundane memories are representative and more general (Conway, 2009).

Some researchers believe that memories are linked to other systems in the brain related to the self and goal attainment, and that memories may stored or remembered depending on whether they relate to goal attainment (Conway and Pleydell-Pearce, 2000; Conway, 2009). An event such as carrying out an experiment in a lab may be better or more accurately recalled if the event is related to some goal of the individual. For example, an experiment may be better remembered by students if they plan to continue study in the subject or work in that field, or by researchers if it is significant to the core of their research. Memory may also be affected based on the perceived success or failure of the experiment. If the experiment failed and the researchers did not see the results they were expecting this may change what information they recall or what they choose to report.

2.2.2.6 Remembering

In the past it was believed that the amount of practice or rehearsal was the primary factor that influenced whether information passed into long-term memory, and therefore whether the information would be remembered. The reality is much more complicated than that – there are no definitive laws of memory (Roediger, 2008). Prior knowledge, conditions at encoding (when the memory was generated), influences after encoding, and conditions at retrieval, all can have an effect on performance at recall time.

2.2.2.7 Free recall

Free recall is believed to allow information to be recalled in the form closest to the structure that it is actually held in memory. Studies have found that under free recall memories are recalled based on their distinctive detail or temporal order, suggesting that these memories are organised both temporally and thematically (Anderson & Conway, 1993).

Making use of these temporal and distinctive detail elements may help improve better recall of what happened during the experiment. For capturing the experiment record we can ask for the information in temporal order, what happened (or what did you do) to capture information about the procedure and elicit script related actions. Subsequently questions can be asked related to the significance of the events. What were the interesting observations, what worked or did not work as expected, how did it meet expectations, or how did it meet the goals?

2.2.2.8 Cued recall

Recollection of an event depends upon the information stored in memory about the event, but also due to the information available at the point of retrieval in the form of cues. Without the right cues you might have trouble remembering something, but given the right cue you are much more likely to be able to recall the information you are trying to remember. Cues act as a prompt for memory recall. The questions ‘who’, ‘what’, ‘where’ and ‘when’ can make effective cues, although they are not all equally as effective; the cueing order also makes a difference, with ‘What’ on its own most effective, followed by ‘where’, ‘who’ and ‘when’ (Wagenaar, 1986).

Cues can only be effective if the appropriate semantic links exist in memory, but of course an inappropriate cue is unlikely to retrieve relevant information or memories, and may instead cue the recall or retrieval of irrelevant information.

Various studies have shown that cues that are ‘self-generated’ are more effective for recall than those generated by someone else. These cues may be individual words to describe a set of information (Sharifian, 2001), or through the use of drawing diagrams or plans that can be used to recall specific details of a related event Dando et al., 2009; Brewer & Treyens, 1981). This finding suggests that self-generating metadata to describe an experiment, and drawing observations or a flow chart of the experiment procedure may help to improve memory and recall for what was seen and done for a specific experiment.

2.2.2.9 Everyday memory and knowledge

Memory errors in the real world can have real consequences. Eyewitness testimony is one of the main areas of memory and recall research outside of the psychology lab. Techniques for improving memory retrieval derived from cognitive studies and research such as those detailed above have been utilised in the collection of eyewitness testimony through the cognitive interview. The cognitive interview focuses on more than just recall; it is also designed to help minimise memory errors that might lead a police investigation in the wrong direction or lead to innocent people being found guilty. Eyewitness testimony can be subject to a number of different memory errors in addition to forgetting. Not only can witnesses have incomplete memory, but their memory can be misled and influenced by their own biases, and even the questions they are asked. For example, when subjects in a classic study, by Loftus and Palmer (1974), were asked to estimate the speed two cars crashed in a video, the estimated speed was significantly higher when the word ‘smashed’ was used in the question compared to less suggestive words such as ‘contacted’ and ‘bumped’. Not only did this affect their original answer but also their longer term memory of the event, with the use of the word “smashed” leading to significantly higher estimates of speed, and to many more subjects claiming to have seen broken glass a week later, when in fact there was none.

Other studies have shown that eyewitness memory can be influenced by post-event information and even be affected by information prior to observing an event, with interference from earlier memories where prior experiences or knowledge are thematically similar. Memories created by misleading information may be very similar to those produced by the actual event, making it difficult for the witness to distinguish between them (Allen and Lindsay, 1998). For memories of completing experiments, memories of previous or subsequently conducted experiments could interfere with memory for a specific experiment, and therefore cause inaccurate information to be recorded or to make it difficult to relocate the details.

As mentioned before, our previous experiences and knowledge influence what we understand and remember. Confirmation bias and social bias are commonly seen in eyewitness testimony where witnesses remember and report what fits with their knowledge and perceptions of crimes and criminals, rather than what actually happened. Information that matches our expectations is easily remembered overtime, but information that is inconsistent with expectations or perceived as irrelevant is more rapidly forgotten (Tuckey and Brewer, 2003a; Stangor & McMillan, 1992). There is also evidence that information that is inconsistent with expectations can become modified to fit the expectations, as a result these changes will be remembered as though they really happened (Tuckey and Brewer, 2003b). This again raises the question whether researchers recording experiments are recording what really happened, or whether they are recording what they think should have happened.

2.2.2.10 The cognitive interview

Despite the importance of complete and accurate eyewitness testimony the traditional police interview has been found to be relatively poor at generating correct information compared to the cognitive interview, which generates significantly more information and introduces less incorrect information (Memon et al., 2010; Fisher et al., 1989; Geiselman et al., 1985). Techniques used in the cognitive interview includes context reinstatement, 'report everything', recall from a variety of perspectives and recall in a different temporal order. In context reinstatement the witness is encouraged to mentally reconstruct the situation from the event, what they could see, who was there, how they felt and so on. Asking the witness to report everything they can remember, even if it the memory is only partial or incomplete, follows this. Next the witnesses are asked to adopt the perspective of other people at the scene, what might those people have seen or felt? The final stage encourages the witness to recall events in a different temporal order, for example running the event backwards, or starting from a point in the middle that has relevance for the witness.

A framework for effectively communicating with witnesses was also introduced with the enhanced cognitive interview to prevent biasing the witness or providing them with misleading information and ensuring they have the maximum chance to recall and communicate their testimony (Memon et al., 2010).

Unfortunately the cognitive interview has provided to be quite demanding and application of the procedure by the police after training is low (Dando et al., 2009). However a simplified cognitive interview has been devised, that is less demanding and shorter, makes use of a free recall and sketch reinstatement of context where the witness is asked to draw a sketch or plan of the event, followed by the witness using the plan to explain what happened. This enables the witness to generate their own retrieval cues, another retrieval technique derived from the psychology studies discussed earlier.

Running a cognitive interview protocol for experiment recording would be excessive, but the methods used are based on the memory retrieval theories described in the previous sections. Elements of the modified cognitive interview such as the reinstatement of context through the use of diagrams, a reminder to report everything, changing the perspective and recalling information in the correct temporal order may all assist recall in the lab.

Similar techniques based on the cognitive interview are used in survey and questionnaire designs to elicit effective recall from students and members of the public (Desimone & Le Floch, 2004; Schwarz, 2007; Holleman & Murre, 2008; Tourangeau et al., 2000; Jabine et al., 1984).

2.2.3 Templates for recording experiments

Using templates or other structured interfaces provides the opportunity to make use of the various techniques for eliciting information from users described in the previous section. These techniques encourage the recall of information that will be of use to the author and others who seek to make use of the information in the future. Templates in ELNs could provide cues that might act as reminders to increase the completeness of information recorded and encourage authors to enter more information. Templates could, for example, include sections to encourage students to record important information that they might otherwise forget to record, or to capture information that will be helpful when they come to look up the experiment later. The majority of ELNs make use of formal structures for recording experiments, but others, such as the LabTrove ELN developed at Southampton University, do not mandate any specific structure and enable users to create their own templates if they want to (Milsted et al., 2013).

However, there is evidence that the inflexibility and formality of templates might restrict the content that the author enters (McGlade et al., 2012) and that important information is not recorded if the template does not specifically ask for it (Swinglehurst, Greenhalgh & Roberts, 2012). In theory the researcher has ultimate authority over what they record and can therefore choose to improvise if the system does not allow them to capture everything they want (Shankar, 2007). But in reality the researcher may not realise that important information is not requested because the design of the system is not providing any cue or other reminder for it. Using templates might lead to the loss of important experimental information or details of the personal experience of conducting the experiment, such as observations or knowledge that was learnt as part of the experiment process.

2.3 Managing the experiment record

Once the experiment record has been created it needs to be effectively managed. The captured information needs to be preserved, but to make the best use of the information it also needs to be retrievable and reusable. Digital curation is the term used for the actions involved in the process of preserving, maintaining, and adding value to these digital assets over their entire lifecycle (Beagrie, 2006). The ingredients of curation are digital objects in the form of data and information, together with descriptive and other metadata that can facilitate efficient and effective access (Bird et al., 2013b). One of the most important elements that metadata can provide for experiments is context, without which digitally captured data in particular becomes meaningless (Borgman, 2008; Frey, 2008). Data is 'connected' to other information produced by researchers and without access to knowledge about the experiment or other activity that generated it you cannot make sense of the data (Fear, 2011). Without formal curation as part of managing scientific records there is a loss of data that would have been useful if it had been properly preserved (Downing et al., 2008).

Although some metadata can be created automatically, some tasks require 'human-generated' or 'user-defined' metadata to capture the full context (Currier et al., 2004; Greenberg and Robertson, 2002). Curation is often thought to be a task for librarians and information specialists, but the researchers themselves are best placed to provide the context for the experiment and therefore generate the experiment metadata (Bird et al., 2013b; Frey, 2008). Previous studies of researcher behaviour and attitudes have identified the so-called 'burden of curation' that leads to difficulties encouraging researchers to generate appropriate metadata for their data (Borgman, 2007, 2008; Crystal and Greenberg, 2005; Ryan and Walmsley, 2003; Frey, 2008). Content creators see metadata creation as extra work, owing to limited understanding of the rationale and value of adding metadata, combined with a lack of incentives for creating it (Greenberg, 2004; Currier et al., 2004). In addition to a lack of incentive for data management comes a reluctance to share research data in the first place, particularly by chemists whose data has high value to industrial funders (Borgman, 2010).

The expanding use of electronic laboratory notebooks (ELNs) provides an opportunity to assist the creators of scientific records by providing a mechanism for the capture and creation of metadata (Bird et al., 2013b; Milsted et al 2013). Providing this mechanism designs curation into the experiment process and encourages 'curation at source', making metadata creation more effective, efficient, and less error-prone (Frey, 2008). However, just providing the capability to add metadata within experiment records is not sufficient to ensure that researchers will add it or that it will be useful for curation.

Although researchers are subject matter experts, they do not have the skills of information professionals and most will not have specialized metadata creation expertise (Crystal & Greenberg 2005). Research suggests that relatively few students actually know what metadata is or how to use it (Mitchell, 2009) and knowledge of data management and curation is uncommon among researchers across all disciplines (Whitton & Takeda 2011; Whitton 2011; Ward et al., 2011). The metadata terms that are chosen and the volume of metadata added will have a significant impact on the retrievability and usefulness of the notebook data in the future (Bell & Gemmell, 2007). Quality issues are often seen in projects involving user-generated metadata, with problems including misspellings, inconsistent use of terminology, and choice of properties to describe, demonstrating that resource creators often do not understand the purpose of metadata or have an appreciation of its value (Currier et al 2004; Ryan & Walmsley 2003, Pancerella et al 2003).

Metadata can be automatically generated or 'machine-generated'. Examples of metadata that are commonly machine-generated include timestamps, locations, authors and user identifiers, and information derived from files such as EXIF information from photographs, and instrument settings from data files. However, information that is most valuable for providing context is likely to be most effectively generated by humans, in particular the creator of the original record. Tools for capturing and making use of metadata therefore need to encourage researchers to add metadata at the point of creation of the experiment record. These tools also need to support the researchers to help minimise the effort involved in creation and to minimise the number of errors and inconsistencies made.

2.4 Sharing and collaboration

Collaboration and sharing is fundamental to science (Bird et al., 2013a). In order to make use of data and to be able to reproduce results, it is necessary to be able to readily access both the data and the context provided by the experiment record. Digital technologies including the digitisation of instruments, digital storage, data repositories, and the Internet have led to significant changes in the way that researchers work. Communications have become more open and less formal; collaboration has become more common and extensive; research has become faster and more global; and projects have become larger and more multidisciplinary (Bird et al., 2013a; Tabard et al., 2008). As a result more data is being preserved and utilised - typically in open, well-established domain and shared repositories, and global databases (Bird et al., 2013a; Amorim et al., 2014; Tabard et al., 2008).

These repositories and databases typically store data rather than the associated 'knowledge' and background that is frequently captured in a personal notebook but rarely shared (Tabard et al., 2008). Data is also typically only deposited in these data repositories at the point of publication, and therefore not available during the collaborative data production phase. This lack of availability means that that researchers have to find other tools to act as a collaborative environment for sharing such as email, Dropbox, or Google Drive (Amorim et al., 2014). Some examples of ‘know-how’ are shared through protocol and workflow sharing services such as myExperiment[7] (De Roure et al., 2009) and open OpenWetWare[8].

Whether or not researchers are willing to share research data is a complex interplay of issues including (Borgman, 2012):

- ethics,
- nature and format of the data itself,
- economics and intellectual property,
- expertise,
- incentives and rewards,
- public policy,
- regulatory requirements,
- and funding body pressure.

Discussions about whether or not data should be shared and how is beyond the scope of this thesis, but these discussions highlight the need to provide granularity in access control to balance the sometimes conflicting needs of data protection and open collaboration.

Beyond the well-known issues of data sharing and open science, there are other barriers that may discourage researchers from participating in online information exchange, in particular the time and effort needed to contribute and a fear of personal feedback (Matschke et al., 2014). However, well-designed tools can successfully support virtual collaborations and relationships between individuals and research groups (Chin et al., 2002). The lack of physical and temporal proximity does mean that the materials and projects in such collaboratories need to be appropriately organised and provide additional context for collaborative materials and projects for researchers to be able to make effective use and understand the relationships between them (Chin & Lansing, 2004).

---

[7] http://www.myexperiment.org
[8] http://openwetware.org

Chapter 3: Keeping and managing records

3.1 Introduction

This chapter provides a review, based on interviews and user studies with researchers in our community, of how researchers keep and manage their records. These findings are compared with similar activities from the literature to build up a view of common recordkeeping behaviours and attitudes to ELNs amongst chemists and other researchers. These findings also inform requirements for the design and development of tools such as ELNs.

The findings included in this chapter were derived from a variety of semi-structured interviews completed with members of the LabTrove ELN community and other groups interested in electronic notebook adoption. A semi-structured interview is an open technique for conducting interviews that uses some structured questions to initiate discussion, but enables new ideas to be discussed based on the responses of the interviewees. Details of the different groups included in this research and the activities undertaken can be found in Table 3-1. Information is also included from direct examination of a number of notebooks created by physical chemistry researchers and students at Southampton University. The results of the formal studies carried out as part of this thesis are discussed in the following chapters.

Researchers make notes in a wide range of locations and environments, and scientific records are captured by both field and laboratory scientists. Some researchers conduct their research in both of these environments, for example, archaeologists, biologists, environmental, and earth scientists. Lab scientists themselves may be conducting activities in different locations within the lab; in different rooms, often moving between the lab and the office, or when different instruments are located in different rooms; and even different buildings (Tabard et al., 2012; Klokmose & Zander, 2010; Arnstein et al., 2002). Although the focus of this thesis is primarily on laboratory scientists, it is appropriate to consider the activities and needs of a variety of scientists to understand their similarities and differences.

Table 3-1: User research activities

Communities	User research activities
Analytical Chemistry; University of New South Wales, Australia	Semi-structured interviews with PI, research students, and IT support staff on ELN use
Organic Chemistry; University of Sydney, Australia	Semi-structured interview with group of research students on ELN use
Organic Chemistry; University of Bath, UK	User testing with several groups of researchers and semi-structured interviews on ELN adoption
Molecular Informatics and IT in Chemistry; Cambridge University, UK	Semi-structured interviews with staff and research students on ELN use
Archaeology; University of Southampton; UK	Semi-structured interviews with staff and research student on note-taking behaviour and technology use
Synthetic Chemistry; University of Southampton; UK	Semi-structured interviews with research students on note-taking behaviour, daily routine, and technology use
Environmental Chemistry; Penn State University; USA	Semi-structured interviews with staff and research students on ELN use
Chemistry; KTH Royal Institute Of Technology, Sweden	Note-taking workshop – activity on note-taking practices and attitudes, and experiment used for a study

The needs and behaviours of these different groups are not substantially different at a high-level; although the nature of the data, the structure of the records, and terminology used differs between domains. The structure of notes differs depending on the purpose of the notes. For example, a laboratory or formal field notebook differs in structure from an informal notebook used for personal 'notes to self' (Van Kleek et al., 2009).

3.2 The recordkeeping process

The majority of researchers go through the same or similar basic processes in the construction of their records as illustrated by the examples provided in Table 3-2.

Table 3-2: High-level processes in creating the experiment record

Discipline & source	High-level processes & descriptions
Synthetic biology (Tabard et al., 2012)	**Design:** Sketching & designing a structure, defining the procedure (materials & protocols to use). **Construction:** Mixing, heating, cooling, data recording, changes to protocol. **Analysis:** Visual image (gels), quantitative, statistical
Cell biology (Arnstein et al., 2002)	**Preparation:** Creation of working document (experiment protocol & observation sheet) **Execution:** Combination, incubation, dispensing, separation, detection (recording & data capture), storage and retrieval **Documentation:** Formal record creation, transcription from working document, pasting printed results, adding references
Life sciences (Scholl & van Laerhoven, 2014)	**Experiment:** Combination, Dispensing, Separation, Incubation, Detection, Labelling and Identification.
Chemists (Interview with researchers & staff)	**Planning:** What to do, materials and safety **Experiment:** Alter method to what did do, Observations. Yields. Deviations from plan **Analysis**
Environmental chemistry (Interviews with researchers)	**Planning:** Literature, reports, talking to collaborators; Figure out protocol, materials, and tool requirements **Experiment:** Preparation, washing, melting, milling, washing, leaching, filtering, and drying. **Analysis**: instrument capture, processing, report generation

Synthetic Chemists (Interviews with researchers)	**Planning:** Literature, reaction scheme, materials and safety, record planned procedure **Experiment:** Record actual procedure, measurements, times, and observations **Analysis:** Processing, record analysis of results

In the lifecycle of the scientific record, individual experiments are used to refine methods and published in some form such as group presentations, quarterly reports, and papers; and then finally preserved or archived in some form. Figure 3-1 shows a visualisation of this process as a circular flow, because the data preserved from the experiment contributes to the design of future experiments. This can be seen in practice from the interviews with the chemistry researchers, who make use of online search and repositories to locate literature and data for developing their protocols; and also for comparison for their results to verify the success of their experiments. The activities that are performed by researchers on a day-to-day basis are focused on the beginning of the cycle, with preservation rarely mentioned as a consideration for most researchers when thinking about their research.

Figure 3-1: The recordkeeping process

Planning involves pulling information from other locations, typically literature and data from repositories and other sources on the Internet. This activity is usually performed at a computer in the office. Some researchers also communicate with collaborators at the planning stage either locally or remotely. Once information has been gathered and evaluated it is used to put together protocols for the experiment and also details of the materials to be used.

Capture of information is associated with the ‘enactment’ of the experiment, where protocols are followed and materials manipulated. Observations, measurements, and the details of instrument or equipment settings or configurations are captured. On the fly changes to the protocols may be made as the experiment progresses depending on the observations made, and these changes are recorded. Data may also be captured through the use of a variety of instruments, and this information may be recorded through notes or via the generation of digital files.

Analysis of the observations, measurements, and results data is performed, usually using some additional software support. Raw data is processed and evaluated, and often compared to data generated previously, either by the researcher, collaborators, or from the data repositories on the Internet.

Refining involves the preparation of materials for sharing or publication. For researchers this may involve the preparation of presentations and reports for members of the research team and collaborators, and scientific papers for the wider scientific community. The refining stage is where much of the creative thought and development of ideas occurs.

Publication and preservation are the later stages of the process where the record and the final products of the research are deposited for access to the wider scientific community, or alternatively locked away in a filing cabinet and forgotten.

The following cartoons provide some examples from our user research activities of how different researchers construct their experiment records and organise their work, including this element of reusing data from other scientists in their activities. The researchers are using differing combinations of paper-based and computer-based methods for capturing their records.

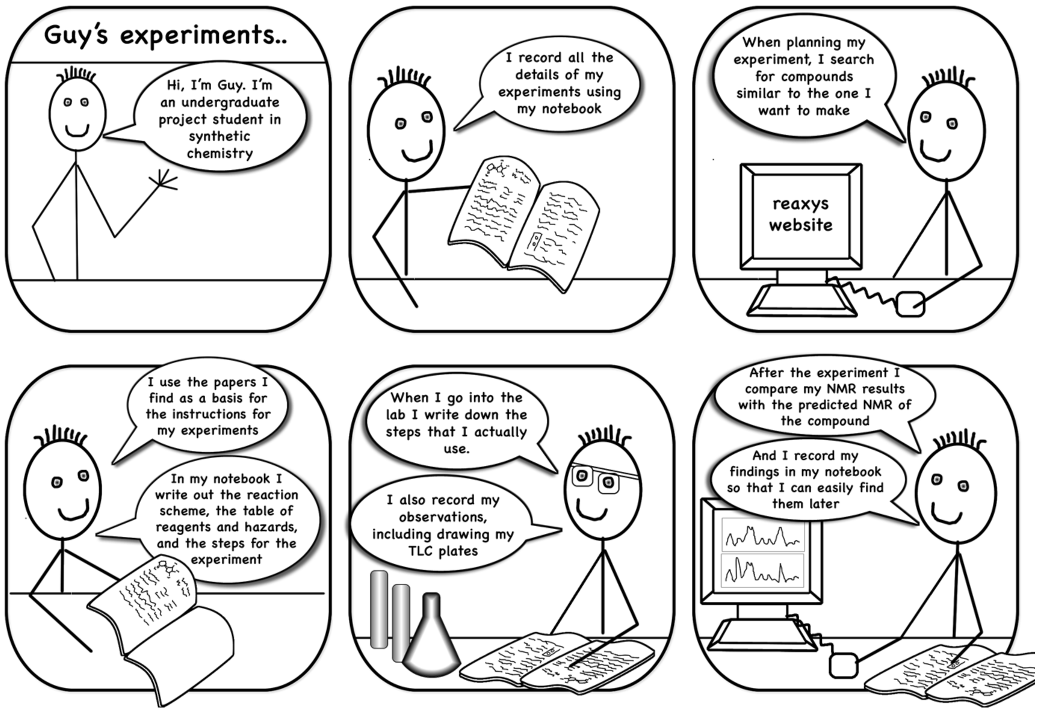


Figure 3-2: An example of the experiment process recording by paper (Synthetic Chemist)

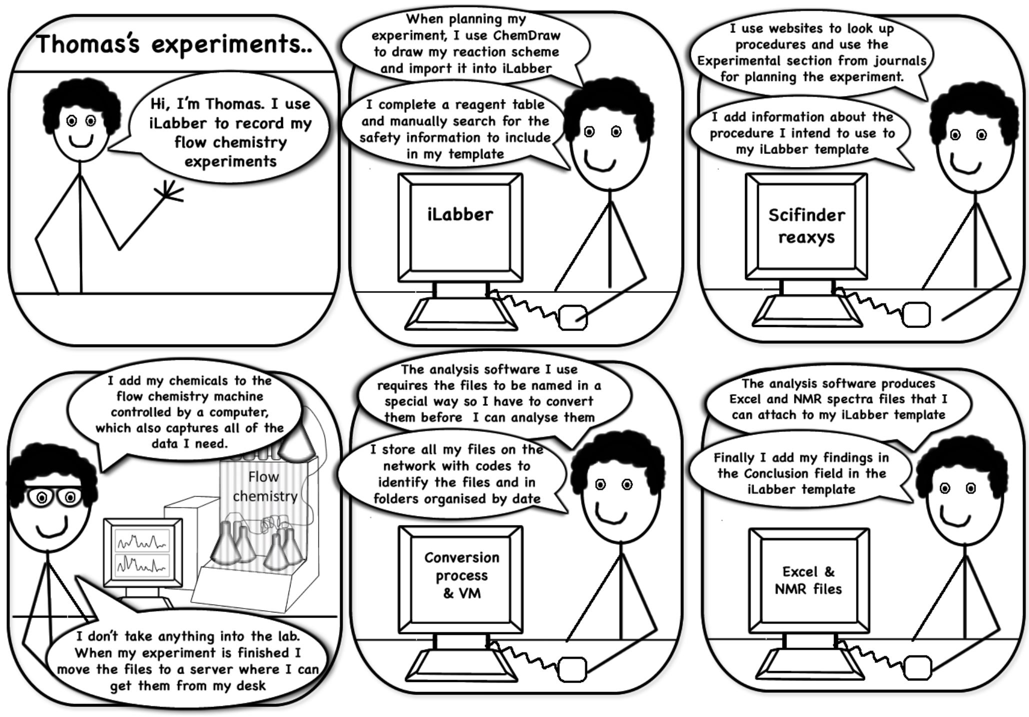


Figure 3-3: An example of a flow chemistry process using iLabber (Synthetic Chemist)

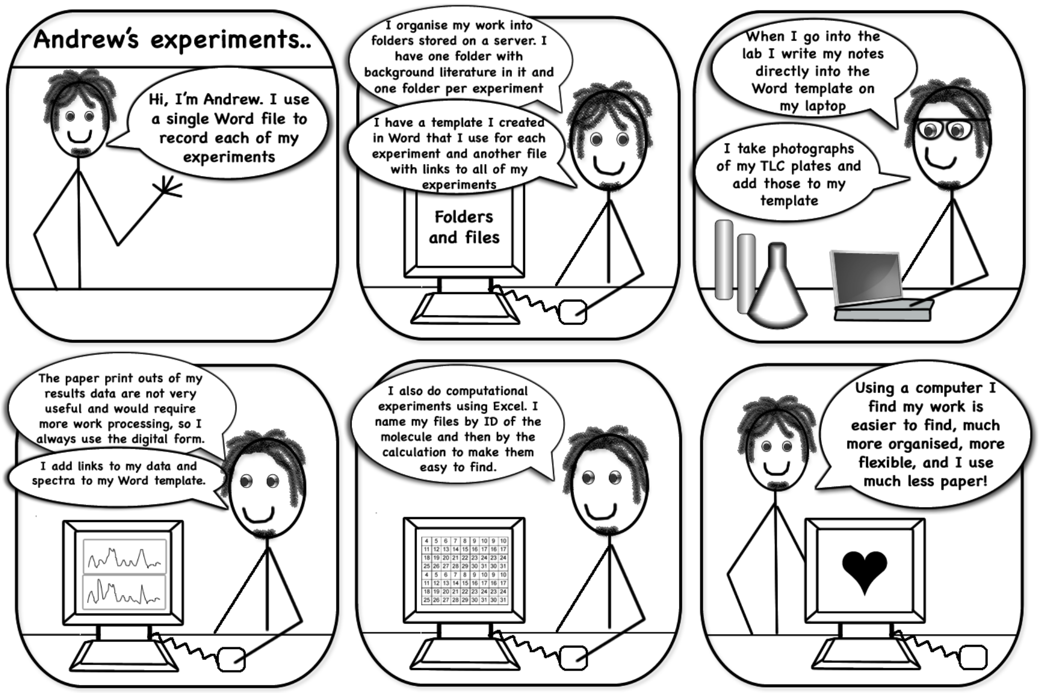


Figure 3-4: An example of the experiment process recorded using a customised Word template (Synthetic Chemist)

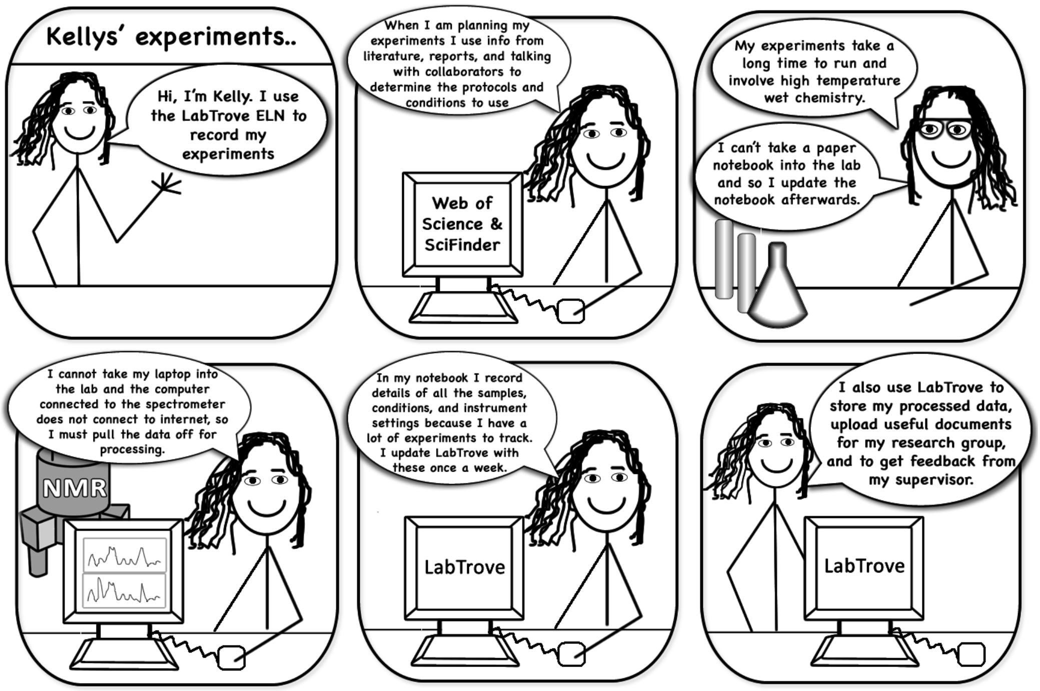


Figure 3-5: An example of the experiment process recorded using LabTrove (Environmental Chemist)

3.3 Learning to keep records

As we have established in the first two chapters of this thesis, it is important that records of scientific activities are accurate, complete, and accessible. If the record is missing information, contains incorrect data, or is difficult to understand, then it cannot serve it's function as a tool of learning and support for the scientist to whom it belongs, to members of the research group who may need to make sense of it, and a broader audience if the ownership or reputation of the research is at stake. Unfortunately many students and researchers often receive little or no formal training in techniques for recording, analysing, storing, managing, or sharing their data (COSEPUP, 2009; Mitchell, 2009; Swann & Brown, 2008; Shankar, 2009).

It is common for undergraduates to be provided with guidance on what information should be included within their laboratory notebooks and the structure of their formal reports as part of classroom learning or as formal guidelines for laboratory work. In teaching environments the experiment record provides an insight into the effectiveness of learning, whether the student has understood the theory of the experiment, accurately followed the procedures, and understood why some aspect of the experiment may not have behaved as expected. The experiment record must therefore be an accurate and complete reflection of what actually happened, and not simply a regurgitation of the intended plan. The notes taken during experiments for learning may also form the basis for future research or instructions for techniques in future work and therefore has a longer lifespan than may be anticipated at the time of creation.

Outside of the educational environment formal guidance is less common, and in fact many research laboratories do not provide in-house standards or recommended best practices for recording experimental data in laboratory notebooks (Wilson, 2011). The learning of recordkeeping processes may be developed as part of membership of a scientific community of practice; some research groups share best practices or standards amongst themselves with the development of local practices. However most scientists will develop their own personal style of recording by trial and error (Shankar, 2007, 2009). The development of additional management, such as a coherent numbering system, table of contents or other indexing system, which are often imposed retrospectively, may only come through learning from earlier mistakes (Shankar, 2004).

From our own discussions with researchers, many of them mentioned the difficulty of taking over another researcher's work and how a lot of training was required before they were able to repeat the work. A part of the reason for this is that a researcher has a standard operating procedure, and repeats the same procedure again and again. But the consequence of this behaviour is that not everything is written down in the record because this is seen as unnecessary duplication and because the researchers themselves already 'know' how to do the procedure after their training and practice. When one researcher takes over another researcher's work the process is hard because the record is missing significant information. New PhD students at Southampton University who were trying to repeat experiments from the literature also raised the same issues. The methods and protocols provided in the literature were missing critical information because standard operating procedures vary between different groups and laboratories, but these differences are not documented. These problems confirm that important aspects of personal and organisational memory are not externalised in the scientific record when these details are 'routine' and 'obvious' to the author.

None of the researchers that we interviewed had received any formal data management training.

3.4 How and why do we take notes?

As part of a workshop on the topic of note-taking we asked academic researchers and staff from the KTH Royal Institute Of Technology to provide responses to the following questions:

- Why do we take notes?
- How do we take notes?
- Who uses the notes and why?

These questions were designed to get the researchers and staff to think about note-taking and the responsibilities of researchers, but the responses have also given us the opportunity to check our assumptions about attitudes towards note-taking with researchers.

Table 3-3: Why do we take notes?

Theme	Response	
Memory	* Help us remember * So we can remember what we did * Remember what you've done * To remember what you have done in reality	* Remember * Because you do not remember what you have done otherwise * Remember * To remember
Proof / IP / Legal	* To be able to prove when and what was done * Prove when you discover for IP * Legal reasons * documentation	* Protect IP (proof of doc) * Have a record of what has been done * For others to know what you have done * To prove that you have actually done an experiment
Organisation	* To organise data * Highlight important information	* To find hidden patterns, unknown variables
Reproducibility	* Reproduce * To remember how we did the experiment do that we can repeat it.	* To assure reproducibility * Other may use it for instruction. * Be able to repeat
Troubleshooting	* Analyse what could have gone wrong * Comparing with former experiments.	* Troubleshooting weird results (Felsôkning)
Other	* Impress your supervisor * Observations * As background for discussions	* For reports * To document for posterity * Intellectual continuity

Table 3-3 shows the themes and responses derived from the question ‘Why do we take notes?’. The importance of note taking as a form of memory is clearly reflected in the responses, together with ‘proof’ of conducting the work, providing the mechanism to be able to reproduce experiments, and to understand and troubleshoot problems.

Table 3-4: How do we take notes?

Theme	Response	
When?	* During experiment	* as you go
Computer vs. notebook	* Writing in computer * Writing in notebook * In a notebook * Pen, paper, sketches * Lab book * hand written Lab book * Smart phone => electronic book * Electronic records	* Pen and paper * Writing on pieces of paper * Write in a notebook * Both in a notebook (paper) and on the computer * Manually in a lab note book * iPad -> cloud *Electronic * One note
Detail and accuracy	* As detailed and accurate as possible * As clear as possible	* Clearly so people can read * Detailed information on the experiments
Content	* With schemes and tables * Summarise data tables with comments and explanations	* Title, Reaction scheme, reactant amount, equality parameters, result, characterisation of the compound
Dates and times	* With date so you can do follow ups. In the best case in chronological order	* Time log of activities an observation
Other	* In English * Remember conclusions * Photos of screen	* Too few notes * "The email Trail"

Table 3-4 shows the themes and responses for the question ‘How do we take notes?’. Interestingly the majority of responses were focused on the medium that the notes are produced in – the majority describe using paper, but several also mention using computers or devices.

Table 3-5: Who uses notes and why?

Theme	Response	
Who?	* Supervisors * Colleagues * Next generation of students could use my notes * Other people in the research group * we use the notes ourselves * Me! Memory * I use the notes	* Colleagues & ourselves * Coworkers * Colleagues * Coworkers use my notes * "the next generation" * you yourself * yourself
Why?	* Impress supervisors * we use notes to investigate procedures in case something goes wrong * (Colleagues) to repeat the work * Don't have to repeat existing experiments * For accurate presentation of your data and procedures	* myself and co-workers to remember * Repeat, improve, or refer to it for manuscripts * The writer to remember and publish * to see what other people have done
Others	* Show your wife (prove you work where you say) * Financiers "due diligence"	* A lazy competitor (if they could)

Table 3-5 shows the themes and responses for the question ‘Who uses notes and why?’. The responses indicate that the researchers themselves and their colleagues are the primary users of their notes with several indicating that others do make use of their notes for memory and to repeat, improve or avoid duplication of experiments.

3.5 Notebook structure

Although researchers have great individual freedom to choose how they create their records to best represent the objects and actions of their research, there are also disciplinary, institutional, and professional norms in how they undertake to do so (Shankar, 2004). Guidelines are often provided to students and researchers on how to structure their notebooks and reports; and what each discipline records in their notebook can differ depending on the task that they are completing and how the data for the activity is captured. Research notebooks include not only data, but also details of ideas, conjectures, plans, equipment and the corresponding apparatus set-ups, methods, observations, analyses, and results (Bird et al., 2013a; Wilson, 2011). Historical notebooks also notably contain periods of reflective thinking, refinement of scientific theory, and careful journaling of the ideas leading up to scientific innovation (Bird et al., 2013a; Wilson, 2011). The information contained within is often, but not always, in temporal order. In either case the structure at the higher level provides an insight into the evolution of the research trail (Holmes, 2004).

As mentioned previously, paper notebooks are very flexible and can be used for different purposes simultaneously (Brandl et al., 2010). As a part of the record the notebook can contain any information that the author considers to be useful, thoughts, descriptions, and observations relating to the research activity; but also more personal information, such as contact details, web addresses, reminders, and 'to-dos'. The information included in the notebook may be in any format the author chooses, and may include text, symbolic abbreviations, bulleted lists, sketches and diagrams, charts, flow charts, maps and plans, tabular data, and calculations (Yeh et al., 2006).

Table 3-6 shows the structure of notebooks for a number of different disciplines. The entries in the table include examples of different structures that we have directly observed in use by scientists or that have been derived from the student guidelines mentioned. It can be seen that there are many similarities between the different disciplines, but field sciences and lab sciences show a number of key differences. For example, in the field sciences 'location' is important, but the date is more significant to the lab-based notebooks, with the timestamp having a long history of importance in scientific and engineering work (Cardenas, 2014).

Table 3-6: Structure of Notebooks

Discipline & source	Structure	
Life sciences (Scholl & van Laerhoven, 2014)	* title * date * goal * abstract * protocol overview * protocol steps (workflow)	* results * involved persons (experimenter & supervisor)
Field Biology[9,10] (Guidelines)	* locality * route of travel + map * hours of observation * weather conditions * participants * habitats and topography * list of species seen and numbers * general impressions	**Catalogue** section of notebook includes locality and samples collected. **Species account** section of notebook includes species identification, where found, what is was doing, how it looks and sounds.
Field Geophysics[11] (Guidelines)	**Field:** * Name/Date * Diagram * Location Map * Identification of apparatus * Tabular data	**Lab Reports:** * Objective of experiment * Diagram * Location Map * Identification of apparatus * Tabular data / Spreadsheet * Calculations * Graphs * Discussion
Geology[12]	* Excursion title	* Locality description / large

[9] http://gk12calbio.berkeley.edu/lessons/fieldnoteguide.pdf
[10] http://instruct.uwo.ca/biology/320y/fj.html
[11] http://cires.colorado.edu/people/sheehan.anne/lab.html
[12] http://www.geos.ed.ac.uk/undergraduate/field/fnb/

(Guidelines)	* Excursion objectives * Weather / conditions * Locality number * Location / grid reference	and small scale observations, incl. measurements * Sketches with scale
Biology (Lab)[13] (Guidelines)	**Title page:** Name, address, course, phone number, project title or description **Table of Contents pages:** Date of experiment, descriptive title, page, signatures **Table of Protocols pages:** Protocol, date, page, signatures **Glossary:** Term, definition **Experiment records:** * Title * Date	* Notes, reflections and literature references * Objectives (why, testing hypothesis, repeating old experiment, who proposed experiment?) * Hypothesis * Protocol details (materials, chemical manufacturer, equipment, photos, drawing, experiment set-up) * Results, raw data, observations * Discussion, interpretation, and conclusions * Notes and reflections on outcome.
Physics[14] (Guidelines)	* Title * Date * Name of lab partner * Aims/background * Method and explanation of experimental decisions * Sketch of experimental equipment	* Tabular data * Graph of data * Comments on the Physics * Conclusions * Review of your work
Chemistry (Guidelines from Southampton University)	* Title * Date * Short description of the experiment * Risk Assessment	* Data collected * Results, analysis, and conclusion (included in a formal report)

[13] http://blog.lib.umn.edu/jveldof/calculator/BiologyLaboratoryNotebooks.pdf
[14] http://labs.physics.dur.ac.uk/skills/skills/labbook.php

	* Procedure including key observations	
Synthetic Chemistry (Paper notebook) (Interview with researcher)	* Reaction scheme * Table of reagents * Hazards * COSHH entries for the significant or unusual hazards	* Planned procedure * Actual procedure with weights, times and observations * Analysis of the results
Synthetic Chemistry (i-Labber template) (Interview with researcher)	* Author * Experiment ID * Title * Batch ID * Start Date * Reaction Scheme	* Reagents * Safety Info (manual) * Procedure * Processed results
Synthetic Chemistry (Word-based notebook) (Interview with researcher)	**Table of contents:** * ID of experiment * Title of experiment * Date * Reaction scheme * Links to the experiment file * Links to the other folders * Comments	**Experiment files:** * Experiment reference * Date * Title * Reaction Scheme * Reagents table (incl. risks) * Literature references (links) * Experiment procedure * TLC plate * Links to data and spectra
Environmental Chemistry (Interviews with researchers & notebook)	* Title * Date * Composition * Form * Synthesis Location	* Conditions * Tools * Processing * Description/observations * Measurements * Data
Physical Chemistry (from notebooks)	* Date * Title * Materials table (Name, Quantity, Hazard)	* Precautions * Method * Disposal * Spillage Control

The notebooks investigated from the physical chemists at Southampton University showed very little formal structure at all, although all included dates and often included titles. Rarely a formal record of an experiment, as illustrated in the final row of Table 3-6, was included within the notebook or as a loose leaf.

3.5.1 Retrieval aids – table of contents and indexes

Faraday used many retrieval aids or reminders to locate the contents of his notebooks, many of which could not have been created until after the notebooks were completed (Tweney, 1991). These retrieval aids include numbered records, retrieval sheets with a short description and reference to the numbered records, contents lists, and even indexes of indexes. However in more contemporary notebooks the use of contents and indexes seems to be quite rare and often imposed after the creation of the contents (Shankar, 2004); although some of the guidelines in Table 3-6 mention creating a table of contents. Many researchers start recording new experiments on a blank page, leaving space for any additional information to be added to an experiment currently in progress (Wilson, 2011); this would make it reasonably easy to create a table of contents after the fact.

In our own discussions with researchers, only one researcher mentioned creating a table of contents within a paper notebook, whilst the synthetic chemist using Word to capture experiment records also created a table of contents document to keep a track of his experiment documents and provide a simple way to retrieve them. Many ELNs provide a table of contents facility, but may only display a small amount of information about the experiment making it difficult to use. Using the reaction scheme or structure of the materials used in the experiment may help make the table of contents more useful for chemists (Cooke & schraefel, 2004). The ability to be able to use the chemical structure for discovering information is particularly important for synthetic chemists.

3.5.2 Linking, finding, and metadata in the notebook

Various forms of structuring and highlighting are commonly seen in paper-based notes that are used for linking together different content throughout the notebook - highlighting content as important and making it easier to find, together with the use of other markings and symbols annotating the content that could be considered to be metadata. Examples of structuring and highlighting for linking and indicating relationships between content across multiple pages are demarcations, such as horizontal lines and different pen colours, text started on a new page, categorization with topic and date, and through the use of symbols; the symbols relate to the subject of the content and link together information at different locations in the notebook (Brandl et al., 2010). To assist with finding information later, various ‘tags’ and symbols are use to distinguish different snippets of the content. (Van Kleek, 2009; Brandl et al., 2010). Highlighting with coloured markers, circling or framing regions, dog ears, and paper clips are also used to make content easier to find (Brandl et al., 2010).

These mechanisms may be more likely to be used in informal notebooks perhaps, but guidance on ‘Keeping a field journal’[15] from The University of Western Ontario provides examples of highlighting and symbols that could be considered metadata, specifying the use of underlining to highlight important information, for example wavy lines for common names and straight lines for the scientific names of species described in the journal. New locations and dates are also underlined. These markings enable the key information to be picked out quickly without needing to read the entire text. The guidelines also show an example of ‘marginal symbols’ for annotating text and linking, for example a symbol with a dot within a square to indicate a photograph was taken.

Amorium et al (2014), also describe more formal capture of metadata within notebooks that is then translated into digital format at a later stage in a chemical engineering and biodiversity lab. The metadata used by these groups

---

[15] http://instruct.uwo.ca/biology/320y/fj.html

included word or phrases to describe samples, descriptions of analysed materials, geographic locations, photos, and sketches.

Field archaeologists historically capture most of their excavation data on paper with the use of context sheets that contain numerical identifiers for places and archaeological contexts. The notebooks may provide references to other information such as photo IDs and drawings in context to other elements of interest such as stratigraphic context, colour descriptions, and interpretations. The majority of the information relating to the excavation record is recorded on various structured forms, and every element of the excavation has an ID, for example the locality, the trench, each layer in the soil, and each archaeological find, plus other metadata such as a geographic reference. Metadata is therefore stored both on the structured forms and within the notebooks.

Our investigation of the note-taking behaviour in notebooks from chemists at Southampton University suggests that there is some evidence that inclusion of more formatting and 'marking up' content may develop with expertise or experience in note-taking. When the paper notebooks of 3rd year undergraduate students and PhD students are compared, in general the PhD students included more highlighting and formatting. The PhD students frequently included the following types of highlighting and information:

- underlining of titles and individual elements such as important information or results
- horizontal lines
- names of files
- tabular information / data
- bullet points
- different coloured inks
- double underlines and double horizontal lines
- circling and framing
- use of all capitals for emphasis
- annotations
- time stamps
- sketches, graphs, and diagrams

The undergraduates in contrast used much less emphasis in their notebooks but titles and 'metadata' style information such as date and name were often underlined. Numbered and bulleted lists were also used by some of the students, but tabular data and sketches were rarely included. Some symbols were used, but these were typically 'arrow' types symbols indicating flow or conclusion from a previous statement in the text.

The structure used in the notebooks varied from student to student with some of the undergraduates having a very formal style containing Name, Date, ID, Title, Aims, Procedures, Observations, and Results, whilst other undergraduates included their name, date, and procedure only. Interestingly the notebooks belonging to the PhD students had very little formal structure within them, were much more 'diary' style, and included much more personal information with detailed descriptions of set-ups, ideas, observations, and other reminders, and often details of 'internal discussion' – or thinking. The lack of formal style may reflect the topic of research and the norms of the field, as evidenced by the researchers in synthetic chemistry at Southampton all using a very formal structure for recording their experiments. Different interfaces may suit these closely related, but clearly distinct, disciplines because of their different approaches and ways of working.

3.6 Recordkeeping behaviour

The details about the structure of experiment records, processes involved, and different notations described in the previous sections provide some insights into recordkeeping behaviour and how they differ amongst researchers. However, there is more to recordkeeping than just the nature of the 'inscriptions' themselves. The more personal elements of recordkeeping are discussed in this section.

3.6.1 Recordkeeping as ritual

Shankar regards engagement with the record as a mental and physical activity, and the act of capturing notes and the daily routine of 'writing up' as a ritual norm of scientific society (Shankar, 2009). The following of a prescribed structure within a notebook could be considered a part of that ritual, together with the details of when and where the record is constructed. The specifics will depend on the nature of the work and the habits of the particular discipline; research in recordkeeping has demonstrated that there is wide variation in personal and local practice with individual micro-cultures within science (Shankar, 2004). For example, researchers who run one experiment per day may record every step in the notebook as they go along, but those whose experiments are performed over longer periods of time may update their notebooks on a weekly basis instead (Jones & Nemeth, 2005). Recordkeeping may be a ritual practice, but that does not mean it is consistent.

From our own experiences of note taking we would agree that there is an element of ritual involved. In laboratory work this comes from the formal structure and the process of planning, conducting, and analysing the results of an experiment. In geology, recordkeeping definitely has a ritual element to it, with notes, sketches, and measurements captured in the field during the day and each night a process of ensuring that the notes are correct and legible, and also 'inking in' of details captured in pencil in the notebook (if it was raining) and on maps during the day. No sleep is allowed until the details have been properly preserved in ink. In archaeology note-taking seems to be less of a formal task with the notebook there to capture the details that you individually feel are important and to act as reminders of tasks. The formal recording takes place on the centrally stored forms such as the context sheets and small finds records. Details captured in the notebook, however, are there to track down the correct information if something has gone wrong, for example the same context ID has been used twice. The archaeologists at Southampton asserted that the paper notebook was not such an important record of an excavation, unless you were the PI, but that they were more of a trophy to be stored on the shelf – proof that you were there and a part of the experience. We would agree with this as an assessment for the archaeology notebook, and to some extent for the geology notebooks. Geology notebooks also have a second life as a reference for future excursions; lecturers frequently use their own notebooks as guides for fieldwork excursions when educating their students.

3.6.2 The private nature of notebooks

Whether you consider notebooks to be private or not may be influenced by your role and the nature of your research. For example, Wilson (2011) found that only principle investigators viewed laboratory records as a shared resource, whereas others relied on information to be converted into public resources such as talks, articles, and reports; in other labs supervisors and other research project members reviewing the notebooks (Jones & Nemeth, 2005).

We tend to think of the notebook as being the place where all notes and associated activities are recorded in the laboratory or in the field. In reality though many researchers may actually be recording a lot of their work on scraps of paper or in ‘rough’ notebooks and then later discarding them or transcribing them when a better or more complete story has formed in the researcher’s mind. For example, Shankar (2007) describes the note-taking behaviour of one researcher who uses a rough notebook for writing equations and taking notes during an experiment, and then later writes-up the ‘main record’ into his ‘main notebook’ in the correct order once everything is understood. Neatness of handwriting may be an indicator that researchers are creating a ‘main notebook’ because the record was created from notes or from memory after the event compared to rougher notes made during an activity (Hoffmann, 2013).

When provided with the ELN for recording their work in the user tests, both sets of chemistry students at Bath University frequently used scraps of paper to make notes whilst they were carrying out the experiments. Many of the students did some write up of the background of the experiment in the ELN first, but the majority captured the actual measurements on paper before adding it to the ELN. This behaviour was quite consistent with the described habits of archaeologists in the lab where the students tend to take notes as they perform the analysis only temporarily on scraps of paper, which are then discarded at the end of the process when the structured data has been digitally captured. There is a risk with ELNs that the computer becomes the ‘main notebook’ supplemented by additional scraps that are later discarded.

There was an almost hostile reaction from some of the Bath students to the idea of sharing the information in their notebooks in an ELN. One of the students even said that if it could be accessed by anyone else then they would simply not be allowed to use it. There were two main concerns expressed by the students about sharing of their notes; the first around the issue of intellectual property and the second around the issue of ‘readiness’ of the information. A number of the students commented that they would not want to let anyone else see their notebook because they view it as a personal record and a ‘work in progress’. The formal reports are the finished items that could be exposed to others. Only a very small number of students thought that being able to share the content and receive comments on their notes would be valuable. These were students that were actively working with other groups on their research, and therefore already shared their work and needed to receive input from an external group to help them. One of the Bath students was a supervisor for other students, and although he was hostile to the idea of using the ELN for his own work, said that he would be happy for his students to use it so that he could see what they were working on. This statement confirms how role changes how personal the notebook is perceived to be and how different user roles may have contradictory requirements.

Some groups expect their colleagues, students, or other collaborators to view and make use of their notes, whilst many researchers only share their notebooks with their supervisors. It may be that the researchers who expect to share their notes with others produce better quality notes than those who are keeping the notes only for themselves. Adherence to defined guidelines or standards may also improve note-taking quality. For example, one of the researchers from Penn State University is working on a project funded by a grant body that requires adherence to precisely defined guidelines for keeping notebooks - including dating every page, providing lot numbers for every compound, and for each page to be signed. A representative from the grant body visits to check that the notebook is being kept correctly.

3.6.3 Transcribing, transformation and creative thought

As mentioned previously, research notebooks contain reflective thinking, ideas, conjectures, plans and refinement of scientific thinking, in addition to data, methods, and observations; paper itself having been credited with being an effective medium for creativity (Brandl, 2010). Notes may be expanded and developed through transformation through reports, draft papers, and ultimately scientific papers. The process of writing and transforming the information itself can help with externalising memories, filtering, organising information, and thought development (Tabard et al., 2008; Solomon, 1997; Yeh, 1996). Draft papers can also provide an insight into creative thought with important theories, opinions, and ideas evolving and growing over time (Holmes, 1987). This refining stage is an important part of the creative process, enabling the researcher to think and process the information contained within their notes and memories.

The archaeology students produce models and representations of what an archaeological site would have looked like. They are also actively encouraged to take notes on this interpretation process using blogs, diaries, or as a linear unstructured word-processed document. They include notes in their diaries about the versions of the models and the methods they use. This explicit association to the 'what' and 'how' is important, but the structure of the document for recording is not considered important because they believe that the act of making the notes helps to consolidate knowledge, understanding, and theories. As a result of this consolidation exercise the researchers are thought to not need to go back to refer to the notebooks directly.

Currently many researchers transcribe materials from their paper notebooks into digital formats; an activity that is both time-consuming and error prone (Weng et al., 2012; Yeh et al., 2006). It is unlikely that all forms of transcribing will have benefits for creative thinking, for example, transcriptions that are a word for word conversion from one format to another, or the conversion of tabular data into a spreadsheet. The majority of researchers that we have spoken to have data produced by instruments digitally and either use an ELN or a paper notebook to capture their experiment plan and notes. Some researchers use both ELN and paper notebook, and transcribe information from the paper notebook to the ELN, but this does not seem to be a reflective process, and the need for duplicated effort in this regard has turned many off using the ELN entirely.

There is a risk that using digital tools for the capture and management of scientific records may reduce the information that is manually created by researchers and therefore reduce the benefit of 'writing' for reflection and creative thinking. However, there are also opportunities for designing in functions to encourage creativity, and digital technologies, together with effective curation, provide new ways of organizing and presenting the information than would be practically impossible through manual means.

3.7 Using digital tools and ELNs

Many researchers are now working with a mixture of paper and digital materials. In addition to the digital materials in the form of raw and processed data from instruments and other sources, researchers are using tools such as word processors, spreadsheets, drawing, modelling, and online tools in their daily work.

Linking to data is perhaps the one area where we found a majority of users had a positive response to an ELN. Many of the communities that are successfully using the LabTrove ELN, including the universities in Sydney, make use of 'auto-blogging', where sensors or instruments automatically upload new data to a notebook. The uploaded data can then be linked to from another notebook. Many of the notebook entries in these communities also include data uploaded to entries including photos and other data files. The entries can also include links to data on other notebooks, repositories, and web pages.

In the user studies at Bath, some of the students recognised the usefulness of being able to link to or include data from instruments, particularly NMR data. Many of the students did attempt to record their experiment data in Excel and then copy it in, or even drag the file into the ELN. A few of the students mentioned currently using Dropbox as a solution for storing data.

For the archaeologists creating links between the notebooks and the real data is a very difficult problem. There is a vast amount of data in various formats from an excavation and associated surveying and lab analysis, for example a single excavation may generate 30,000 photographs, geophysical reports, character data, reports, text descriptions in a database, and structured data such as spread-sheets, databases, x-ray, multi-spectral, 3D database, point cloud, and geometry data. The data may be stored on file systems or in databases or other repositories. Whilst some of the data may be stored online, much of it is distributed in inaccessible locations including personal computers, in private repositories, and even on memory sticks.

Although the archaeologists feel that being able to create documents that contain links to the data would be valuable, there are a lack of tools that enable this to be done effectively. It is important to be able to make a link to the data at the time the document is created because you are unlikely to return to do this after the event. Having a link can also be used as a way to highlight that a piece of information has value. It is also important that links do not get broken if the data moves, ideally the link should also be two-way, like Microcosm for example, which provides a link base with bidirectional links (Fountain et al., 1990).

Those researchers we spoke to commonly mention Word and Excel and integration with Microsoft as a common expectation or desire for ELNs. This expectation was also clear during our user testing at Bath where many of the researchers captured their numeric data in Excel or Word and tried to copy it into the ELN, and many others commented about the lack of value of a system that is less effective than the tools that they already use and get value from. The ability to be able to move content between these applications is important to researchers. This may be to fit in with their current workflow, but probably is also because of the familiarity and perceived ease of use of these tools.

Researchers in the Chemistry department at Cambridge use an Excel spreadsheet to manage their inventory of chemicals. Each group holds around 1-2,000 chemicals and the spreadsheet acts as an inventory of the stock, but is a standalone document that can only be used by a single user at a time. Even using the spreadsheet is a very manual process and requires information from outside to be manually pulled in such as densities of chemicals, and hazard and risk information. Some of the researchers at Bath were using Excel or Google docs for COSHH reports, capturing tabular results data, and for performing calculations.

Archaeologists use blogs and word-processed documents to record their thought processes. When asked whether they had experience of using Electronic Lab Notebooks, both the archaeology student and the supervisor said only that they had tried to use digital pens. This suggests strongly that the archaeologist view of an ELN is a tool that exactly replicates the paper notebook by creating an electronic version of what is written in such a notebook, rather than replacing the paper notebook. They already use other tools like the chemists do to produce formal reports. The chemistry researchers at Cambridge also mentioned the use of digital pens to capture information in the paper notebook and then later upload the image to the ELN.

3.7.1 The Blank Canvas

Aside from the limitations of the ELN software compared with Word and Excel, one of the difficulties experienced by the Bath students in the first user session was the effect of the 'blank page' or 'blank canvas' because the users were given a brand new notebook to work with. With the blank notebook many of the students did not even know where to start with the ELN and needed to be directed at the "New Post" button. Many also commented they were unsure what they needed to record in an electronic notebook, not viewing it as an equivalent to the paper notebook.

Within the ELN there is access to other notebooks, but the link is hidden well away from the 'working area'. Once the researchers were within their own notebook they aren't really aware of anything else or the context that they are in. The majority of the students in the study did not attempt to navigate away from their own notebook, and therefore had no additional contextual information to work from, even though they were encouraged to add links to related information in their notebook entries. A small number of students did have a look at the other notebooks visible to them, but only to "see what you can do in a notebook” and saw no reason to link to anything, even though the information was directly relevant to their own work.

The blank canvas effect of lack of context and examples was also seen to influence the effective or ineffective use of metadata. Both of the universities in Sydney indicated they had difficulty in creating a consistent metadata schema, and that the students were put off from using any metadata at all because they didn’t know what to use, although a small number of students seemed to have a good understanding of how to develop their own schema that was useful for their own information retrieval. The University of Sydney commented that they did not use metadata very much because of the difficulty of retrospectively adding metadata to a notebook where it had not previously been used. They were also nervous that if they added metadata they might ‘mess things up’. Having a different way to add links to important posts and structure the homepage was felt to be a more valuable requirement than adding metadata with the current interface. Both of the Sydney universities expressed a preference for tags and tag clouds rather then the “key/value pairs” used in the ELN.

Only one of the Bath participants showed any interest in the metadata asking what the interface elements were for and why they should use it. With explanation and examples from a notebook with exemplar metadata she understood the value, but did not manage to use metadata appropriately in her own entries. Use of a single ‘section’ metadata is required in the notebook, and so all of the other participants added this section metadata but only after prompting by an error message from the software.

The discussion with the archaeologists highlighted that they have an appreciation of the value of metadata. Reports are organized in a folder hierarchy, but there is no document management metadata. A standard indexed search is used to find data. A very methodological person is used to organise the vast quantities of GIS data. XML metadata files are structured according to GIS best practice, which is described as being ‘beautiful and clean’. However, it took months and months of work to organize it effectively and continues to need attention throughout the whole project to guarantee consistency of the metadata. For the geophysics work the team started off using a folksonomy based on an existing taxonomy. The folksonomy was used initially to generate project metadata and once it was used in a consistent way they then transitioned to using a rigid taxonomy based on the folksonomy. The archaeology student indicated that he has tried to add some metadata to his own notes in tools such as Evernote and suggested that the Wordpress functions for both tags and categories encourage people to add metadata. He also makes use of a faceted classification plugin for Wordpress that enables the user to choose to classify using tags or categories and see both side by side. He does however admit that he does not get much use from the metadata he has added, and therefore does not add it consistently.

3.7.2 Ease of use and desired features

As discussed in **Chapter 2: Digital tools for capturing the scientific record**, there are barriers to the adoption of ELNs and ease of use is one of the most significant of these. For those researchers already using ELNs there was a familiar story from each about the limitations and instabilities of the software, and for those that were new to ELNs a great deal of anxiety was expressed about the anticipated usability and stability problems. It is hard to challenge the usability of a paper-notebook.

Typical comments made include the difficulty in configuring and maintain the ELNs, problems with networks and accessibility, lack of an intuitive interface, and poor integration with other essential tools. Other concerning issues are frequent crashing, corruption of data, and an inability to extract vital data from the ELN when the researcher moves on. These concerns and experiences are in parallel with more publicised concerns about data protection, intellectual property, and costs. However, as mentioned previously, some researchers do see the value of ELNs and do make effective use of them (Badiola et al., 2015; Carpenter, 2012).

Interviewing both experienced and inexperienced ELN users has provided the opportunity to gather a list of features that researchers expect or would like to see in an ELN, as well as problems that need to be fixed. The themes, requirements, and sources of these requirements are provided in Table 3-7.

Table 3-7: Requirements for ELNs from user interviews and studies

Theme	Requirement	Source
Administrative	* Simplify install and day-to-day management * Online configuration of systems * Migration tools * Export	UNSW, Cambridge, Sydney
Data-mining	* Live data mining for chemical names and associated links – recognise a chemical when it has been entered	UNSW
Voice capture	* Voice may be more intuitive in the lab than typing and means you don't have to put everything down to record what you are doing	UNSW
Search	* Search by chemical structure * Include reaction schemes in search results * Voice search * Advanced search * Filtering * Sort search results * Search by success / ratings * Organise by project	Sydney, UNSW, Cambridge
Collaboration	* Configurability of stand-alone pages to act as portals for projects and landing pages for collaborators. * Sign-up and get involved pages * A place to enable coordination for open source and access. * Moderation of comments * Bulletin board / forum for discussions * Notifications * Page statistics * Mechanism to find out who is working on similar molecules or reactions (requires a built in understanding of molecules)	Sydney, UNSW, Bath, Penn State

	* Function like Google+ circles for linking together people and related notebooks. * Automatic tweeting of new entries * Share standard list of instruments, reagents, materials * Shared files * Shared notebooks	
Access & Security	* Secure access wherever you need it * Granularity in security for users, roles, and individual items within a notebook * A place to enable coordination for open source and access. * By organizational structure	Sydney, Cambridge, Penn State
Activity indicators	* Need to be able to see “Recent Activity” * Notifications for comments and responses to comments * Activity feed from project students	Sydney
Linking	* Linking between different notebooks * Automatic update of data without intervention * Linking to external resources such as CAS registry number or catalogue numbers * Automatic linking of chemistry terms to the definitions in the IUPAC Gold Book * Table of contents page / overview screen * Organise by project	Sydney, UNSW, Cambridge
Archive	Archiving functions for long-term preservation	Sydney
Safety	* Risk assessment template * Badges / Flags for hazardous chemicals * Have chemicals/values/measures in database so they can be globally edited * Automatic notifications for approvals * Electronic signature for risk assessments * Index of COSHH materials and import	Sydney, UNSW, Penn State, Bath, Cambridge

Sign-off	* Sign off button that also makes the entry non-editable * Review and approval options for entries	Sydney
Metadata	* Prefer tag clouds * Autocomplete tags to avoid different case/misspelling * Spellchecker for chemical names * Use of standard vocabularies (Properties Ontology (ChemAxiomProp), Measurement Techniques (ChemAxiomMetrology), Analytical Information Markup Language (AnIML)) * Success and ratings indicators * Highlighting experiments – Like function	UNSW, Bath, Cambridge
Calculations	* Calculator/calculations in the notebook * Equation editor	Bath, Penn State
Integration	* Copy sketches from elsewhere into Notebook * Copy from Word and Excel into Notebook * Integration with paper notebook * Integration and automatic upload from digital pen	Bath, Penn State, Cambridge
Personalisation	* Ways to personalise the notebook * Post it notes or ‘to-do’ list in notebook * Own default values	Bath, Cambridge
Publication	* Thesis button	Penn State
Other	* Timeline feature * Integrated user assistance * Fully featured sketch tool * Tablet and smart phone applications with text recognition, drawing, and photo capabilities that could be taken into the lab to record experiments in the same way that a paper notebook is used * Built-in language for extensibility * Validation	Bath, Cambridge

	* Work on all platforms (web-based)	

3.7.3 Encouraging ELN adoption through teaching

We have found that there is some evidence that attitudes towards ELNs from students can be improved by making use of ELNs in teaching. For example, in one third-year undergraduate course at Southampton University, students were provided with instructions for an experiment and associated data analysis within the LabTrove ELN and asked to record their results using the ELN. The students were also provided with feedback on the experiment and their results through the ELN.

After completion of the experiment, the students were surveyed about their attitudes towards ELNs before and after the activity. Of those that responded to the survey, 65% had negative opinions compared to 35% positive opinions about ELNs before they used one for the course, but only 28% had a negative opinion compared to 72% with a positive opinion after their experience. When comparing the ELN to paper the majority felt that both had pros and cons, but the majority felt that the ELNs were more effective for sharing work and getting feedback. However, many of the students had not realised that they had been provided with feedback in the ELN, indicating the importance of ‘activity feeds’ and notifications for collaboration, raised as important requirements in 3.7.2.

Chapter 4: Interaction design and tools for keeping and managing records

Tools, and the interfaces to work with them, are required to support researchers in keeping and managing their research records. Interaction design is an umbrella term describing a variety of processes and activities associated with designing user experiences of software, interfaces, and products (Rogers et al., 2011). In this chapter we explore some interaction design techniques to assist in the design of tools for keeping and managing records.

4.1 Conceptual models and metaphors

Conceptual models are like the scripts and schemas that we discussed in 2.2.2 - easy to understand paradigms and metaphors for the objects and activities that we encounter in everyday life applied to computer-based interfaces. For example, the majority of computer file systems are based on the conceptual model of the office filing cabinet, where documents are arranged in folders and sub-folders; the majority of shopping websites provide an approximation of the experience of visiting a supermarket by enabling us to browse items for sale, add them to our basket, and then go to the checkout to make a payment. Metaphors and patterns like these can be used to conceptualise what we are doing (e.g. recording an experiment), to construct the interface we are using (e.g., the notebook), and to visualise an action (take a photograph).

In **Chapter 3: Keeping and managing records**, we explored how researchers capture and manage their recordings, including how they plan their experiments, record them, analyse their results, and other ways they use the materials they generate, such as:

- sharing them with colleagues and collaborators
- turning them into presentations, theses or papers
- making use of them to design new experiments, refine a technique, solve a problem, or look for patterns within the results

At some point down the line their work is preserved, either directly by themselves or perhaps more commonly in the past by the organisation for which they conducted the work. Reflection and refinement can occur at any point in the process, and researchers might choose to update their plans, record ideas, and engage in discussions as a result of these reflections.

In the physical world researchers perform the following activities in developing plans and capturing experiments:

- **Search** for and **locate** *information*
- **Extract**, **refine**, and **write** *information* to make *plans*
- **Identify** and **write down** *materials*, *measurements*, and *hazards*
- **Design** the *experiment protocol* or *instructions*
- **Read** *information* and **follow instructions** from the *plan*
- **Write** *notes* about *ideas* and **record** *observations*
- **Draw** *sketches* and **take** *photographs*
- **Make** *annotations* and **highlight** *important information*
- **Draw** *tables* and *graphs*, and **record** *results*
- **Allocate** *IDs* and **label** *materials*, *samples* and other *resources*
- **Record** *IDs* in the *notebook*
- **Paste** printed *data* or *photographs* into the notebook
- **Flick** through a *notebook* to **locate** *information*
- **Explore** *information* in a *notebook* or a *specific experiment record*
- **Compare** *information* and *results* from different *experiments*
- **Extract** *information* from the *notebook* and **transcribe** into different *formats*
- **Share** the *notebook* and **discuss** the *content* with *collaborators*

The **verbs** and *nouns* in these activities have been highlighted to demonstrate how familiar terminology could be translated to a virtual environment. Other subtasks could be identified within this process to increase the level of detail. For example for ‘Write notes’ we might also include subtasks and other decisions that the user might make:

- **Start** a *new page*
- **Choose** a *pen colour*
- **Record** the *date*
- **Start** a *new line*
- **Write** a *bulleted list*

There are other tasks that are impossible to do in the real world with a paper-notebook, or at least impractical. Many of the requirements that were identified by researchers in 3.7.2 are requirements to add activities that they already do in the real world such as completing forms for risk assessments and acquiring sign-off, providing a table of contents, and performing a structure search. Others are requests for something that is not possible with an actual notebook such as spell checking, 'liking' experiments, automatic tweeting of new entries, different ways of organising information, and identifying future collaborators.

Tools to support researchers need to be able to do the tasks that researchers perform in the physical world within the virtual environment, and supply 'added value' by including these extended capabilities.

We should also think beyond the desktop and consider the benefits that ubiquitous technology can bring to support researchers in the lab and throughout the research lifecycle; although clearly we need to get the basics right first!

4.2 Research roles

In order to design tools for researchers we first need to have an understanding of our users, their goals, and needs. When we spoke with researchers we identified that there are at least seven roles associated with the use of an ELN. It is entirely possible for a single person to belong to multiple roles, and for more than one person to be needed to fulfil a single role.

Although it is possible that an individual may choose to install an ELN on their own machine, the most likely scenario is that a research organisation or research group will have a shared ELN. In this scenario, an important (and often forgotten) user role is the administrator of the system, who is likely to be the person responsible for installing, configuring, maintaining, and updating the system. Day to day administration activities may include user management, security, backup, database administration, training, and other support tasks. The 'administrator role' may be held by a member of the research team, an internal IT-support team, or even an external contractor, depending on where the ELN system is hosted and who is responsible for its set up and maintenance. The interfaces to support this role are outside of the scope of this thesis, and so this role will not be considered in any further detail here (sorry Adrian).

The second role identified for the ELN is the role of principal investigator or supervisor, depending on the nature of the research group. The 'supervisor role' is responsible for monitoring the research activities of the team and providing feedback to individuals. They may also be interested in exploring the research of their team, including comparing experiments and results, and looking for patterns. The supervisor may want to be able to extract information from notebooks and other resources and then use that information to hold discussions, create presentations, and include within other publications. The supervisor is likely to be involved in the coordination and relationship with external collaborators. Some supervisors may be interacting with their researchers remotely. We found in our survey of LabTrove, and with discussions with researchers, that a large percentage of notebooks were used by supervisors for project and team management purposes, acting as a tool for organizing meetings, distributing information to research teams, and for capturing personal thoughts and activities. The prevalence of this use for the ELN suggests that project and team management are also important goals for the supervisor role. Personas will be created to explore the requirements of this role in more detail below.

The third role identified for the ELN is the role of researcher, acting as the author of the research information that is to be captured and utilized in the system. The tasks performed by the researcher have already been documented above, but this role will also be examined in more detail through the use of personas.

The fourth role of the ELN to be considered in this thesis is the role of collaborator. Collaborators want to be able to access the research of staff using the ELN and to be able to collaborate in a variety of different ways. Collaborators may be other researchers, second supervisors, collaborators from other disciplines, or active members of the wider research community. Although in open science projects the notebooks may be visible to the entire scientific community, an individual is not counted as a collaborator unless there is a two-way dialog between the members. An individual who merely observes the research without engaging in dialog is a fifth role 'observer', for example, interested members of the scientific community and members of funding bodies may observe the research process and the results obtained without directly contributing to the conversation or the research effort. The collaborator and observer roles will also be investigated through the use of personas.

Open research and the potential for public interaction through an ELN system understandably raise concerns about individuals engaging in nefarious activities, with the possible abuse of facilities provided. With any system accessible outside a firewall, the potential exists for spamming, trolling, denial of service, data theft and corruption. These are legitimate concerns and mechanisms are required to ensure security of data, enable moderation of comments, blocking of users, and blocking of external inputs to systems. The scenarios and use cases investigated here will focus only on ‘happy day scenarios’ where everything works as expected; errors, exceptional conditions, and misuse or abuses of the are system are considered to be beyond the scope of this thesis.

The final roles of the ELN to be considered in this thesis are the roles of teacher and student. ELNs can be used as tools for both teaching science and teaching data management skills that will be required in industrial laboratories. Teaching can also be used to encourage the adoption of ELNs as described in 3.7.3. Teachers will need to be able to plan activities, create and share materials, monitor and provide feedback to students. The student role is likely to be similar to the researcher role, although the information that they capture and how they interact with other materials and other students may be different. Personas will also be created for these roles to explore their goals and needs.

The personas will be described together with a brief narrative of the scenarios for which they have been developed in 4.3 below.

4.3 Scenarios

Four typical scenarios have been derived from our research with members of the LabTrove research community. These scenarios can be used to help elicit information about the requirements for an ELN for academic environments. The following scenarios have been developed from our user research:

1. Closed research with interaction with a funding body, based on the research group at Penn State University.
2. Fully open research with interaction with global scientific collaborators, based on the research group at University of Sydney.
3. Closed research in a local research group, based on the research groups based at UNSW, Southampton, and Bath.
4. Undergraduate practical course, based on a third-year undergraduate practical experiment at Southampton University

For each of these scenarios a brief narrative will be provided to describe the scenario, together with the roles and identities of the personas that will be used to represent them.

4.3.1 Scenario 1: Closed research including interaction with a funding body

At the Mountain University, Professor Felix runs an environmental and analytical chemistry research group including eight research students. Felix spends a large amount of his time at a different location making it a challenge to manage the research and interact with his team at Mountain University. Much of the research is funded by external funding bodies who enforce detailed quality standards for research documentation; periodically the funding body needs to have access to the notes kept by the students to check that the documentation matches the standards, and that the research is progressing as planned. Felix's students, once trained, are able to work on their experiments autonomously, but need to have sign-off from Felix before they start experiments. The students also like to receive feedback to reassure them that they are on the right track. The experiments that they perform involve intensive activities across different buildings on the campus, meaning they need to take notes in several different locations. Their experiments also typically take several weeks or months to perform, and they are often working on multiple experiments at the same time. The team are interested in using an ELN to help Felix with managing the research project and providing support to the researchers remotely, and also to enable the researchers to work more effectively with their funding bodies.

In Scenario 1, Professor Felix (Figure 4-1) has the Supervisor role, research student Alejandra (Figure 4-2) has the Researcher role, and auditor Alvin (Figure 4-3) from the funding body ‘Central Environmental Services’ has the Observer collaboration role.

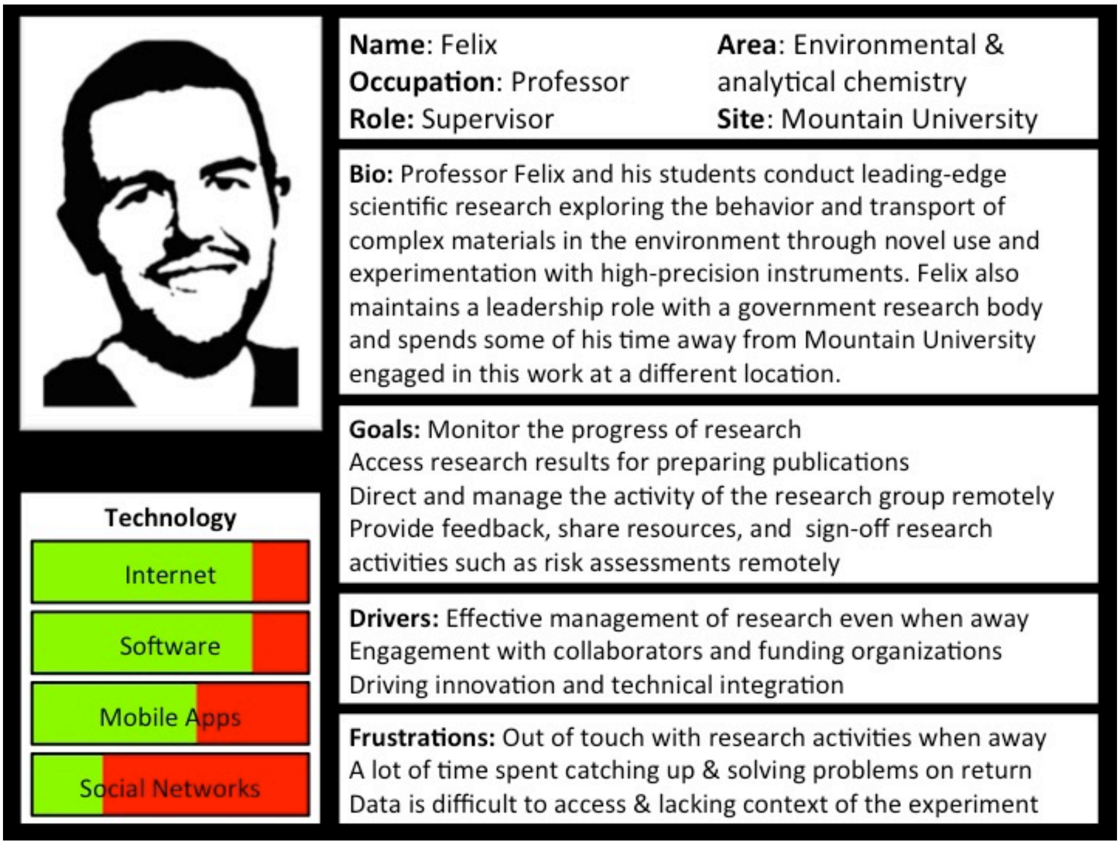
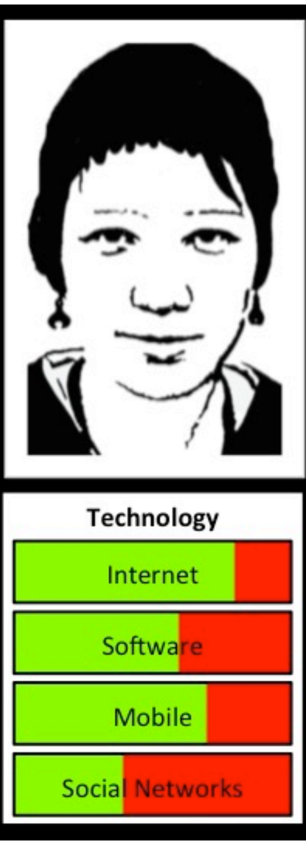


Figure 4-1: Supervisor persona 'Felix'



Name: Alejandra
Occupation: Researcher
Role: Researcher

Area: Environmental & analytical chemistry
Site: Mountain University

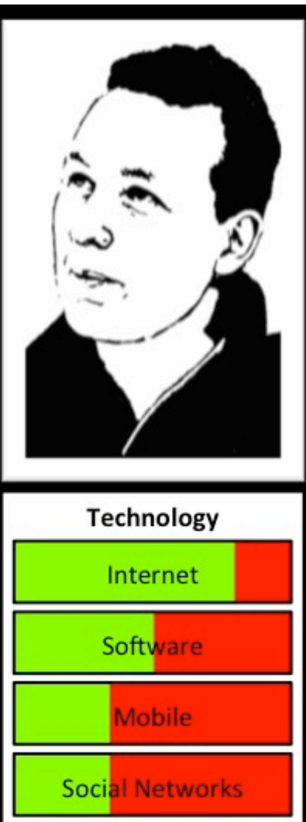
Bio: Alejandra is a research student in Felix's research team at Mountain University. She is in the third year of an externally funded PhD project investigating properties of materials in the environment using NMR analysis. Her experiments involve working in several different buildings and take a long time to run, so she may be running several experiments at a time. She helps Felix manage the research project and team while he is away.

Goals: Successfully complete experiments for her research
Produce results for publication
Get feedback, direction, and sign-off from Felix
Share findings and work with other researchers in Felix's team
Document & complete activities in line with funder standards

Drivers: Produce high quality research for PhD and future career
Meet the requirements of the funding body
Make writing of PhD thesis less painful

Frustrations: Funding body requirements means every page and lot number has to be signed & funder wants to check
Harder to get feedback and approval when Felix is away

Figure 4-2: Researcher persona 'Alejandra'



Name: Alvin
Occupation: Auditor
Role: Observer

Area: Funding body
Site: Central Environmental Services

Bio: Alvin is an auditor for Central Environmental Services, one of the funding bodies for Felix's research activities. He is responsible for checking that groups funded by the body are following the standards that are set for documenting their research & recording their results, and that the research is going to schedule. This involves visiting research locations & checking documentation for compliance and research progress.

Goals: Verify researchers are adhering to prescribed standards
Examine documentation produced by researchers
Examine results to determine if research is on schedule
Communicate and resolve any issues with the project lead and other funding body members

Drivers: Ensure research is conducted to high standards
Ensure research is progressing on schedule
Effective communication with stakeholders

Frustrations: Access to research data and documentation is restricted by time and geographic constraints
Compliance checking is manual and time-consuming

Figure 4-3: Observer persona 'Alvin'

4.3.2 Scenario 2: Fully open research including interaction with global scientific collaborators

At the City University, Professor Guy runs an organic chemistry research group including seven research students and two post-doctorate researchers. Guy is an evangelist for Open Science and works with a variety of international collaborators on projects to improve techniques for building organic molecules for use in the treatment of disease. Managing the interaction with the collaborators and trying to develop new relationships with relevant research groups is a challenge. The nature of the experiments in Guy's research group mean that the students are repeating very similar processes, and they save duplication by not recording information that they have already learnt. This makes it a challenge for new students to the group to understand the working practices of the team and to be able to take on the work of experimenters that have moved on. Guy's students find the risk assessment process to be onerous and involve unnecessary duplication. Information has to be manually extracted from databases or records, transcribed and then countersigned by Guy; the students would like to streamline this process to make it easier and less error prone. Students need to work with collaborators but find it difficult to keep up to date with their activities and to find specific information in their records. The team are interested in using an ELN to help with interacting with collaborators, to make some processes easier, and reduce duplication. They also want to use an ELN to provide a platform to encourage open science and generate new collaborations.

In Scenario 2, Professor Guy (Figure 4-4) has the Supervisor role, research students Ollie (Figure 4-5) and Lukas (Figure 4-6) have Researcher roles, and Professor Yuki (Figure 4-7) from the Island University, across the other side of the world, has the Collaborator role.

Name: Guy
Occupation: Professor
Role: Observer
Area: Organic chemistry and Open Science
Site: City University

Bio: Professor Guy and his students conduct research into developing ways to build molecules and to develop drugs to help treat tropical diseases. Guy believes strongly in the value of open science and collaboration to share techniques and speed up the process of drug discovery. Guy fosters relationships with collaborators across the world and is always looking for ways to engage new collaborators and share research.

Goals: Develop and refine new techniques to make molecules
Encourage and manage members of his research team
Share the techniques and research findings of the group
Identify new projects and engage new collaborators
Effectively co-ordinate research projects and collaborators

Drivers: Developing better techniques
Encouraging research sharing and collaboration
Demonstrating and improving effectiveness of open science

Frustrations: Methods & data are inaccessible in paper archives
Difficulty searching for relevant data by chemical structure
Difficult to identify and engage with potential collaborators

Technology

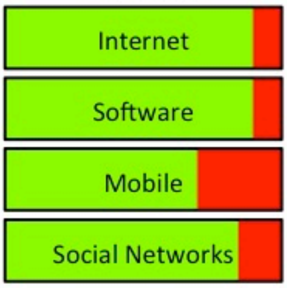


Figure 4-4: Supervisor persona 'Guy'

Name: Ollie
Occupation: Research student
Role: Researcher
Area: Organic chemistry and Open Science
Site: City University

Bio: Ollie is a research student in Guy's organic chemistry group. He is in the third year of his research project investigating better techniques to build a molecule that is used for treating disease. His experiments are completed in the lab and involve repeating similar experiments using different conditions, reagents, or steps to identify which yield the best results. Ollie will soon be moving on to a new project and will need to hand his work over to Lukas.

Goals: Successful completion of experiments to design a better method for development of the molecule under investigation
Share findings with collaborators and learn from their work
Access and make use of results outside of the lab for planning
Minimize effort for writing thesis and handing over to Lukas

Drivers: Contribute to important research methods
Improve reputation as a researcher through collaboration
Complete thesis and handover quickly to start new project

Frustrations: Repetitive process for risk assessment & planning
Difficulty accessing and using data from collaborators
Difficult handing over because own notes are not easy to use

Technology

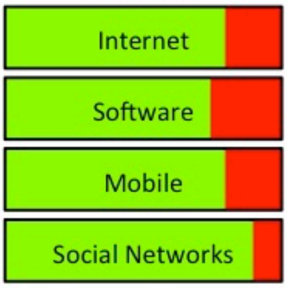


Figure 4-5: Researcher persona 'Ollie'

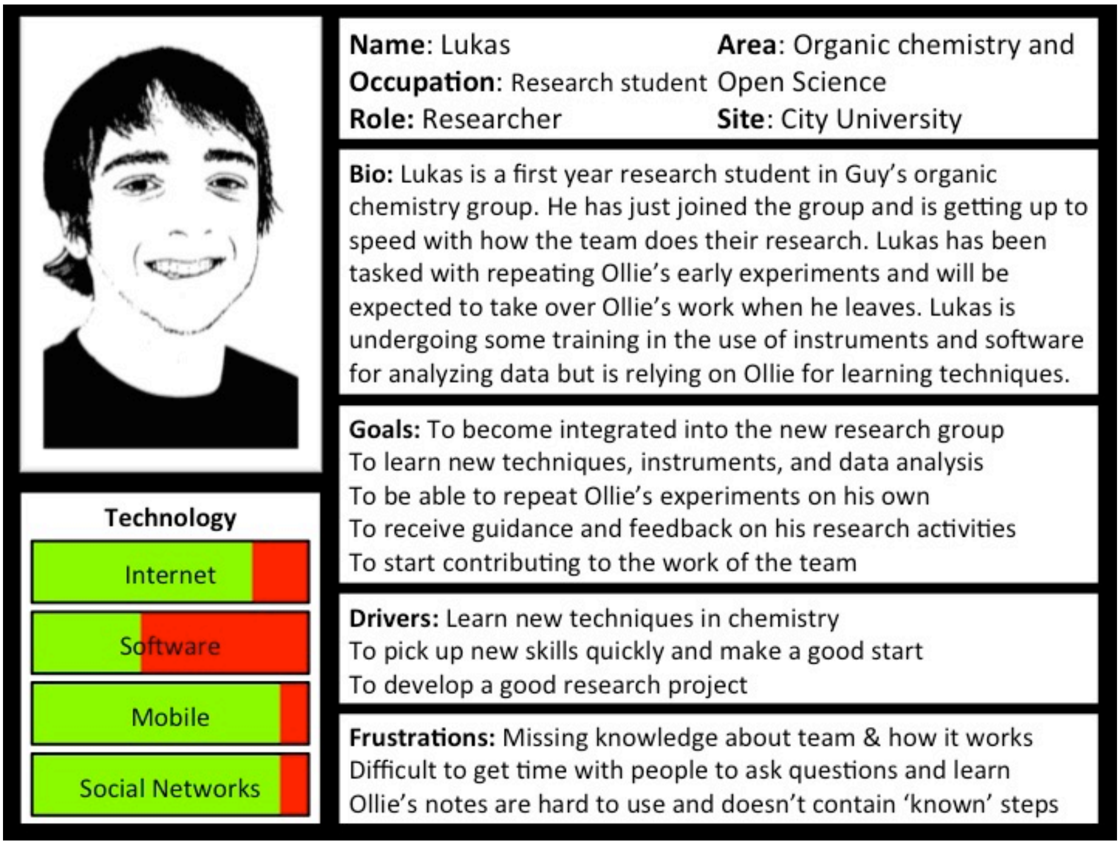


Figure 4-6: Researcher persona 'Lukas'

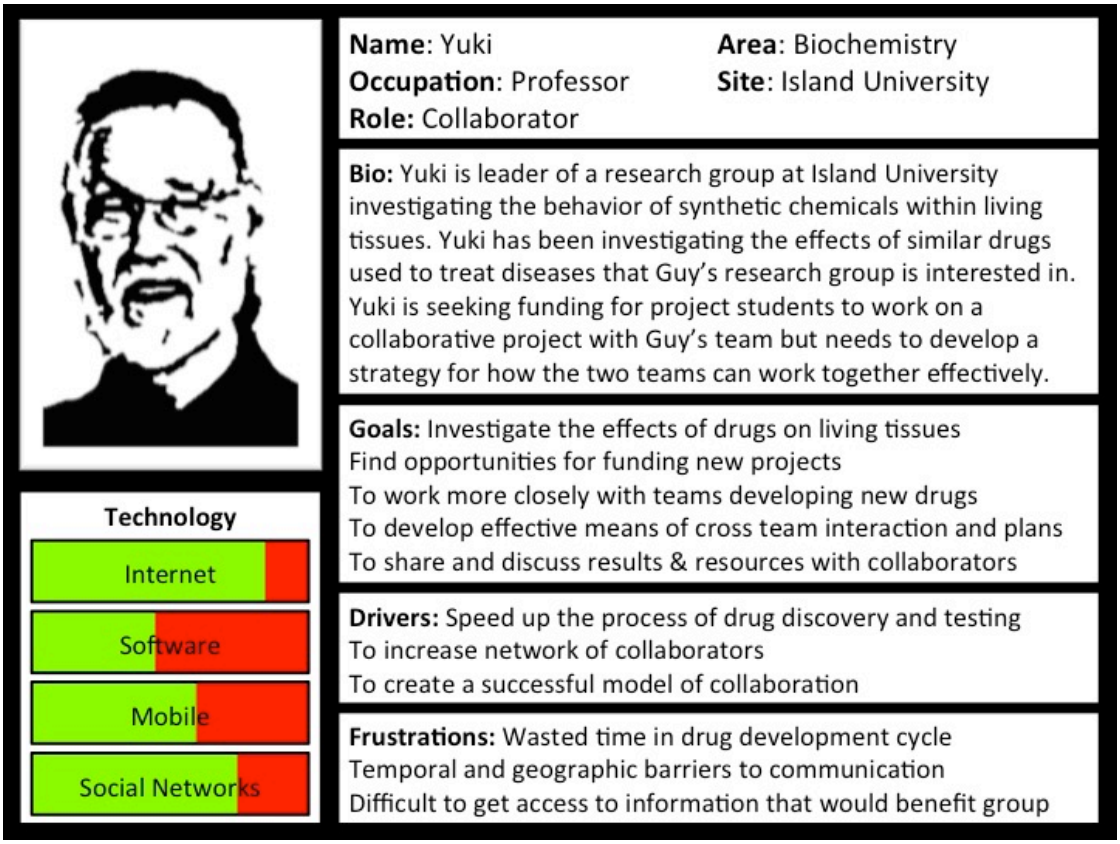


Figure 4-7: Collaborator persona 'Yuki'

4.3.3 Scenario 3: Closed research in a local research group

Dr Pal is a postdoctoral research student in the Synthetic Chemistry research group at Castle University. Pal's research is funded by an industrial funder and involves the development of novel techniques for synthesizing molecules. Pal's research has significant commercial value for the funder, and the resulting intellectual property and results must be protected from competitors. As a result Pal works independently and shares his results once a quarter with his managers who are located elsewhere. Pal is also responsible for supervising a small group of research students who also work independently on their own projects. Pal finds it difficult to keep a track of what the students are doing and making time to check their work; to do this he needs to sit with them to manually go through their notebooks and data, and to provide feedback. Pal would like to find a way to improve how he can manage both his own research and that of his students, so that both he and the students can get more regular and faster feedback on their work. An ELN has been suggested to Pal as a way that he might be able to improve the management of his own and his students' research. Although he is suspicious of using an ELN for his sensitive data, he likes the idea of being able to instantly access information about his students' activities.

Leo is one of the Pal's research students. He is new to the team and is struggling with the new techniques he has to learn, and with repeating experiments from the literature. The more experienced members of the research team have all gone through the experiments he is trying to replicate when they first joined, but he finds it hard to get time with them to ask questions and learn from their mistakes. Leo spends a lot of time trying different things to solve his problems and feels frustrated because he doesn't know if he is proceeding in the right direction. He doesn't want to waste Pal's time by asking him simple questions. In his undergraduate course at Highway University, Leo liked using a VLE system with an area where the students could ask teachers questions and get quick responses on their practical courses. Leo would like something similar so that he and the other new researchers could help each other and learn from more experienced staff.

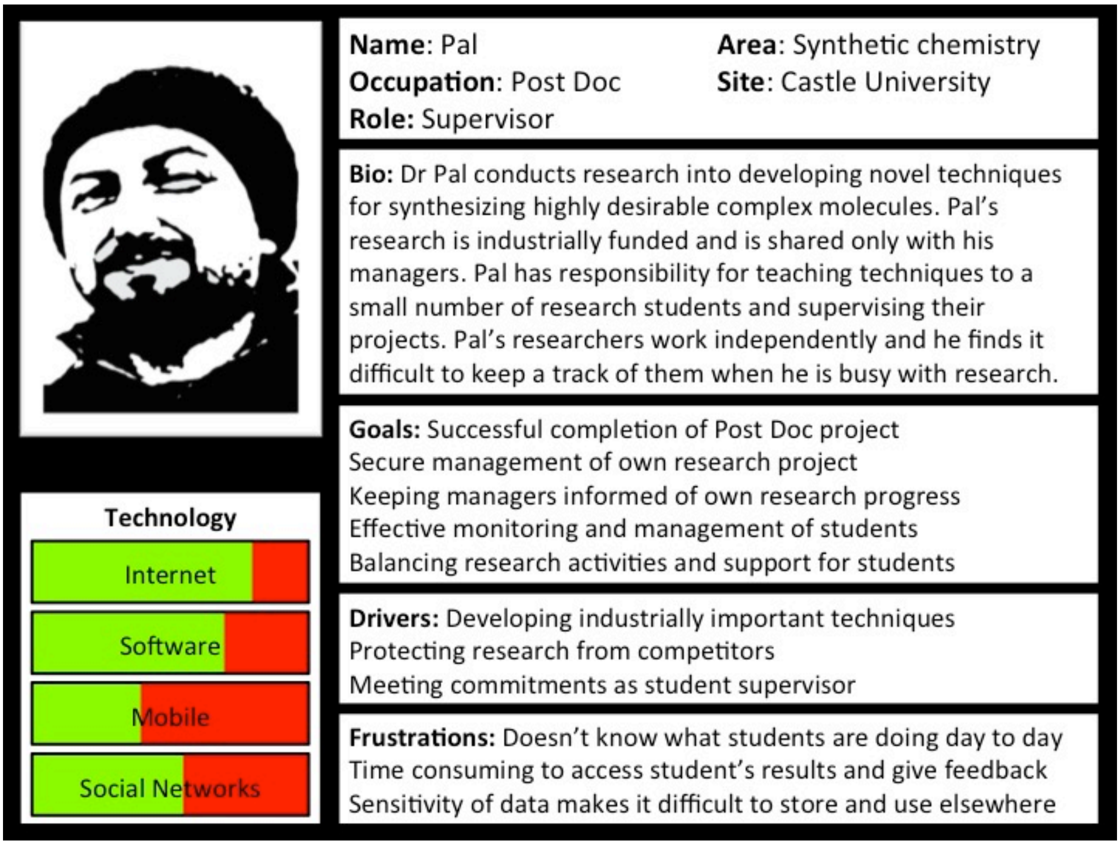


Figure 4-8: Supervisor persona 'Pal'

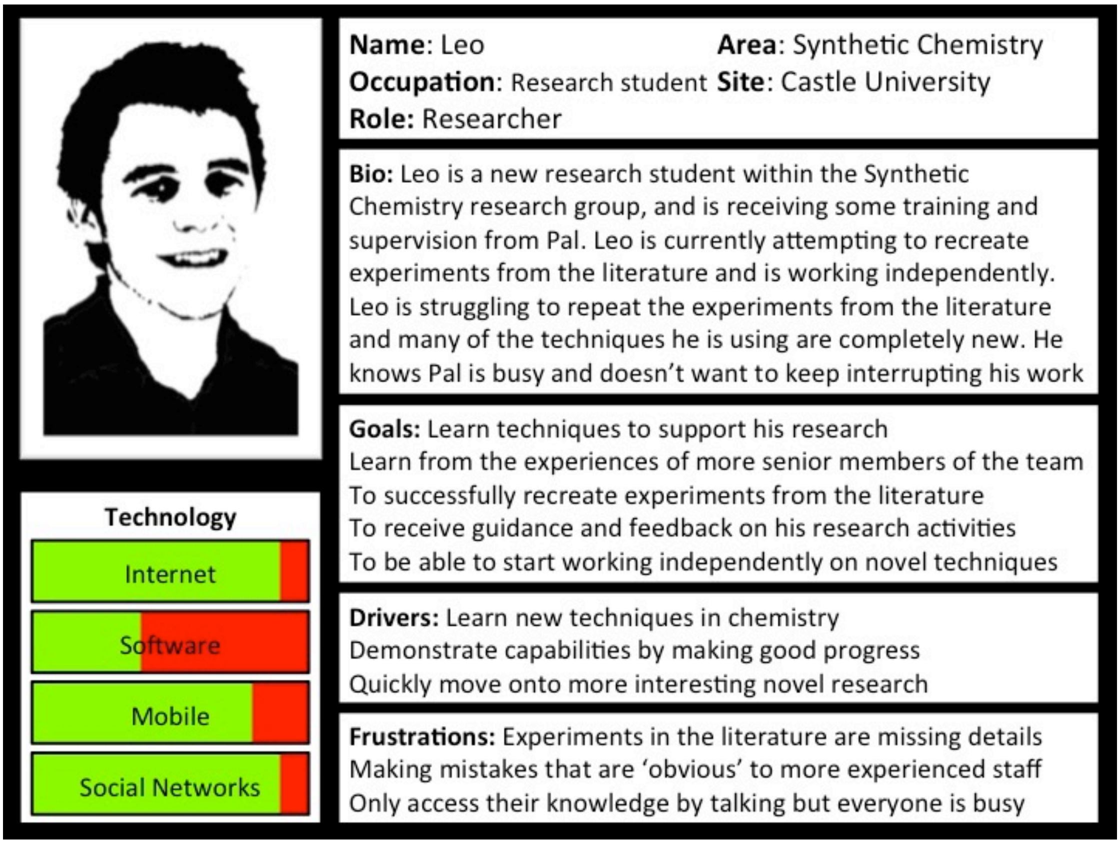


Figure 4-9: Researcher persona 'Leo"

In Scenario 3, Dr Pal (Figure 4-8) has the Supervisor role and research student Leo (Figure 4-9) has the Researcher role.

4.3.4 Scenario 4: Undergraduate practical course

Dr Jenny is a lecturer in the Physical and Analytical Chemistry group at Seaside University. Jenny is a technology enthusiast and is actively involved in projects to assess and implement technologies applicable to chemistry within the department, together with other collaborators outside the university. Jenny is keen to try out new things to help support her students and teach them analytical techniques. Although Seaside University has trialled ELNs with researchers, Jenny is keen to see the undergraduates use an ELN in their courses to teach them about data management and prepare them for future research or careers in industry. Jenny also hopes that an ELN system might streamline the student marking process, which is currently manual and very time consuming.

Dan is one of the ninety students on Jenny's Advanced Analytical Techniques course. Dan is a bright student and although he enjoys the practical work, he gets frustrated having to wait for the others students to finish before he can move on with tasks. He also gets frustrated with the feedback process, which takes a very long time and involves handing in his notebook after the practical sessions. He would prefer to get immediate feedback and to be able to use his notebook between practical sessions to remind himself about the previous experiment.

In Scenario 4, Dr Jenny (Figure 4-10) has the Teacher role and undergraduate student Dan (Figure 4-11) has the Student role.

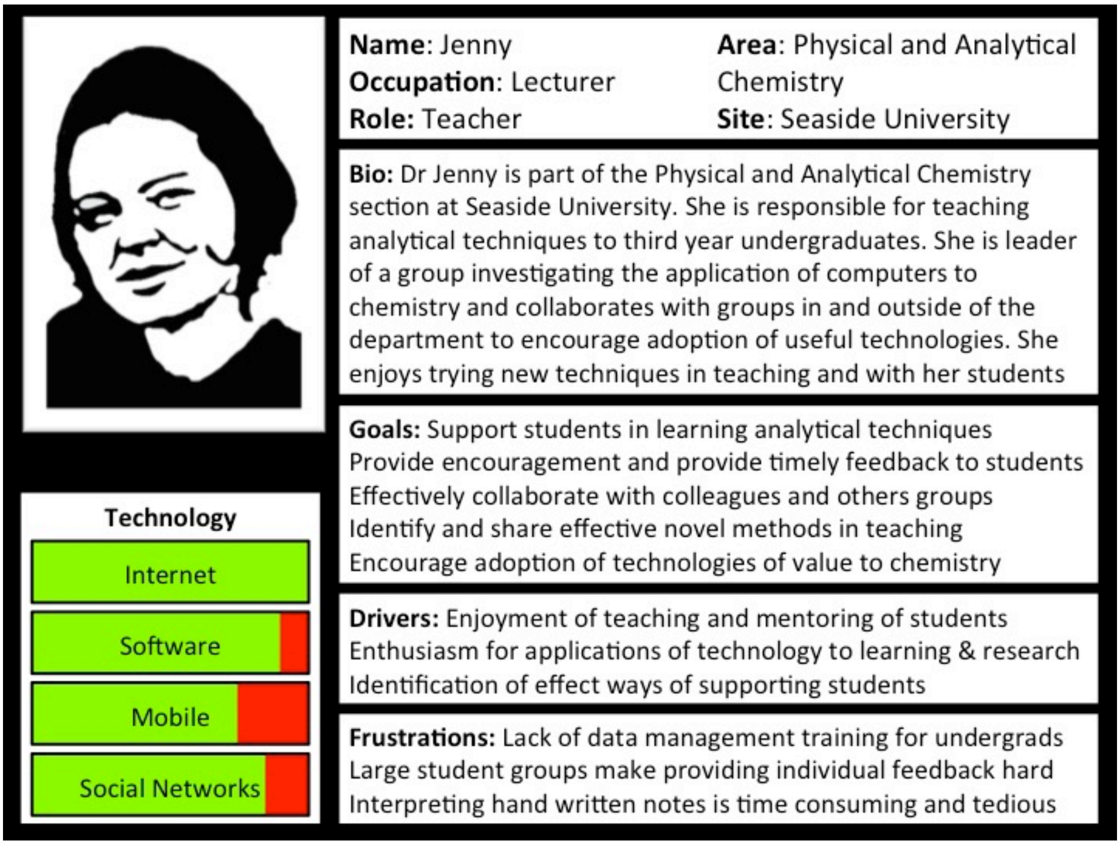


Figure 4-10: Teacher persona 'Jenny'

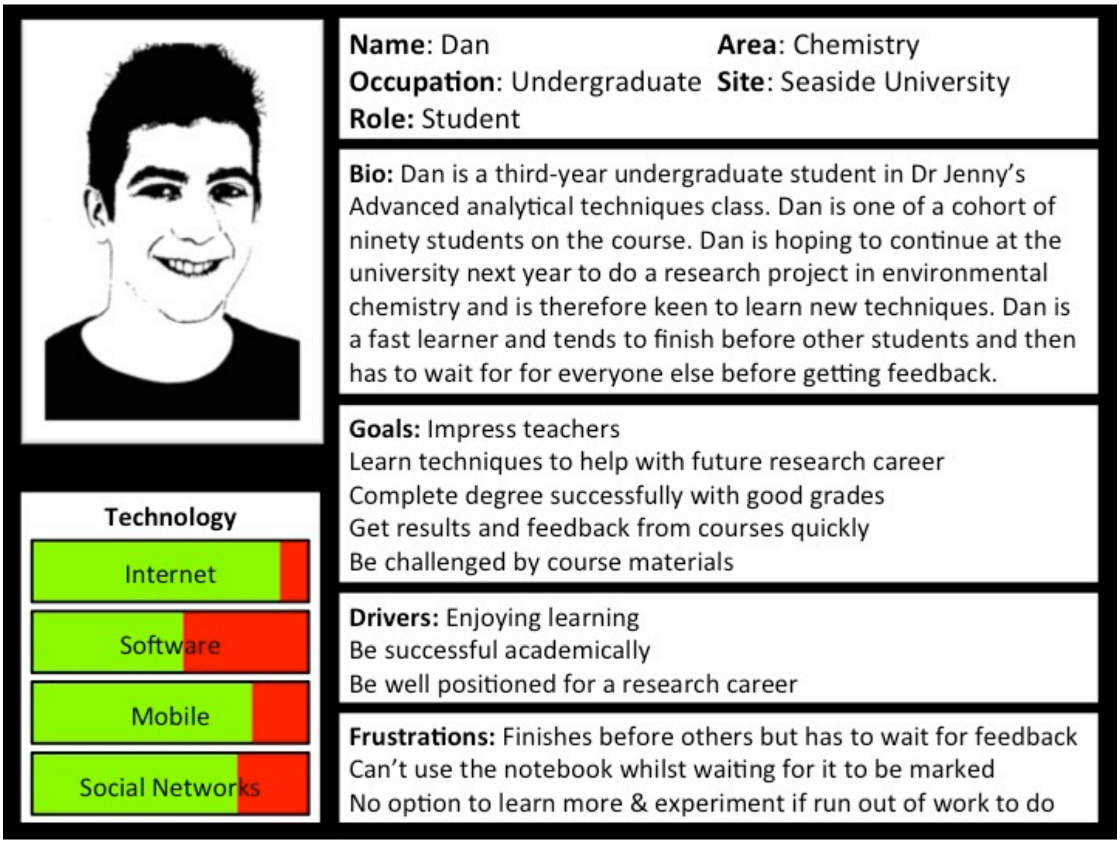


Figure 4-11: Student persona 'Dan'

4.4 Storyboards

Storyboarding is a useful way to transform scenarios and personas into a format to help develop specifications from tools, by providing examples of how tools might be used, and what they might be used to accomplish. Storyboards can help to highlight features that would provide a significant benefit to users, but that have not previously been considered, as well as how to handle user expectations about more obvious tasks. Figure 4-12 to Figure 4-18 show some example storyboards expanding on the scenarios and personas detailed in 4.3.

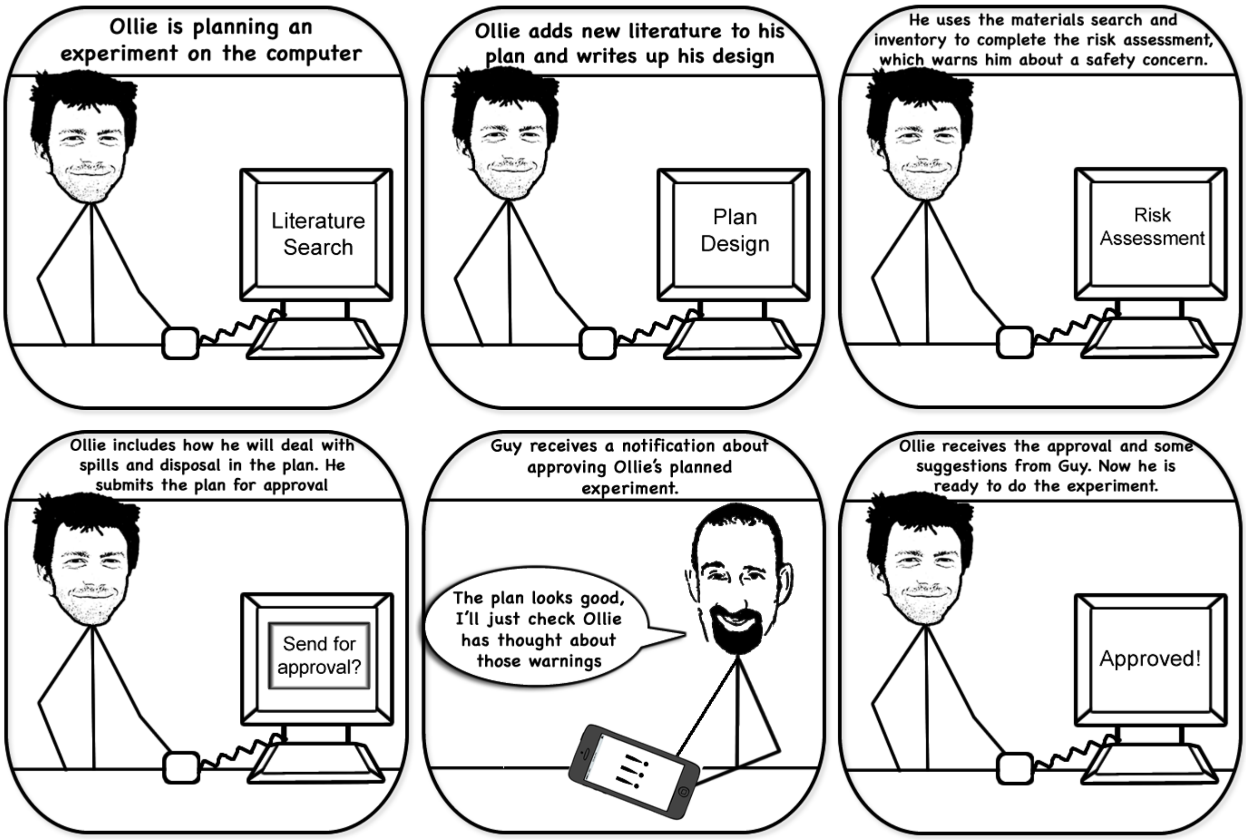


Figure 4-12: Ollie & Guy storyboard - planning an experiment

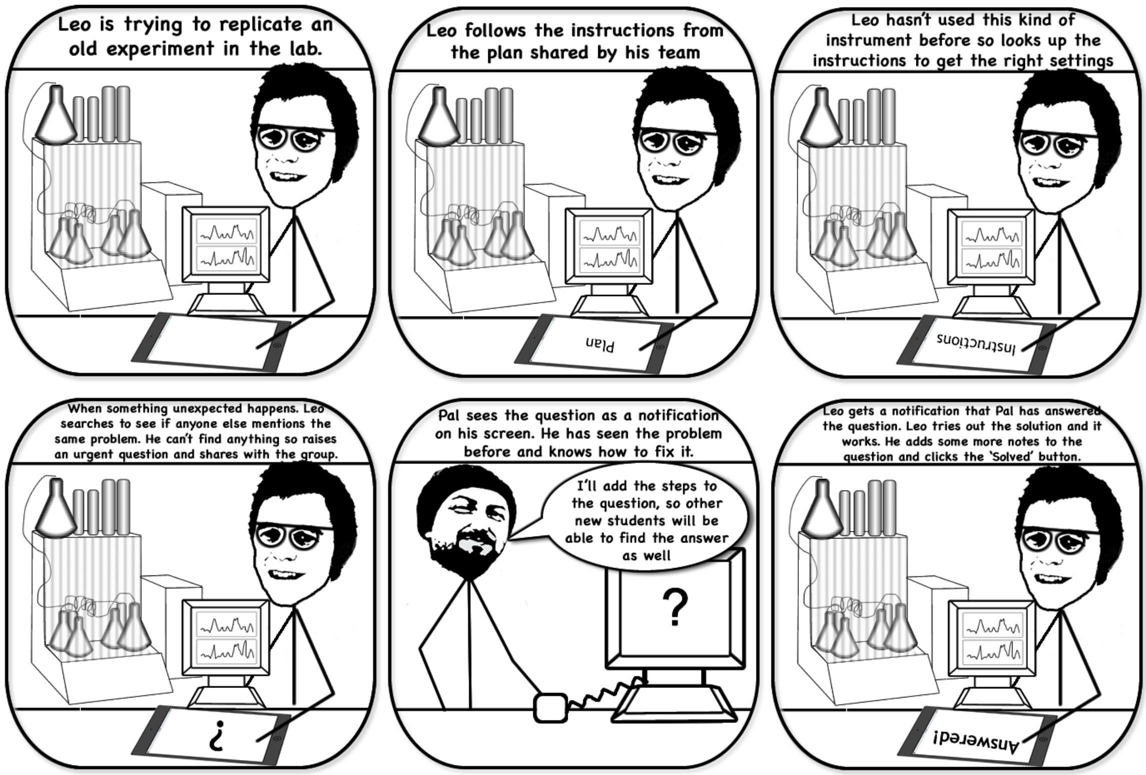


Figure 4-13: Leo & Pal storyboard - running an experiment

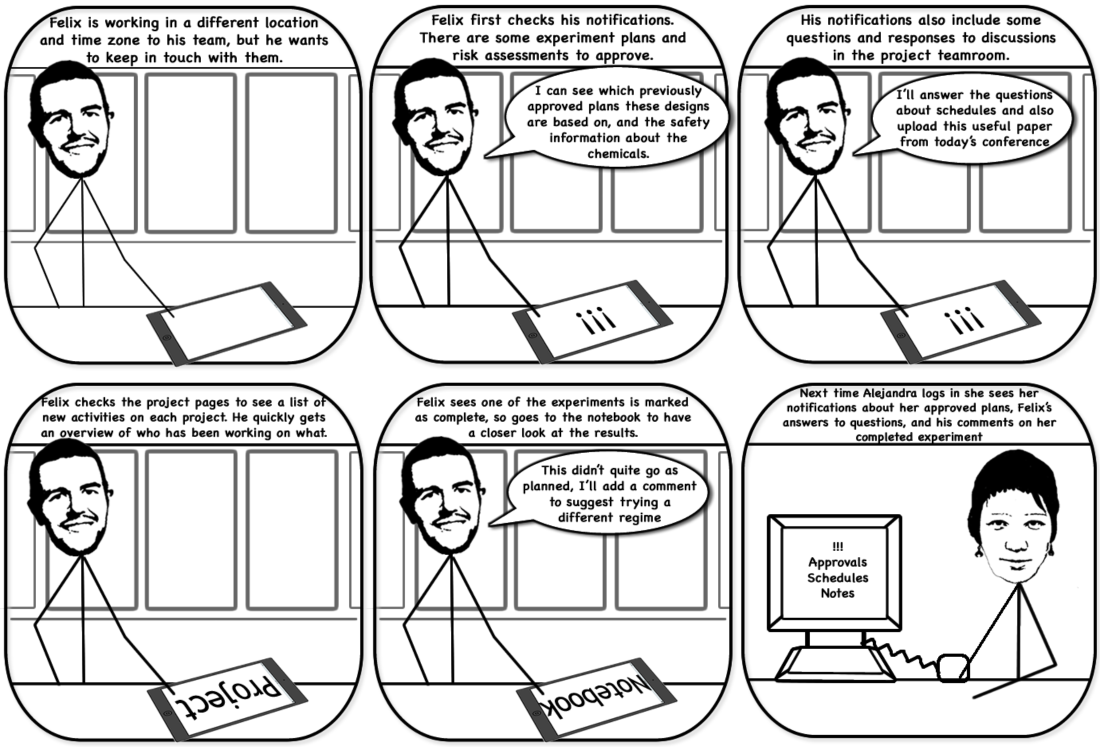


Figure 4-14: Felix & Alejandra storyboard – managing a team remotely

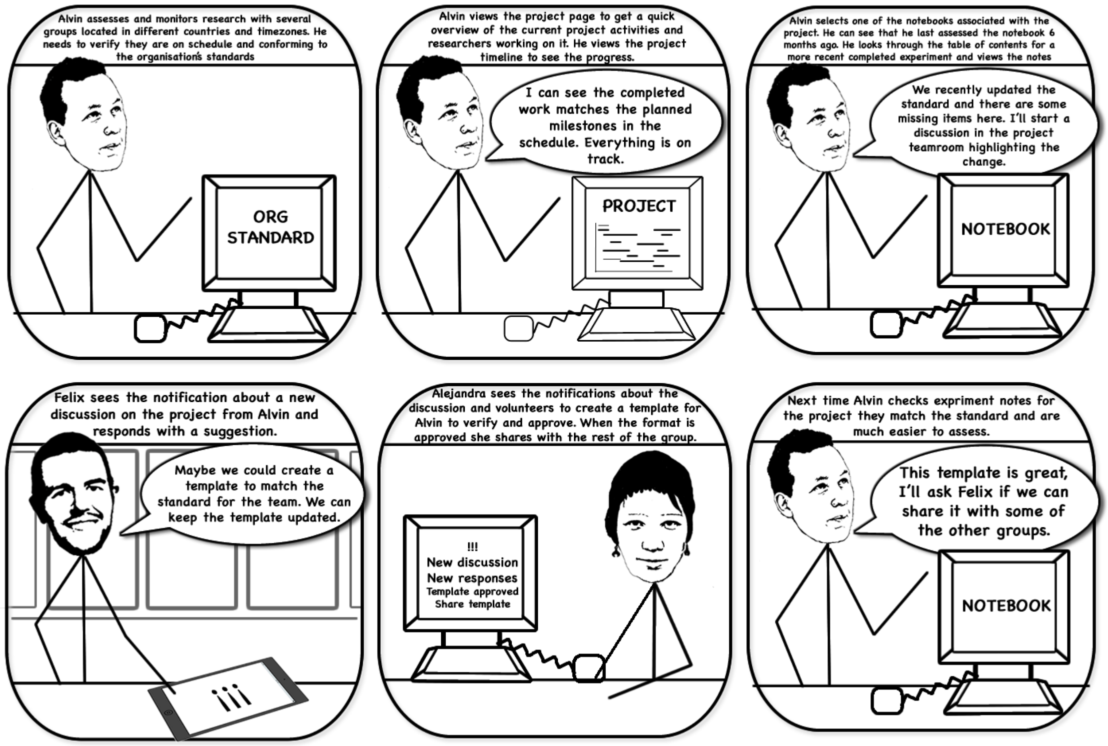


Figure 4-15: Alvin, Felix, and Alejandra storyboard – monitoring a project remotely

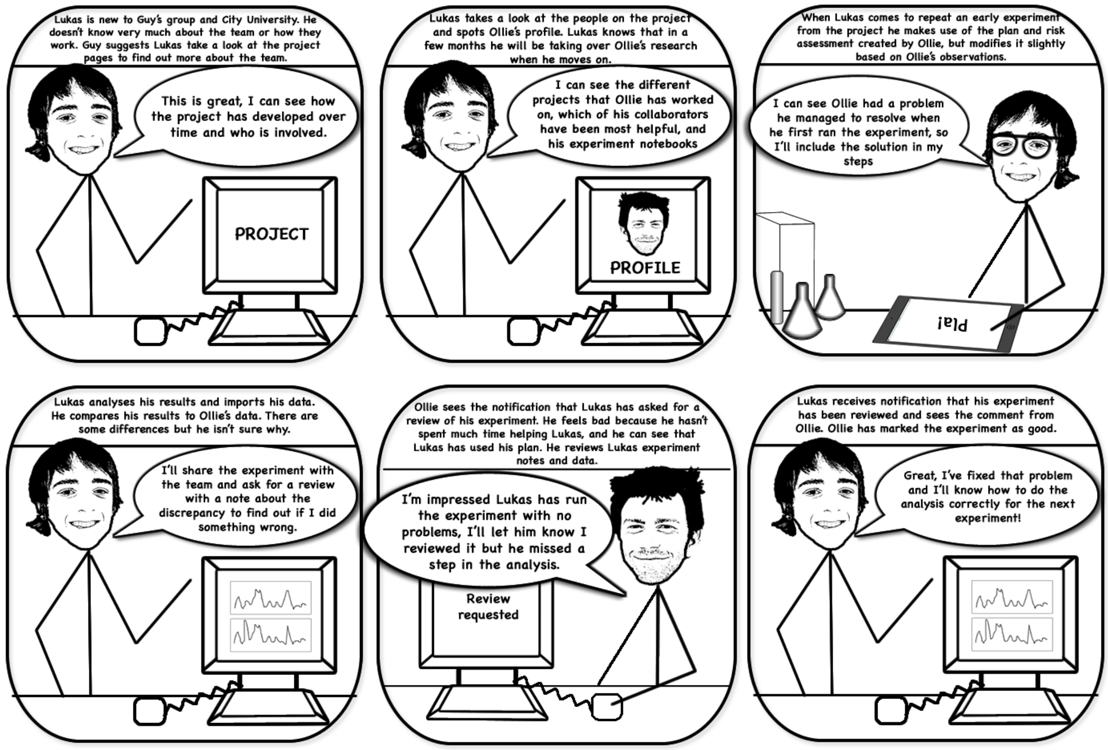


Figure 4-16: Lukas & Ollie storyboard - integrating a new team member

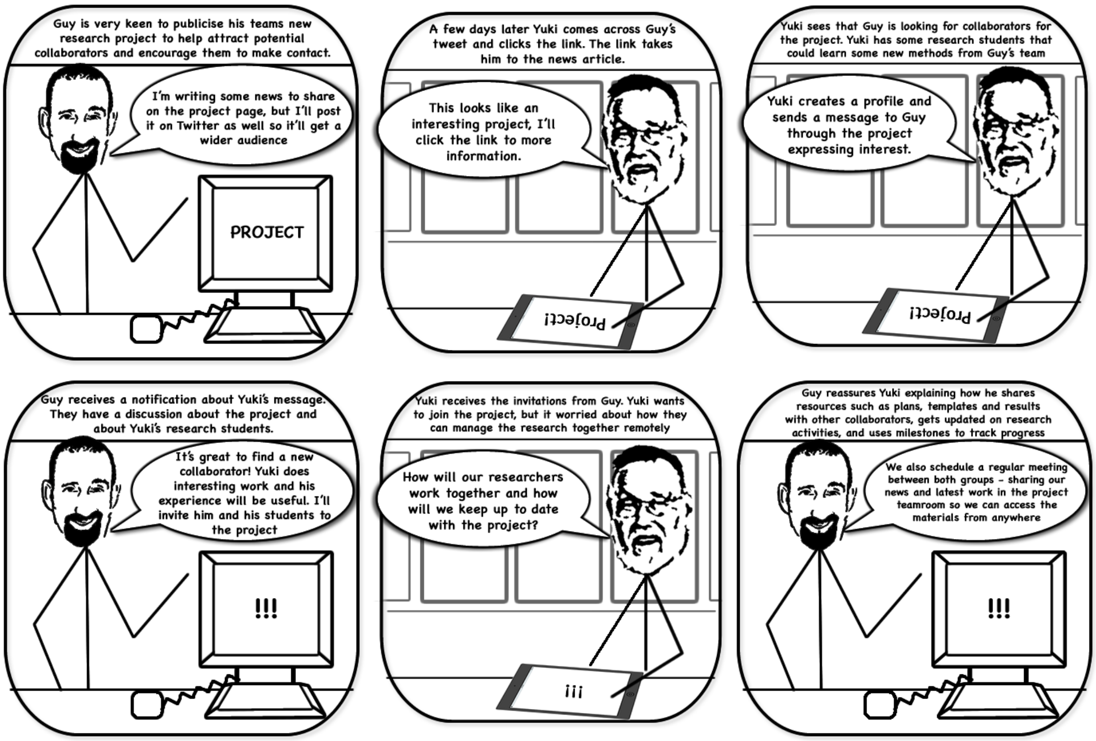


Figure 4-17: Yuki and Guy storyboard – research collaboration and project management

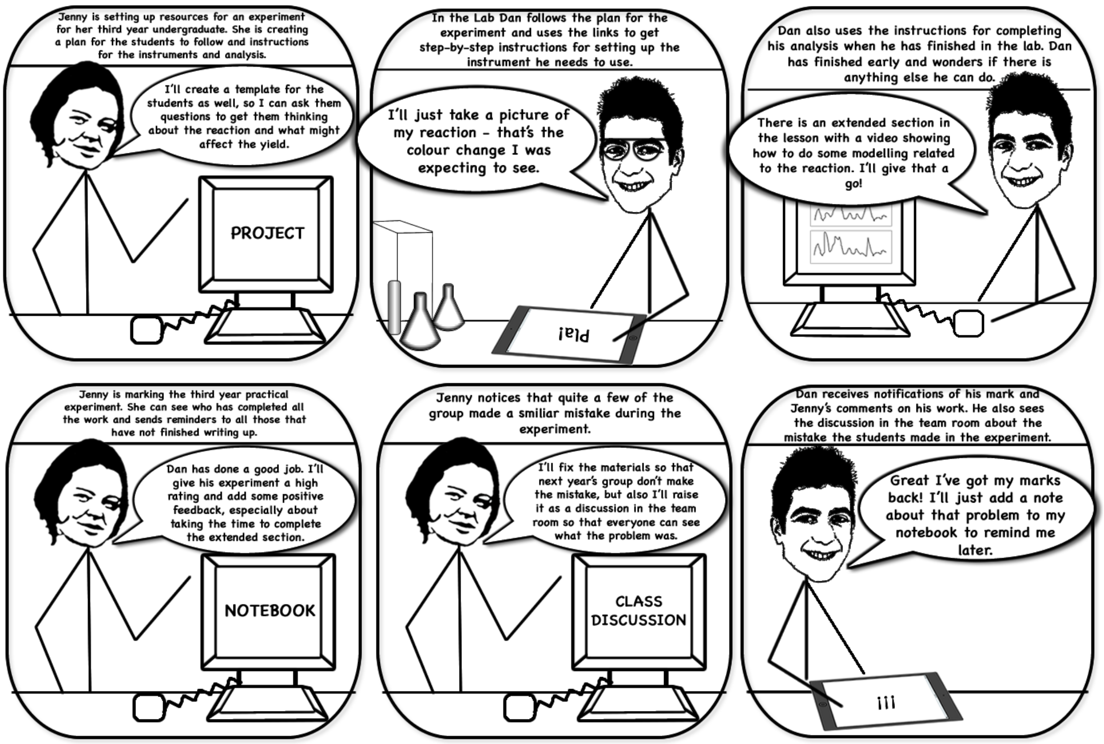


Figure 4-18: Jenny & Dan storyboard - undergraduate practical

Unexpected requirements that were generated from the storyboards above are described in Table 4-1, together with a reference to the associated storyboards.

Table 4-1: Unexpected requirements from storyboards

Unexpected requirements	Storyboards
Highlight if a plan has been created from a plan that has been approved, what the changes are, and who was the approver.	Figure 4-14
Include a facility to ask questions, mark questions as 'solved' so they can provide tips to others, and have an urgency or importance associated with a question.	Figure 4-13
Mark a notebook as assessed and enable notification when a new assessment is due.	Figure 4-15
Include an approval method for templates as well as plans and risk assessments.	Figure 4-15
Include gamification elements (e.g. trophies, badges, & ratings) to encourage helpfulness, interaction, answers to questions, and to encourage the quality capture and design of plans, experiments, & templates.	Figure 4-13, Figure 4-16, Figure 4-18
Include buddy checking and rating of experiments	Figure 4-16, Figure 4-18
Enable team members to create project updates that can be shared as news privately and publicly, and enable sharing on social media.	Figure 4-17
Enable private and public discussion on a project/lesson.	Figure 4-17, Figure 4-18
Enable templates to be used to ask students questions as well as capture records for practical assignments.	Figure 4-18
Send students reminders and automate reminders.	Figure 4-18
Enable linking between experiments, 'annotations', and project discussions and resources.	Figure 4-18

4.5 Next steps in designing tools

Further activities within interaction design are focused on narrowing down on specific requirements, for example making use of essential use cases to describe the idealized requirements of a software system without the constraints of any particularly technology for implementation. Use cases also identify the anticipated interactions between actual users and the system under design. The final stage before implementation is prototyping, where interface designs of varying fidelities can be developed to try out the interactions derived from the use cases.

At a high level it is relatively easy to come up with straightforward designs for many of these features and interactions, borrowing from standard interfaces that researchers are likely to be familiar with, particularly those for social networking that include many of the kinds of feature that would be appropriate for the scenarios and storyboards included above. The difficulty comes with designing interfaces for capturing the experiment records themselves.

The studies in **Chapter 5: Asking the right questions: Effects of Using Structured Templates for Recalling Chemistry Experiments** investigate how the format of interfaces might impact on the information that is captured in an experiment**. Chapter 6: Use of user-defined metadata in ELNs and other platforms** investigates how users in the LabTrove ELN and other platforms currently use metadata to describe experiments, and **Chapter 7: Generating metadata by using cues and changing perspectives** investigates whether they are better ways to capture metadata. In **Chapter 9: Notelus: An Exemplar Mobile ELN** we make use of our findings from **Chapter 3: Keeping and managing records** to build a mobile ELN for supporting researchers in their experiments, capturing experiment information in a variety of ways, and capturing relevant metadata.

Chapter 5: Asking the right questions: Effects of Using Structured Templates for Recalling Chemistry Experiments

5.1 Introduction

The primary function of an Electronic Laboratory Notebook is to capture the experiment record. As mentioned in Section 2.2.3, templates could be used to standardise the information that is capture within an ELN, but there is some evidence that the relative inflexibility and formality of templates may have a detrimental effect on the information that is captured. Before the studies described in this chapter, it was unknown whether using templates would have a positive or negative impact on the recording of experiment-related information. We had expectations that the information we specifically asked for in the templates would be recorded, but that some information might be lost. The research from the studies discussed in this chapter has been prepared for submission in the following paper:

> Willoughby, C., & Frey, J.G. Effects of Using Structured Templates for Recalling Chemistry Experiments (In preparation, April 2015)

This chapter provides a brief background to the research included in the paper, the methods that were used, a brief discussion of the major findings, and the significance of the findings in terms of designing tools for scientific researchers.

5.1.1 Research Questions

The following research questions were investigated as part of these studies:

- What information about an experiment do students and researchers record on a blank piece of paper?
- Does using a structured format change what information about an experiment is recorded?
- Does providing cues using a template lead to the recording of the information prompted by the cues?

Using structured templates

- Does providing cues using a template lead to the recording of less personal information about the experiment experience?
- Does asking personal questions increase the amount of personal information about the experiment experience that is recorded?
- Is the user's experience of using different templates to record information different?

5.2 Background

As we have discussed in Section 3.3, Learning to keep records, what researchers record in their notebooks is likely to be a combination of what they have learnt, the professional practices they have been exposed to, and their own personal style. However, when it comes to ELNs, it is likely that the interfaces in the ELN will influence what is recorded. In a flexible interface where the user is provided with a simple text box, we might expect that researchers will record much the same information as they would record in their paper notebook. If the interface is instead a formal 'form' or template style interface, then we would expect different information to be recorded. Clearly, for an ELN to be fit for purpose, it must be able to capture the experiment record at least as well as a paper notebook. Providing a structure for capturing the experiment might provide an advantage over paper notebooks by encouraging consistency, promoting good record-keeping practices, and reducing the risk of mistakes (Milsted et al., 2013; Elliot, 2004).

The findings of psychology and related memory studies, such as those detailed in Section 2.2.2, The role of memory and knowledge in capturing the experiment record, suggest that using a template to provide cues, would act as reminders to the researchers, resulting in an increased chance that the cued information would be remembered and recorded. However, we had concerns about possible negative impacts of using templates stemming from reports by researchers that ELNs which enforce a rigid structure for capturing experiments were too cumbersome and inflexible, and not suited to the less formal and more haphazard nature of academic research (Milsted et al., 2013).

A number of studies were undertaken in order to investigate whether using templates made any difference to the information that was captured. The details of the studies are provided in Table 5-1, and the methods used in each study are detailed below.

Table 5-1: Conditions, participants, and settings for studies investigating using structured templates

	Setting	Participants	Conditions
Study 1	Organic Chemistry Summer School, University of Southampton	20 chemistry undergraduates	• 'No Template' • 'Template'
Study 2	Organic Chemistry Summer School, University of Southampton	20 chemistry undergraduates	• 'No Template' • 'Titles Template' • 'Profile Template'
Study 3	KTH note-taking workshop and team-building day	~65 chemistry researchers & staff	• 'No Template' • 'Titles Template' • 'Profile Template'

5.3 Study 1: Organic Chemistry Summer School and paper-based templates

The first study was carried out using students taking part in an Organic Chemistry Summer School (OCSS) at Southampton University. 20 undergraduate students are invited to take part in the OCSS each year. In the first week of the summer school the students complete the following 3 experiments:

Reaction 1: The Preparation of an Organoboronic Acid
Reaction 2: Radical Benzylic Bromination
Reaction 3: Carbon-Carbon Bond Formation using the Suzuki Reaction

On the first day the students complete either Reaction 1 or Reaction 2, on the second day they all swap over to complete the other reaction, and on the third and fourth day all complete Reaction 3 using the products generated in the first two reactions. In Study 1, the students were asked to produce a write-up of their experiment for just the first two reactions, but all three reactions were included in Study 2 described below.

5.3.1 Methods

At the end of Reaction 1 & 2 all the students were asked to write a report of their experiments in one of two conditions: 'No Template' and 'Template'. For the second experiment they were swapped to the other condition. Both conditions used paper questionnaires, with the 'No Template' questionnaire essentially a blank piece of paper, whilst the 'Template' questionnaire contained high-level headings based on the Southampton University guidelines; the titles acting as cues for information that we might be expect to be recorded in an experiment, based on the student guidelines. The questionnaires that were used in Study 1 are available in Appendix A.1.

Each report was marked and graded by an independent marker familiar with the experiments, and the contents of each report were analysed using quantitative and qualitative analysis. For example, the number of words in each report was calculated, the mark for each report was converted into numerical form, and the number of types of information in each report were counted, such as the number of aims, reaction schemes, relative molecular masses (RMMs), materials, equipment, actions, explanations, observations, results, and measurements. Other information, such as whether sections were missing from the template, the writing style, and whether the steps were recorded in the correct order were also noted.

5.3.2 Results

What was perhaps most surprising about the results of Study 1 was that such a simple template using a few headings made such a difference to the information that was captured. The most significant findings from Study 1 were as follows:

- On average more words are used in the 'Template' condition reflecting a greater amount of information captured.
- The majority of students received a higher grade in the 'Template' condition.
- The reports in the 'No Template' condition focused on the details of the procedure conducted in the experiment.

- Explanations and observations are more common in the ‘No Template’ condition; and around half of these occur outside of the experiment procedure.
- The percentage of students that included Aims, Reaction Schemes, RMMs, and Results in the their report was significantly higher in the ‘Template’ condition.
- Significantly more theoretical information about the experiment is present in the ‘Template’ condition.
- Learnt information, perceived success of the experiment, and discussions were only present in the ‘Template’ condition.
- A large number of the students switched style between the two templates, with the majority of students using the past tense in the ‘No Template’ condition and half of those using a ‘command’ or instructional style in the ‘Template’ condition

The results of the study indicated that templates were successful at standardizing and increasing the amount of information that was captured by the students, at least in certain areas such as aims, reaction schemes, and results, but that there was a reduction in the amount of observations that were made and explanations about actions taken during the experiments. Although observations and explanations were captured, they were generally fewer in number in the template, and were not captured within the procedure of the experiment and were therefore disconnected with the actions and events that generated them. The study also revealed that the template was more effective at capturing information that would be useful in academic environments, with more learnt and theory about the experiment captured in the Template condition. Study 2 was devised as a follow-on study to investigate if the results could be replicated, and to add a new style of template asking more personal questions as cues to see if these questions might help prompt the recall of more personal information to recapture the information lost in the title-based template.

5.4 Study 2: Organic Chemistry Summer School and computer-based templates

The second study was also carried out using students taking part in the OCSS in a subsequent year using 20 undergraduate students. This study was conducted to verify the findings of the first study, and also to determine whether a different format style might influence the information that was captured in the template. Based on the findings in Study 1 the new template included a series of questions to investigate whether using personal questions would result in the capture of more personal responses and overcome some of the drawbacks of the original title-based template. In this study the questionnaires used for the different conditions were computer-based rather than paper based to assess whether the computer-based format created any limitations for recording.

5.4.1 Methods

This study investigated three conditions, the same 'No Template' condition as Study 1, the template from Study 1 one became the 'Titles Template' condition for Study 2, and a new 'Profile Template' condition with the addition of the new template style using questions as cues instead of headings. The 'Profile Template' also included some additional fields to investigate the capture of metadata relevant to experiments and the results of the metadata additions, but these elements are discussed in more details in **Chapter 7: Generating metadata by using cues and changing perspectives**.

The 3 reactions from the first week of the OCSS, as described above, were used to compare the different conditions, with Reaction 1 or 2 recorded using the No Template or Titles Template questionnaires, and all the Reaction 3 experiments recorded using the Profile Template questionnaire. All the questionnaires that were used in Study 2 are available in Appendix A.2.

Each report was marked and graded by the same independent marker as used in Study 1, and the contents of each report were analysed using quantitative and qualitative analysis in the same way as before.

5.4.2 Results

Overall the results of Study 2 backed up the important findings of Study 1 – that using a template changes the information that is recorded compared to a blank piece of paper, and that using cues encourages the recall and recording of the prompted information. An advantage of the computer-based templates was that completing fields was mandatory, and as a result almost all of the templates were completed in full, in comparison to Study 1 where more than half failed to complete all of the sections in the 'Template condition'.

The most significant findings from Study 2 were as follows:

- On average more words are used in both of the template conditions compared to the 'No Template' condition.
- The grades of the majority of students are higher in the two template conditions.
- Students only included Reaction Schemes and RMMs in the their report in the Titles template condition; with no reaction schemes and RMMs included in the 'Profile Template' or 'No Template' conditions.
- All of the students in the template conditions included Aims and Results; in the No Template condition a minority of students included Aims and Results.
- In the 'Profile Template' condition the amount of detail about the experiment procedure is significantly reduced compared to the other two conditions; and often contains a list of techniques rather than specific steps
- The number of students recording explanations about why they performed actions in the experiment is much higher in the No Template condition, but the number of explanations related to analysis and learning are much higher in the template conditions.
- More observations are seen in the template conditions, in contrast to Study 1, but the majority of these are not included as part of the experiment procedure description in these conditions.
- The vast majority of conclusions or 'success statements' are recorded in the template conditions, together with significantly more information about theory and learning recorded in these conditions.

- No clear pattern of style change is seen between the different conditions, in contrast to Study 1, with the majority of reports written in the past tense; but a new 'list-like' style is seen in the computer-based templates

5.5 Study 3: Lego rocket cars templates and group discussion

The final study had a different format than the first two studies, and was carried out using groups of academic staff and research students from KTH in Stockholm. These researchers and academics in chemistry were taking part in a note-taking workshop as part of a team-building day. We have already mentioned the results of some of the other activities that were included in the note-taking workshop in 3.4. For the purposes of this study the staff and students were randomly divided into one of 15 mixed groups and asked to complete an experiment using a Lego car powered by an Alka-Seltzer reaction, and asked to record their experiments within one of three Google Docs[16] templates allocated to each group. The instructions used in the experiment are available on-line[17].

This study was conducted to verify the findings of the first two studies with a different experiment and a different type of subject. Although this study was less formal than the first two, the workshop provided a valuable opportunity to get feedback on the experience of using the different templates through a post-experiment discussion, where groups using different templates were brought together to discuss and share their experiences.

5.5.1 Methods

This study investigated the same three conditions as Study 2, a 'No Template' condition, a 'Titles Template' condition, and a 'Profile Template' condition, although there were some differences in the format of the templates. An addition was made to the 'Titles Template' to include a Results table with column headings appropriate to the experiment in the study. An addition to

[16] http://www.google.co.uk/docs/about/
[17] https://legorocketcars.wordpress.com/instructions/

both the Titles and Profile templates was a section about the “Plan” for the experiment. All of the templates also included some additional fields to investigate the capture of metadata relevant to experiments and the results of the metadata additions are discussed in **Chapter 7: Generating metadata by using cues and changing perspectives**.

The templates used in Study 3 were in the form of Google Docs, and are available in Appendix A.3. Each group was given a URL to one of the templates based on their group number. The contents of each report were analysed to compare them to Study 1 and Study 2. No grading was assigned to the reports.

After the experiment was finished, the groups were combined into 5 groups with representation from each of the different template styles to discuss the results from using the different template conditions. After the groups had a chance to discuss their findings between themselves, each group was asked to present their findings to us. Notes were taken of the reported findings from each group that provided an insight into the usability of the different template types and issues encountered by the teams.

5.5.2 Results

The results of Study 3 do confirm the results seen in studies 1 & 2 in that cues were effective for encouraging the recall of the prompted information, but this study also high-lighted some differences in behaviour with the ‘No Template’ condition, and the ‘Titles Template’ condition in particular. The nature of the experiment was rather different with the KTH subjects repeating the same experiment under different conditions and making measurements to compare the effects of the changed conditions. This difference between the experiment formats appears to have produced some ‘self-structuring’ within the ‘No Template’ condition where teams have created their own structure typically containing method, results, measurements, and conclusions.

The addition of the ‘Results table’ in the ‘Titles Template’ condition had a significant undesirable effect that had not been anticipated. The presence of the table resulted in less information recorded in this template compared to the other two conditions. The presence of the table also had a negative impact on the experiment itself, with teams commenting that they had less time to do the experiment because they spent most of their time completing the table.

Using structured templates

The other significant findings from Study 3 were as follows:

- The majority of teams included photographs and/or videos in their reports in the ‘No Template’ and ‘Profile Template’ conditions, but none were included in the ‘Titles Template’ condition.
- In general the word count and the amount of information captured in the ‘Profile Template’ condition was higher than the other conditions.
- More information about significant results and what was learnt was captured in the ‘Profile Template’ condition than in the other conditions.
- The experiment procedure was less detailed in the ‘Profile Template’ condition than in the ‘No Template’ condition.
- Including a field for planned activities revealed differences between the design of the experiment and what was actually done.

The group discussion provided some valuable insights into the experience of using the different templates from the researcher’s perspective. The following findings are the main observations made by the majority of the groups:

- They were surprised at how different the reports were when the different templates were used.
- The ‘No Template’ condition provided more flexibility and freedom.
- The ‘Profile Template’ condition led to the capture of more information.
- Information captured in the ‘Profile Template’ condition had been forgotten in the ‘No Template’ condition.
- The ‘Profile Template’ was laborious to complete.
- The ‘Profile Template’ condition would be a good starting point for developing personal templates.
- They would prefer to develop their own template based on personal needs.

5.6 Discussion

The studies described in this chapter provide strong evidence that the design of interfaces for capturing experiments will have an impact on the information recorded. The findings also indicate that relatively subtle differences in designs may have a large impact on both the information that is recorded and that some designs could have a detrimental impact on the way an experiment is actually performed.

Providing a formal structure for recording experiments has some benefits. The amount of information captured is greater and information that would otherwise be forgotten is recorded and is available for later use. The cues that are used do influence the information that is recorded, with the presence of a cue encouraging the recall and recording of that information, and often additional information such as the theory or learnt information related to the experiment or the perceived success of the experiment. For students, the grades that they obtain could be improved by using a template rather than a blank sheet of paper.

Different types of experiments may lead to different types of information being captured without a template. For example, in the synthesis reactions the main information that was recorded was the personal memory of the procedure of the experiment, whereas the Lego Rocket Cars experiment generated 'self-structuring', with more information recorded about results and conclusions.

The use of cues does have some unexpected effects. Sometimes a simple prompt like the heading 'Aims' is all that is required to get the information requested, but the studies also demonstrated that the precise wording can make a significant difference to what is recorded and how. This effect is seen most clearly for the reaction schemes and the experiment procedure. In the 'no template' condition, in the paper-based Study 1, around 50% of the students included a reaction scheme in their reports, but none included a reaction scheme in the computer-based Study 2 for this condition. Almost all of the students included a reaction scheme in the Titles Template, where the title specifically asked for it. However, in the Profile Template none of the students included a reaction scheme for 'What reactions were involved in the experiment?' and instead provided a wide variety of answers describing the theory of the experiment. In Study 3, the responses to this question were simpler, more consistent, and more like reaction schemes, perhaps reflecting the simpler nature of the reaction involved in that experiment compared to the more complex synthetic experiment. The goals of the experiments were also quite different and this may have influenced what memory the cue triggered in the participants - prompting the recall of memories related to the theory of the experiment rather than the reaction scheme itself. Both these types of information have value to the experiment record, but it would be better to be able to capture them separately and unambiguously.

Another area where the wording of the cue seemed to cause significant differences was in the recall and recording of the procedure of the experiment. On the blank templates, all of the participants recorded information about the procedure of the experiment, and often the observations and explanations associated with the actions taken in the experiment. Often no other information was included. In the Titles Template condition, the cue 'Step by step experiment procedure' was used. This cue seemed to change the recording of the procedure for many of the students from a personal record of what was done in the experiment to a more formal list of instructions of what to do. The result of this was a loss of detail about why actions were performed and observations made, together with a switch in style for the paper-based template. This change in style, from 'What I did' to 'How to do it', influenced the design for the profile template to try and elicit information about the personal experience and also prompted the idea to investigate changing perspectives in **Chapter 7: Generating metadata by using cues and changing perspectives**. It was hoped that using the question "What did you do in the experiment?" for the Profile Template condition would lead to the inclusion of more personal detail and explanations about what was done in the experiment, but in fact this cue resulted in a much briefer description of the steps of the experiment, in some cases just a list of short phrases or 'techniques' describing the steps. Details of specific materials and the equipment used in the procedure were also omitted in this template.

The briefer answers recorded in the Profile Template condition may be a result of the larger number of questions to answer or that some of the questions were more open, for example "What did you do in the experiment?" is actually more open than "Step by step experiment procedure". Some studies have suggested that open-ended questions may lead to a more superficial search of memory, and therefore using closed questions can generate more results (Tourangeau et al, 2004). This may be because the more closed questions contain more specific cues, for example "What chemicals or other materials did you use in the experiment?", "What instruments or equipment did you use in the experiment?", and "What were the aims of the experiment?" produced sensible and specific responses from all of the students in the Profile Template condition. In contrast the more open question “What reactions were involved in the experiment?” worked less effectively than the more specific “Balanced equation with relative molecular masses (RMM)”, even though this was a difficult question to respond to on a computer. Care therefore needs to be taken with selecting the most appropriate cues to ensure the capture of relevant, useful, and accurate information.

The loss of explanations and observations seen in Study 1 were reinstated by asking specific personal questions such as “What observations did you make in the experiment?” and “What did you learn in the experiment?”. This increased the number of observations and explanations in the Profile Template, and a similar effect was also seen in the Titles Template, where more explanations and observations associated with the experiment were captured in the Results, Discussion, and Conclusions sections. A problem, which is not overcome by using these cues, is the separation of the observations and explanations from events and actions within the procedure of the experiment where they occurred – losing some of the temporal and contextual information from the experiment, that may help the researcher to remember what happens, but also be less useful for someone else trying to repeat the experiment.

There are benefits to all the styles of templates; although as shown in Study 3, including the wrong kind of cue in the template can have significant negative consequences. The inclusion of the results table in the template was not expected to generate negative effects – however the table had the result of constraining the information that was recorded by causing the researchers to diligently focus on completing the table at the expense of both completing the other fields in the template and even in completing the experiment itself.

Including a table of results as standard in a structured interface is problematic. We know that within the LabTrove community, the majority of templates that are used and shared by research groups do include tables for results or for describing the properties of materials or instrument settings. Researchers find tables particularly useful for capturing information for repeating experiments. However, designing a standard table for an experiment is not practical, and even a customised table may cause problems. Providing tools for researchers to design their own tables and include them in templates may be more beneficial than enforcing a standard table structure. A cue to 'Create a results table', with perhaps a choice of some table templates, or even some intelligence in the interface to suggest columns that might be used based on the content of the experiment record, may be more useful than directly including a table in the template.

Using the Profile style template is more onerous than using the blank piece of paper, and also provides less flexibility and freedom. A balance needs to be struck between capturing 'the right' information, and making the process too much of a chore for the researchers. The question of 'What should be recorded for an experiment?' can probably only be 'it depends'; requirements for education probably differ from requirements for researchers, and the requirements probably differ between disciplines, as indicated **in 3.5** Notebook structure, where each discipline has it's own guidelines and standard practices. Providing the researchers with ways to design, configure, and share their own templates may help to mitigate some of these problems. Using the Profile template as a basis for user-generated templates may be an appropriate strategy for achieving a balance between flexibility and capturing the right information that can be customised for the needs of individual researchers and research groups.

Chapter 6: Use of user-defined metadata in ELNs and other platforms

6.1 Introduction

Electronic Laboratory Notebooks provide an opportunity to both capture and make use of metadata related to experiments and their data. Surveying the LabTrove ELN provided an opportunity to investigate what metadata is used within an ELN, how effective that use is, and whether lessons can be learnt to improve the quality or quantity of metadata captured. The results of the survey and subsequent research work have been published in the following paper:

> Willoughby, C., Bird, C. L., Coles, S. J., & Frey, J. G. (2014). Creating Context for the Experiment Record. User-Defined Metadata: Investigations into Metadata Usage in the LabTrove ELN. Journal of Chemical Information and Modeling, 54(12), 3268–83. doi:10.1021/ci500469f

This chapter provides a brief background to the research, the methods that were used, a brief discussion of the major findings, and the significance of the findings in terms of designing tools for scientific researchers. The research included in the paper and this chapter was conducted by C. Willoughby. C. Bird provided editorial assistance on the paper and S. Coles initiated the contact and helped arrange the user-testing sessions at Bath University, described below. Those user-testing sessions were also designed and run with the help of E. Tonkin from UKOLN.

6.1.1 Research Questions

The following research questions were investigated as part of these studies:

- How is metadata used within the LabTrove ELN?
- Does the way that metadata is used change with the number of authors of a notebook?
- Does the way metadata is used change with the privacy settings of the notebook?

User-defined metadata

- Is the way metadata is used in LabTrove unusual compared to other platforms that provide user-defined metadata?
- Are there differences between metadata used on different types of information?
- Are there lessons that can be learnt for capturing user-defined metadata more effectively from the design of other platforms?

6.2 Background

The LabTrove ELN provides two different mechanisms for adding metadata to notebook entries: a single tag like ‘section’ metadata field, and multiple key-value pairs, which provide a more powerful mechanism for categorising notebook entries than simple tags (Milsted et al., 2013). The metadata enables the characterisation of notebook entries, enabling grouping of related experiments, aiding in search and linking, and providing the visible navigation within the interface. However, in our early discussions with researchers using LabTrove, it became apparent that the interface for creating metadata was difficult to understand and many researchers were anxious about using it. Anxieties ranged from not knowing where to start with metadata to actively fearing the consequences of ‘getting it wrong’, and the amount of effort that would be involved in fixing any mistakes. As a consequence many of the researchers who were current users of LabTrove were struggling to add metadata for their experiments. The lack of metadata on their entries meant they were also finding it increasingly difficult to locate specific notebook entries and the details for particular experiments.

At around the same time we had the opportunity to run some user testing sessions on LabTrove with several groups of researchers at Bath University who were planning to adopt the ELN. We included metadata as one of the elements of the testing to investigate how the researchers used the metadata features within LabTrove. Two different styles of study were conducted. In the first the researchers were provided with an empty notebook on LabTrove and asked to conduct a simple experiment and record their results in the ELN. In the second study the researchers were also given a blank notebook, but were asked to make use of an exemplar notebook, containing sensible metadata, to design their own experiment; again recording the results in their notebook in the ELN.

The instructions and guidance given to the students in the Bath studies are available in Appendix B. In the first study we found that the researchers only added a value for 'Section' because they could not save their work without doing so. In the second study we found that the researchers avoided adding metadata unless prompted, and then failed to use appropriate terms for the key-value pairs even though examples of good practice were provided. Figure 6-1 shows that the terms entered for the 'Section' were sensible values. When the researchers were interviewed, it was clear that they did not have a good understanding of what metadata was or how it might be useful.

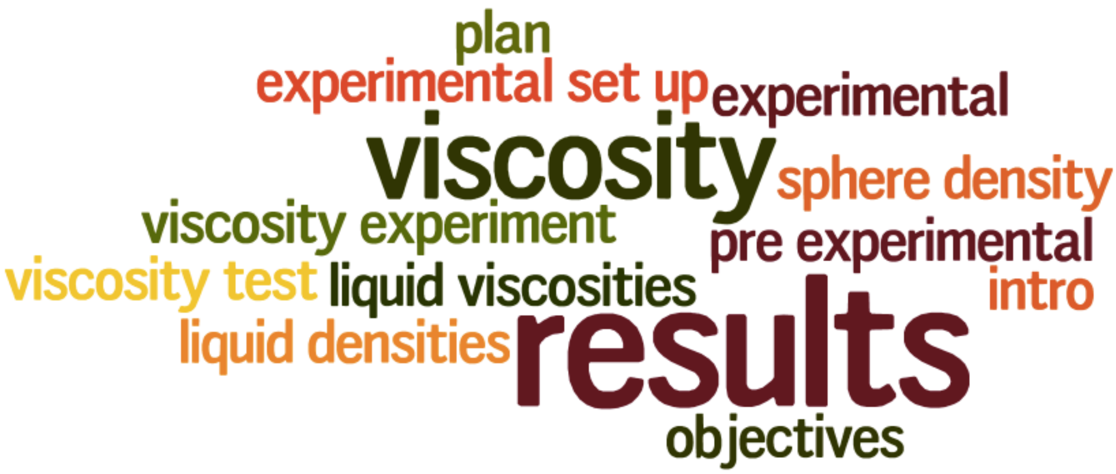


Figure 6-1: Word cloud visualisation of 'Sections' from the Bath user studies

To investigate whether the issues of minimal metadata use and perhaps inappropriate metadata values were common in LabTrove, we conducted a survey of metadata use across over 100 LabTrove notebooks. The survey enables us to investigate how metadata was being used within LabTrove notebooks across the whole community.

6.3 Methods

As mentioned in the paper, for each of the notebooks that were analysed metadata was extracted from the navigation menu, which provides the terms used for 'Section' metadata and the key-value pairs, together with the number of entries, as shown in Figure 6-2.

User-defined metadata

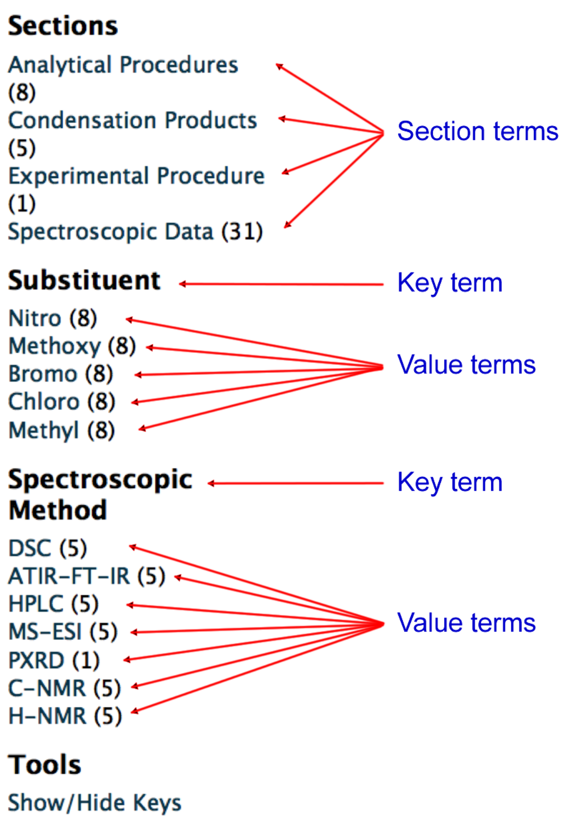


Figure 6-2: Example navigation menu from LabTrove with metadata terms

Other information was also recorded about each notebook, for example the number of authors, the number of entries in the notebook, whether the content of the notebook was primarily produced by an 'auto-blogging' script, and what the primary use of the notebook appeared to be.

The terms were analysed based upon the numbers used within each notebook to analyse what percentage of notebooks were using metadata and how many different terms were being used within each one. The numbers of the different metadata types being used were also analysed by the number of authors and privacy settings of each notebook to determine whether an increase in authors or more public access to the notebook were associated with an increase in metadata used.

Each of the metadata terms extracted from the notebook was classified in a number of different ways to analyse what kinds of terms were being used within the ELN:

- Whether the term was high-level or specific
- By word or phrase type - Noun-type, Verb-type, and Adjective-type

- Noun-type classified by topic

After the LabTrove survey was completed, similar surveys were conducted on a variety of other platforms where users are able to create user-defined metadata. Surveys of the other platforms were conducted in order to investigate if the pattern of metadata used within LabTrove was unusual. The following platforms were investigated:

- Flickr
- 53 NASA blogs
- 50 chemistry-related blogs
- myExperiment

The metadata terms extracted from these platforms were also analysed by word or phrase type and topics.

6.4 Findings

From the Bath user studies the following findings were made:

- Students that split their experiment reports into stages used terms for Section that reflected the appropriate stage; in contrast, those who used a single entry tended to use a catch-all category
- When exemplar notebooks with metadata were included in the study students were more likely to use metadata, but the combinations chosen were inappropriate and they did not add their own more relevant terms
- Researchers reported feeling more comfortable with metadata in tag form or for making links

The most significant findings of the LabTrove survey are as follows:

- The majority of users are not really using metadata at all, although some groups are using it effectively
- The biggest inhibitor to adding useful metadata is the “blank canvas” effect, where the users may be willing to add metadata but do not know where to start
- A large percentage use a ‘catch-all’ term when metadata is mandated but no assistance is provided

User-defined metadata

- The amount of metadata used increases with the number of authors
- There appears to be no correlation between the amount of metadata used and the privacy settings of the notebook
- The metadata terms used are mostly high-level labels, rather than specific terms that identify specific experiments or elements
- Terms used are overwhelmingly nouns, with <5% verbs, and 2% adjectives
- Activities, Codes, Dates and Values, Equipment and Instruments, Labels, Materials, and Project Information represent the dominant subjects for the metadata terms

The most significant findings of the surveys on the other platforms are as follows:

- 80% of NASA blogs and 30% of the chemistry blogs use a 'catch-all' value, either a default value such as 'General', or 'uncategorized' or a user-defined value such as 'miscellaneous'.
- A default value is not effective for encouraging appropriate metadata use.
- A much larger number of adjectives are used to describe photographs than any of the text-based information types explored in the surveys
- Different communities use different terminologies. The metadata can be used to infer the community using the platform.
- No one set of terminology is likely to work for all users of an ELN, but commonalities do exist at a high level between these different communities, particularly for the "label"-type information.
- Providing prompts or invitations to use or reuse metadata in interfaces appears to be effective at encouraging metadata curation.

6.5 Discussion

The findings from these surveys and the associated user research indicate that researchers do not necessarily understand the benefits of metadata. Although some individuals and groups are successful in designing metadata schemes for their notebooks, the majority of users are using a 'minimum required' approach and avoiding metadata use. Support is clearly needed to encourage researchers to create metadata that describes their experiments.

6.5.1 Social considerations – multiple authors and privacy settings

The number of authors that use a notebook is in general positively correlated with the amount of metadata suggesting that the metadata facilitates better organisation. The presentation of metadata in the interface must act as an aid and not as a hindrance, as demonstrated by the way that Key and Value metadata numbers reduce compared to the Section metadata as shown in Figure 6-3. A large number of Keys and Values in the LabTrove interface can make it more difficult to locate items rather than easier. Any interface designs for tools need to be effective whether they are displaying small quantities of metadata or large amounts.

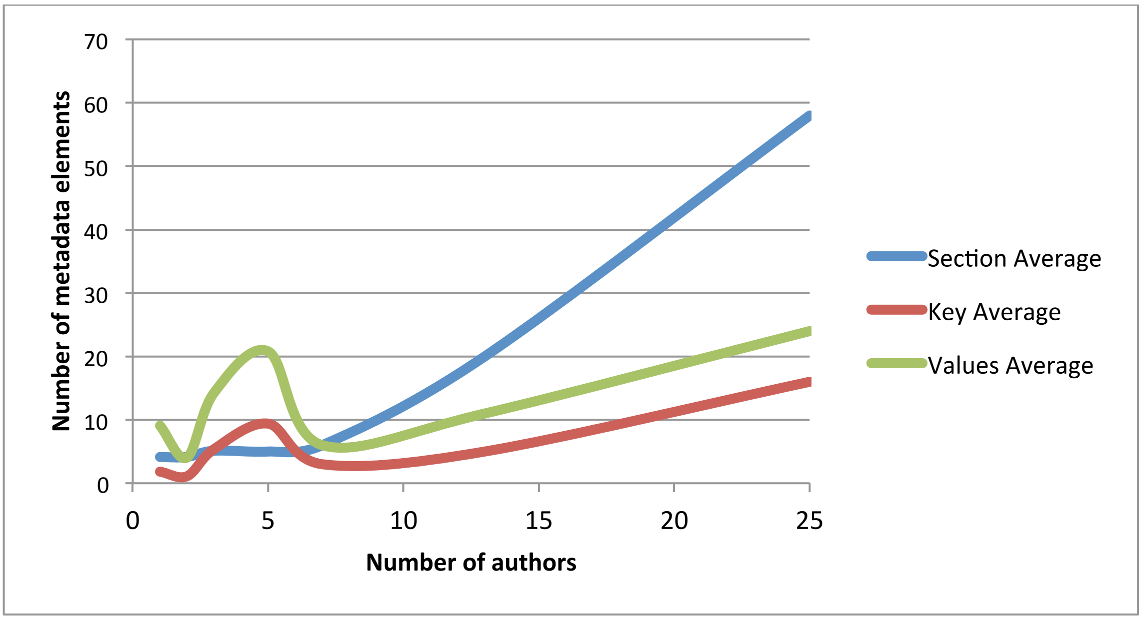


Figure 6-3: Average metadata by number of authors

Metadata does not appear to be correlated with the privacy settings of the notebooks in LabTrove strongly suggesting that users are adding metadata for both social reasons and personal organization or future retrieval, in common with behaviour observed on other platforms (Ames & Naaman, 2007; Hammond et al., 2005). Tools for adding and displaying metadata must support both of these uses and there may be a case for distinguishing between metadata that is generated for personal organisation and for social reasons.

User-defined metadata

6.5.2 Terms used

The metadata used in LabTrove is similar in many respects to the other platforms that support user-defined metadata in that nouns or ‘things’, with relatively few verbs and adjectives, dominate the metadata terms used, but the exact terms that are used depend on the platform and the community. The terms that are used in LabTrove, particularly the high-level labels provide an insight into the types of metadata that are most useful to capture for an experiment. The fact that some types of metadata are added more readily than others suggested that some useful metadata is not being captured, such as information about experiment activities and conditions. Techniques may be needed to specifically encourage researchers to add this useful but underrepresented context information.

6.5.3 Overcoming the ‘blank canvas’ effect

As described in 3.7.1 and in 6.4, one of the biggest inhibitors to effective use of the LabTrove ELN, particularly for metadata, is the ‘blank canvas’ effect, in which users do not know how to start creating a metadata scheme. This effect most likely contributes in part to the problem of metadata avoidance seen in the user studies and the LabTrove studies. A number of different strategies could be used to tackle these problems, both through the use of tools and appropriate education, as detailed in the paper:

- Developing and including standard schemas in tools
- Adding invitations to add and reuse metadata in tools
- Including metadata as part of an experiment or notebook entry ‘profile’
- Automatic creation of metadata and ‘smart’ assistance
- Providing visible benefits to adding metadata
- Educating researchers

6.5.3.1 Using schemas

Different communities use different terms, but there are also higher-level terms and topics that are consistent between the communities such as materials, locations, data types, people, events, and equipment. These findings indicate that there are terms that could be used across communities, but that different disciplines or domains would benefit from a tailored or specific schema. A good example of a high-level term that could be used across communities is ‘Materials’, but one or more of the following terms may be more appropriate depending on discipline: “Inputs”, “Reagents”, “Strain”, “Products”, “Substance”, “Compound”, “Chemical”, “Molecule type”, “Batch”, and “Sample”.

Basic schemas could be developed using these common elements and different communities could develop their own extended schemas, which could be shared. Alternatively, and perhaps more appropriately, standard schemas and ontologies could be incorporated into the tools. These schemas and ontologies could provide a valuable starting point for users and ensure consistent terminology and interoperability with other systems. The problem comes with choosing which schema is most appropriate from the hundreds of alternatives available, the majority of which are designed for describing scientific data rather than describing experiment records. Providing the flexibility to incorporate a schema or ontology of choice, or even providing tools to design and extend existing schemas could be included. However, decisions about which schemas to include and the design of new or customisation of existing schemas are likely to be specialist tasks for research groups with the assistance of information professionals, rather than tasks undertaken by individual researchers. Individual researchers will be able to make use of such schemas once included within a tool.

6.5.3.2 Adding invitations, using profiles, and making use of metadata in the interface

Many of the other platforms included in the surveys use a variety of prompts and ‘invitations’ to encourage users to add tags, categories or labels to their items. The metadata that the user adds is also displayed prominently in the interface, for example the tags in Flickr, tag clouds in WordPress, and using tags for searching in MyExperiment.

Invitations are a way to encourage users of software to provide information or to perform tasks. They often highlight a part of the interface and invite you to use it, provide a link or a button such as a plus to encourage you to explore, or ask you a question to elicit particular information. Invitations act in part to attract attention, but also act as prompts to record information. Other social networking platforms also use these kinds of invitations to capture metadata from users. For example, Figure 6-4 provides an invitation example from Facebook[18], where the user is asked a question about what they are thinking. Users are also encouraged to add more detailed metadata about how they are feeling, what they are doing, where they are, and who they are with. ‘What’s on your mind?’ could be considered a prompt for free recall of information, whereas the other questions are specific and can be considered as cues.

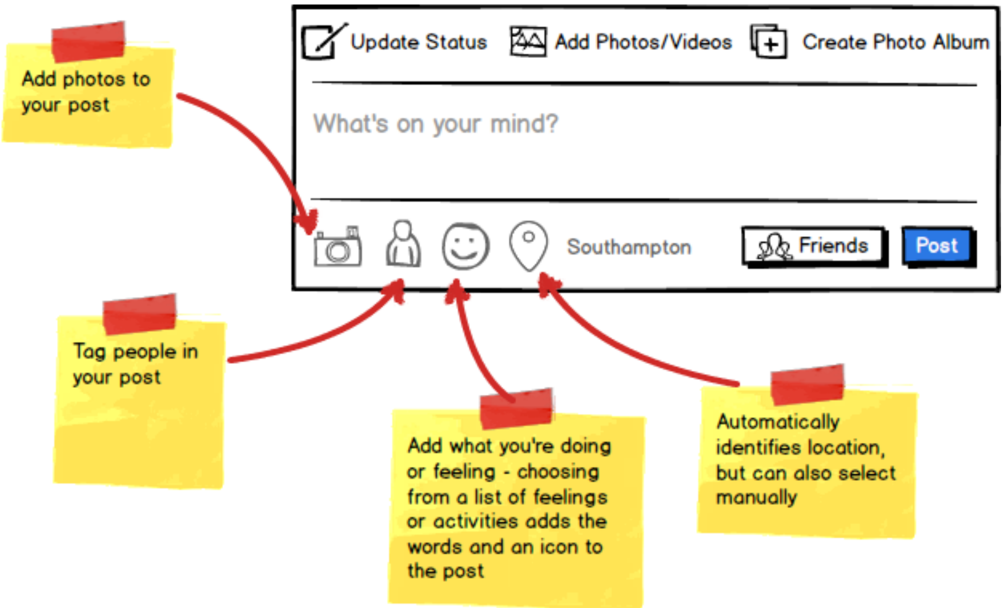


Figure 6-4: Asking questions and adding metadata from Facebook (Mock up)

Figure 6-5 shows an example of a LinkedIn[19] invitation with a question used to elicit metadata about skills. The user is invited to add their skills in the textbox and autocomplete will match them to skills that other users have chosen. The skills that are added are displayed prominently within the editing interface shown, indicating to the user the number of endorsements they have for each skill, but are also prominently displayed for other users in their profile and included in search results, making it more likely that the user will be found and matched with appropriate contacts or career opportunities.

[18] https://www.facebook.com
[19] https://www.linkedin.com

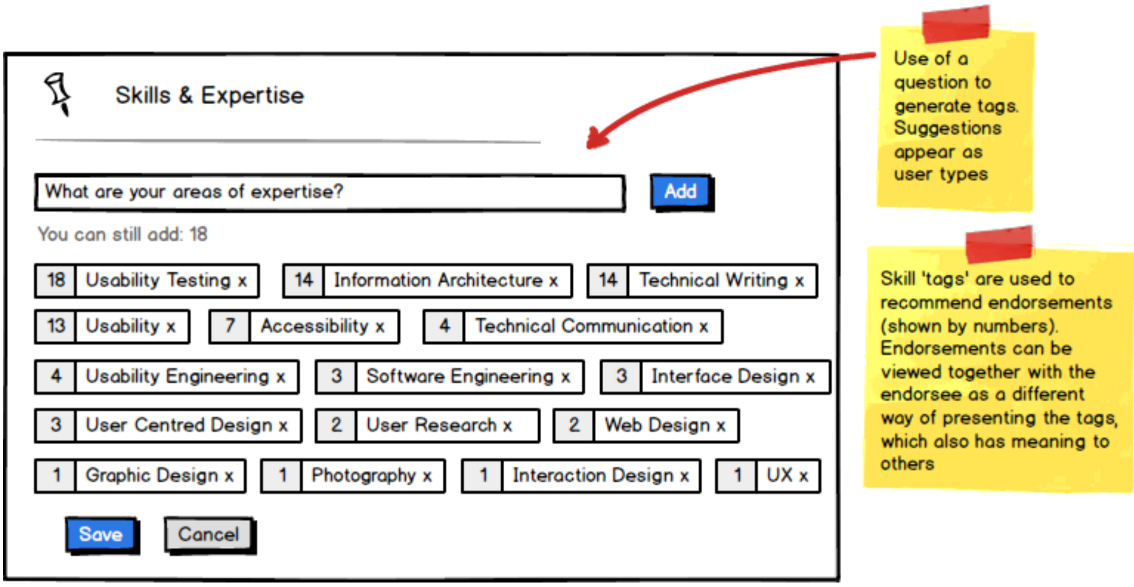


Figure 6-5: Asking questions and tags from LinkedIn (Mock up)

Metadata is also collected in many of these social networking platforms in the form of 'profiles' where users are encouraged to complete a series of fields, usually about themselves or to describe an object. For example, Figure 6-6 shows a mock up of an invitation from LinkedIn encouraging users to add information that has not yet been included in a personal profile. Invitations include the phrase 'Recommended for you', an arrow pointing at the interface, and the 'plus' buttons. Clicking on the plus symbol reveals a form to capture the metadata for this particular content. Placeholders are left within the profile for sections without content as a reminder to complete them.

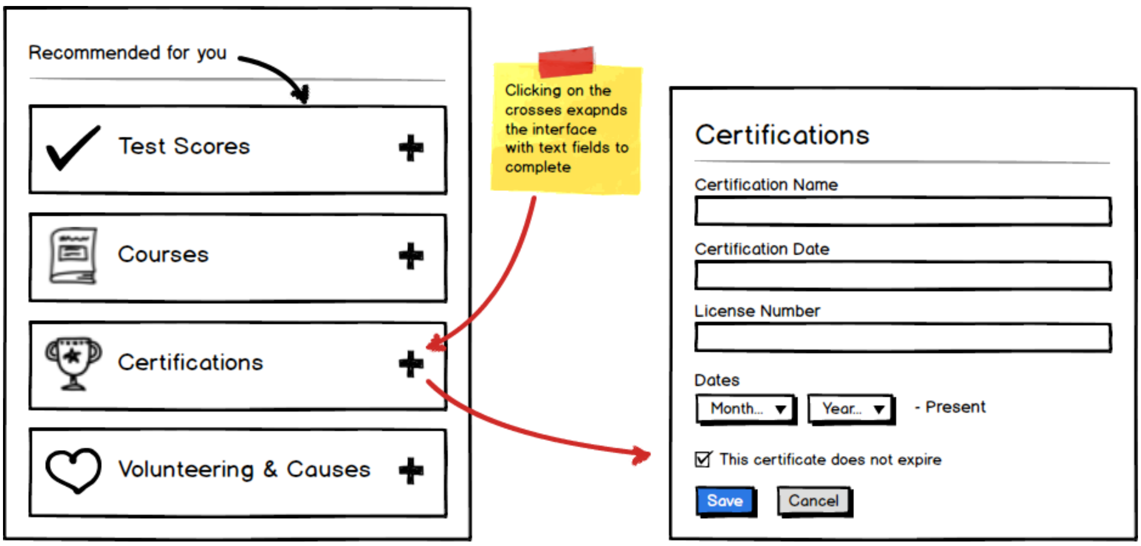


Figure 6-6: Invitations and metadata elicitation example from LinkedIn (Mock up)

User-defined metadata

Figure 6-7 shows a mock up of the Flickr photograph interface showing the ‘Profile’ for a photograph containing a number of invitations to perform actions and create new content. The user is encouraged to add a title, comment, tag, and add people in the photograph through placeholders with instruction text, perform actions such as add the photograph to an album or group, and through examples provided in the interface. If information is missing, such as no tags, or no location invites are still visible in the interface to encourage you to add that additional metadata.

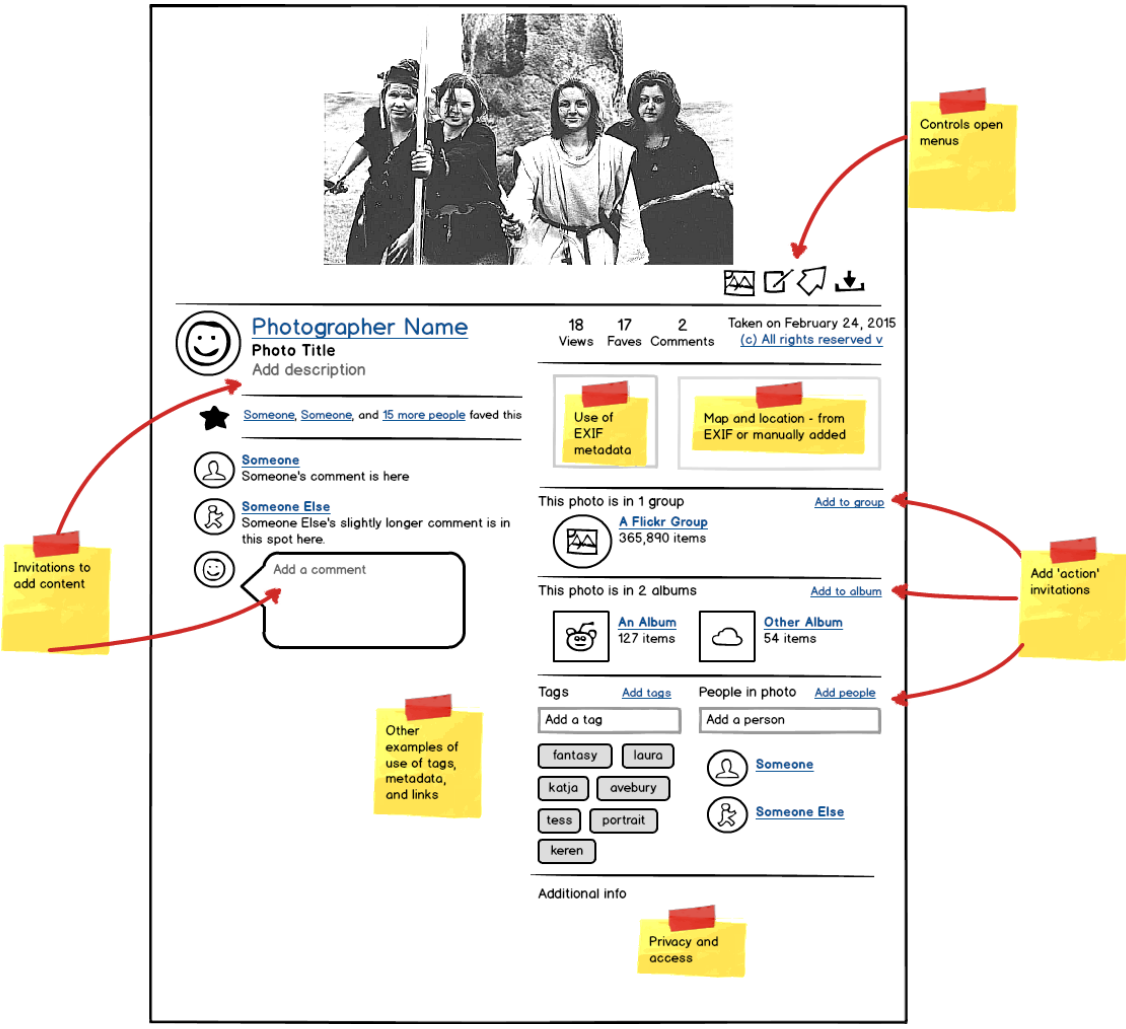


Figure 6-7: Photo 'Profile' from Flickr (Mock up)

The Flickr interface makes effective use of metadata, for example EXIF data is used to display information about when the picture was taken, the camera and settings for the photo; geotags are used to display a map of where the photo was taken; statistics about views and activity are displayed; what groups, albums, and galleries the photo belongs to; people in the photograph; in addition to the tags added by the user. Most metadata are displayed as links that can be used for searching, grouping, and filtering content.

In **Chapter 7: Generating metadata by using cues and changing perspectives**, we investigate different strategies of capturing metadata to inform effective designs for invitations and profiles.

6.5.3.3 Automatic creation of metadata and smart assistance

Metadata can be generated automatically based on information that a system knows. For example, the LabTrove ELN automatically generates data based on author, date and timestamps, versions and revisions, and file types for data. Additional metadata could be derived by using data mining techniques to extract meaning from the content of experiment records or associated data files. Some work has already been done for matching chemical names in LabTrove entries with records from the ChemSpider database[20](Day, 2013). The matched names are converted into links to the ChemSpider database for the appropriate chemical.

Automatic creation of metadata relies on appropriate information existing within the content of a record and in an appropriate format for the system to make sense of it. Any potential terms that could be used as metadata could be used to support researchers by providing them as 'suggested' values to add. For example, WordPress suggests tags that could be added to newly published blog posts based on other tags added to the post.

6.5.3.4 Educating researchers

Contrary to what might be expected, research has found that even researchers in physical sciences do not have a good understanding of metadata and data management practices in general (Bird et al., 2013a; Goodger & Worthington, 2013; Whitton & Takeda, 2013; Whitton, 2013; Ward et al., 2011 Rowlands et al., 2008).

[20] http://www.chemspider.com

User-defined metadata

In addition to data management training, various educational resources and user assistance could be provided with tools for researchers, such as:

- Improved documentation
- Video tutorials
- Exemplar notebooks and case studies
- Metadata checklists (Higdon et al., 2013)

A subject librarian or other information professional, particularly in the early stages of a project, could also be recruited to provide hands-on support to individual researchers, small groups, or a whole research team.

Chapter 7: Generating metadata by using cues and changing perspectives

7.1 Introduction

Inclusion of metadata on records helps with providing access, effective curation, improving search, and providing context. As described in **Chapter 6: Use of user-defined metadata in ELNs and other platforms**, just providing researchers with the facility to add metadata to their experiment records does not mean that they will make use of it, or if they do that the metadata they add will be relevant and useful. Our findings in the LabTrove survey and other user research indicated clearly that researchers need support to create effective and meaningful metadata. Tools, such as ELNs, provide an opportunity to encourage researchers to include metadata in their records, but also add extra value by making effective use of the metadata that is provided. Our first study in **Chapter 5: Asking the right questions: Effects of Using Structured Templates for Recalling Chemistry Experiments**, demonstrated that using cues was an effective way to encourage researchers to remember and record specific information. As detailed in 2.2.2.3 and 2.2.2.10 previous research has demonstrated that 'changing perspective' is another technique that could be used to help recall additional information. In this chapter, we detail the results of studies to investigate whether using cues and changing perspectives could be used to capture useful metadata for different kinds of information. The results of most of these studies have been published in the following paper:

> Willoughby, C., Bird, C.L., & Frey, J.G. User-Defined Metadata: Using Cues and Changing Perspectives. *International Journal of Digital Curation*. *10*(1), 18-47

This chapter provides a brief background to the research, the methods that were used, a brief discussion of the major findings, and the significance of the findings in terms of designing tools for scientific researchers. This chapter also includes data from a new study not covered in the paper above. The research detailed in the paper and in this chapter was undertaken by C. Willoughby. C. Bird provided editorial assistance in the published paper.

Using cues and changing perspectives

7.1.1 Research Questions

The following research questions were investigated as part of these studies:

- How does the information type affect the metadata that is generated?
- Are different metadata generated using the following strategies:
  - Free recall
  - Cued recall
  - Changing perspectives
- Is the metadata generated appropriate?
- Would an information specialist produce the same metadata?
- What metadata is generated for experiments using the different strategies?
- Can templates be used to generate metadata?

7.2 Background

We know from in **Chapter 6: Use of user-defined metadata in ELNs and other platforms,** that researchers need support to help them create metadata to describe their experiments. There are a number of ways that metadata could be automatically generated or researchers could be asked to select metadata from a predetermined list of terms, but these methods are unlikely to be as reliable, specific, or flexible as terms that can be provided by the researchers themselves.

Metadata on social networking platforms, such as those investigated with LabTrove, typically collect metadata using a 'free recall' method – where invitations are used to encourage the addition of tags, categories, or other metadata types. These invitations act as a reminder to add the metadata, but are not really acting as cues to prompt the capture of any information in particular. Metadata can be used for the dual purposes of describing content by providing context and aiding retrieval by providing terms that might be used as search keywords. Different terms may be needed to fulfil these different needs, and different techniques may be more successful at generating terms appropriate for these different needs. In the studies described in this chapter, three different techniques were used to generate metadata. The metadata terms generated were analysed to assess differences between the techniques and usefulness of the terms generated.

7.3 Study 1: Online metadata survey

The first study was designed as a standalone investigation to find out whether using cues or changing perspectives would generate different and more useful metadata than free recall. The study was also designed to investigate whether different metadata was generated for three different information types:

- Photographs
- Structured textual information
- Unstructured textual information

Generating metadata for photographs was used to establish whether more adjective-like words and phrases would be produced for photographs as we observed for the photographs on Flickr in **Chapter 6: Use of user-defined metadata in ELNs and other platforms**. The use of structured and unstructured text was to determine whether metadata generated for these two information types differed. As discussed in 2.2.3 and **Chapter 5: Asking the right questions: Effects of Using Structured Templates for Recalling Chemistry Experiments**, experiment records can be captured in ELNs using structured templates or unstructured text. We expected that using structured text might help to make it easier for subjects to generate more metadata or more relevant metadata because key elements are separated and highlighted compared to the unstructured text. If our expectations are correct, then using templates may be more effective at encouraging the addition of metadata than interfaces that use more flexible unstructured text. All of the information objects used included activities to reflect the procedural nature of experiments:

- Photographs of people
- Cooking recipes (structured text)
- Instruction for performing a task (unstructured text)

7.3.1 Methods

The study sought to elicit responses from subjects using an online survey with the following three conditions:

Free recall: The user is asked to write up to ten words or phrases that come to mind when they look at the photograph or read the text in a questionnaire.

Audience (change perspective): The user is asked to write up to ten words or phrases that come to mind for use in a search engine when they imagine helping someone else to find a similar picture or piece of text on the internet.

Profile (cued recall): The user is asked to provide word or short phrase answers to a number of questions to elicit specific information about the photograph or text, including locations, people, equipment, activities or actions, other objects, and any other words or phrases of their choice to describe the material.

Each survey was divided into three sections, each containing a photograph, recipe, and unstructured instruction text. Subjects were randomly assigned to one of two versions of the survey. The first section of the survey elicited responses using the Free recall condition, whilst sections two and three of the survey elicited responses using the Audience condition or the Profile condition depending on the version of the survey, as shown in Figure 7-1. In this way responses for section two and three were captured in both Audience and Profile conditions to enable a comparison between the terms generated in the two conditions.

The content of the study was designed to be familiar for most people and to not require specialist technical or scientific knowledge to maximise participation. The information objects used in the survey are available in Appendix C.1 and the questions used in each condition for the image and text information types are available in Appendix C.2.

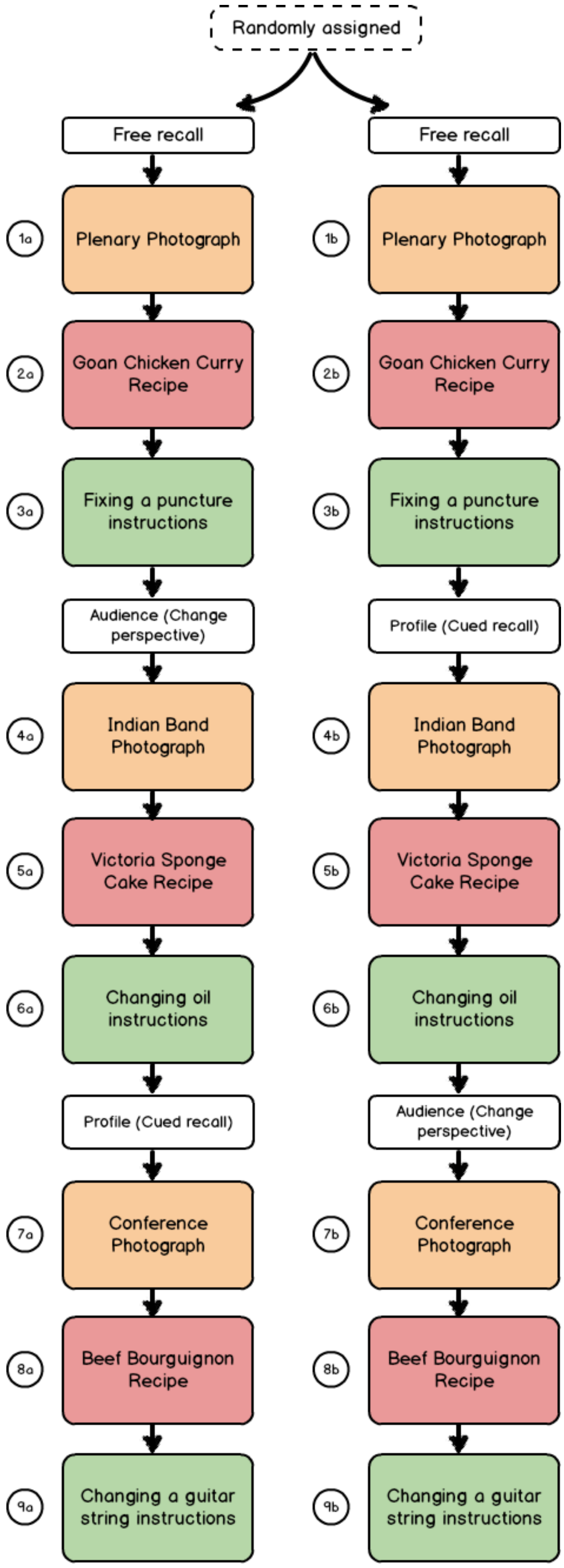


Figure 7-1: Online metadata survey structure

Participants for this study were anonymous and were recruited via Twitter and from attendees at an ACS conference. 47 subjects took part in the survey, although not all completed every question. Although the number of subjects varied between different sections in the survey, this did not affect the results because the responses to each information object and condition were analysed independently. In total 2408 responses were received across all the conditions and information objects.

The responses that were received were analysed by word-type in the same ways as the metadata in the LabTrove and other platforms surveys described in **Chapter 6: Use of user-defined metadata in ELNs and other platforms,** although the additional types ‘statements’ and ‘questions’ were also needed for those responses that did not fit into the other word types. Our interest in the word-type stems from the observation that in LabTrove relatively few verbs are being used, and this suggests that important information may be missing from the metadata. If we can generate more verbs and adjectives then we may capture more useful contextual metadata.

The responses were also categorised into the following topics based on the topics that were commonly found in the metadata analysed in those surveys:

- People
- Location
- Events
- Materials
- Activities
- Equipment
- Other

Finally the responses were analysed for common themes. In the Profile condition the themes matched the questions that were asked, but they were also analysed for any sub-themes. For example ‘appearance’, ‘feelings’, and ‘activities’ could be sub-themes for the question “What list of words or short phrases would you use to describe the people in this text?”. Also for the Profile condition the relevance of each of the terms included in the responses for each of the profile questions was also analysed to determine whether appropriate metadata values would be generated for these questions, and if any confusion arose from those questions that were less relevant to the information object. For example, the People and Location questions were expected to cause some unexpected effects because neither the recipes nor instructions mention people or locations.

As the analysis progressed it became apparent that it was also necessary to note whether the response described something personal (such as personal opinion or memory), something external to the content provided (something not mentioned in the text, for example, a mechanic for changing the oil), or something about the information format itself (for example whether it was easy to understand or badly structured).

An expert information professional was asked to classify the different information objects in order to produce a list of appropriate terms that would be used by a professional curator to enable a comparison with the responses from the study. The results of this exercise are available in Appendix C.4.

7.3.2 Findings

The major findings from Study 1 were as follows:

- The Profile condition produced a significantly higher average number of responses compared to the other two conditions.
- In the Audience and Profile conditions more responses were produced for the photograph than the structured and unstructured text, suggesting that it is easier to generate metadata for photographs
- In the Free recall condition the responses are mostly noun-type with few verb-types as seen in the previous surveys in **Chapter 6: Use of user-defined metadata in ELNs and other platforms**, but more adjectives are generated than we expected. The Audience and Profile conditions are also dominated by noun-type phrases, but the Audience condition has fewer verbs and adjectives than the Profile condition.
- The Profile condition has the most consistent distribution of word-types and also includes the highest number of verbs – primarily generated by asking about ‘Activities’.
- Photographs are associated with more adjectives as expected.
- The highest number of statements is produced in the Free recall conditions for the text-based information objects; these are most commonly associated with comments about the information objects themselves.
- The pattern of topics observed is based on both the content of the information objects and the recall condition.
- The Profile condition provides the most consistent and balanced distribution of topics because cues were provided for many of these topics.

- However, the topics for which cues were not provided - 'Materials' and 'Events' - are significantly underrepresented or even missing in the responses in the Profile condition. This is particularly noticeable for the recipes where very few ingredients are recorded in the Profile condition, but is one of the main themes in the other two conditions.
- Themes also match the content of the information objects and are similar across the Free recall and Audience conditions.
- The Profile condition produces a greater diversity of terms than the other two conditions
- The Profile condition produces more personal opinion, more guesses, more descriptive information, and more external information – or information about things not mentioned in the content – especially for People and Location.
- Personal opinions and external elements are even more common in the Free recall condition.
- The Audience condition produces useful search terms and also recommendations for searching such as information sources and types (e.g. websites, manuals, search terms, videos).
- The terms generated by the formal curation activity by the information specialist did not match any of the conditions closely, although more of the terms were generated in the free recall condition.

Visualisations of the results for each information object in the different conditions are available in Appendix C.3.

The results of Study 1 suggested that using these different methods to generate metadata could be useful, but that the addition of personal experiences and opinions, and information that was external to the content of the information object may lead to the generation of metadata that was irrelevant; although it might also add useful context that could not be derived directly from the content of the information object itself.

7.4 Study 2: Organic Chemistry Summer School - computer-based templates

As part of the studies into using templates to capture experiment records in **Chapter 5: Asking the right questions: Effects of Using Structured**

Templates for Recalling Chemistry Experiments, fields to capture metadata were added to the templates described in Study 2: Organic Chemistry Summer School and computer-based templates. A Free recall field was added to each of the three templates used in the study, and a set of 'profile-style' question fields were added to the Profile Template only. These additional fields were based upon Free and Cued recall conditions from Study 1: Online metadata survey described above. At the time this study was put together, it was thought the 'Change perspective' questions from Study 1 were generating terms that could be derived from a sensible title for an information object, and this condition was therefore not included in this study.

7.4.1 Methods

The templates were used to capture reports of the experiments from 20 undergraduate students taking part in an Organic Chemistry Summer School as described in Study 2: Organic Chemistry Summer School and computer-based templates. Each template included a 'Free recall' section asking the students to:

- Write down up to ten things that come to mind when you think about the experiment that you completed today.

The following profile-type questions were included only in the Profile Template:

- What chemicals or other materials did you use in the experiment?
- What instruments or equipment did you use in the experiment?
- Where did you do the experiment?
- What activities or techniques did you use in the experiment?
- What other information might be useful to remember about the experiment? For example: sample identifiers, safety information, and settings for the analysis?

The responses generated for the questions were analysed by word-type and themes as in Study 1 above.

7.4.2 Findings

The major findings from Study 2 were as follows:

- An average of 6 responses were generated by each student for the Free recall question; 67% of these were statements.
- The responses for Free recall varied by student, with some choosing mostly personal responses about their experiences, and others providing a mixture of chemicals, equipment, and techniques.
- The Free recall responses were dominated by activities, techniques, and conditions (40%); followed by chemicals and reactions (25%); then by personal responses such as anxieties, feelings, and learnt information (14%).
- Equipment (9%), safety (6%), and results (6%) were relatively underrepresented in the Free recall responses.
- Relevant responses were provided for the Cued recall questions, except for the ‘other information’ question where ideas for what could be recorded were provided based on the suggestions given with the question.
- Cued recall generated mostly nouns, but as seen in Study 1 the ‘Activities’ question generated verbs as well, with many nominalized activities.
- Cued recall responses are reasonably consistent between subject as shown in Figure 7-2; although the chemicals cue generated a diverse range of naming conventions, with a mixture of chemical names, common names, and molecular formula.

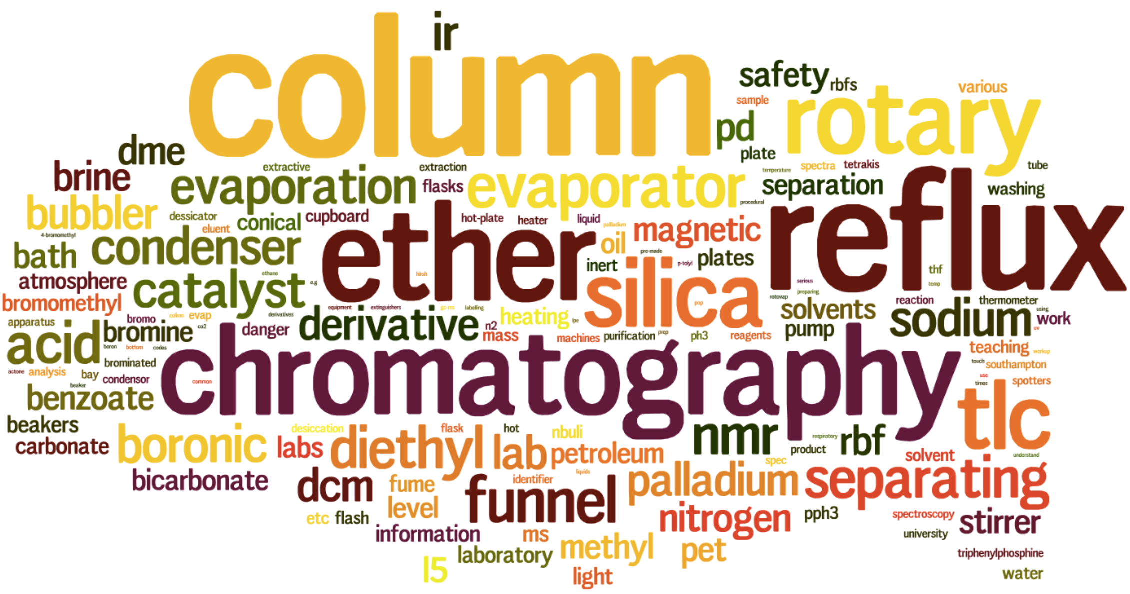


Figure 7-2: Word cloud visualisation of cued recall responses

7.5 Study 3: Lego rocket cars templates

Study 3 makes use of the templates devised for recording the Lego rocket car experiment from the note-taking workshop with KTH in Stockholm, as described in Section 5.5. These additional fields were based upon the Free and Cued recall conditions from Study 1: Online metadata survey described above. Instead of using the same ‘Change perspective’ questions, a ‘Keywords’ field was included to assess whether this term would generate similar ‘search terms’ in the template.

7.5.1 Methods

A section for capturing metadata related to the experiments was added to each template as detailed below.

For the No Template condition the following field (Free recall) is included:

> Write down as many words or short phrases as you can that come to mind when you think about the experiment

For the Titles Template condition the following field is included:

> Keywords

For the Profile Template condition the following fields (Cued recall) are included:

- What chemicals or other materials did you use in the experiment?
- What instruments or equipment did you use in the experiment?
- Where did you do the experiment?
- What activities or techniques did you use in the experiment?
- What other information might be useful to remember about the experiment?

The responses generated for the questions in the templates were analysed by word-type and themes as in Study 1 and 2 above.

7.5.2 Findings

As can be seen in Figure 7-3 to Figure 7-5, the different templates produced quite different patterns of metadata, with the cued recall in the Profile template generating the most terms. The Keywords produced a useful overview of the experiment, even though few terms were produced. The Free recall condition produced the most personal terms but also some useful terms describing the experiment that were not captured in the other conditions, such as ‘pushing’, ‘propulsion’, ‘aerodynamics’ and ‘gas’.

Figure 7-3: Word cloud visualization of responses from the No Template (Free recall) condition

Figure 7-4: Word cloud visualization of responses from the Titles Template (Keywords) condition

Figure 7-5: Word cloud visualization of responses from the Profile Template (Cued recall) condition

The major findings from Study 3 were as follows:

- The Free recall word-types were similar to those seen in Study 1 with 62% nouns, 15% adjectives, and the remainder questions and statements.

- Free recall themes include: conditions of the experiment, components of the experiment, results, and personal elements such as 'exciting' and 'childhood memories'.
- Keyword responses were more limited and produced terms similar terms to the Audience condition in Study 1 as expected. These terms represented the main components of the experiment such as 'Lego' and 'car' that could have been derived from the activity title.
- All of the groups produced useful information for all the Cued recall questions, including those that responded to the Other information question.
- There was some confusion between Materials and Equipment, with different teams including the same terms in the different cued categories.
- Unexpectedly teams included their own recording equipment in the responses, in addition to the expected materials and equipment provided for the experiment.
- Naming of chemicals was again variable as seen in Study 2, with a mixture of specific chemical names, formulae and common names. One team represented the reaction itself using "H2O + aspirin".

7.6 Study 4: Remote experiments study

A further study was devised to investigate the 'Free recall', 'Audience', and 'Profile' conditions with a larger set of subjects to assess the value of these strategies for capturing metadata related to experiments. The study was in preparation at the time the 'User-Defined Metadata: Using Cues and Changing Perspectives' paper was being written, and the results of the study were collected and analysed after publication of the paper.

7.6.1 Background

Undergraduate Chemistry students at the University of Southampton are required to complete a series of 3 'remote experiments' as part of their first year practical chemistry module. The students are asked to complete a post-lab survey after completing each remote experiment. A total of 117 students took part in the study.

The details of the remote experiments included in the study are as follows:

- **Viscosity experiment**[21]: In this experiment, a form of falling sphere apparatus was set-up and run for various liquids. Videos taken of the experiments have been set up online enabling students to make appropriate measurements to calculate the viscosity of the various liquids. Conducting the experiment in reality requires space, access to equipment, and is very messy.
- **Gas Law Experiment**: In this experiment the students can make measurements of pressure, temperature, and water levels in the equipment set-up to investigate how the pressure and volume of a gas change at different temperatures. The equipment set-up automatically resets itself by heating the vessel at fixed times allowing measurements to be taken in real time at any time of the day.
- **Beers Law Experiment**: In this experiment the student can remotely control a Class 3B laser, which they would otherwise not have access to, making measurements as the beam passes through a solution in order to make calculations about the absorbance properties of the solution.

The viscosity experiment is directly based on the viscosity experiment used at Bath University, as described in 6.2. We created this version of the viscosity experiment specifically for our studies on using different techniques to elicit information from students and researchers about experiments. Other members of the Chemistry department at Southampton developed the remaining experiments specifically as remote experiments for use in education.

7.6.2 Methods

There are three versions of the post-lab surveys to generate responses for the three conditions based on Study 1 in this chapter:

- **Survey 1: 'Cued recall' / Profile questions**:
  - What list of words or short phrases would you use to describe the chemicals or other materials used in this experiment?

[21] http://www.l4labs.soton.ac.uk/practicals/remote/viscosity/index.html

    - What list of words or short phrases would you use to describe the instruments or equipment used in this experiment?
    - What list of words or short phrases would you use to describe the locations used in this experiment?
    - What other information would you use to describe the experiment?
- **Survey 2: 'Change perspective' / Audience question:**
    - Imagine you are helping someone else to find information about a similar experiment to this on the internet, what words or phrases come to mind for use in a search engine? Write as many as you can below.
- **Survey 3: Free recall question:**
    - Write down up to ten things that come to mind when you think about the experiment that you completed today. Write down these thoughts in the order that you think of them.

The students are asked to complete one of the surveys after they complete each experiment. The surveys they are asked to complete are randomly assigned based on surname.

The responses to the surveys have been analysed by word-type, topic classification, and also for relevancy and personal opinion in the same way as Study 1.

7.6.3 Results

As can be seen in Table 7-1, in common with the other studies, the greatest number of responses are generated in the Cued recall condition, and the fewest are generated in the Change perspective condition.

Table 7-1: Study 4 - Average and total number of responses for each condition

Condition	**Free recall**	**Change perspective**	**Cued recall**
Total responses	750	620	1043
Average responses	6.41	5.49	9.07

Word cloud visualisations of the responses for each of the experiments in each condition can be seen from Figure 7-6 to Figure 7-14. These visualisations show how the Change perspective condition has the fewest, but most consistent use of terminology between the conditions, as was seen in Study 1.

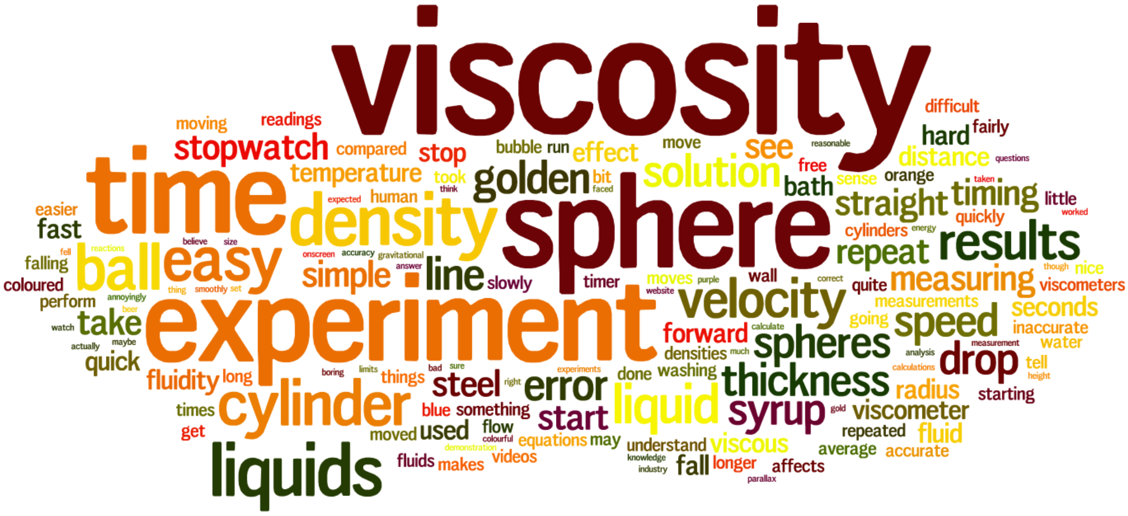


Figure 7-6: Word cloud visualisation of the Viscosity experiment in the Free recall condition

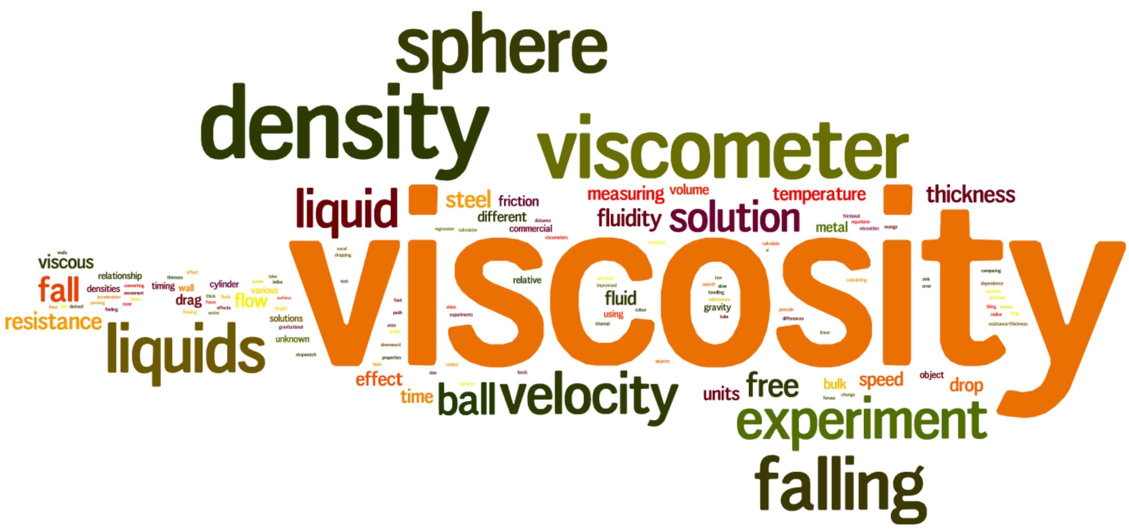


Figure 7-7: Word cloud visualisation of the Viscosity experiment in the 'Change perspective' condition

Figure 7-8: Word cloud visualisation of the Viscosity experiment in the 'Cued recall' condition

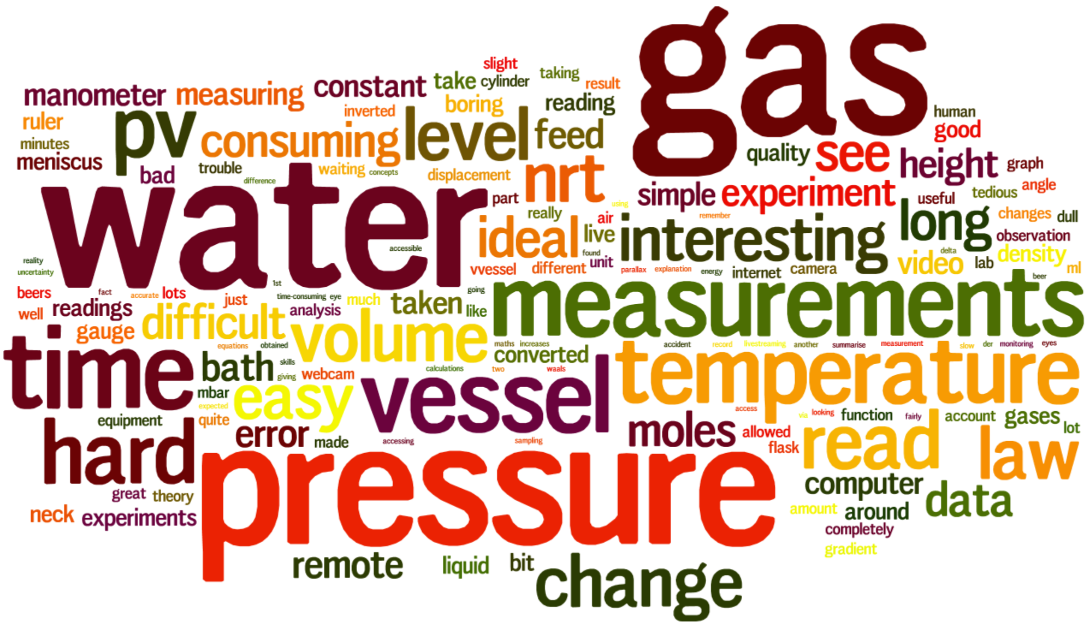


Figure 7-9: Word cloud visualisation of the Gas Law experiment in the Free recall condition

Figure 7-10: Word cloud visualisation of the Gas Law experiment in the 'Change perspective' condition

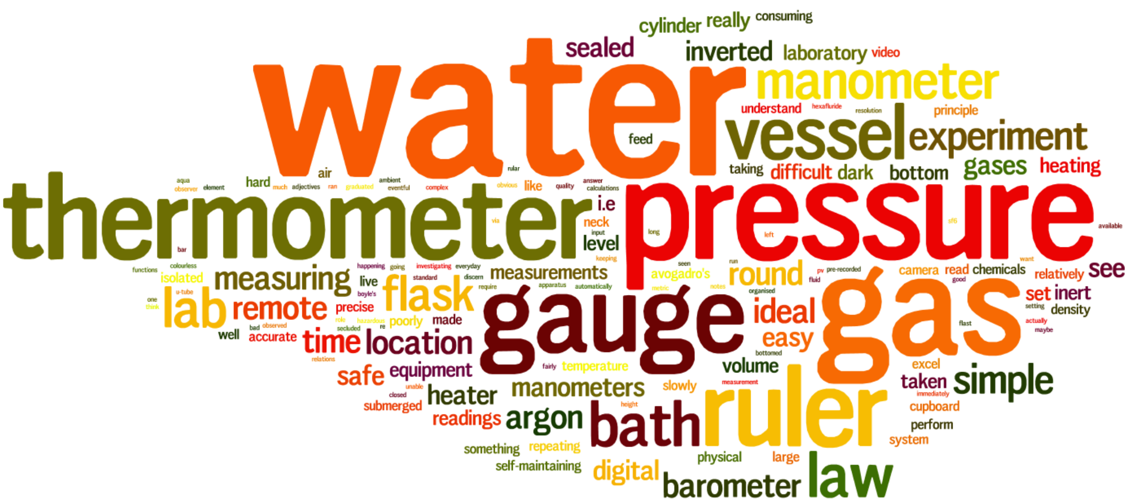


Figure 7-11: Word cloud visualisation of the Gas Law experiment in the 'Cued recall' condition

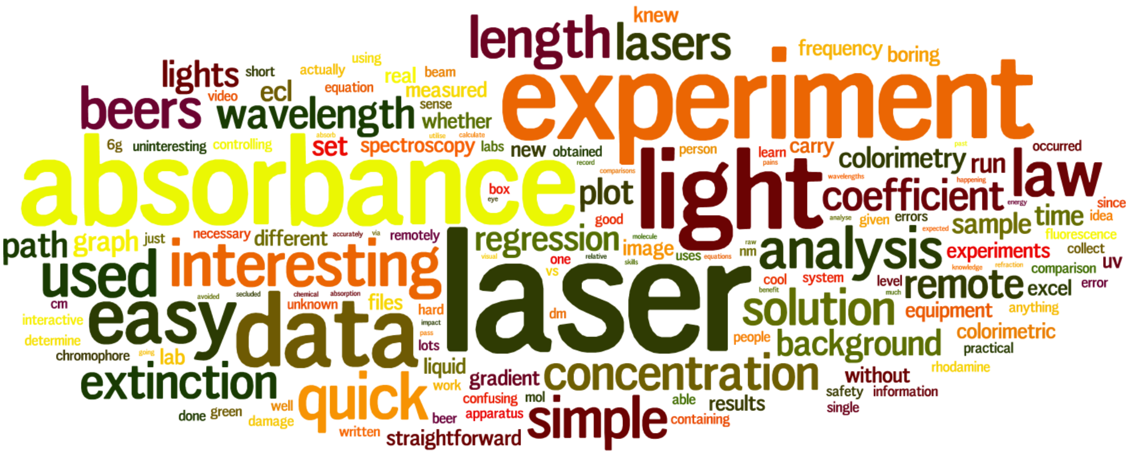


Figure 7-12: Word cloud visualisation of the Beers Law experiment in the Free recall condition

Figure 7-13: Word cloud visualisation of the Beers Law experiment in the 'Change perspective' condition

Figure 7-14: Word cloud visualisation of the Beers Law experiment in the 'Cued recall' condition

Across all conditions, the responses are dominated by nouns and statements similar to Study 1, although the number of adjectives is lower in the Free recall and Change perspective condition, and the number of verbs is reduced across all conditions. In the Change perspective condition the number of nouns is higher and with fewer statements compared to the other conditions.

The topic distributions for the conditions are similar in both the Free recall and Change perspective conditions, with responses in both conditions primarily classified as 'Other', followed by Equipment, Materials, and Activities, with a very small number of responses about the Location. The topics in the Cued recall condition are more evenly distributed, with a reduced number of responses classified as 'Other', and the majority of other responses matching the cues in the survey; Equipment has the highest number of responses in this condition. The full break-downs are shown in Figure 7-15.

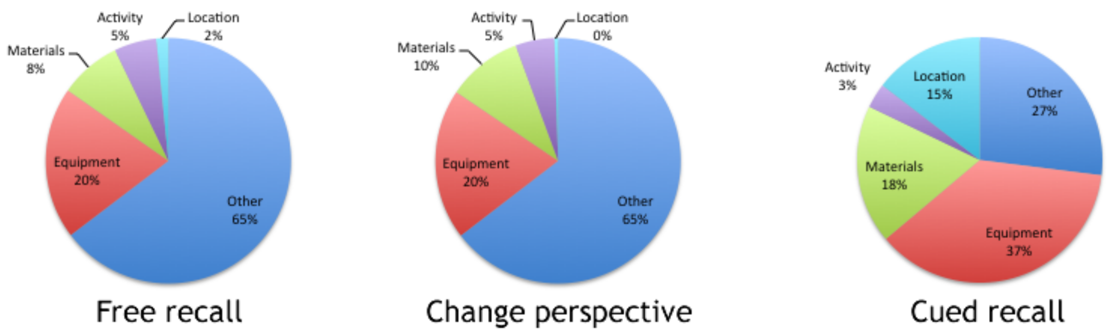


Figure 7-15: Study 4 - Topic classifications

The Change perspective condition contains almost no personal or irrelevant information at all, compared to the Free recall condition where around a third of all of the responses are not specifically about the experiment or represent an opinion. In the Cued recall less than 10% of the responses are irrelevant to each of the cued categories; however 50% of the responses in the 'Other information' section of the survey are personal opinion and not particularly useful as metadata, for example: 'easy to complete' and 'boring'.

Statements are frequently used to either provide comments that are not very relevant to the experiment (e.g. 'Not sure what i'll have for dinner'), comments about them as remote experiments (e.g. 'Interesting Concept of Livestreaming the Setup' and 'The experiment was fairly convenient and quick to complete'), or comments about problems with the experiment format, particularly relating to difficulties making measurements or network connectivity. A number of statements also detailed the steps of the experiment or the theoretical background. In the Free recall condition, almost all of the irrelevant or unhelpful responses come from statements. In the Cued recall condition, adjectives such as 'fun' and 'sufficient' account for a large number of responses that were not particularly relevant to the question cue.

These results are particularly useful in terms of evaluating the usefulness of the Change perspective for generating metadata for experiments. For these experiments, the Change perspective condition actually captured some of the most useful responses, generating the fewest number of statements, the most consistency of terms, and producing almost no irrelevant information.

7.7 Discussion

We knew from the surveys in **Chapter 6: Use of user-defined metadata in ELNs and other platforms** that some differences were seen between metadata describing photographs compared to textual information such as notebook entries, blog posts, and workflow descriptions. In Study 1 these differences were confirmed to some extent, with on average more adjectives used to describe photographs in all conditions, and a higher average number of terms generated, suggesting that photographs are easier to produce metadata for than textual information. Providing functionality to enable users to add metadata to both their experiment records and other media generated during recording may provide the opportunity to capture valuable metadata to provide additional context. This may be particularly true for images such as photographs, graphs, and diagrams.

As expected, adopting different strategies or asking the subjects different questions results in the generation of different metadata. Each method of generating metadata has advantages and disadvantages. In general, Free recall appears to be effective for generating a diverse range of terms that could provide context to the experiment, but also produces a lot of irrelevant and personal information that is probably not valuable for use as metadata. These personal or opinion terms are probably of least value to others trying to make use of the experiment record. Changing perspective is effective for producing terms that would be relevant for search, reducing the amount of irrelevant information that is captured, and producing consistent values that are likely to be meaningful by different users. This condition produces the fewest responses, but specifically asking for search terms is much more effective than asking just for keywords for generating a higher number of terms. The Cued recall condition is very effective for producing high numbers of both relevant and diverse responses for the cued categories. However, two major downsides affect cued recall, if a cue is missing, important information may not be captured at all, and any ambiguity over the meaning of the cue can lead to confusion and a loss of information or the generation of irrelevant terms.

In Study 1, an information specialist was asked to curate the information objects used in the study to enable a comparison between expert generated terms and those of the study subjects. The librarian was also asked to provide terms for the viscosity experiment materials from the Bath study (similar to Study 4) and also the notes from a real organic chemistry experiment from a LabTrove notebook. These are the terms generated by the information specialist:

- Viscosity experiment terms: Instruction, Science, Experiment, Viscosity, Fluids
- Science, Organic chemistry write-up terms: Chemistry, Experiment, <Chemical Name>

The terms generated by the information specialist are generally at a higher level than the responses generated in the studies, and again suggest that including a specialist to add metadata at a later stage would be very beneficial for consistency in curation, but the metadata generated by the non-specialists using the methods in the studies is much richer in contextual detail.

7.7.1 Data mining

Studies 2 and 3 captured metadata at the same time as the experiment data, also described in 5.4 and 5.5. A useful and unexpected side effect of attempting to capture metadata at the same time as capturing the experiment record in this way was that it provided the opportunity to compare the relative effectiveness of asking the user to provide metadata versus automatic metadata creation using data mining. The major example is the materials used in the experiment, which were more specific and higher in number in the cued section, than in the procedure of the experiment. The same finding is also true for information for the other ‘metadata cues’. For example, no information about the location or people involved in the experiment would have been provided without asking these questions, and as mentioned only 50% of students included equipment in the same write-up, yet the responses to the appropriate cue in the Profile template generated an average of more than 5 items of equipment. The cue asking about chemicals and materials in the profile templates generated a range of chemicals with a mixture of different chemical names, common names and molecular formula. However, this variation in terminology is also common across all of the experiment reports and reflects the familiar problems of consistency of chemical naming in general. Our findings suggest that actually asking the researchers for information about the materials they are using in this way is actually more reliable than data mining because more terms are included in the cued response and they are not subject to the unrecognised abbreviations, typos and ambiguous values actually used within the reports. As described in 6.5.3.3, tools such as ChemSpider could be used to convert the metadata responses (and any recognised values within the remainder of the text) into unambiguous values such as the InChi standard or the SMILES specification for the purposes of metadata use and linking.

Chapter 8: Mobile Electronic Notebooks

8.1 Introduction

Electronic Laboratory Notebooks have the potential to be a revolutionary replacement of paper notebooks in the laboratory for the digital capture of experiments (Borman, 1994). They can serve a variety of roles and provide a variety of functions, such as facilitating the long-term storage and preservation of experiment records (Myers, 2003); ensuring standard operating procedure compliance and providing interfaces to instrumentation (Hice, 2009); and the apparently contradictory functions of supporting IP protection, collaboration, and open science (Taylor, 2011; De Roure & Frey, 2007). While some systems are repositories of raw data and results, others have the potential to support researchers through the whole experiment lifecycle (Fakas et al., 2005; Hughes et al., 2004).

Although Electronic Laboratory Notebooks are being increasingly used for scientific research, their use is still highly limited in academic environments (Rudolphi & Goossen, 2012; Goddard et al., 2009; Shankar, 2007). At an organisational level costs, licenses, and availability of appropriate IT support and infrastructure play a role in the adoption of ELNs (Coles et al., 2015; Bird et al., 2013a; Rudolphi & Goossen, 2012; Goddard et al., 2009). However, for the researchers that actually use the tools the largest barrier to overcome is the relative 'ease of use' of paper notebooks compared to electronic systems making paper notebooks easier to use, easier to read, easier to transport, inexpensive, flexible, annotatable, persistent, readily available, requires no training, and minimal IT support (Bird et al., 2013; Mackay et al., 2002; Goddard et al., 2009; Luff et al., 2004). On the other-hand, ELN software for taking notes is more difficult to use, takes more time, and is less flexible. Related to this are anxieties that researchers have about the stability - "computers crash much more than books!" - and the accessibility of ELNs; in terms of availability both when and where the researcher needs it (Coles et al., 2015; Drake, 2007).

ELN tools that are developed need to address these concerns and take account of the affordances that paper provide so that they are not a revolutionary change from the day-to-day experience that the scientists are used to (Lass, 2011). They also need to provide proven value and meet the needs of their intended user-base (Bird et al., 2013a).

Although there is a resistance to the adoption of Electronic Laboratory Notebooks in research environments, scientists nonetheless have high expectations about their capabilities including functionality, portability, and ease of use (Elliott, 2012).

Mobile device ownership is increasing. In the 2014 ECAR Study of Undergraduate Students and Information Technology it was found that 86% of undergraduates own a smartphone, 47% own a tablet, and 90% of students own a laptop (Dahlstrom & Bichsel, 2004). People are becoming more accustomed to easy, intuitive interfaces through the use of products such as smartphones and tablet computers; a trend that is also being noted by a number of ELN vendors leading to more focus on making ELN interfaces simpler to use (King, 2013). Tablets provide the opportunity to both develop apps that work in the way that researchers are familiar with and design interfaces that can take advantage of the strengths of the platform (Elliott, 2012).

While vendors are busy improving their interfaces and developing tablet and smartphone clients for their enterprise ELNs, other groups are taking advantage of free or cheap, easy to use technologies such as electronic notebooks available for these platforms. Many of these applications perform a subset of modern ELN capabilities, with note-taking applications such as Microsoft OneNote[22], Evernote[23], GoodNotes[24], and WritePad[25]; together with cloud backend applications such as Dropbox[26] and Box.net providing data management and collaboration capabilities; and image capture and barcode scanners providing input sources (Elliot 2012). These technologies do not necessarily provide all the functions that researchers desire and there is therefore still an opportunity to create a good digital notebook that is flexible

[22] http://www.onenote.com
[23] https://evernote.com
[24] http://www.goodnotesapp.com
[25] http://www.phatware.com/index.php?q=page/writepad/ipad#
[26] https://www.dropbox.com/about

and useful enough for researchers (Giles, 2012). To survive, any tool must be usable as well as functional.

8.2 Potential benefits of a mobile ELN

Mobile ELNs have the same advantages for researchers as ELNs, but also include additional advantages that help to mitigate some of the perceived drawbacks in comparison to paper notebooks. For example mobile ELNs on tablet type devices are relatively small size and have long-lasting batteries, making them considerably more flexible than ELNs on desktop computers, enabling them to be can be carried and used wherever they are needed. The use is not restricted to just the capture of experiment information in the lab, but can also be used to provide remote access to information that would not normally be contained within the paper notebook. This information could, for example, be the details of previous experiment results or instrument settings, research papers, notes from other researchers, or the experiment protocol.

In a traditional environment the scientist is often limited because information is tied to physical objects such a notebooks and working documents, which support a limited set of interaction models and data types, making it difficult to both access and share information (Arnstein et al., 2002). Digital capture methods, in particular the advanced capabilities of smart mobile devices, enable recorded information to be captured in a way that they cannot be on paper or easily using a desktop computer. For example it is common for lab scientists to draw pictures of their equipment set-ups or their results data such as TLC plates or gel electrophoresis (Frey et al., 2004), but a more effective way to record this information is to take a photograph of it. Mobile devices can also capture free-hand sketches, audio files, and videos, enabling information to be captured in the most appropriate way (Arnstein et al., 2002). Using these methods has the advantage that the data, such as a photograph, can be placed directly into the notebook and then exported as a resource that can be stored elsewhere, and used in other materials such as publications. Researchers that we have spoken to also indicate that they would like to be able to record sound, enabling them to be able to go ‘hands free’ in the laboratory.

A common complaint we have heard from chemists about existing ELN systems is their reliance on a network or internet connection and associated problems of accessing experiment information whilst in the laboratory, which often do not have a data connection or even disallow computers entirely. Researchers also want to able to access their research in other locations, for example, accessing their experiments at home or whilst travelling, which is difficult when the enterprise ELN system is hidden behind a firewall. As well as wireless access, smart phones and many tablet devices have access to data through the cellular data network, enabling access to data stored at other locations as even without a local network connection.

Having a connection to data on a network reduces the need to return to the office, and enables the researcher to look up information as they go during an experiment, for example looking up results or settings from previous studies, that are not necessarily their own (Arnstein et al., 2002). Another advantage of a mobile platform is that the data can be stored on the device and then uploaded to a server or other ELN service at a later time. This means that the data is accessible, and more information can be entered, wherever you want it in the lab, at home, or on the train, but that you can still ensure it is safety stored by exporting the data after you have finished in the lab.

Having the ELN on a tablet might also give more opportunity for researchers and students to explore the interface and to learn the system (UWM, 2012).

Although not every lab can afford to give its researchers laptops or iPads, it may be that making use of these devices is more cost effective than desktop computers and servers, for example 'The Cloud' may help lower administrative costs and infrastructure support requirements (Giles, 2012).

8.2.1 Note-taking in the Lab

The use of both paper and digital notebooks can be challenging in cleanroom or hazardous environments where neither may be permitted (Scholl & Van Laerhoven, 2014; Myers et al., 2001) and fear of contamination may make it undesirable to move the record between the office and the lab (Badiola, 2015). In such environments the protocol for the experiment and any notes need to be remembered by the researcher (Scholl & Van Laerhoven, 2014) and then written up in the office at a later point. Digital solutions that provide information access and notes capture in the lab using distributed tablet or interactive table top interfaces, such as LabScape (Arnstein et al., 2002), eLabBench (Tabard et al., 2012), and Biotisch (Echtler et al., 2010), or wearable technology (Scholl & Van Laerhoven, 2014) are more appropriate for these environments than mobile devices.

In environments without these issues mobile ELNs could have value, as described above. Many of the researchers that we have talked to expressed an interest in having access to electronic note-taking facilities in the laboratory, including access to LabTrove for those researchers already using the system.

8.3 Existing mobile digital notebook solutions

As can be seen in Table 8-1, there are a large number of projects that have attempted to develop digital systems to capture the notes made by researcher in the lab or field, in addition to those developed by commercial ELN providers or specialised electronic notebook applications such as Microsoft OneNote. Not all mobile notebooks rely on tablets or other smart devices for data entry, and many of the example solutions described in the table are in fact hybrid forms that make use of digital pens and paper notebooks, whilst others use interactive screens located in labs rather than mobile devices. Examples of digital note capture are included from both inside and outside the laboratory environment because many parallels can be drawn between note-taking requirements for researchers in both lab and field-working environments, with some researchers, for example environmental archaeologists working in both locations. Additionally some more generic digital capture systems are included because they capture and manage content that could be useful during the preparation, execution, and documentation of an experiment in the lab.

For the mobile ELN solutions described, several anticipate or require that the researcher will have created a plan for their experiment in advance of entering the lab, enabling progress of the experiment to displayed and often annotated to assist in the execution and recording of the experiment. For example, the Smart Tea system makes use of the requirement for researchers to complete a COSHH form listing the materials they intend to use in advance of starting the experiment. A web-based form based on the elements of the COSHH form asks the researchers to include extended descriptions of the steps for the experiment process, together with materials and quantities to create a plan (Hughes et al., 2004). Although this is extra effort, the benefits include helping the researchers to remember what to do next in the experiment and ensuring persistently legible results are available, demonstrating that ELNs can serve to augment the memory researchers, both in the lab whilst executing experiments and in the future through the storage and access to notes and information about a previous experiment and results (Scholl & Van Laerhoven, 2014).

Table 8-1: Digital solutions for capturing research notes

ELN name	Description
Paper ++ (Luff et al., 2004)	Paper ++ is a hybrid system making use of a digital pen and paper. Printed documents are printed with a pattern that uniquely encodes the x-y and page location on the document. The code is interpreted by a pen-like device, which then initiates some action by a connected digital device enabling the paper to behave like a touch screen. The tool allows associations to be made between printed materials and electronic resources, but focusing on reading rather than writing.
LabScape (Arnstein et al., 2002)	A smart environment for biologists in a cell laboratory to help them produce a more complete record of their research with less effort and improve access to information within the lab to benefit collaboration. A plan is created in the form of an interactive flowchart diagram which is used to visualize the procedure and it's progress during execution, and enables details of the experiment to be captured through annotations on a touch-interface, keyboard, and from instruments via ID tags and URLs. The tool also provides the ability to search database to answer queries during experiment execution. The touch interfaces are static, but connected to a back-end database positioned in different places in the lab so that the notebook and other information can be accessed where needed, augmenting the laboratory.

The Pacific Northwest National Laboratory (PNNL) Electronic Laboratory Notebook (Myers et al., 2001)	Forum like web-based interface for the laboratory, with chapters, pages and chronologically ordered notes with the capability for threaded discussions. Notes can be text, html, rtf, equations, sketches, screen images, and arbitrary files. There is also a form-based editor that allows users to gather pre-defined parameters and design the presentation on the page. Notes are tagged with the author and time-stamped. A search capability is also provided.
Smart Tea, myTea (Hughes et al., 2004)	A Web-based planning tool and tablet software for capturing data during experiment execution in the lab. The SmartTea system enables the capture of the process of a chemistry experiment from plan to execution. The planning part of the system is a web-form representing an extended COSHH form that additionally enables the definition of the steps of the experiment, as well as information about the materials to be used. The recording part of the system comprises an electronic lab book replacement on a tablet computer, pre-populated with the elements of the plan, stored in a flexible back-end storage system. Users can capture their wet and dry measures through a numeric keypad, and additional context by providing annotations in the form of hand-written notes or drawings for each step of the experiment.
eCat (Goddard et al., 2009)	eCat is an online ELN with a simple and customisable template-based data entry system. It enables the creation and linking of constructs, protocols and records. And the ability to perform searches and share records with others (MacNeil, 2011b). Tools are also provided for sample management. The majority of ELN functions are available for use on the iPhone or iPad[27].

[27] http://www.researchspace.com/electronic-lab-notebook/ipad_inventory.html#

NiCEBook (Brandl et al., 2010)	A Hybrid generic note-taking system. NiCEBook is a prototype paper notebook that uses an Anoto pen and paper to enable users to capture notes on paper as usual, while the notes are digitised in the background. Special control elements are included on the notebook pages to enable users to assign notes to categories, define topics, and also provide the capability to 'dogear' pages. The physical notebook can be synchronised with a digital notebook that can then be used to search for content, view the content in different ways, and share with other programs or users.
GIG (Reimer et al., 2011)	Generic desktop note-taking tool for students. Global Information Gatherer (GIG) is a system designed to help students in higher education to manage, understand, and keep their academic work by providing an integrative interface through which students can access commonly used programs, record notes, and organize files. Users can record notes alongside the files they are working with a rich text format editor, in which they can also include web links and 'snippets', and copy and paste content from other files.
Prism (Tabard et al., 2008)	Prism is a hybrid system for biologists that makes use of the Anoto pen and paper to capture hand-written text and drawings as time stamped images, an online journal and an activity log capturing xml links to emails, web pages, and other online documents.

A-book (Mackay et al., 2002; Mackay & Pothier, 2001).	The A-book project developed a number of hybrid prototypes to attempt to capture notes made by biologists. A tablet prototype making use of a graphics tablet (rather than a tablet computer or smart device), a paper notebook, and an inking pen connected to a computer. The notebook included a command menu attached to the side, and a pocket reader to detect the page number. The captured strokes are stored online as an image. The system recognises underlines and boxes; an underline indicates to the system that the information is text and can be used to open a URL or include details of a procedure; a box drawn around a sketch will save it as an image. Printed documents can be scanned using the pocket reader for tracking online and offline documents. A table of contents can be generated automatically. A prototype using a portable graphics tablet called cross pad was also used so that notes could be captured away from the desk, with a facility that also includes hand-writing recognition. The A-book itself was designed to be a portable device that would enable the notebook to be linked to physical objects and to enable the retrieval of information in different places as needed, and to enter data in locations when they don't have the physical notebook. The notebook is stuck onto a graphics tablet, and a small board computer is attached and runs a database and other software. The system detects the biologists gestures and uses them to reconstruct the notebook pages. An 'Interaction Lens', which is also a stand-alone PDA that can be used to browse the digitised notebook, is used to annotate or create links to physical objects by placing the filter on top of the notebook and using a non-inking pen to select items. These items can then be used to create links, add items to a table of contents, or for other actions within the PDA software such as tracking specimens.

ButterflyNet (Yeh et al., 2006)	ButterflyNet is a hybrid system to help support biologists in the field capture heterogeneous information that previously used a lot of manual labour to coordinate and distil. The system uses an Anoto pen and paper, a smart camera that can read barcodes, and 2D barcode tagged envelopes to collect specimens. These are used to link together to associates notes, specimens, and photos, including facilitating the transfer of captured content into spreadsheets for easier collaboration. Other sensors can be added to the system. Paper is still the central media for capturing notes.
eLabBench (Tabard et al., 2012)	The eLabBench is a digital table top system supporting biologists in the organization of personal information, capture of experimental work for later access and management in the office, and providing access to a variety of computational resources directly at the lab bench. The design of the system supports an open way of organizing information rather than requiring a strict sequential or time based organisation seen in workflow-based or notebook systems. It supports on the fly changes to the experiment, enables lightweight capture of information during execution, and enables computations to made directly at the bench rather than needing to go back and forwards to the lab and the office. The system comprises an interactive table top with access to the kinds of tools that the biologists would have access to on their computers in the office as well as a web-based notebook and an area for making rough notes. Interaction mechanisms include keyboard, mouse, pen, tagged physical objects such as racks and instruments, and buttons, for example to activate an overhead camera.

BioTISCH (Echtler et al., 2010)	The BioTISCH system replaces the traditional workbench with a glass covered table top system that presents information on the bench's surface and includes some augmentation of objects in the form of reagent sensors. The system allows the biologists to view plasmid charts from a database, displays an experiment log that allows the user to see the steps in the experiment and mark off each step as it is completed, a concentration calculator is used to perform conversions, and the reaction sensors highlight the next reagent to be used.
CERF for iPad[28][29]	CERF for iPad is an example of a native tablet client from a commercial ELN developer. It is intended for researchers to use in the lab but connects to an enterprise CERF ELN for storage. There is an offline function to overcome network issues, and offers web browsing, text entry, audio recording, camera, specialized science app functionality, and pasting files.
MyLifeBits (Gemmell et al., 2006)	MyLifeBits is a life-logging system developed by Microsoft Research to manage a vast database of captured materials. The system enables users to organize, search, annotate and make use of the content of the database, which includes stored digitised archival material (e.g. articles, books, music, photos, and video) and natively digital documents (e.g. office files, email, digital photos), web pages, phone calls, meetings, room conversations, keystrokes and mouse clicks for every active screen or document. The system also captures content from the wearable SenseCam camera that automatically takes 1-2 thousand photos a day, and records environmental and location information.

[28] http://irisnote.com/wp-content/uploads/2013/02/iPad-Case-Study-2013.pdf

[29] http://irisnote.com/ipad-client/

RAMSES (Ancona et al., 1999)	A field-based system for the digital capture of notes during archaeological excavations and surveys. RAMSES or Remote Archaeological Mobile Support Enhanced System is composed of a fixed station that acts as an active object repository, whilst mobile units input archaeological evidence by means of an electromagnetic pen.
The Living Cookbook (Terrenghi et al., 2007)	A home-based activity recording and playback system for collaboration and learning. This is included as similar technologies could easily be utilised in a laboratory for capturing an experiment procedure, and the playback function would be an alternative way to deliver teaching within the lab or across distributed locations. The Living Cookbook enables people to share their cooking experiences, to educate others in the cooking practice, and to suggest a sense of presence and sociability. A touch screen tablet PC with a simple user interface implemented in and mounted to the door of a kitchen cabinet. To this is attached a camera and a projector. The system enables the user to both record their activities with the camera using a simple interface, while they speak aloud their 'kitchen story', and also to playback and control other recordings in order to learn from others.
iTrench (Warwick et al., 2009)	A hybrid field system trial for capturing digital notes. A trial to investigate how archaeologists work with IT on an archaeological site using digital pens and paper and Nokia N800 handheld PDAs to record their work. The digital pens were successful, but the PDAs were not because of the extreme conditions. Problems included the screen being too reflective to be seen in bright sunlight, font size too small, the stylus was uncomfortable to use, and the sketch function was awkward and inadequate quality. In contrast the digital pens were seen as easy to use and actively useful. Other examples of digital notes recording and their uses can be found in Cianciarulo & Guerra, 2007, and Jewell et al., 2012.

Wearable digitization of life science experiments. (Scholl & Van Laerhoven, 2014)	The Wearable digitisation system augments the researcher in the lab. The system is designed for biologists to help in both the capture of experiments and in accessing information during the execution of an experiment. The system comprises a head mounted display (HMD, Google Glass), a Smart watch that records the wearers motion and includes an RFID-reader can support the scientist during the experiment by providing access to information, and by recording elements of the experiment. For example, the watch can record wrist motions to use for activity detection, the RFID reader with tagged containers can identify materials used, the Google Glass can take pictures and record and comment key observations.

8.4 Challenges of developing for mobile devices

There are challenges in developing applications for mobile devices. It is not simply a case of taking an existing solution and compiling it for use on mobile platforms. The market for ELNs is relatively small, and development of applications is expensive. It is not appropriate to port a desktop application that is not designed for use in a 'touch-interface' environment, to a mobile device. Applications usually need to be developed from scratch, and costs further increase with each different device you attempt to support. Each brand of device has it's own programming language and applications need to tested across the range of different devices and different operating systems (Elliot, 2012). Other challenges for mobile devices include dealing with presenting information in a relatively small space and providing enough information to help researchers navigate the system and find the information or screen that they are seeking (Cooke and schraefel, 2004).

Chapter 9: Notelus: An Exemplar Mobile ELN

9.1 Introduction

Notelus[30] is an Electronic Laboratory Notebook developed as a native application for the iPad. Notelus is designed to support researchers, in particular synthetic chemists, to complete their experiments and capture their results digitally in the lab. The project was a proof of concept activity to rapidly produce an initial prototype, and not a fully featured mobile ELN, but the app provides basic note-taking and sketching capabilities and makes use of the extended features of the platform to enable the capture and inclusion of photographs and audio recordings in the notebook.

A look-up function enables researchers to search for materials from the free ChemSpider[31] chemical database using the ChemSpider API[32], and then include an appropriate image within their experiment. All data is stored locally on the iPad, so there is no requirement for network or data access within the lab or other locations where notes may be captured. An experiment notebook can be exported in a number of formats to enable permanent storage and sharing. The primary method of export is to a LabTrove Electronic Laboratory Notebook (Milsted et al., 2013), but researchers that do not use LabTrove also have the option to export their notebooks using email, enabling them to save the resources to a commercial ELN solution, or their file system.

From discussions with users of LabTrove and researchers that had little or no experience of ELNs we were aware that researchers have a significant attachment to paper notebooks and reservations about the usefulness of ELNs (Bird et al., 2013a). One of the main influences on the design of Notelus was to try and utilise the notebook metaphor to provide a bridge between the familiar and comfortable paper notebook interface and the digital ELN. We used the notebook metaphor to provide a book-style interface with pages that can be drawn on and turned like a book. Studies investigating digitally captured notes have found that researchers desire to be able to ‘flip through pages’ in a

[30] https://itunes.apple.com/us/app/notelus/id593269701?mt=8
[31] http://www.chemspider.com/
[32] http://www.chemspider.com/AboutServices.aspx

notebook (Cooke & schraefel, 2004; Mackay et al., 2002). The use of an actual digital book interface where the user can flip through pages to see data and the steps of the experiment should help provide a bridge to the paper notebook experience.

A secondary aim of the project was to provide support for running the experiment in the lab, by providing a way to step through an experiment plan and record observations using the plan as a skeleton. As a result of this activity researchers were given the mechanism to create a simple plan for their experiment that defined the basic steps in the experiment, the materials and quantities to be used, an image for the proposed reaction scheme and the equipment to be used. This plan could be used to pre-populate a notebook with additional information. Each page in the book is titled with one of the steps in the correct sequence enabling the researcher to see the next step in the experiment; whilst information about the materials, hazards, quantities, and equipment are accessible from every page.

The experiment plan can be created by hand using a very simple XML format to describe the steps, materials, and equipment in the experiment, together with details about the plan such as a title, author, and keywords. It is unreasonable to expect that chemists would have experience of working with XML and therefore a second app was developed as a companion application to enable the creation of experiment plans using an easy to use interface. Plan Buddy for Notelus[33] generates an XML file in the appropriate format that can then be opened with Notelus, ready to create a new pre-populated experiment notebook, or opened in another editor for viewing. Experiment plans can be reused with Notelus as many times as required, and copied within Plan Buddy to enable researchers to create new plans based on similar experiments they have planned before, enabling effective reuse and saving time.

9.2 Implementing a mobile ELN

The design and requirements for Notelus were developed based upon the knowledge gathered from the user research activities described in **Chapter 3: Keeping and managing records** and by talking directly with synthetic

[33] https://itunes.apple.com/tw/app/plan-buddy-for-notelus/id663881731?mt=8

chemists about how they work, as described in more detail below. Issues raised by user testing with researchers during the development process also contributed to the final design of the ELN. The design and development of Notelus was a parallel stream to the research activities included in this thesis, and therefore the interaction design artefacts generated in **Chapter 4: Interaction design and tools for keeping and managing records**, were not included in the process, but some of the design decisions were particularly influenced by the findings from **Chapter 6: Use of user-defined metadata in ELNs and other platforms**.

The Notelus app was specifically targeted for use by synthetic chemists, and it was therefore important to understand the requirements for a mobile system from these researchers. In order to gain an understanding of their workflows for planning and executing experiments and how they currently work with their data and organise their workflows we interviewed and observed members of a synthetic chemistry group at Southampton University. Each of the researchers were engaged in a variety of different research, all involving practical laboratory work and one that also included experiments ‘in silico’.

Three of the researchers were interviewed and observed in detail, while the others were involved in more open discussions about how they worked and their experiences with ELNs. For the in-depth interview, the questions and observations covered the whole experiment lifecycle, from initial background research and planning of an experiment, through what they do and how they capture their observations in the lab, how their data is captured and analysed, and how they prepare their reports. Observations included how they use computers and software, organise and store their work, and how they use paper notebooks or ELNs currently. Each researcher also provided a walk through of the areas that they work in the lab and the actions that they would take within those areas, returning to their desks to show examples of their data from a typical experiment and how they processed and organised it. Examples of the current task flows and information organisation used by the observed synthetic chemists are shown in Figure 3-2, Figure 3-3, and Figure 3-4 on pages 38-39. Each researcher was performing different experiments and each had their own different ways of working and organising the information they produced: recording everything in a paper-notebook, use of a commercial ELN, and a self-designed personal information system using Microsoft Office files.

9.3 Use cases for a mobile ELN

The following use cases were derived for the Notelus project from our observations of the synthetic chemists and our previous user research. Items marked with ★ were deemed out of scope for this project and those marked with § were considered partially in scope.

General:

- As a chemist I want to know that any content I create is safely stored (that I won't lose it, and no-one can access it without my permission)
- I want to be able to view any content that I produce in a list by experiment name, or by date
- I want to be able to search any content that I produce by experiment name, materials, equipment, or procedures §

Planning:

- As a chemist I want to be able to create a plan for my experiment
    - I want to define my reaction scheme §
- I want to define the materials, equipment, and procedures that I intend to use
- I want to link to create links to background material for the experiment ★
- I want to look up safety information §
- I want to create a table of amounts and concentrations of materials that I will use in the experiment
- I want to share my experiment design (with my supervisor, with my peers, with the community)
- I want to be able to take a copy and modify an existing plan

Recording the experiment:

- As a chemist I want to be able to record my experiment as it progresses
    - I want to follow the instructions that I created in my plan

    - I want to see the safety information about the materials I am using
- I want to complete the measurements I actually used in my table of materials
- I want to record what I did, and what happened in my experiment using notes, sketches, and calculations.
- I might also want to record what I did using photographs, videos, and sound recordings
- I want to be able to amend my notes §
- I want to be able to add links to any analysis of materials from my experiment to my notes ★
- I want to record my findings
- I want to share my record of the experiment (with my supervisor, with my peers, with the community)

Writing up and publishing the experiment:

- As a chemist I want to be able to write up my experiment (for my supervisor, for publication) ★
- I want to be able to use my experiment plan, record, and results in a report writing tool
- I want to be able to extract any sketches, photographs, videos, sounds, results as required to use in my report
- I want to package my raw data and analysed results with my report for distribution
- I want to distribute or publish my report ★

9.4 Testing and user verification

During the development of Notelus it was important to get some feedback from a real user in the lab. One of the synthetic chemistry researchers, not included in the initial information seeking exercise, conducted testing of Notelus during the development process using the application to run and record his experiments. This researcher was already familiar with a number of ELNs that he had also been testing during his experiments, although he was new to LabTrove. He manually created a plan for his experiments and imported them into Notelus. The feedback from this researcher was extremely valuable and directly contributed to the decision to add a number of features before release that were either not considered in the original plans or had been considered to be less important than they were found to be in practice.

Although not all of the suggestions were possible to include in the proof of concept, features implemented based on the feedback included the creation and use of folders for organising Notebooks, providing a properties tab for viewing and editing descriptive information about the experiment, and providing the option to delete imported plans. The decision to create the Plan Buddy tool to coincide with the App Store release of Notelus was also made because of the difficulties that the researcher had with understanding and creating an XML plan.

9.5 Using Notelus

From discussions with researchers it was clear that the application should fulfil the role of replacing the paper notebook in the task flow that was being used, in order to be of most value. A plan for the experiment could optionally be created so that information about the steps of the procedure, equipment to be used, and measurements to be taken, would be available to help run the experiment, in particular reminding the researchers about where they have got to and what the next step will be. The creation of the plan would not be an additional step and would in fact save effort in the long run, because multiple runs of the same experiment could use the same plan and plans could be reused as a starting point for planning a similar experiment. There would also be no need to write down the procedures and materials in the notebook because these elements could be imported automatically.

The task flow for an experiment using Notelus with a plan can be seen in Figure 9-1.

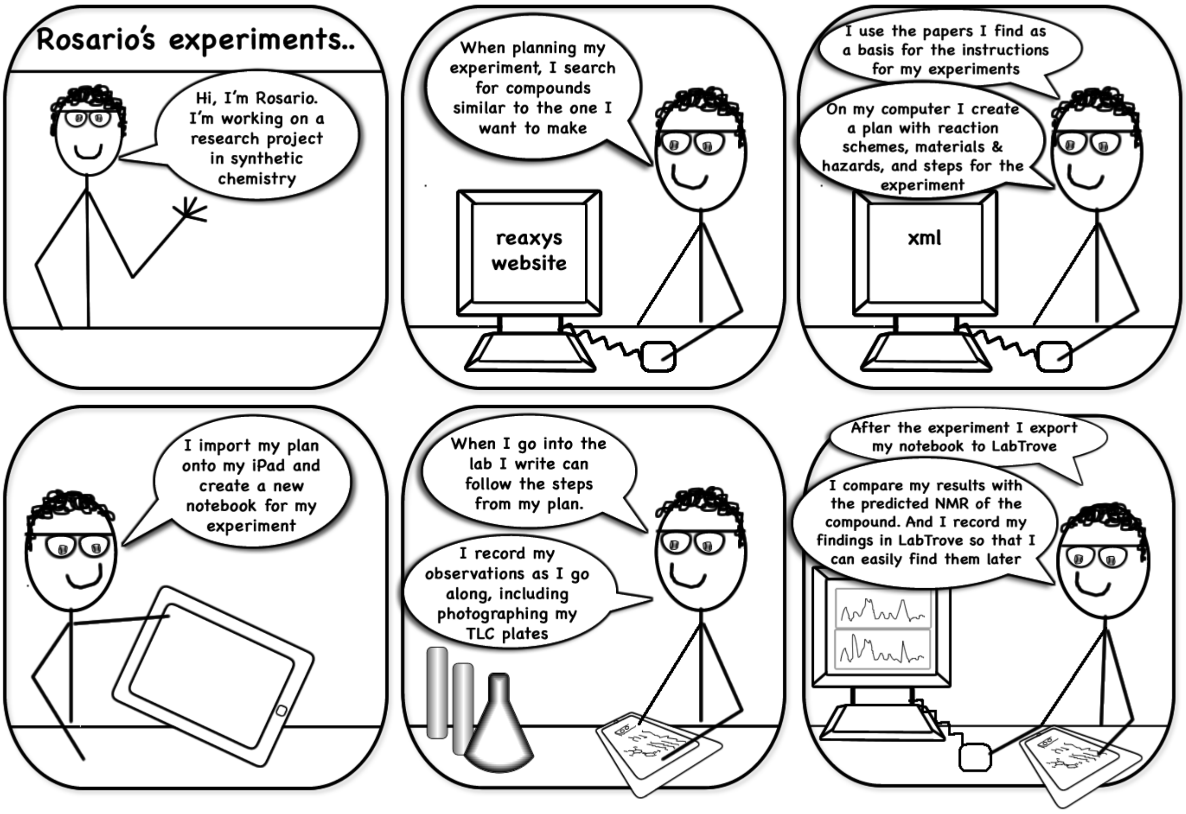


Figure 9-1: An example of the experiment process using Notelus and LabTrove

9.5.1 Notebooks in Notelus

A single Notelus notebook represents a single experiment record, and a new notebook should be created for each new experiment. There is no restriction on the number of pages in a notebook that is not created from a plan (a blank notebook), but the higher the number of pages, the more difficult it will become to locate information within it. A notebook created from a plan contains the same number of pages as steps defined in the experiment plan. Each page contains the appropriate experiment step as the title to the page, guiding the user in the next task to complete and allowing notes and observations to be linked to the specific step. It is possible to create additional pages for any experiment step if more space is required to enter observations and images.

9.5.2 Creating and organising notebooks

The first task when using Notelus is to create a new notebook and to provide it with a name. There are three options for creating a notebook, to create a new blank notebook, to create a tailored notebook from an imported plan, or to create a new customised notebook from a plan on the web. The home screen of the Notelus application provides buttons for these actions and also displays a gallery of the existing notebooks. A single notebook is displayed in the gallery when the app is first installed on an iPad providing some basic user assistance to get started using Notelus.

To help the researcher organise their notebooks when they have multiple experiments, they can create and name folders into which notebooks can be moved. The notebooks and folders are ordered by the creation date as a default, but buttons are available in the gallery view to switch between a date and alphabetical sort order for the notebooks.

To open a notebook (or folder) the user can tap on the icon to open it. Other options for interacting with individual notebooks can be displayed by performing a ‘long tap’, that is holding the finger on the notebook for a second. A menu of available actions is displayed. The available actions include ways to export the notebook, move the notebook to one of a list of folders, and an option to delete the notebook. The notebook gallery and action menu are shown in Figure 9-2.

Figure 9-2: Notebook gallery with notebook actions menu displayed

The delete action is available to help the researcher organise their notebooks and to provide a way to control the amount of space used on the device. Many notebooks with large numbers of images and sound recordings will take up a lot of space on the device. The intention is that the researchers will keep the notebooks on the device for as long as they need them for, but will ultimately export the captured information to a different location such as Dropbox or another ELN so they retain access to them. Once the researcher has finished updating their experiment, exported and safely stored the notebook files in another location, they can go ahead and delete the notebook. A confirmation dialog is presented to the researcher when they tap the Delete notebook action, so it is not possible to accidentally delete notebooks.

Notelus captures some basic metadata about the notebook when it is created, such as the date and time of creation, name, whether the notebook is associated with a plan and the details of the plan.

9.5.3 Recording an experiment

When a user taps on a notebook in the gallery it is opened. The user can using the ‘pan’ gesture to turn the pages of the notebook. In a blank notebook, two pages are added to the notebook by default. When the researcher wants to add more content, they just turn the page and a new page is added automatically.

The buttons available in the notebook depends upon whether the notebook is created from a plan or is a ‘blank notebook’. At the top of the notebook are various options for creating and manipulating content in the notebook. These buttons are available in both types of notebook. In notebooks created from a plan an additional set of buttons are provided at the bottom of the screen to provide functions for viewing and interacting with information from the plan, and also navigating through the steps of the experiment. Figure 9-3 and Figure 9-4 show examples of pages in a ‘blank notebook’; and Figure 9-5 shows an example of pages in a notebook created from an experiment plan.

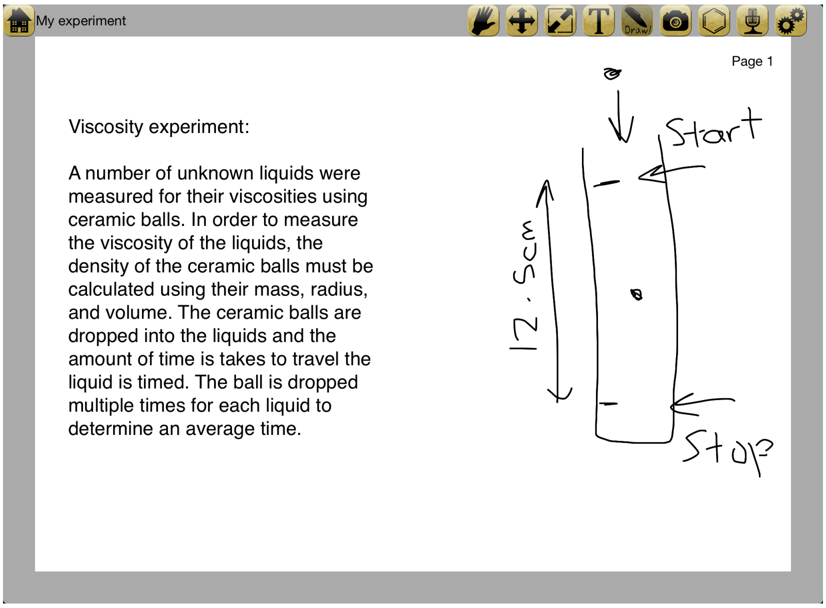


Figure 9-3: Page from a blank notebook showing a sketch and text

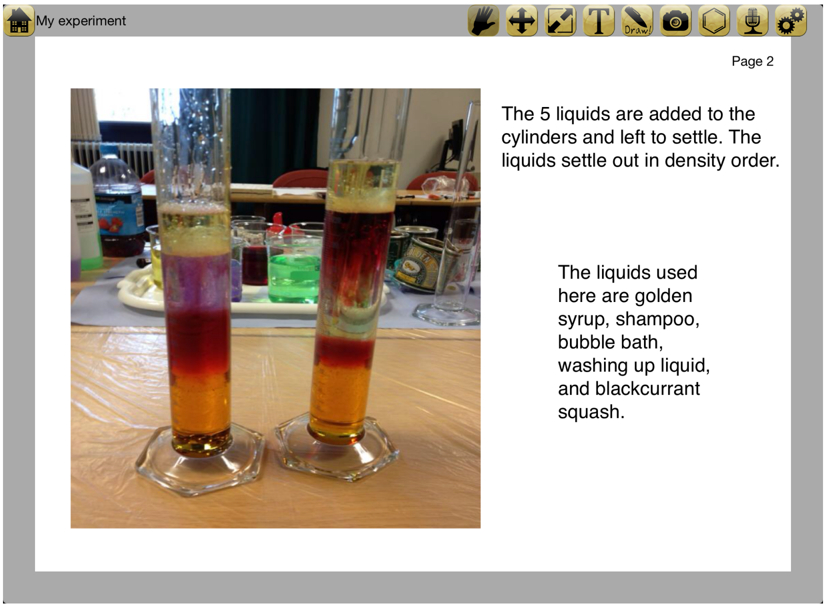


Figure 9-4: Page from a blank notebook showing a photo and text

The Draw feature enables the researcher to draw sketches on the screen and take rough notes using their finger or a stylus. These sketches are static and cannot be manipulated once they have been added. The Text feature enables the researcher to type text onto the screen using the on-screen keyboard or a separate blue tooth keyboard if this is attached. The text is added in a selectable text box that can be moved around the page and resized to enable it to be positioned on the page. The Photo feature enables the researcher to take a photograph or to select an existing image from the iPad photo gallery. The selected photo is added to the page and can be moved around or resized.

A tool is provided to enable the researcher to perform a search for a material from the ChemSpider database. This tool requires an Internet connection to be available and uses the ChemSpider Web service API to perform the search and retrieve information and a chemical structure for any matched materials. The researcher can choose to add the material to their notebook. This action will add an image of the chemical structure to the current page in the notebook and can then be positioned and resized as required. The user must be a registered member of ChemSpider to use the search because ChemSpider requires a user security token to be sent when requesting information using the Web service API. Entering a user security token enables the search. The security token is saved in Notelus, but there is an option to edit it if needed. If the user is not already registered an option is available to open the

ChemSpider website within Notelus to register without needing to leave the app.

Another way that the user can capture their observations and notes is through sound recording. The Sound function enables the user to make multiple sound recordings and play them back. Each sound recording is tagged with the date and time of creation and is available from any of the pages of the notebook.

The Properties popover enables the user to configure information about their experiment such as the name of the project, the name of a sub-project, a description of the experiment, the conditions of the experiment, and also a list of keywords that are used to describe the experiment. If the notebook is created from a plan the author can choose to provide this information as a part of the plan. The values for the properties are editable for both types of notebook.

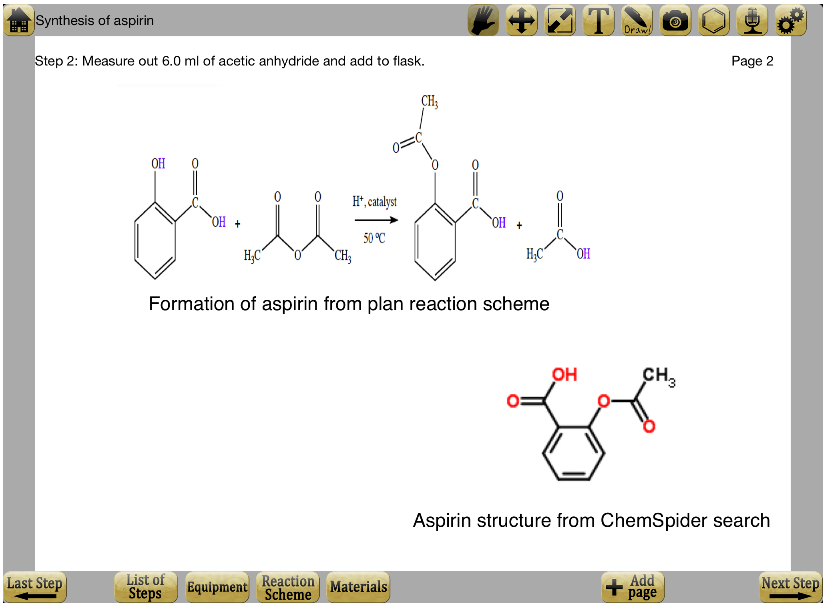


Figure 9-5: Notebook created from a plan with a reaction scheme and chemical structure from ChemSpider

For notebooks created from a plan the buttons at the bottom of the notebook help the user to navigate between steps in the experiment, but providing options to go to the last or next step, or jump directly to a step from the list of steps.

A list of equipment is visible from the Equipment button. If images of the reaction scheme were included in the plan, then these can be viewed and optionally added to the current page of the notebook. Details of the materials that are used in the experiment can be seen in the Materials table. The table includes the name of the material, the amount to be used, and hazard information. The user can update the table with the actual amounts of material that were used in the experiment by editing the value in the table. Updated values are included with the notebook resources when the notebook is exported.

For both types of notebook, any content that is added to the notebook is automatically saved onto the device as they go along.

9.5.4 Exporting a notebook

Once an experiment has been completed the resources relating to the experiment can be exported for reuse, manipulation, and storage elsewhere. There are a number of choices available on where and how to export the notebook. As mentioned previously, options are provided to enable a use without LabTrove to export their notebooks so that they can use a different ELN or storage service. To do this the notebook can be exported via email. This is also useful as a mechanism to share the notebook with others, for example a supervisor. Two options are provided for export via email, to send a notebook as a PDF or to send a zip of all the resources. The alternative option is to export the notebook to LabTrove. The options for exporting to LabTrove will be discussed in more detail below, but include the options to export as a PDF, as a zip, or individual files.

The PDF of the notebook captures the notebook exactly as it looks on the screen, so is essentially a copy of the paper notebook in digital form. The intention is that in this format it is readable but cannot be manipulated. If the notebook were created from a plan, then information from the plan is included in the first few pages of the PDF, this includes the reaction scheme images, equipment list, materials table, and the steps of the experiment.

The zip file splits the individual elements of the notebook up into the appropriate formats. Any text added through the text tool is added to a text file, along with the equipment list, information from the materials table, and the steps of the experiment if the notebook was created from a file. Any notes or sketches made with the Draw tool are converted into images, one for each page. Any chemical structure, photo, or reaction scheme included in the notebook has an image created for it and included in the zip file. Finally any sound files are included in the zip. The option to export only certain files exists in the export to LabTrove option. The user can choose to optionally add all images, text, sketches, and sound files to the zip. Alternatively, the user can choose to export the files as individual files. The advantage of using individual files is that it enables the user to display thumbnails and links to the individual files within their LabTrove notebook entry. Choosing to use a PDF, zip, or individual files depends on the intended use of the materials by the user. If LabTrove is used as a repository then using a PDF is suitable if the user does not intend to do any further additions or manipulations of the material in the notebook. The zip file would be the more suitable option if the user intends to add material or to take the material for use in a different tool, for example, to write a formal report or publication. If LabTrove is used as a sharing and collaboration environment choosing to export both as a PDF and as individual files may be appropriate.

In order to use LabTrove, details of at least one LabTrove account must be configured. An account is needed for each notebook within a LabTrove instance, but any user can have multiple notebooks on multiple LabTrove instances. When a Notelus notebook is exported to LabTrove a new entry is created within the selected LabTrove notebook with the Notelus files are added to that entry. The user can choose to compose text content for the LabTrove entry and optionally to automatically include metadata to the entry. The metadata is described in more detail below.

9.5.5 Designing and importing a plan

A key function of Notelus is to provide information to the user to help them run their experiment as well as to capture information as they enact the experiment. In order to have this information included in a notebook the user must make a plan in advance of the experiment. This experiment can then be imported into Notelus and used to create a custom Notebook, complete with information about the steps to follow, the equipment to be used, reaction schemes, a table of materials and hazards, and properties about the experiment. Information that is included in the plan is also included in the output formats for the notebook when it is exported, for example the materials, steps, and equipment

A ‘Plan parser’ is implemented in Notelus that extracts the elements of a plan from an XML file. Users can choose to create an XML file using the appropriate format. Two examples can be found at the end of this chapter, one is the example plan from the Notelus help and the other is a plan from an actual synthetic chemistry experiment used by a researcher to test Notelus. The XML file can be imported into Notelus from the web if it is stored as a file on the web, for example shared on Dropbox or on LabTrove, or by emailing it to an account on the iPad device. Alternatively a user can use "Plan Buddy for Notelus" app from the Apple App Store to create the plan and import it in Notelus.

The XML plan format contains the following sections that contain individual elements that describe the details of the experiment:

about-plan This section contains information about the plan. A unique identifier for the plan, a name for the plan, and the plan author must be provided.

about-experiment This section is optional, but is used to populate the notebook settings in Notelus with the project name, a sub-project name, conditions, a description, and keywords. These values can be updated within Notelus, and are also used as metadata if the notebook is imported into LabTrove.

reaction-schemes This section is optional. Links can be added to images of the reaction schemes for the experiment. A name must also be provided for the reaction scheme in this section. The links to the images must be publically accessible.

materials This section details the materials that are used in the experiment, and provides the content for the 'Materials" table in Notelus. Each material must have a material-name, material-description, material-safety, molecular-weight, density, ratio, planned-amount, and a planned-amount unit. “-“ can be used where ratio is not relevant. At the current time, the planned-amount-unit must be g or mL.

equipment This section lists equipment that will be used in the experiment.

procedures This section lists the steps in the experiment, and is used to create a page for each step in the experiment. A step-number and a procedure-description must be provided.

The plan structure is quite simple, making is easy for any users that do already have an understanding of XML to put one together themselves quickly and to create new plans based on existing XML. You can reuse a plan over and over to create new notebooks in Notelus, so it is possible to run the same experiment repeatedly. Existing plans are easy to adapt, and XML has the advantage of being human readable as well as machine readable making it possible to share the plans with the community, for example using myExperiment (De Roure, 2009). There is also scope to develop the plan XML in the future to tailor it for particular disciplines or to handle more complex cases. Providing a digital form of a plan that can be cloned, adapted, and used to create a customized notebook helps to reduce the duplication of effort and reduces the risk of errors compared to using a paper notebook for recording (Taylor, 2011).

Plan Buddy was created in order to make it easy for scientists unfamiliar with XML to create plans on the iPad and import them directly into Notelus. Plan Buddy creates the same structure XML as described in the section above. The user essentially completes a form, adding the elements to include in the plan. The Plan Buddy app is very basic and has only minimal verification of the formatting in order to support the use of the proof of concept Notelus app.

In the future a more sophisticated plan schema can be developed, and a more sophisticated plan creation tool can be created to support it. There is some value in having an iPad based tool, as it means that a plan can be created on the move and without the need to have an internet connection, but the new applications could be created could be web or desktop based. There has been some discussion on developing a plan creation function within LabTrove in the future.

9.5.6 Metadata – annotating the experiment

Metadata is useful for access, search and retrieval, reuse, and providing context for documents including experiment records (Gilliland, 2008; Zeng &

Qin, 2008; Bird et al., 2013b) LabTrove provides a flexible framework for metadata, including both machine-generated and user-generated metadata types (Willoughby et al., 2014; Milsted et al., 2013).

For the user generated metadata there are two types, a Section, best used to describe the content of the entries, and key-value pair data useful for describing specific elements of the experiment such as codes, instruments, equipment and materials. We felt that it was important to make the best use of the metadata capabilities of LabTrove by providing the user with the opportunity to add metadata to their LabTrove entries based on Notelus notebooks. The metadata is automatically generated based on information provided by the user and the plan author if a plan is used.

Notelus is used as the default value for Section. It is straightforward for the user to change this value within LabTrove, but this value could be customisable within Notelus in the future. For notebooks that are not generated from a plan, the *Project name*, *Sub-project name*, *Conditions*, and *Keywords* information from the experiment properties are used to create key-value pair metadata. The user is able to create a list of keywords separated by commas, with each keyword a value for the keywords key.

Notebooks that are created from a plan have additional metadata derived from the plan. Key-value pairs are created for each item of equipment and material included in the plan using the keys *Equipment* and *Material*. Key-value pairs also created for the *Plan ID*, *Plan Name*, and *Plan Author*.

9.6 Implementation of Notelus

Notelus is implemented using Xcode[34] and Objective C[35] for the iOS (iPhone, iPad, and iPod) operating system[36] (v6.0). The app uses some features that only became available in v6.0, so requires the iPad to be on at least that version of

[34] https://developer.apple.com/technologies/tools/

[35] http://developer.apple.com/library/mac/documentation/Cocoa/Conceptual/ProgrammingWithObjectiveC/Introduction/Introduction.html

[36] http://www.apple.com/ios/what-is/

the iOS operating system. A single Storyboard[37] is used for the majority of the interface designs with segues[38] to provide the task flow between the different screens. A Collection View[39] is used for the notebook gallery (Notelus home screen) because it provides functions for displaying, adding, and deleting icons in a gallery style layout. A Page View Controller[40] is used to control the notebook part of the application because it provides the interface for working with turning pages and handling navigation between an array of pages.

Objects in the internal data model include collections of notebook references and folders; notebooks containing references to plans and arrays of page references; plans containing steps, materials, and equipment; and pages containing paths (sketches and notes created with the Draw tool), text, and references to images. The objects in the internal data model are stored in files on the device using NSKeyedArchiver[41]. Images that are added to the notebook as photographs or chemical structures are stored together with the notebook as image files. When a plan is imported with reaction schemes, the images are stored locally in the Documents[42] folder of the Notelus app.

Web services calls are used to communicate with ChemSpider and LabTrove. In ChemSpider a number of calls are made to perform a chemical search and to retrieve information and a chemical structure that can be added to the notebook. The user enters a search term and a call is made to SimpleSearch[43]. If a match is found, then two further calls are made, one to

---

37 http://developer.apple.com/library/ios/documentation/General/Conceptual/Devpedia-CocoaApp/Storyboard.html#//apple_ref/doc/uid/TP40009071-CH99

38 https://developer.apple.com/library/ios/documentation/UIKit/Reference/UIStoryboardSegue_Class/

39 http://developer.apple.com/library/ios/documentation/WindowsViews/Conceptual/CollectionViewPGforIOS/Introduction/Introduction.html

40 http://developer.apple.com/library/ios/documentation/WindowsViews/Conceptual/ViewControllerCatalog/Chapters/PageViewControllers.html

41 https://developer.apple.com/library/mac/documentation/Cocoa/Reference/Foundation/Classes/NSKeyedArchiver_Class/Reference/Reference.html

42 http://developer.apple.com/library/ios/documentation/FileManagement/Conceptual/FileSystemProgrammingGuide/FileSystemOverview/FileSystemOverview.html

43 http://www.chemspider.com/Search.asmx?op=SimpleSearch

GetCompoundInfo[44], which returns strings for CSID, StdInChIKey, StdInChI, and SMILES; and one to GetCompoundThumbnail[45], which returns an image of the chemical structure in PNG format. The ChemSpider Web service calls require a registered user's security token to be sent as a part of the call. The security token can be seen on the profile page of the registered user on the ChemSpider site.

In LabTrove a number of Web service calls are made depending upon the options that the user chooses when they export their notebook. Data must be uploaded before the LabTrove notebook entry is created using the adddata call[46]. For the PDF and zip file export options the data file is attached to a single adddata call. The adddata call responds with the ID of the data item that has been added from the LabTrove database. This dataid is then used in the addpost call to create a new entry in the selected LabTrove notebook including the dataid of the PDF or zip file to link them together. The addpost[47] call returns the url of the newly created notebook entry. When individual files are requested as the export option a adddata request is sent for each file. An array is created to keep track of the returned dataids. When all of the files have been sent, an addpost request creates a new entry in the LabTrove notebook with the list of individual files attached.

9.7 Limitations and future directions

Notelus is not a perfect application but it does serve the purpose of demonstrating what could be done to bridge the gap between the flexibility and familiarity of paper notebooks and the more formal and restrictive nature of ELNs. The use of mobile applications is the logical step in bringing about the reality of a paperless lab. It is essential to ensure that usability is at the forefront of such developments to encourage adoption and also to ensure that the users get the best value out of the functions that are provided (Bird et al., 2013).

---

[44] http://www.chemspider.com/Search.asmx?op=GetCompoundInfo
[45] http://www.chemspider.com/Search.asmx?op=GetCompoundThumbnail
[46] http://docs.labtrove.org/dev/lt/Adding_data
[47] http://docs.labtrove.org/dev/lt/Adding_a_post

As mentioned previously, the scope of ELN software is potentially very broad, but Notelus was designed to fulfil only one a slice of this scope. Notelus was designed to assist the user in conducting their experiments and to provide them with a variety of mechanisms to record their observations and notes, for use during the experiment and for future preservation and manipulation. Notelus is not attempting to replace the desktop or online ELN, but rather to supplement it in the lab. Notelus is also not attempting to be a reporting tool within itself, but the various resources that are captured within the notebooks can be exported to other systems as raw files that can be used to construct reports; these files can also be converted to PDF or zipped as part of a package of supplementary materials.

9.7.1 Extensions to Notelus

Notelus actually provides functionality for the majority of the use cases that were gathered as a result of the discussions and feedback from the user community. As mentioned above, Notelus was not designed as a report-making tool, and it makes more sense for this activity to be completed in the usual tools that researchers currently use for the purpose. For the other 'out of scope' use cases an extension to Notelus could provide access to background materials and results through the use of a Web view[48] to access web-based resources and a PDF renderer[49] to open PDF files from within Notelus. Options could be made to embed the URLs of web-based resources, copy and paste text and images into the notebook, and include these resources in the 'notebook bundle' for later export and archive. It would also be possible to integrate Notelus with other systems that provide an API for connection to an IOS apps, for example Dropbox[50], Box[51], and Mendeley[52].

---

48 https://developer.apple.com/library/ios/documentation/Cocoa/Reference/WebKit/ObjC_classic/index.html

49 https://developer.apple.com/library/ios/documentation/GraphicsImaging/Conceptual/drawingwithquartz2d/dq_pdf/dq_pdf.html

50 https://www.dropbox.com/developers/core/sdks/ios

51 https://developers.box.com/sdks/

52 https://github.com/Mendeley/mendeleykit

For those use cases that are partially fulfilled in this version of Notelus there are also possible extensions that could be made. Currently if a notebook is exported to LabTrove, then is it possible to search the text content in any posts and associated metadata if this has been added. Currently it is possible to locate an individual notebook by scrolling through the list of notebooks and sorting by name or date, but there is no option to search for a particular notebook or search the contents. Notelus could be extended to add a search facility that searches by the name of notebook, text content within the notebook, and through the contents of plans associated with notebooks in order to generate a filtered subset of notebooks for the researcher to choose from. In addition the search and navigation of experiments could be improved by incorporating an experiment signature into the home screen. The experiment signature would be based on key attributes of the experiment, especially the reaction scheme, that make it easier to recognise the content of the associated record and find the experiment they are looking for (Cooke and schraefel, 2004). The menu style in Notelus would enable the researchers to 'slide' through the experiments and provide a similar mechanism to flipping through the pages of a notebook.

Currently in order to include a reaction scheme in the notebook it is necessary to create an image of the reaction scheme and either save it on the device in Photos or to include a web accessible URL in the plan, where it will downloaded and made accessible within the notebook. Creating a chemical structure drawing tool or equation editor was felt to be outside of the scope of Notelus because the user tasks for these tools are most likely to be performed as a part of the planning of the experiment, rather than in the lab. A built-in drawing tool would enable researchers to create a reaction scheme within a notebook. This would require a significant effort, and it may actually be more sensible to allow the researcher to make use of already existing applications, such as the ChemDraw for iPad app[53] that have the capability to save images to the iPad or to enable the ChemDraw files to be exported from within that application to Notelus using the same hooks that Plan Buddy uses to enable the XML plan files to be opened in Notelus.

53 http://www.cambridgesoft.com/Ensemble_for_Chemistry/ChemDraw/ChemDrawforiPad/

At the present time the materials table functionality is limited for the proof of concept and only performs a small amount of maths to calculate the number of moles of the material. More sophisticated calculations could be included, for example, stoichiometry and a predicted yield based on the actual amounts used. An equation editor is another feature that we have heard requested for an ELN, and this is something that could be considered for Notelus.

It is not possible to look up safety information in Notelus directly, but safety information is displayed if it is included in the plan. There are challenges with providing access to a service from within Notelus to look up hazard information, but if an API to a system with a suitable standard format then this would be possible, if it were thought to be a useful feature. It is possible to derive some of this information from the ChemSpider database, but ideally a more sophisticated planning tool would fetch this information for them, and could be imported into Notelus as a part of the experiment plan. A more useful extension might be to make more use of the hazard information included in the plan. The researcher could be asked to identify the hazard using an appropriate code in the plan, and this could then be used to generate visual warnings within the notebook. With more sophisticated parsing of the content of the steps, hazardous materials could be identified and appropriate visual warnings displayed when the researcher turns to the page with the hazardous step identified. This would support the researcher without running the risk that they will fail to process hazards they are likely to encounter.

At the moment researchers can amend notes created on the keyboard, but there is no facility to edit hand drawn notes or sketches. The majority of users would expect the ability to be able to undo actions in order to remove a mistake. However, it is standard practice within scientific paper notebooks to ‘show your working’ (Bird et al., 2013a) and cross out any mistakes. The use of tipex, or loose bindings is not allowed. If no way is provided to remove content, then there is no way to cover up mistakes or change results after the event. However, from our user research activities with LabTrove we believe that users may be surprised by the lack of this capability. In LabTrove when a user makes a change to an entry, even if this is to add more content rather than change the existing content, they are asked to provide a ‘Reason for change’ and every version is kept so that the changes between versions can be examined. However, we found the majority of users are confused by the request to provide a reason and fail to adequately explain why they made a change. Notelus could be extended to provide an 'Undo' function, a variety of scales of delete, or an eraser function.

Notelus could be extended to integrate with one or more ELN systems. LabTrove does not currently provide an API for retrieving a specific entry at the current time or for creating or accessing an experiment plan, but if it did Notelus could be more deeply integrated in it. Notelus could also be developed to interact with other ELNs using APIs to integrate more closely with those. It would probably be better to develop different flavours for different parent ELNs in order to prevent the user interface from becoming too cluttered, and also to tailor the functions and terminology to the specific ELN.

Such integration could enable a user to access plans and post-experiment write-ups from within the application, and to create a single record for an experiment all in a single location. Although the single record has its merits, there are drawbacks in such an approach. The interface for the mobile application is likely to be more restrictive, it requires a user to have access to a specific ELN (and all the problems of potential software and data compatibility issues in the future (Whitton & Takeda, 2011)), as well as the significant issues of synchronisation, network, and firewall issues mentioned previously.

Providing a mobile notebook that does not require an ELN to be used provides flexibility, and is also likely to lead to earlier adoption. We had originally hoped that we would be able to add Dropbox as an option for exporting notebook files and importing plans, but at the point of development Dropbox was in the middle of changing their API. A decision was made to not include this functionality in the proof of concept of Notelus because of the significant performance issues experienced with connecting via the API. From our discussions with a variety of researchers handling digital data it is clear that Dropbox is widely used in the academic community for data storage and also for collaboration (Sam et el, 2011), and would be beneficial to include in future releases.

Integration with instruments and sensors would also be possible with an application that provided an interface to instruments, in a similar manner to the CombeChem weigh station application (Hughes et al., 2004). Numerous iPhone and iPad applications make use of Bluetooth and Wi-Fi to connect to sensors and other devices, for example, the E-paper Pebble watch[54] and the new Apple Watch[55], biometric sensors for health[56,57] and fitness[58], and bicycle cadence, speed, and power sensors [59,60]. External devices can also be attached directly to the iPhone or iPad and applications programmed to use the External Accessories Framework[61], for example a telescope controller[62], microphones and musical instruments[63,64,65], environmental sensors[66], and even a personal Geiger counter![67]

The tablet and mobile interface is not yet at a point where it could be considered a replacement for pen and paper. The drawing resolution through the touch interface is not as good as pen and paper, meaning that although you can make some basic sketches and rough notes, it is not yet possible to capture detailed movements. This means that writing must be much larger than it would be if you were using a pen and paper, and therefore much less information can be fitted onto a page than the equivalent size pad of paper. Some applications use a ‘zoom box’ on the screen where the characters are written and then shrunk for display on the screen, but this is fiddly to use and takes away from the more natural experience of just writing where you want the characters to appear on the page. Touch gesture technology is improving all of the time, and in the future we can expect to see increased touch gesture resolution that will help with such problems.

Another possible future development, particularly when the resolution of touch gestures is improved to create more accurate character reproduction, is the

---

[54] http://getpebble.com
[55] https://www.apple.com/uk/watch/overview/
[56] http://www.scanadu.com/scout/
[57] http://www.withings.com/us/smart-body-analyzer.html
[58] http://www.polar.com/beat/
[59] http://www.wahoofitness.com/subpages/product.asp?prod=262&dept=173
[60] http://www.stagescycling.com/stagespower-tech-specs
[61] http://developer.apple.com/library/ios/featuredarticles/ExternalAccessoryPT/Introduction/Introduction.html
[62] http://www.southernstars.com/products/skywire/index.html
[63] http://www.ikmultimedia.com/products/irigmic/
[64] http://www.ikmultimedia.com/products/irighd/
[65] http://www.ikmultimedia.com/products/irigkeys/
[66] http://cubesensors.com/#features
[67] https://mylapka.com/pem

use of handwriting recognition technology. There are two approaches to performing handwriting recognition, one is to use classes to carry out the recognition tasks on the device itself, and the other is to use Web services to do the translation. Although some free and open source options exist, for example services and classes using tesseract-ocr[68], these are not very effective compared to the commercial options available. Combining the use of a 'zoom box' together with commercially available handwriting recognition classes for iOS such as writepad by Phatware[69] or MyScript Builder by Visionobjects[70] may provide a way to add on device handwriting recognition. Ways would need to be provided to enable the user to change any characters that were not correctly interpreted, a process that is likely to slow down the process of note taking considerably in a lab environment compared to just writing on the screen. The use of a Web service to translate written characters on the fly could also be possible, but this adds the requirement to have access to a data connection and is likely to significantly slow down the process. At the present time it is possible to use online handwriting recognition services to translate the images of the characters captured in the notebook manually.

Voice recognition would also be a useful feature to capture notes. Off line and online options are also available for voice recognition software such as OpenEars[71] and iSpeech[72]. In many respects including offline voice recognition would be superior to the use of handwriting recognition, because less work is required from the users perspective to capture the information that they would like to record.

Additionally functionality could be used to automatically create more metadata to assist with describing the experiment and creating provenance information. For example, at the current time information about when the notebook is created is captured that could be used to create a start timestamp for the experiment, similarly the export date and time could be used to generate a stop timestamp. Both could be useful metadata for inclusion in an ELN. A

---

[68] http://code.google.com/p/tesseract-ocr/
[69] http://www.phatware.com/index.php?q=page/writepadsdk/ios
[70] http://www.visionobjects.com/en/myscript/software-development-kits/myscript-builder-embedded/description/
[71] http://www.politepix.com/openears/
[72] https://www.ispeech.org/developers

unique experiment ID could be created for each notebook, including details of the experiment plan if included. It would also be possible to ask the user to create properties for their version of Notelus including their name, institution and other values that could be used to create a unique ID for each of their notebooks for provenance purposes.

If a standard format was defined for experiment notebooks for consumption by ELNs or by data preservation or publication systems, such as ChemSpider[73], then Notelus could provide an option to export the information in the appropriate format, and use Web service calls like the export to LabTrove to upload the record and the associated notebook to that particular system.

9.8 Conclusions

Notelus is by no means a perfect ELN, and there are many improvements that could be made to both performance and stability; and additional functions could be added to improve the capabilities as detailed in 9.7. Limitations to the LabTrove API contribute to the difficulty in developing Notelus further as a client to LabTrove. For example, it is not possible to programmatically retrieve the content of a post and update it, or to create the plan as a template within LabTrove. Notelus is also in need of proper user testing 'in the wild', but as we have found throughout our studies, getting willing human participants is difficult especially where time constraints or interference with learning or research is a concern.

The Notelus project does remind us however that it is still likely that we need to bridge the gap between paper notebooks and digital tools, and the familiar interface and powerful capabilities of tablet and mobile interfaces are the most likely way to accomplish this challenging goal. We must seriously consider capturing information in the lab using an interface that allows handwriting and drawing if we are to encourage widespread adoption of ELNs in the academic environment. There are a number of ways that this could be attempted, and many of the references included in **Chapter 8: Mobile Electronic Notebook**, have tried alternate methods with some degree of success. However the ideal

73 http://aurora.labtrove.org/rsc_chemspider_test/724/ChemSpider_SyntheticPages_style_reaction_template__help_and_further_information.html

development of such interfaces requires a smooth interaction between hardware, software, and transcription services in order to make the best use of the information that can be captured through them. Although there are still challenges with capturing handwriting and sketches on the iPad, the technologies are changing and improving all the time.

Chapter 10: Human and computer interfaces for digital tools: The bigger picture

10.1 Introduction

Memories fade and change. If we fail to capture the information inside of researchers' heads, it will be lost forever. The scientific record is the externalization of these memories for use by the individual researcher, their local research group, wider institution, and ultimately the rest of the scientific community. If important information is not captured, an individual researcher may be able to rack their brains to try and remember the information, but this option is not available a few months after the event once the researcher has moved on, and certainly not in a hundred years. The more information we can capture in the record, the less reliance there is on fallible memory and the more opportunities there are for reusing the information for search, sense-making and gaining new perspectives (Gemmell et al., 2006).

The tool of choice for information capture is still the paper notebook for the majority of researchers. Although paper is a powerfully flexible way of capturing information, it comes with a number of problems, some of which at least can be mitigated by using digital formats for capture. Despite the potential benefits of digital tools and electronic capture, we know that researchers are sceptical of the technology and have a variety of anxieties. Adoption of any digital system will only happen if the system meets the needs and requirements of the creators of the scientific records (Bird et al., 2013a). Benefits of a system must be clear, and it must be easy to use. To quote a researcher from one of the Bath user tests:

> *This software needs to make my life easier.*

In this thesis we have explored how researchers and students currently record their experiments, how they use ELNs and other systems to capture and organise their experiments, and how different interface designs might impact the information that is recorded in digital systems. Although the various experimental studies have been focused primarily on the capture of information relating to experimental records, along the way scientists have provided an insight into their desires for what capabilities digital tools might be able to provide for them; not just replicating and making the ‘traditional’ way of working more effective, but actual to change and improve the way they work. These findings provide a valuable glimpse of ‘the bigger picture’ that extends beyond merely the need for data archives or electronic notebooks.

10.2 Research summary

10.2.1 Recordkeeping

In **Chapter 3: Keeping and managing records**, we investigated scientific recordkeeping behaviour, by both examining the experiences of others in the literature and also through studies and interviews with a range of scientists – both researchers and staff. These investigations confirmed that note-taking is primarily an act of memory for both the benefit of the scientist and for the benefit of others – directly for the ‘next generation’ or local research group, and indirectly through contribution to the scientific community as a whole through published works. Different researchers have different ways of working, different styles and structures for their notes, and different needs when it comes to sharing their records with others. A vast gulf can exist in ways of working and needs within the same community and indicates that a ‘one size fits all’ approach to tool design is unlikely to be effective. Tools need to provide considerable flexibility and customisation to accommodate these different needs.

Specific suggestions for the capabilities that such tools should include have been provided by both new and experienced users, as described in 3.7.2, but their experiences have also been captured as interaction design artefacts, detailed in **Chapter 4: Interaction design and tools for keeping and managing records**, that can be used to inform the designs of new tools. In particular, the Roles, Personas, and Scenarios developed based on the activities and needs of the real researchers have been used to generate a number of unexpected requirements for tools for supporting researchers, as described in more detail in 10.3.

10.2.2 Capturing the experiment

In addition to legal concerns for the contents of notebooks, reproducibility is an essential element of the scientific record. Although we may discuss and share our findings, the credibility of our research is limited if insufficient information is provided that enables research or experiments to be repeated. As discussed in 2.2.3, one way of ensuring that sufficient and appropriate information is included in the experiment record is to use ‘cued recall’ through the use of a template for capturing the information. From our studies, described in **Chapter 5: Asking the right questions: Effects of Using Structured Templates for Recalling Chemistry Experiments**, we found that templates are in fact an effective way to ensure that the information we request is captured. We also found that computer-based templates are more effective than paper-based templates for capturing some of this information because completion of sections can be mandated in a way that is impossible with a paper-based template.

A number of difficulties do arise from the use of templates, particularly computer-based templates. These difficulties do need to be addressed in the design of tools that make use of them for capturing the experiment record. A problem that can be solved in a real-world design is managing the capture of chemical information, particularly the capture of reaction schemes and materials used in the experiment, together with their related properties. We found that students do not readily remember the properties of materials such as the RMMs, and reaction schemes were more effectively captured on paper because they are most effectively and most consistently represented by drawing. We also know from our discussions with synthetic chemists that being able to draw and make use of chemical structures in an ELN is essential. We should therefore design tools that enable researchers to draw reaction schemes and chemical structures, search for materials, and automatically include the relevant properties in their experiment record, as shown in the proposed interface design example in Figure 10-1.

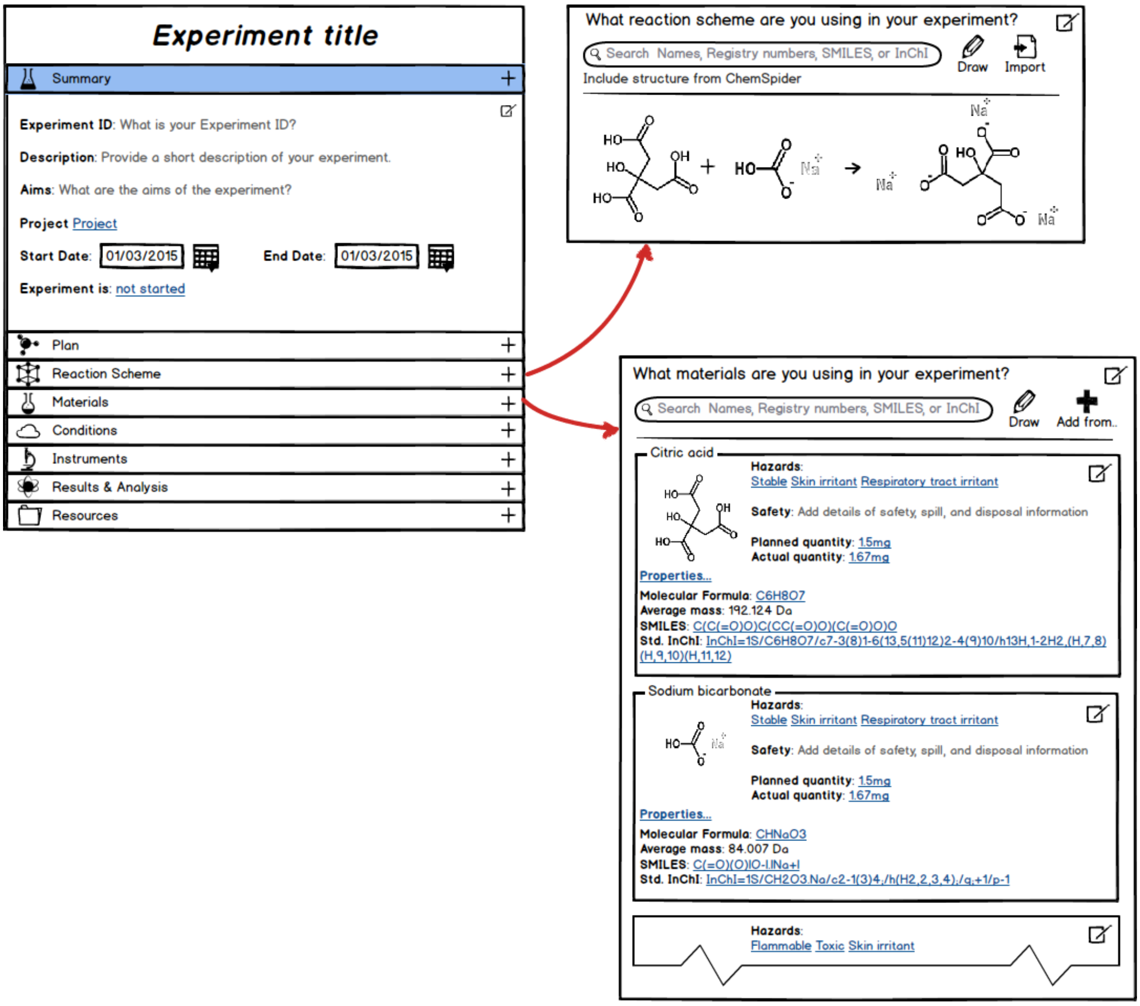


Figure 10-1: Example interface for capturing chemical information

The biggest challenge comes from attempting to capture the step-by-step procedure of the experiment. Typically without a template the procedure of the experiment is captured in the style of a personal diary that provides information about what was done in the experiment and why, together with the observations that are made. Within the template the step-by-step procedure or 'what I did' information can be drastically changed to become just a brief overview of the procedure or even a list of techniques with no further details. The observations and explanations typically become separated from the procedure information, taking away the temporal relationships between the events that would enable another researcher to repeat the experiment from the notes alone. The template style of capture is extremely effective for the capture of certain information, but for capturing this journal style recording of the personal experience, it is ineffective. The interfaces for tools need to enable both of these styles, and more work is needed in this area to understand how effective different designs may be. Interfaces for capturing information for a specific experiment also need to be customisable – different disciplines capture different information, as described in 3.5, and researchers want to be able to record the information that is most relevant for their experiment, as described in 5.5.

10.2.3 Curating experiment records

Our investigation into user-defined metadata use within the LabTrove ELN highlighted that researchers need support with the creation and curation of their digital notebooks, as described in **Chapter 6: Use of user-defined metadata in ELNs and other platforms**. Although some metadata can be automatically generated by systems to help curate research, data, and experiment, many of these methods require that appropriate information is entered into the system in the first place, as described in 7.7.1. We believe that the researcher is in the best position to determine the most appropriate metadata to describe their experiments. Researchers need to be supported through effective interface designs to encourage the capture of appropriate metadata for curation, in addition to other methods, such as data management education and the use of discipline-appropriate metadata schemes. Our investigations of user-defined metadata use and capture on other platforms provided some suggestions of effective metadata interfaces that could be used in an ELN including the use of cues and questions, as described in 6.5.3.2.

Strategies for capturing metadata were investigated in **Chapter 7: Generating metadata by using cues and changing perspectives**, making use of different strategies for recalling information, as described in 2.2.2. The findings of the studies indicated that cues are an effective way of capturing metadata, but information that is not cued for will not be captured. Free recall can generate some useful context information, but will also contain a lot of irrelevant information that may need to be cleaned. Asking the researcher to change perspective by asking a question about terms useful to someone else provides the most consistent metadata terms and is an effective way to generate metadata that can be used for search purposes. Invitations and context assistance can be added to interfaces to encourage the capture of metadata that can form ‘an experiment profile’ to describe experiments and other related resources and support other tool capabilities.

10.2.4 Implementing a mobile electronic notebook

As discussed in **Chapter 8: Mobile Electronic Notebooks**, a number of researchers have attempted to create digital solutions for capturing and working with research information that are portable or designed for working within the lab. The advent of smart phone and other mobile technologies have provides an opportunity to bridge the gap between paper-based notebooks and desktop ELNs through the use of familiar and intuitive interfaces. The Notelus ELN was designed and implemented as purpose-built tool to support synthetic chemists during their experiments, to enable them to capture their experiment records and associated metadata, and to export their experiments for sharing and archive. Notelus is by no means a perfect ELN, and there are many improvements that could be made, as detailed in 9.7, but the project does remind us that it is still likely that we need to bridge the gap between paper notebooks and digital tools. Although the familiar interfaces and powerful capabilities of tablet and mobile interfaces are certainly an important contender for accomplishing this challenging goal, at the current time the technology is still insufficient in this area.

10.3 The bigger picture: a Digital Research Management platform

Although much of the research carried out as part of this thesis is concerned with the relatively micro-view specific to the capture of experiment records, this research has also identified that requirements for tools to support researchers go far beyond the capture of notes in a digital format. As discussed in **Chapter 3: Keeping and managing records** and **Chapter 4: Interaction design and tools for keeping and managing records**, researchers also need help managing resources, monitoring and teaching their students, sharing information and gaining approvals from supervisors, getting access to and making sense of other researchers experiments, and engaging with collaborators.

We must not forget that capturing accurate, appropriate and sufficient information for our experiments is essential and ensuring our interfaces achieve these goals is critically important; any tool that fails in this task is worthless. Capturing all of this information about experiments, in particular the automatically created and user-defined metadata associated with them, enables us to add value within the ELN far beyond what is possible to do with a paper notebook.

10.3.1 Managing experiments with plans

Performing an experiment is more than just walking into the lab with a notebook, doing some actions, and recording some results. Although not applicable to all scientists, typically some kind of plan is required that detail materials to be used and actions to be performed. For many researchers a plan is required as part of the risk assessment process where the safety and treatment of materials need to be considered and 'signed off' before the experiment can be started. A plan can be considered as a stand-alone artefact or workflow that can be used to develop new plans, to conduct repeating experiments. Plans can also be shared amongst a research group or research community. As indicated by the discussions with researchers (and as highlighted in the scenarios) a DRM platform needs to provide effective management of the creation and sharing of plans and risk assessments. This management can provide opportunities for effective reuse of plans by tracking provenance, details of approvals, information about changes that have been made, and maintaining or forging new links to relevant literature and materials repositories and inventories, for repeating experiment runs, and sharing within communities.

The creation of plans for experiments provides other advantages. The metadata and descriptive content developed as part of a plan can be automatically added to an experiment that is created from the plan to minimise duplication and to create useful links and relationships. As discussed in 9.5.5 for Notelus, if you construct a plan in advance that includes the methods, materials, and equipment that you intend to use in the experiment then that information can be used to actively support the researcher by providing them with access to the information they need about the experiment and acting as a prompt for the next step to take. Warnings can also be provided for hazards in the experiment.

10.3.2 Capturing experiments with flexible templates

Effectively capturing information produced during an experiment is vital. As described in 3.5, note-taking styles and structures vary enormously, ranging between rigid formal structures that are repeated for every experiment and more diary-like formats composed of reflections, sketches, and personal information, but no formal structure at all. Tools for capturing the experiment record need to be able to accommodate both these styles of capture, and possibly even encourage more of the valuable ideas, reflection, and personal information.

Our research has demonstrated that using templates increases the quality of the record through encouraging the recording of specific information that we requested. Using templates provides a valuable starting point acting as a reminder to researchers to record specific information. Templates, in the form of 'experiment profiles' may in fact go some way towards mitigating the 'blank canvas' effect. Templates could also be beneficial for researchers and organisations that need to adhere to strict guidelines on recording from funding and other stakeholder bodies; templates could be developed and even certified to meet particular standards.

However, using templates does change the information that is recorded and can result in a loss of the specific details of the procedure and the personal narrative of the experience of the experiment, critical for reproducibility and future understanding. Two approaches are proposed to overcome the loss of the personal experience of the experiment and to provide the flexibility required due to differences between individuals and disciplines in what they record. The first proposal is to make use of a 'hybrid template' that provides a combination of cues for recording and a formal structure, together with 'free space' for recording the experiment procedure and observations during the experiment. The second proposal is to make use of customisable templates so that researchers can select to record the information that is most relevant to them, and not be 'forced' to record information that is not useful or confusing. Customised templates can extend to the capture of metadata, where scientists could perhaps import their preferred metadata schema and an appropriate interface for the template could be automatically generated. Customised templates could easily be developed manually for particular disciplines, research groups, and for repeating experiments, in the same way that templates are used in LabTrove currently. However, the interfaces need to be somewhat more sophisticated and extremely easy to use.

10.3.3 Activities, timelines, and enhancing search

For the information that we capture within a DRM system we can create table of contents, indexes, and perform searches. We can add filtering to searches so that we can find information about what experiments have been conducted on particular materials and related materials, information produced within a particular project, or by a particular institution, we can limit our search to experiments conducted during a specific week 4 years ago, and potentially even more obscure searches such as the results of experiments using a certain batch of a reagent, when the temperature was above a certain value, and two specific researchers were working in a specific lab at the same time!

For any specific information that is captured we can produce logs of activities and timelines of events. We can visualise the research trail and any element involved within it. For an individual experiment for example we could view a timeline of updates made to the experiment, for example measurements, observations, or information from an analysis running as the experiment proceeds. The changing conditions and data points could be plotted on a timeline, as a graph or even as an animation enabling the experiment to be 'replayed'. If the experiment is one of a sequence, a repeat, or just a related experiment they could be plotted together with others and compared and analysed to find new patterns or examined to explain differences between them or their results.

Timelines and visualisations are not just limited to experiments and results data, but also to projects and people. For example, the activities of a project such as discussion topics and participants could be visualised or activities on a project could be added to a timeline together with the milestones for the project to provide an overview of the progress of the project for collaborators. Relationships between different people could be examined through connections, for example, connections to other researchers, institution, projects, materials and instruments in common, and through shared topics of interests. Visualisations could show the network of relationships, but could also even create recommendations for inviting collaborators to projects by highlighting common connections; such tools could help those in the open science community to identify future collaborators as requested by the researchers at Sydney University. The career of an individual could be visualised on a timeline or through animations, showing their 'trail of research' by showing how the topics and elements of their experiments change and mutate over time. Their projects, institutions, and research connections could also be visualised to provide a unique tool for viewing the research paths and career evolution of an individual.

For a more everyday level the activities and actions that occur throughout the system such as creation of objects, updates, new relationships, events, and annotations, comments, and discussions, can all be visualised in 'recent event' style timeline, or as more commonly known on social media sites, within a newsfeed. Each user can have their own newsfeed to remind them about their own activities and also show activities that are relevant to them, such as updates to their projects, approvals, invitations and so on; projects and notebooks can have their own news feeds providing a collection of recent events of relevance to that particular object that can be viewed by anyone with the appropriate access rights.

10.3.4 Teams, collaborations, and fostering communities

One of strong themes that came out from the interviews with researchers and from the work on the Scenarios in **Chapter 4: Interaction design and tools for keeping and managing records** was around collaboration and connectedness. Even the researcher, who was horrified by the idea of sharing their notebook in an ELN, saw the value in such capabilities for monitoring the activities of his students. For teams actively engaged in Open Science a platform that provides ways to attract, inform, update, and interact with collaborators would be immensely valuable. Collaboration can be on the scale of a supervisor and his students, a local research group, or a global research project. Such capabilities could include enabling:

- the creation of project 'homes' with access to notebooks, experiments, other resources, and people involved in the project

- team members to create project updates that can be shared as news privately and publicly
- manual and automatic sharing of content on social media
- interfaces for public and private discussions including interfaces for asking questions and sharing tips
- encouragement to develop new connections and get involved with open projects, for example collaborator recommendations based on similar research interests derived from the person profile, metadata in similar projects and experiments, and common connections.
- Search, hit, and visitor statistics to inform PIs and researchers about the community of interest around their work

Community engagement can potentially be fostered through a variety of means that can encourage interaction between collaborators, provide connection and support for researchers who may be new to a team or remote from their colleagues, and through the construction of personal reputation. Examples of capabilities that could be provided include:

- Personal researcher profiles with research interests and organisational information
- Buddy checking of experiments, plans, and analyses
- Indicate success status of experiment
- Comments, annotations, and ratings of experiments, plans, and templates
- Social curation
- Ability to ask and answer questions

Developing a positive online community requires effort and responsibility. Researchers are unlikely to engage in activities that are perceived as a waste of time or lacking personal value. One way to encourage engagement and effort by users is to provide a mechanism for recognition and reward. Gamification elements, such as trophies, badges, & ratings, could be used to encourage helpfulness, interaction, answers to questions, and improved design of plans, experiments, & templates by boosting the perceived reputation or standing of the researcher within the community. Getting the balance correct, however, is unlikely to be a trivial task.

It may be possible encourage greater participation in open science with by reinforcing the value of research through effective and responsible online communities where participants are perceived as individuals with personal research stories, rather than the comparative anonymity of archiving of data in a repository.

10.4 Paper notebooks or digital notebooks?

Although we can propose a richer, more interactive, and more intelligent environment for working with research and experiment information, there is an issue that remains unresolved. As we have discussed in detail elsewhere (Sections 1.2 and 2.1) there are advantages and disadvantages to both paper and electronic notebooks. The question is, do we need to make a choice or to enforce one or the other option upon researchers?

As discussed in **Chapter 8: Mobile Electronic Notebooks**, a number of different hybrid and mobile systems have been developed and investigated. Such system vary in complexity from simple digital pen and paper systems or graphics tablets where writing and sketches are captured as images (e.g. Luff et al., 2004; Brandl et al., 2010; and Tabard et al., 2008), to much more complex systems involving touchscreens, cameras, barcodes and RFID tags, reagent sensors, and even motion tracking with smart watches (e.g. Yeh et al., 2006; Arnstein et al., 2002; Echtler et al., 2010; Tabard et al., 2012; and Scholl & Van Laerhoven, 2014).

These hybrid systems can be relatively smart or dumb; systems that capture without context are relatively dumb. For example, the typical digital pen solutions and simple drawing tools on tablet applications such as Notelus are dumb – they capture the content so it is available to be preserved, protected, and accessed electronically – but no information can be extracted from the content; it does not know if you have written text, drawn a reaction scheme or sketched a daisy. The A-book system is smarter, recognising some basic mark-up enabling some interactivity between handwritten notes and sketches (Mackay et al., 2002). Tablet or mobile-based ELNs, including Notelus, can be classed as hybrid systems that combine hand written notes and sketches with the other features that digital systems provide.

Although graphics tablets are currently still better than mobile tablets in terms of the accuracy and resolution of content capture for handwritten and hand drawn materials, they are not very practical for use in either a lab or field because of their need to be connected to a computer; also they usually require a connected or powered stylus to work. However technology for tablets is improving all the time; high-end graphics tablets currently available provide an experience with the same feel and accuracy as paper. The manufacturers of these high-end graphics tablets are already developing solutions for mobile platforms, for example the Cintiq Companion and Hybrid tablets for Windows and Android platforms[74,75].

We would argue that when these ‘feels like paper’ mobile technologies become ubiquitous, it will be harder for researchers to resist the call of digital tools for supporting all aspects of their research. The challenge will be for developers to create compelling and easy to use tools that make best use of the inscriptions created by researchers.

---

[74] http://cintiqcompanion.wacom.com/CintiqCompanion/en/
[75] http://cintiqcompanion.wacom.com/CintiqCompanionHybrid/en/

Published papers

This section of this thesis includes the published and draft papers summarised in Chapters 5 to 7, plus a list of other publications that have been contributed to as part of the research of this thesis and other activities undertaken as part of the PhD project.

Effects of Using Structured Templates for Recalling Chemistry Experiments

Willoughby, C., & Frey, J.G. Effects of Using Structured Templates for Recalling Chemistry Experiments (In preparation), as described in **Chapter 5: Asking the right questions: Effects of Using Structured Templates for Recalling Chemistry Experiments.**

Effects of Using Structured Templates for Recalling Chemistry Experiments

Cerys Willoughby[a] and Jeremy G. Frey

The way that we recall information is dependent upon both the knowledge in our memories and the conditions under which we recall the information. Electronic Laboratory Notebooks can provide a structured interface for the capture of experiment records through the use of forms and templates. These templates can be useful by providing cues to help students to remember to record particular aspects of their experiment, but they may also constrain the information that is recorded by encouraging students to record only what is asked for. It is therefore unknown whether using structured templates for capturing experiment records will have positive or negative effects on the quality and usefulness of the records for assessment and future use. In this paper we report on the results of a set of studies investigating the effects of different template designs on the recording of experiments by undergraduate students and academic researchers. The results indicate that using structured templates to write up experiments does make a significant difference to the information that is recalled and recorded. These differences have both positive and negative effects, with templates prompting the capture of specific information that is otherwise forgotten, but also apparently losing some of the personal elements of the experiment experience. Other unexpected effects were seen with templates that can change the information that is captured, but also interfere with the way an experiment is conducted. The results suggest that interfaces for recording information about chemistry experiments, whether paper-based questionnaires or templates in Electronic Laboratory Notebooks, can be an effective way to improve the quality of experiment write-ups, but that care needs to be taken to ensure that the correct cues are provided.

Introduction

For a scientist, the structure and information contained in the scribbles in our notebooks are likely to be a combination of the result of classroom learning, socialisation to professional practice, and our own personal style (Shankar, 2009). When it comes to entering information on a computer, however, what we record is likely to be influenced by the design of the interfaces we are using. In particular the questions we are asked and the prompts that we are given will change what information we provide. Electronic Laboratory Notebooks (ELNs) frequently make use of forms or structured templates to capture information. What is not known is whether providing a formal structure for the recording of experiments will be beneficial or if they may present negative consequences by potentially constraining or changing the information that is recorded.

The keeping of good records is essential in laboratory science, to capture both the thoughts of the experimenter and the detailed procedures of the experiments (Eisenberg, 1982). In both education and research settings, capturing the experiment record is more than just capturing the experiment procedure. In addition to recording "What did I do?" information, the record must capture quality information about what was observed, decisions that were made, explanations, and how observations and events relate to the experimenter's knowledge of chemistry. Without these observations, explanations, and decisions, then the full context will not available for future reference, either for the benefit of the researcher or for assessment of understanding and learning by the teacher or supervisor. Similarly students, future researchers and the wider scientific community. will have difficulty justifying arguments based on the data and to critically evaluate their own performance in the experiment.

The traditional medium for the capture of scientific records by students and researchers alike is the paper-notebook. While ELNs are still relatively rare in academic environments with a variety of challenges to overcome (Rudolphi & Goossen, 2012; Goddard, Macneil & Ritchie, 2009), the advent of computers, and the digital capture of data in particular, has begun to change the way many laboratory scientists record their experiments (Bird et al, 2013). ELNs provide a range of features that can help to improve data management, data retrieval, and collaboration providing positive benefits for teaching, learning, and research (Badiola et al, 2015).

Paper notebooks are by nature blank, and offer no guidance on what information should be recorded within them. Consistency in recording comes from standard guidelines for recording experiments used in both academia or formalised by regulatory organisations, for example the Good Laboratory Practice Handbook (Good Laboratory Practice, 2009; Eisenberg, 1982). Many academic researchers are still keeping their paper notebooks in exactly the same way they were when they were taught decades ago (Carpenter, 2012), and many still believe learning how to keep a paper notebook is a vital skill for students and researchers (Coles et al., 2015).

The proliferation of instruments producing data in a digital format means that even researchers still exclusively using paper notebooks have had to change their behaviour, either by printing out electronic data and pasting it into their paper-notebooks, or storing digital files and including a reference to the storage location and file names within their notebooks (Carpenter, 2012). Some researchers have begun to record their experiments in digital format using tools including Word and Excel, or digital notebook such as Microsoft OneNote and Evernote (Hayes, 2012). ELNs have the advantage of being able to automatically capture provenance and audit trail in a way that Microsoft Word or Excel files cannot (Carpenter, 2012).

ELNs can help to enable efficient and consistent recording of experiments through the use of templates that provide formal structure for data entry (Milsted et al, 2013). The vast majority of ELNs surveyed use templates or forms to enforce a standardised structure for the capture of scientific experiments. [1]

Studies in cognitive psychology have demonstrated that what we remember or choose to recall when asked is dependant upon our knowledge and previous experiences (Tourangeau et al, 2000; Jabine et al, 1984), but other factors, in particular the use of cues, can influence and shape what is recalled (Tourangeau et al, 2014; Higham & Tam, 2005; Marian & Neisser, 2000). The results of such studies have been used to design effective methods of information collection such as improving the design of surveys and questionnaires (Schwarz, 2007; Tourangeau et al, 2000; Jobe & Mingay, 1989) and questioning techniques used in eye-witness interviews (Memon et al, 2010). However, it is possible that using templates may have a negative impact. For example, other studies have indicated that the inflexibility and formality of templates might restrict the content that the author enters (McGlade et al, 2012) and that important information is not recorded if the template does not specifically ask for it (Swinglehurst, Greenhalgh & Roberts, 2012).

Table 1: Participants and conditions for the template studies

	Participants	Conditions
Study 1	20 chemistry undergraduates	'No Template' and 'Template'
Study 2	20 chemistry undergraduates	'No Template', 'Titles Template', and 'Profile Template'
Study 3	~65 chemistry researchers & staff	'No Template', 'Titles Template', and 'Profile Template'

In this paper we discuss the results from a series of studies looking at the effect of using templates to record experiments with UG/PG students and academic researchers. The studies investigate whether using templates improves or impairs the quality of information captured by students undertaking scientific experiments. We do not believe that such a formal investigation has been applied to information capture and recall in a laboratory setting. The conditions, participants, and settings for the three studies can be viewed in Table 1.

Study 1: University of Southampton Organic Chemistry Summer School (OCSS) and paper-based templates

The initial study was carried out during an annual Organic Chemistry (OCSS) Summer School run at the University of Southampton. The Summer School is run over 3 weeks after the end of term with a small group of second-year undergraduate students, enrolled on a variety of BSc and MChem programmes. The students were selected for the Summer School based on previous academic achievements. The purpose of the Summer School is to give the selected students experience of advanced lab techniques and working with industry between their second and third undergraduate year. The Summer School itself does not involve any formal assessment and therefore provided an opportunity to perform a study where different conditions could be assessed without any impact on the students' course marks. Taking part in this study was a mandatory part of the Summer School.

Two experiments from the first week of the Summer School were included in this study: preparation of an Organoboronic Acid and a Radical Benzylic Bromination. Two experiments from the Summer School were used to enable each student to generate a write-up in each of two study conditions: the *No Template condition* and the *Template condition.* The students were randomly allocated to one condition for the first experiment on one day and then swapping condition for their second experiment on the second day. In each condition the students were asked to complete a paper questionnaire independently after they had completed their experiment.

The *No Template condition* was effectively a blank piece of paper, whilst the *Template condition* included a number of sections for the students to complete, each section title acting as a cue to remind the students what to record. The titles for each section were based upon the laboratory notebook and report writing guidelines given to all students within the chemistry department at Southampton University. The questionnaires used in the study can be found in the supplementary information / appendix.

Results and discussion for Study 1

In general the average number of words used by the students to record their experiments is higher in the *Template condition* with 145.4 compared to 121.6 in the *No Template condition* consistent with the expectation that the *Template condition* would lead to the capture of more information. However, nearly

[1] Of 28 ELNs reviewed in 2011, only 29% explicitly mention the use of templates (Rubacha et al, 2011), but of the 20 that still exist and have sufficient product information, 85% of them use templates.

half of the students actually reduced their word count in the *Template condition* because the majority of these students did not complete the Discussion and Conclusion section of the template.

In order to investigate if the different conditions would have a positive or negative effect on student grades, each of the reports for both conditions was given a grade by an independent marker. These grades were converted to a numeric value so that they could be compared across the conditions. On average the *Template condition* resulted in an increase in grade compared to the *No Template condition*, with more than 70% of the subjects receiving a higher grade in the *Template condition* compared to only 13% receiving the same grade, and 13% receiving a lower grade. Of those with a higher grade, 55% had a grade that was significantly higher in the *Template condition* compared to their grade in the *No Template condition*.

One of the expected effects of using the template was to encourage the recording of information requested. The results of the first study demonstrated that this effect did occur, in particular for Aims and Reaction Schemes requested in the template. All of the subjects included at least one aim in their reports in the *Template condition*, whereas less than 20% included the experiment aims in the *No Template condition*. All apart from one of the subjects included a reaction scheme in the *Template condition* compared to around 50% in the *No Template condition*. Although only 61% of the students included the requested relative molecular masses in the *Template condition*, this compares to zero in the *No Template condition*. Almost all of the students included materials, equipment, and actions in their report reflecting the fact that the 'procedure' of the experiment was the dominant information recorded in the write-ups. The steps of the experiment, including collecting analysis information, were always recorded in the correct chronological order, even if steps were missing or the amount of detail was high or low. The average numbers of these elements is similar between both conditions. More students recorded results in the *Template condition* (90%) compared to only 47% in the *No Template* condition. More information about the results was also recorded in the *Template condition*.

We anticipated that using a template for recording an experiment might result in the loss of some information, possibly as a result of not asking for information specifically or a loss of personal information due to the constraints of the template. The results of this first study indicated that some types of information was more common the *No Template condition* than the *Template condition*. For example, more explanations were included in the *No Template condition,* particularly for explanations associated with actions taken in the experiment procedure. Overall, a similar number of students included explanations in their reports, but 65% of students including explanations within their description of the experiment procedure in the *No Template condition*, compared to only 39% of students in the *Template condition, where many* students included explanations about the experiment procedure only in the Discussion section of the template. The number of observations recorded is much higher in the *No template condition* compared to the Template condition, although similar numbers of students include at least one observation in both conditions. A much larger number of observations are recorded associated with their temporal occurrence in the experiment within the procedure in the *No template condition* compared to the *Template condition*. A much larger number of students included observations outside of the 'Step-by-step' section in the *Template condition*, together with a higher number of observations relating to their analysis of the materials created by the experiment – reflecting the fact that more information was included about the results of the experiment in the *Template condition.* In some cases observations such as the colour or state of the product are recorded in the results section of the template, rather than when they occur within the procedure of the experiment. However, some personal information such as the perceived success of the experiment, learnt information, or discussions is only seen recorded in the *Template condition*; particularly within the Conclusion section for those students that completed it. The amount of 'theory' or 'learned' background information about the experiment is much higher in the *Template condition,* and more than twice as many students record this type of information compared to the *No Template* condition.

An unexpected result of the first study was that a significant number of the students changed the style that they use for the experiment report in the *Template condition*. Over 75% of students use the past tense when they write about their experiment in the *No Template condition*, whereas nearly two thirds of the reports in the *Template condition* use imperative sentences or the 'command' tense, compared to only 12% in the *No Template condition*. Of those that used the past tense in the *No Template condition*, 50% of them switched to using the command tense in the *No Template condition*. Additionally, another two of the students started their reports in a command style and the switched back to past tense part way through for the *Template condition*. The subjects that used the command style in the No Template condition used the same style in the Template condition.

The results of the study showed that using a template to record an experiment had both negative and positive effects, with students recording more information with the templates, but also recording less observations and explanations compared to using no template. The original template was very basic, with the use of section titles only, and so a further study was planned to investigate whether the use of specific questions might overcome some of the negative effects of the title-based template.

Study 2: OCSS and computer-based templates

The second study also made use of a subsequent year's Organic Summer School at Southampton University. A number of changes were made for the second study. A new condition with specific questions was added to the previous two conditions, giving three conditions for the study: *No Template*, *Titles*

Template (same as the previous *Template condition*), and the new *Profile Template*. The new template was used to investigate whether asking more specific questions about the students' personal experiences would aid with the loss of personal information seen in the original *Template condition*. The hope was that asking specific questions about their experience would increase the amount of personal information such as observations and what they learnt from the experiment would be recorded. The format was also changed from the paper-based questionnaires used in the first study, to computer-based questionnaires. The change to computer-based questionnaires also enabled investigation of whether using a computer to capture the experiment write-up had any impact on the effects that were seen within the original study.

As before, the Organic Summer School included 20 second-year undergraduates and taking part in the study was a mandatory part of the Summer School. Because there were three conditions for this study, three experiments from the first week of the Summer School were included in the study: the same two as before, and the third experiment in the sequence that made use of the products from the first two experiments: preparation of an Organoboronic Acid; a Radical Benzylic Bromination; and Carbon-Carbon Bond formation using the Suzuki Reaction. The students completed one of the first two experiments on the first day, and then swapped to the other experiment on the second day. All of the students completed the Suzuki reaction and analysis on the following two days. All students completed the questionnaire for the *No Template condition* for their first experiment, the questionnaire for the *Titles Template condition* for their second experiment, and finally all completed the questionnaire for the *Profile Template condition* for the Suzuki reaction experiment. The students were asked to complete the questionnaires after they had completed each experiment. The questionnaires used in the study can be found in the supplementary information / appendix.

Results and discussion for Study 2

One of problems of the paper-based questionnaires was that some of the students chose not to complete all sections. An advantage of the computer-based questionnaire is that the fields in the questionnaire can be made mandatory making it more likely that they will be completed. The effectiveness of the computer-based questionnaire is reflected in the results for this study with the students providing much more complete responses to the templates, with only 3 students missing a section from any of the templates, compared with more than half missing at least one section in the *Template condition* in the paper-based study. The students who missed one or more sections in the *Title* or *Profile* templates did not have a significant drop in word count related to these omitted sections.

In general, the average number of words changes between the *No Template* and the other template conditions as seen in the first study. The average number of words in the *No Template condition* is 142.4 compared to 264.8 in the *Titles Template condition* and 245.1 in the *Profile Template condition*. 95% of the students increase their word count in the *Titles Template condition* and 90% increase their word count in the *Profile Template condition* compared to the *No Template condition,* suggesting that more information was captured in both of these conditions.

The same independent marker as used in the first study was also used to grade every report in the second study. A similar pattern is seen to the paper-based study in that the template conditions resulted in an increase in grade compared to the *No Template condition*. This is true for both forms of the template, with the *Titles template condition* resulting in 70% of students receiving a higher grade, and the *Profile Template condition* resulting in 80% of the subjects receiving a higher grade. Only one student had a lower grade, and that was the in the *Profile Template condition*. Two of the students who had the same grade received the highest grade in all three conditions. Five of the students received the lowest grade in the *No Template condition*, but only 2 of these retain that grade in the *Title Template condition*.

In common with the first study the template conditions do have an effect on the topics that are recorded in the questionnaires. All of the subjects included at least one aim in their reports in the *Title Template* and *Profile Template* conditions, whereas only 10% included the experiment aims in the *No Template condition*, even fewer than with the paper-based study. All apart from one of the students included a reaction scheme in the *Titles Template condition* (although they take different forms compared to the paper-based version). None of the students included a reaction scheme in the *No Template* or *Profile Templates* conditions. Only in the *Titles Template condition* were any relative molecular masses included with 65% providing them, a similar proportion to the paper-based version, with one student admitting to not remembering them.

Almost all of the students included materials, equipment, and actions in their reports, again reflecting the fact that the 'procedure' of the experiment is the dominant information captured in the report, as seen in the first study, although the pattern of information captured is different between the two templates. The steps of the experiment, including collecting analysis information, were always recorded in the correct chronological order, even if steps were missing or the amount of detail was high or low, as before. The average numbers of these elements is similar in the *No Template* and the *Title Template* conditions as seen in the paper-based study. Differences are seen in the *Profile Template condition*, where the average number of materials, equipment, and actions is significantly decreased in the "What did you do in the experiment?" or procedural section of the template. The number of students recording materials, and equipment in particular, are reduced, but all of the students included these elements in specific questions that ask about these elements in the *Profile Template*, for example "What instruments or equipment did you use in the experiment?". If these questions were not included however, the limited responses to the "What did you do in the experiment?" question would have produced significantly fewer of these elements overall for the reports, and the materials and actions described would be more general (and therefore considerably less useful to someone trying to understand or repeat the experiment)

A similar pattern to the paper-based Study 1 is seen for Explanations with more students including them in the *No Template condition* and a higher number associated with actions in the experiment procedure, compared to the *Titles Template condition* and the *Profile Template condition.* For both template conditions, more explanations were provided about information other than steps in the procedure than in the *No Template condition*, and were more likely to be found recorded in a different section of the template Interestingly, in contrast to the first study, the number of observations included in the reports is actually higher in the template conditions than in the *No Template condition*. The procedure information about the experiments in the Titles Template is much more similar to the No Template condition is study 2, but with increased observations related to the analysis, so more of these observations are seen within the procedure section of the experiment. The *Profile Template condition* has the highest number of observations, although the majority of these are recorded outside of the procedure of the experiment, especially in association with the question "What observations did you make in the experiment?" as might be expected.

Only 3 of the students recorded any Results in the *No Template condition*, compared to 100% in both of the template conditions. More information was captured about the results in both of these conditions, with more information about the actual analysis or physical appearance of the product. Although more students include measurements of any sort in the two template conditions, fewer are included in the *Profile Template condition* and these are mostly in the results section. A higher number of measurements are recorded within the procedure section of the *Titles Template* and *No Template* conditions than in the *Profile Template condition* reflecting the difference between the procedure information recorded in the Profile Template condition compared to the other two. The much higher number of measurements in the *Titles Template* is due to the combination of the more detailed procedural section and more results compared to the *No Template condition*.

The majority of "Conclusions" include a 'success statement' about how well the experiment went, or whether the correct product was produced. Only two of the students produced any conclusions in the *No Template condition*. In the *Title Template condition* the Discussion and Conclusion section discuss mostly about the analysis so far, what needs to be next, and some issues around differences in results between the groups. Different groups were trying out different solvents, and some of the group had disappointment with their choice of solvent. As seen in Study 1, more learnt information and theory about the experiment was included in the template conditions, with almost all students including this kind of information in both templates, compared to only 16% in the *No Template condition*.

For this computer-based study, no clear patterns of style change can be seen. 50% of students start with the Past tense in the *No template condition*, and then continue this style in both forms of the template. A change of style is seen most commonly in the *Titles template*. 'List' is a more common style than the command style seen in the first study, although they have some similarities, with both being a briefer style of communication and including less detail.

Study 3: Lego Cars templates and group discussion

The final study was less formal and was carried out as part of a half-day team-building activity. The study included 60 members of Chemistry staff and research students from KTH in Stockholm. The participants were randomly allocated into 15 teams, consisting of both staff and students, and each was provided with equipment and instructions to complete an experiment using Alka-Seltzer to power Lego cars. The experiment suggested that the teams assessed the effects of using different Lego car designs and to evaluate the effects of using different quantities of Alka-Seltzer and water, thereby encouraging the teams to conduct multiple experiments. Each team was asked to record their experiments on one of three different computer-based templates randomly allocated to them. Each computer-based template was a Google word processor document, with different sections and instructions. The templates were similar to those used in Study 2, but with some key differences. The Titles template included a Results table with headings customised to support the participants with the recording of their experiment results. The two template conditions also contained an additional section about the "Plan" for the experiment. Again, the No Template condition was essentially a blank document. Some additional fields were included in each template to investigate the capture of metadata relevant to experiments, but the result of those investigations do not impact this study and are reported elsewhere (IDCC paper).

This third study provided the opportunity to find out the opinions of the subjects on the different template conditions through the use of a group discussion and feedback session held after they had completed their experiments and write-ups. The questionnaires used in the study can be found in the supplementary information / appendix.

Results for Study 3

In common with the first two studies the average number of words used does vary between the different templates, although for this study the average number of words is highest in the *Profile Template condition* with 359.6 words, compared to 251.8 in the *No Template condition*, and only 181 in the *Titles Template condition*. One of the reasons for the low average word count in the *Titles Template condition* is that the teams using this template spent most of their time completing the table and as a result recorded less information in the other sections of the template, with one team completing no other sections. The presence of the table seems to have drawn attention away from the other questions and even 'de-railed' the experiment to some extent, with some of the teams commenting that they had less time to conduct the experiment itself because they spent so much time filling in the table. Interestingly in the

Title Template condition no photographs were provided, and no mentions of videos are made, in contrast to the other two conditions where the majority of teams included photographs or links to photographs in the write-ups. Several teams also mention using video to record the activity.

The majority of the write-ups in the *No Template condition* showed 'self-structuring', possibly because the design of the experiment encouraged running a set of repeated experiments and was also focused on producing measurements. The teams created their own structure that usually contained at least the following elements, with some teams including others:

- Method
- Results
- Observations (measurements)
- Conclusions

The write-ups for the *No Template conditions* typically contained observations, explanations, and conclusions or things that were learnt as a result of doing the experiment. The *Profile Template condition* templates generally captured more information than the other questionnaires, but the actual procedure of the experiment was in less detail than in the *No Template condition*. More information was included about significant results and what was learnt as a result of the experiment in response to the specific questions about these elements. Almost all of the teams working in the *Profile Template condition* completed all of the sections in the template, in contrast to the *Title Template condition* where many sections were not completed, as mentioned above. The inclusion of a section asking for information about what the teams planned or expected to do in the template conditions revealed differences between what they planned and what they actually did. This was particularly important information for this kind of experiment where the procedure to follow was not rigid and each team had the flexibility to perform their own version of the experiment.

Study 3: Group discussion

After each of the teams had completed their experiments and write-ups, teams that had completed different versions of the templates were brought together to discuss the different templates and how their write-ups differed as a result. After their discussion each combined group presented their feelings about the different templates. This activity provided an interesting opportunity to find out whether students and researchers would be comfortable using templates for recording the results of their experiments. The groups were surprised by how different their write-ups turned out when the different templates were used. Table 3 contains the comments made by the different combined teams in the group discussion. The different groups raise a number of common points, in particular that they felt that the *No Template condition* gave more freedom and flexibility, but that the *Profile Template condition* led to the capture of more information, because some information could be forgotten in the *No Template condition*. Several of the groups also mention that the *Profile Template condition* structure was a good starting point for a template, but they would prefer to be able to define their own template based on their personal needs for their own experiments.

Table 2: Comments from the combined teams in the group activity

Group	Comments
Red	The different template styles vary a lot. The team that used the *Profile Template condition* slavishly completed the template, but the other teams wrote notes that would help others to complete the experiment. Using a template produces a minimum level of quality and influences what is recorded. Using the *No Template condition* you can record anything you want – "free to record" – and recorded deviations and other things. The templates enable a minimum level of quality control.
Blue	Using the *No Template condition* risks a lot of information being missed and therefore being unable to find it later. The *Titles Template condition* meant the team didn't think of ideas and ideas were lost. The *Profile Template condition* asked the same questions over again, but would be best in the long term. Suggest that you can start with a template and then develop your own template over time.
Green	(Only No Template and Profile Template conditions) A lot of differences are seen between the notes that were taken. Only 10 lines were written in the *No Template condition* compared to 4 pages in the *Profile Template condition*. Most of the information is the same, but some details are missed in the *No Template condition*. It would be good to have your own template to select the information you want and be assured of getting the same information each time. Neither team checked the template before starting.
Yellow	The *No Template condition* gave many degrees of freedom. More questions is equal to less freedom, but it does add different information. All the templates have their good points. More free in the *No Template condition*. A combination of all 3 is best. Good to have the guidelines to remember what to do in the experiment rather than start with an empty notebook.
Orange	The team in the *Profile Template condition* didn't read the instructions and just made the recording in a (paper-based) table and with video. The team took the approach of "capture the data first and then write up the experiment". The team using the *No Template condition* recorded everything and managed more experiments than the others. The *Title Template condition* team got stressed completing the table and had less time for doing the experiment. Focus on the table caused too much time to be spent on preparation. The *No Template condition* team spent some time coming up with their own structure for the experiment, which they then copied and pasted. Copying the experiment is a template in itself.

Discussion

The guidance and training we give to students encourages them to keep consistent, clear, and accurate recordings of what they actually did in the lab and to generate reports in a standard structure. One possible outcome of the initial study in this paper was that when we handed students a blank piece of paper they would produce a write-up that would match the format that they have been trained to use. In fact, the majority produce a “what I did” description of the procedure, only sometimes including other information such as aims, the reaction scheme, or results. The same result is seen when the students are given a blank textbox on a computer to complete with the same instructions. Although these reports miss important information, they are often detailed procedures that could be repeated by another and do include useful information about the experimenter’s personal experience, such as the observations they made, and explanations about why certain actions were taken.

We were unsure whether providing a formal structure for the recording of experiments would be beneficial, but the use of cues to remind students and researchers to record certain information seems valuable when we note that in these studies important information is frequently missing; an observation also made by the teams from KTH in the third study. However, we were also unsure whether there might be negative consequences of using templates if they were found to constrain or otherwise change the information that is recorded in a negative way.

The impact of the templates is in fact more complicated than we initially imagined. The cues provided by the templates do encourage the recording of the specific information that is prompted for, even when the prompt is simply a heading. For example, Aims are consistently recorded when subjects are asked for it in both the Title and Profile Templates, but are much less consistently recorded in the No Template conditions.

Exactly what is recorded though does depend on both the cue used in the templates and on the experiment itself. In terms of the experiment itself, the Lego Cars experiment in Study 3 was a ‘measuring’ style of experiment, rather than a ‘making’ style of experiment and encouraged the recording of times and distances, and teams recorded ‘Results’ in all of the conditions without the need for a cue. In studies one and two, the experiments involved synthetic chemistry where the aim of the experiments was to ‘make’ specific compounds. For these experiments the cues in the templates led to the recording of significantly more results than in the *No Template condition*, particularly for the computer-based Study 2, where all of the students recorded results in the templates, but only 3 recorded them in the *No Template condition*.

The exact question or nature of the cue in the template does make a difference to the information that is recalled and recorded. This is particularly noticeable for the Reaction Schemes in the experiment write-ups. Although 50% of the reports in the No Template condition produced reactions schemes in the first study, none of the reports in the computer-based studies in the No Template condition recorded reaction schemes, although this may be because capturing this information in a computer is much more difficult than on paper, In Study 2 only one student did not attempt to produce a reaction scheme in the Titles Template condition. In the Profile Template condition the results produced for “What reactions were involved in the experiment?” varied enormously, from very simple to a complex description of the underlying chemistry of the reaction, for example, “A cross-coupling reaction”, “To create a carbon carbon bond using Suzuki coupling”, “Use of the Pd catalyst in a cycle to couple the two compounds together causing the removal of PPH3 ligand”, and “Reaction of the boronic acid with the brominated benzene derivative in the presence of a palladium catalyst”. Interestingly in Study 3, the same question produces more consistent and simpler responses, for example, “HCO3- +H+ = CO2 + H2O plus spectators”, “NaHCO3+H2O -> CO2 (g)”, “Bicarbonate in water reacts with citric acid and forms trisodium citrate and carbon dioxide.”, and “Bicarbonate in water reacts with citric acid and forms trisodium citrate and carbon dioxide.”; these differences perhaps reflect the differences in complexity between these two experiments.

Relative Molecular Masses are only recorded when the questionnaire included the cue to record it, although only about two thirds of the students recorded this information, suggesting these specific details are not well remembered and that the presence of cues alone is not enough to ensure that specific information is recorded. The need for more than the cue for some information is also demonstrated with the Discussion and Conclusion sections in the templates with titles. In the paper-based Study 1 a large percentage of students did not attempt to complete these sections, and many of those who did seemed to be confused, with several students describing ‘discussions’ they had during the experiment, or explanations and observations that were more relevant to the procedural section. In Study 2 the responses to these sections were mandatory, and the responses were more appropriate. The students were also able to see the questions in the computer-based questionnaire on a single page and this may have had an influence on how they responded, perhaps because the combined cues together made the traditional report structure of the sections more obvious.

The original template in Study 1 did appear to have a negative effect of reducing some of the personal information about the experiment as predicted. Students recorded fewer explanations and observations in the *Template condition* than with free recall facilitated by the *No Template condition*. The change in style, particularly produced by the cue “Step by step experiment procedure” seemed to not only reduce the number of explanations and observations recorded, but to entirely change the style of recording by many of the students, with 50% of them changing to an imperative or ‘command’ style of wording. The change in style suggests perhaps a change of perspective of how they were recording the information from a “What I did” style to a “How to do it style”, as though they were writing the instructions for a different audience. Subjects in cognitive psychology studies have been shown to recall different information when asked to recall something from a different perspective (Dudukovic et al, 2004; Tversky & Marsh, 2000) and it may be that different cues might prompt a different perspective to be taken when recording information. In both Study 2 and Study 3 the majority of the write-ups are produced in the past tense, although any change of style in Study 2 is still most commonly seen in the *Titles template*. It was hoped that using the question "What did you do in the experiment?" for the Profile Template condition would lead to the inclusion of more personal detail and explanations about what was done in the experiment, but in fact this cue resulted in a briefer description of the steps of the experiment, in some cases just a list of short phrases or 'techniques' describing the steps. Details of the specific materials and the equipment that was used in the experiment were also omitted from the responses to this question. Many of the other questions in this condition also often generated brief, list-like responses. The briefer answers recorded in this condition may be a result of the larger number of questions to answer or that some of the questions were more open, for example "What did you do in the experiment?" is actually more open than "Step by step experiment procedure". Some studies have suggested that open-ended questions may lead to a more superficial search of memory, and therefore using closed questions can generate more results (Tourangeau et al, 2014). This may be because the more closed questions contain more specific cues, for example "What chemicals or other materials did you use in the experiment?", "What instruments or equipment did you use in the experiment?", and "What were the aims of the experiment?" produced sensible and specific responses from all of the students in the *Profile Template condition*. In contrast the more open question "What reactions were involved in the experiment?" worked less effectively than the more specific "Balanced equation with relative molecular masses (RMM)", even though this was a difficult question to respond to on a computer. Another potential problem with open questions are that chemistry students are less familiar with this style of questioning and as a consequence are not very good at constructing complete and relevant answers that reflect their actual conceptual understanding (Bunce & VandenPlas, 2006). The questions asking about results, observations, what was learnt, and conclusions appeared to be more effective, perhaps because they provided a good balance between providing a specific enough cue and not constraining the response, prompting the recall of information that will be useful for assessing the student's understanding of the topic.

As mentioned by some of the teams in Study 3, there appear to be good things about all of the questionnaire styles. The difference in information that is captured in the different conditions does have an impact on the quality of the write-ups produced as reflected in the grades. In general, the use of the templates leads to higher marks than the free recall of the *No Template condition*. Although some students maintain very high or low grades throughout, a large number of students do see their grades increase significantly with the use of the Template. The increase in grades is likely to be because more information about the theory, purpose, and understanding of the experiment is revealed (and recorded) in these conditions. Including the wrong kind of cue in the template can have significant negative consequences, however, constraining the information that is recorded and even potentially in influencing the way that an experiment is conducted. This was demonstrated dramatically in Study 3 where the inclusion of a 'results table' caused groups to fail to record other important information about the experiment, and also to spend less time performing the experiment compared to the other teams. The teams in this condition also did not record the experiment using photographs or video, unlike the majority of others who chose to record their experiments in this way. This suggests that the use of a template with a very restrictive format may also restrict creativity or change the way that the recording the experiment is approached.

Conclusions

Electronic laboratory notebooks have many potential advantages to academic audiences, but there are a variety of challenges to overcome. These computer-based systems typically make use of forms and templates in order to capture experiment records, and a perhaps unseen challenge is the impact of these templates on the information that is captured and therefore the quality and usefulness of the experiment record. The results of the studies detailed in this paper suggest that using templates, whether paper-based questionnaires or computer-based interfaces can have both positive and negative benefits for capturing experiment records and evaluating student learning. All three of the methods for capturing information investigated in these studies have some benefits, but each also has its problems. There may be no 'one size fits all' template for capturing a chemistry experiment, and different cues may be more useful for different types of experiment. As suggested by the teams in Study 3, the Profile Template may be a good starting point for a template that can then be customised as needed for a particular user, audience, or experiment. Customisable templates are a feature of many of the major ELNs (Rubacha et al, 2011). The questions that were less successful for the Profile Template could be improved in light of the results of these studies, to generate the recall and recording of information that is considered most important, for example, capturing student understanding of the concepts of the experiment for assessment or specific details of a procedure or technique for future reference and reuse.

Acknowledgements

Smart Research Framework, e-Research South. Simon Liversedge, Laura Cowen, Colin Bird. Thomas Logothetis and students, University of Southampton Chemistry Synthetic Organic Chemistry Summer School. Staff and students, KTH, Stockholm.

Notes and references

[a] Chemistry, Faculty of Natural and Environmental Sciences, University of Southampton, Highfield, Southampton, SO17 1BJ, UK.

Badiola, K. A., Bird, C., Brocklesby, W. S., Casson, J., Chapman, R. T., Coles, S. J., Cronshaw, J. R., Fisher, A., Frey, J. G., Gloria, D., Grossel, M. C., Hibbert, D. B., Knight, N., Mapp, L. C., Marazzi, L., Matthews, B., Milsted, A., Minns, R.S., Mueller, K. T., Murphy, K., Parkinson, T., Quinnell, R., Robinson, J. S., Robertson, M. N., Robins, M., Springate, E., Tizzard, G., Todd, M. H., Williamson, A. E., Willoughby, C. Yang, E., & Ylioja, P. M. (2015). Experiences with a researcher-centric ELN. Chem. Sci., 00, -. doi:10.1039/C4SC02128B

Bird, C. Willoughby, C., and Frey, J. (2013) Laboratory notebooks in the digital era: Record keeping in chemical and other science laboratories Chemical Society Reviews 2013, 42, 8157-8175. DOI: 10.1039/c3cs60122f

Bunce, D. M., & VandenPlas, J. R. (2006). Student recognition and construction of quality chemistry essay responses. Chemistry Education Research and Practice, 7(3), 160–169. doi:10.1039/b6rp90007k

Carpenter, S. (2012) Must a Paper Trail Be Paper? Science Careers. http://sciencecareers.sciencemag.org/career_magazine/previous_issues/articles/2012_09_14/caredit.a1200103 (Accessed 9th December 2014)

Coles, S.J., Kilpin, K.J., Willoughby, C., & Frey, J.G. (2015). Dial a Molecule iLabber trial. (In preparation, April 2015).

Dudukovic, N.M., Marsh, E.J., and Tversky, B. (2004) Telling a story or telling it straight: the effects of entertaining versus accurate retellings on memory Applied Cognitive Psychology 18, 125-143.

Eisenberg, A. (1982) “Keeping a Laboratory Notebook.” Journal of Chemical Education 59:12, 1045–1046.

Goddard, N. H., Macneil, R., & Ritchie, J. (2009). eCAT: Online electronic lab notebook for scientific research. Automated Experimentation, 1(1), 4. doi:10.1186/1759-4499-1-4

Good Laboratory Practice (GLP). Second Edition (2009) World Health Organization ISBN-13: 9789241547550

Hayes, M. (2012) Electronic Lab Notebooks http://postdocexperience.scienceblog.com/2012/11/19/electronic-lab-notebooks/ (Accessed 9th December 2014)

Jabine, T. B., Straf, M. L., Tanur, J. M., & Tourangeau, R. (Eds.). (1984). Cognitive aspects of survey methodology: Building a bridge between disciplines. Washington, DC: National Academy Press.

Jobe, J., & Mingay, D. (1989). Cognitive research improves questionnaires. American Journal of Public Health, 79(8), 1053–1055.

McGlade, K., Cargo, C., Fogarty, D., Boohan, M. & McMullin. M. (2012) Handwritten Undergraduate Case Reports. The Clinical Teacher. 9:2: 112–8. doi:10.1111/j.1743-498X.2011.00494.x.

Memon, A., Meissner, C. A., & Fraser, J. (2010). The Cognitive Interview: A meta-analytic review and study space analysis of the past 25 years. Psychology, Public Policy, and Law. 16(4), 340–372. doi:10.1037/a0020518

Milsted, A., Hale, J., Frey, J. and Neylon, C. (2013) LabTrove: A lightweight, web based, laboratory “blog” as a route towards a marked up rec-ord of work in a bioscience research laboratory. PLoS ONE 10.1371/journal.pone.0067460

Rubacha, M., Rattan, A. K., & Hosselet, S. C. (2011). A review of electronic laboratory notebooks available in the market today. Journal of Laboratory Automation, 16(1), 90–8. doi:10.1016/j.jala.2009.01.002

Rudolphi, F., & Goossen, L. J. (2012). Electronic laboratory notebook: the academic point of view. Journal of Chemical Information and Modeling, 52(2), 293–301. doi:10.1021/ci2003895

Schwarz, N. (2007). Cognitive aspects of survey methodology. Applied Cognitive Psychology, 21(2), 277–287. doi:10.1002/acp.1340

Swinglehurst, D., Greenhalgh, T., & Roberts, C. (2012) Computer Templates in Chronic Disease Management : Ethnographic Case Study in General Practice. 1–16. doi:10.1136/bmjopen-2012-001754.

Tourangeau, R., Conrad, F. G., Couper, M. P., & Ye, C. (2014). The Effects of Providing Examples in Survey Questions. Public Opinion Quarterly, 78(1), 100–125. doi:10.1093/poq/nft083

Tourangeau, R., Rips, L. J., & Rasinski, K. (2000). The psychology of survey response. Cambridge University Press.

Tversky, B. and Marsh, E.J. (2000) Biased retellings of events yield biased memories Cognitive psychology 40:1, 1-38.

Creating Context for the Experiment Record. User-Defined Metadata: Investigations into Metadata Usage in the LabTrove ELN

Willoughby, C., Bird, C. L., Coles, S. J., & Frey, J. G. (2014). Creating Context for the Experiment Record. User-Defined Metadata: Investigations into Metadata Usage in the LabTrove ELN. Journal of Chemical Information and Modeling, 54(12), 3268–83. doi:10.1021/ci500469f, as described in **Chapter 6: Use of user-defined metadata in ELNs and other platforms.**

Journal of Chemical Information and Modeling
Article
pubs.acs.org/jcim

Creating Context for the Experiment Record. User-Defined Metadata: Investigations into Metadata Usage in the LabTrove ELN

Cerys Willoughby,* Colin L. Bird, Simon J. Coles, and Jeremy G. Frey

Chemistry, University of Southampton, Southampton SO17 1BJ, United Kingdom

Supporting Information

ABSTRACT: The drive toward more transparency in research, the growing willingness to make data openly available, and the reuse of data to maximize the return on research investment all increase the importance of being able to find information and make links to the underlying data. The use of metadata in Electronic Laboratory Notebooks (ELNs) to curate experiment data is an essential ingredient for facilitating discovery. The University of Southampton has developed a Web browser-based ELN that enables users to add their own metadata to notebook entries. A survey of these notebooks was completed to assess user behavior and patterns of metadata usage within ELNs, while user perceptions and expectations were gathered through interviews and user-testing activities within the community. The findings indicate that while some groups are comfortable with metadata and are able to design a metadata structure that works effectively, many users are making little attempts to use it, thereby endangering their ability to recover data in the future. A survey of patterns of metadata use in these notebooks, together with feedback from the user community, indicated that while a few groups are comfortable with metadata and are able to design a metadata structure that works effectively, many users adopt a "minimum required" approach to metadata. To investigate whether the patterns of metadata use in LabTrove were unusual, a series of surveys were undertaken to investigate metadata usage in a variety of platforms supporting user-defined metadata. These surveys also provided the opportunity to investigate whether interface designs in these other environments might inform strategies for encouraging metadata creation and more effective use of metadata in LabTrove.

1. INTRODUCTION

Metadata is an essential aspect of the preservation of knowledge for future exploitation, and using metadata to provide context to data that would otherwise be difficult to use is recognized as an important component in the creation of intelligent open data[1,2] for science.

The role of metadata in electronic record curation and electronic laboratory notebooks has been described in detail elsewhere.[3] Discussions about metadata in this context typically focus on the "burden of curation" and other difficulties in metadata generation.[4−9] Capturing experiment metadata in an ELN for "curation at source" mitigates some of the burden of curation.[10] Despite the important role of ELNs in capturing metadata, there appears to be little discussion of how metadata should be used and perhaps even more importantly if and how it is being used.

The Chemical Informatics Group at the University of Southampton has developed LabTrove (www.labtrove.org), a researcher-centric Web- and cloud-based ELN that has a blog-style structure. LabTrove contains a facility enabling users to add their own user-defined metadata to describe their entries and experiments.[3,11] In section 2 of this paper, we report an investigation into how researchers use metadata within the LabTrove ELN and provide insights into what types of metadata are used and whether it is used effectively. Although the LabTrove team had some expectations about how metadata might be used and perhaps ought to be used in the ELN, until the survey we had only a limited knowledge of how metadata was being used by the community.[3,11] We also present some of the feedback and observations about metadata and metadata use that we have obtained from undertaking various user research activities with the LabTrove community. The feedback and observations from these activities raised issues that contributed to the decision to carry out this survey.

In section 3 of this paper, we investigate whether the patterns of metadata use in LabTrove are unusual, using a series of surveys to investigate metadata usage across a variety of other platforms that support user-defined metadata. We end this paper with section 4, with observations from the surveys and a discussion of strategies for improving the use and capture of user-defined metadata for researchers in chemistry based on our findings across both sets of surveys.

2. LABTROVE METADATA SURVEY

LabTrove is a blog-based ELN tool that is used by academic researchers in a heterogeneous set of academic laboratories with groups engaged in analytical chemistry, X-ray studies, drug discovery, and biomaterials research; it is used for a range of purposes such as for recording experiments, data archiving, and

Received: July 22, 2014
Published: November 18, 2014

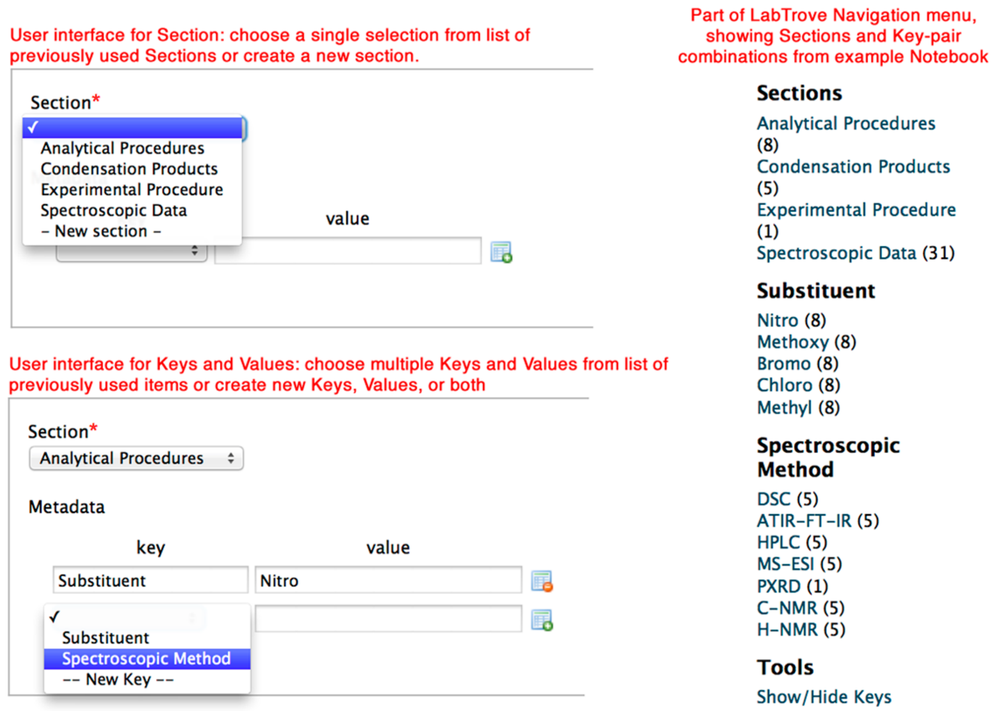


Figure 1. LabTrove interface for adding metadata to a notebook entry and exposure of metadata in the navigation menu (annotated with descriptions of the metadata elements).

project coordination.[3,11,12] Each LabTrove instance, known as a Trove, hosts a collection of notebooks, which may be authored by a single researcher or by a group. Typically, a Trove contains a collection of notebooks belonging to a single research group or a single institution, but some other Troves are shared with notebooks and contributions from researchers from different institutions.[12] Each notebook enables the author to create entries, upload data in the form of files, add user-defined metadata, create links to other materials, and interact with others' notebooks by adding comments.[11] Users can view entries in the notebook in reverse chronological order or can make use of the search facility or a menu generated from added metadata to navigate to specific entries of interest.

Metadata means different things to different people but is usually about more than just "data about data".[9] Metadata can be considered from the perspective of what the metadata describes, such as content, context, structure, preservation, and administration of information,[13] and from the perspective of the function of the metadata, such as allowing resources to be discovered, what the resources are, and how they are organized, which enables use, interoperability, and provides identification and authentication.[14,15] We have previously defined metadata as descriptive information and classification labels that group related items, provide context, and facilitate the reuse of specified research outputs.[16] LabTrove is intended as a "marked up record that can be shared and searched" with the intention that notebook entries record the details of the scientific process together with the data that is produced as a result.[11] The intended functions of the LabTrove metadata features are to describe the context of the research process, enable authors and others to search for and find the research, organize entries by facet, enable reuse both within the ELN and through interoperability with other systems, enable provenance inspection, and archive research. Different styles of metadata and different terms may be more or less effective for each of these metadata purposes.

To investigate the behaviors of research communities and their patterns of metadata usage in LabTrove, a survey of over 100 LabTrove notebooks, covering a variety of disciplines including chemistry, biology, and physics, was conducted.[17] The surveyed notebooks were either entirely public, publicly accessible through an Open ID, or accessible only with an institutional (university) login.

2.1. LabTrove Metadata Elements. The blog-style structure of the LabTrove ELN enables users of the system to record their experiments and activities with individual entries in the notebook. LabTrove provides users with the means to add their own user-defined metadata to their entries in addition to the machine-generated metadata.[11]

The values that are provided for the metadata are utilized internally to create the navigation menu displayed on the right-hand side of the ELN interface. Therefore, the inclusion of metadata makes it easier for both authors and their collaborators to find entries about particular topics or relating to particular experiments. The use of consistent metadata potentially produces a much more effective record than is possible with a paper notebook.[3]

For each entry that a user creates, they are required to specify a value for a Section field. Users can optionally choose to add further metadata to the entries in the form of key-value pairs.[18] The key-value pairs enable the inclusion of metadata that is much richer than could be produced using a simple tagging system and provides a form of classification for the notebook entries. Key-value pairs are used to characterize notebook entries, enable grouping of related experiments, and aid in search and linking.[12] Figure 1 shows the LabTrove user interface for adding sections and key-value pairs and part of the navigation menu.

dx.doi.org/10.1021/ci500469f | J. Chem. Inf. Model. 2014, 54, 3268−3283

2.2. Templates. Templates are a special type of notebook entry that can be used for experiments that are repeated or use parallel procedures to reduce the burden of entering the same information repeatedly. These templates are typically used to provide a structured way to enter details of a procedure or experimental results, for example, through the use of tables and structured input fields.[11] Templates can also be configured to encourage users to add consistent and complete metadata by enabling the template author to define metadata in the template. This metadata is then inserted into newly created posts, although users are free to remove or change the inserted metadata and to add their own instead if they want to. To create a template in LabTrove, the author of the entry sets the Section to "Templates". When a user views the template entry, there is an option to "Use Template" that creates a new notebook entry with the structure provided by the template together with any metadata that has been predefined. The user can then modify the contents of the entry, completing any tables or input fields, and making any appropriate changes to the metadata before saving as a normal notebook entry. Despite the advantages of templates to encourage structured and consistent notebook entries, they are not widely employed, and experiences have been mixed. It is relatively easy to create an entry with a standardized structure without using a template, but we are seeing an increasing recognition of the benefits of templates from users and anticipate their use to increase in the future.[12]

2.3. Survey Method. As shown in Figure 1, the metadata elements used within each of the notebooks is present in the navigation menu on the right of the LabTrove interface. To extract the metadata from each of the notebooks for the survey, each notebook was accessed using the LabTrove Web interface and a copy taken of the Section, Key, and Value elements from the navigation menu. Other information was also recorded about each notebook, including the number of authors, number of entries made in the notebook, whether the notebook was primarily used for "autoblogging", and apparent primary use of the notebook. Table 1 shows the total numbers of notebooks, entries, authors, and different metadata elements that were collected from the notebooks and used in this study.

Table 1. Number of Notebooks, Entries, Authors, and Metadata Elements of Each Type Examined in the Survey

notebooks	entries	authors	sections	keys	values
104	22,818	202	512	263	959

Notebooks associated only with the testing or development of Troves were not included in the survey in order to exclude irrelevant metadata. The notebook survey did include several notebooks that were set up with "auto-blogging", where an instrument or sensor automatically enters data to a LabTrove notebook. The notebooks in the survey also included those with templates and entries created from templates but not notebooks only created for the storage of templates. The results therefore include Section metadata with a value of "Templates" and any metadata that was automatically added to entries that were created from the templates. Templates themselves typically only have Section metadata. Although templates can be used across multiple notebooks within a Trove, the majority are actually used within a single notebook. Although some duplication of terms can be expected by template use across multiple notebooks, the results are the same as seen when consistent metadata is consciously adopted by the group[11] as described in section 2.6.3. The metadata examined in the study are investigated by the terms themselves and not by the number of times they are used; so the metadata terms used on entries created by templates, even if a template has been used dozens of times, does not skew the results of the survey.

2.4. Use of Section Metadata. For each entry in the LabTrove notebooks, it is mandatory for users to enter a value for the Section metadata. No default values for Section are provided, but the user can choose to use values that have previously been used in the same notebook or to create a new value. The Section enables a top-level categorization of all entries. Although the LabTrove team did not prescribe any specific uses for the Section metadata,[11] there has been active debate within the team on how Section should be used and what is the best way to divide up entries in the ELN for an experiment. Two alternatives have been proposed:

1. Multiple entries are used to represent an experiment, with different values for Section that match the stage of the experiment or phase of the research process (i.e., Plan, Materials, Procedure, Results, etc.).
2. A single entry is used to represent the entire experiment, and the value of Section is derived from some other aspect of the experiment.

The survey revealed that a very limited range of values are being used for the Section metadata, with 70% of the notebooks using five Sections or less and around half of those using only one Section; only 10% of notebooks use more than 10 different values for Section. The most frequent type of word or short phrase used as metadata for Section was a "catch-all" phrase, for example, the Section has a blank value such as a space or a dash or only one Section is used with an imprecise value such as *General*. Almost half of the notebooks use a catch-all Section of some sort, with 35% of notebooks using a catch-all Section in addition to other Section values. This frequent use of a catch-all Section, coupled with high numbers of users using only a single Section suggests that that a large percentage of the notebook authors are avoiding adding metadata and are taking a "minimum required" approach, where metadata has only been added because the system requires it.

2.5. Use of Key-Value Metadata. The author of an entry in the ELN can optionally add metadata in the form of key-value pairs. For example, the user may choose metadata to add an experiment ID to their ELN entries. In this case, the chosen Key could be "Experiment ID", and the value could then contain the identifier for that experiment. There is no limit to the number of key-value pairs that can be added to an entry, and each Key can have multiple values. Key-value pairs are not just limited to entering specific identifiers in this way but can be used both to capture the characteristics of individual entries at a variety of levels and to provide the capability to categorize entries using subcategories that are relevant to the individual author. Where the Section metadata key enables a top-level categorization of all entries,[11] the key-value pairs provide a form of two-level classification that complements the flexibility of the notebook entries.

The number of metadata Keys used was counted within each notebook. The figures show that almost half of all notebooks used no Keys at all, with only about a quarter using three or more Keys, indicating that this powerful metadata capability is poorly used within the community. The lack of examples of good practice for this type of metadata probably contributes to this lack of use.

2.6. Classification of Metadata in LabTrove. To investigate the metadata in more detail, the metadata terms were classified in a number of ways. Each metadata item was classified by determining whether the term could be considered high-level or specific, that is, whether the metadata used in the ELN is typically generic or more specialized. More specialized metadata is likely to be more useful for locating entries about specific topics, whereas high-level terms are likely to be applicable to more entries and potentially less useful. The words "Data" or "Enzymes" would be classified as high-level terms, while "Filament" or "Herbal Medicine" would be classified as specific metadata. Figure 2 shows more examples of Sections and their categorization into High-Level and Specific groups.

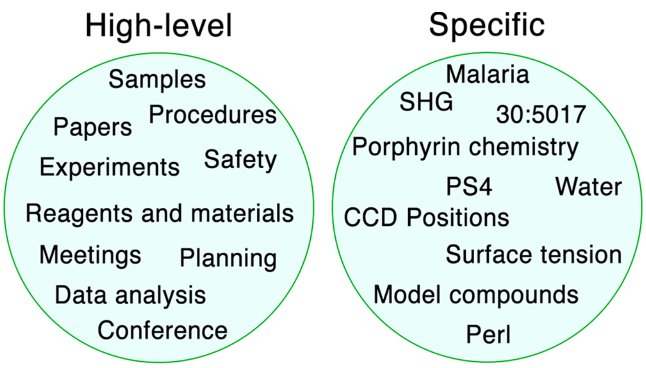


Figure 2. Examples of Section metadata classified as "High-Level" or "Specific".

The LabTrove metadata was also classified by word or phrase type into Noun-type, Verb-type, and Adjective-type. For single words, this is straightforward with the words classified by their dictionary class. For ambiguous words and phrases, the most appropriate class was selected. For example, the terms "Platform management meetings", "Electrochemical properties", and "Interesting new papers" are classified as noun-type phrases. "Freshly crushed", "shade of blue", and "math-related" are classified as adjective-type phrases. "Data (formatting)", "Learning design", and "Electrochemical cleaning" are classified as a verb-type phrases.

Finally, the Noun-type metadata items were classified based on the dominant subjects for each metadata type, as determined by an initial coding exercise. The resulting categories are similar, for each metadata type, with Activities, Codes, Dates and Values, Equipment and Instruments, Labels, and Materials present for each type. Some differences are observed. In particular, a Catch-all category is only seen in the Section metadata, and a Location category is only observed in the Value metadata.

2.6.1. Metadata Is High-Level Rather than Specific. The results of the survey showed that the majority of the metadata used could be classified as high-level rather than specific, with more than two-thirds of the Section metadata and 80% of the Key metadata classified as High-Level. This finding suggests that the capture of metadata may not have been very effective because specific metadata that would help to provide context and identification for individual experiments or records is missing.

2.6.2. "Things" Are More Significant than "Activities". The results of the classification exercise indicate that high-level labels, or descriptions of objects or properties such as materials, data, or instruments, dominate the metadata terms used. The vast majority of metadata terms are Noun-type, including dates, sample numbers, and instrument locations, with Adjective-type often used for the catch-all metadata terms such as miscellaneous and general. Less than 5% of the metadata items are Verb-type, such as modeling, blogging, or monitoring. Nouns that describe activities, such as "purification", "preparation", "analysis", "lithography", "electrophoresis", and "filtration", are used at least twice as often as verbs to represent activities, but even so, these types of nouns that describe techniques, methods, and actions used to complete the experiments are still relatively poorly represented in at less than 10% of the total metadata. There are a couple of potential hypotheses that might explain this observation. It may be easier (or more natural) to classify notebook entries by physical properties or objects rather than activities or it could be that the activity is considered to be less important to record, but the reasons were not investigated.

Table 2 provides a detailed breakdown of the word-type classifications for the different metadata types. The word-type classifications are used to compare the LabTrove metadata with other platforms later in this paper.

Table 2. Breakdown of Word-Type Classifications for Metadata

	Noun-type	Verb-type	Adjective-type
Sections	513	15	3
Keys	263	4	1
Values	893	40	26

2.6.3. Section and Key Metadata Are Dominated by High-Level Labels. The results of the subject classification exercise show that the majority of the Section metadata are high-level labels representing the content of the entries that they characterize, while the remainder describes items such as codes, dates, instruments, materials, topics, and project information. Figure 3 shows a word cloud of Sections classified as "Labels". Similar terms have been combined, for example, material and materials. As mentioned previously, templates are a specific type of notebook entry in LabTrove for which users are required to use the value "templates" for the Section. Of the

Figure 3. Wordle (http://www.wordle.net) visualization showing Section metadata classified as "Labels". Size relates to the relative frequency.

surveyed notebooks, 28% contain templates, and as such, it represents the most frequent Section term used.

The majority of Key metadata are also high-level labels with other Key terms representing items such as specific instruments, materials, documents, and properties. Figure 4 shows a word cloud of Keys classified as Labels. Similar terms have been combined, for example, data type and data. It can be seen that very similar words and phrases are used for the Section and Key metadata, suggesting that these terms are particularly important for recording experiments. This duplication across the Section and Key metadata also shows that users are not differentiating between the metadata types effectively.

Figure 4. Wordle visualization showing Key metadata classified as "Labels". Size relates to the relative frequency.

Figure 5 shows the categorization of the Values metadata in the LabTrove entries and shows that these metadata refer primarily to materials, activities, and various codes, such as notebook and experiment identifiers, dates, and activities. Examples of Values metadata categorized as "Other" include "Alpha", "computational science", "E-notebooks", "Forms", "Introduction", and "Matlab".

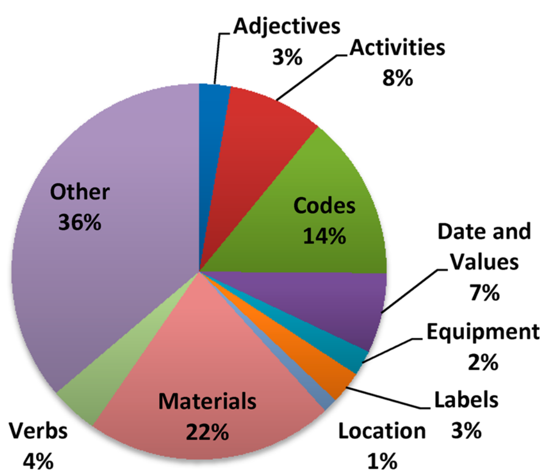


Figure 5. Categorization of Values metadata.

The "post-type" Key name is used primarily, but not exclusively, by a set of notebooks related to bioscience experiments. The group conducting these experiments has consciously made a decision to use "post-type" as a defined a set of values associated with the Key, including material, data, sample, procedure, safety, template, and note.[11] This group is in fact the main group that uses templates across multiple notebooks, but the "post-type" Key and its associated value, describing the type of information captured by the entry, is typically the only automatically added metadata, while entries not created from a template are assigned an appropriate "post-type" value or Section by the author. Additional Sections, "post-type" values, and other key-value pairs are assigned personal terms by individual authors based on their own needs.[11]

A full list of the metadata categories and numeric breakdowns are provided in the Supporting Information.

2.7. Impact on Metadata of Privacy and Number of Notebook Authors. The majority of notebooks in the survey (65%) have only a single author, but 8% of notebooks have five or more, with one notebook having 25 authors. The number of authors on a particular notebook appears to vary by the function of the notebook, with a higher average number of authors seen in notebooks where the primary function is recording group and project activities (2.95) or for discussions about the ELN itself (2.46) compared to notebooks primarily used for recording experiments (1.40). The majority of notebooks with two or more authors are notebooks available only through institutional logins. The authors on these notebooks are typically collaborators belonging to the same group and institution, but even for the publicly available notebooks, the majority of authors are co-located and belong to the same group.

The highest numbers of Sections are observed in notebooks with the highest numbers of authors. There is also a trend toward increased key-value metadata use with an increase in notebook authors. There appear to be two reasons for this increase depending upon the function of the notebook. In notebooks with multiple authors, which have relatively high numbers of Sections or Keys, and that are used primarily for formal recording of experiments or project activities, the metadata terms used relate to many different topics, with no duplication of meaning. For those notebooks that are used primarily for discussing and experimenting with the ELN itself, or "sandpit" notebooks, the metadata terms used are much less broad and often include terms that have the same meaning or are very similar, for example, "procedure" and "procedures", "API testing" and "testing API", "test" and "testing". Discussion with users has also highlighted a problem where multiple authors use the same metadata term but may use different capitalisation leading to different metadata stored in the notebook. The example given was the term "NMR" used as a Key, where different authors had entered "NMR", "Nmr" or "nmr", leading to multiple Keys being created. Although removing case sensitivity is one option, a merging process to replace or standardize metadata values may be useful where duplication occurs, such as with abbreviations, synonyms, homonyms, and misspellings.[19−21]

The notebooks with the highest numbers of authors also have a higher number of entries, which may in turn lead to the creation of more metadata to describe the additional activities. When the number of entries in a notebook is high, it may be more apparent that better organization is required, which would drive the production of metadata in those notebooks. This is supported by the broader range of terms used for metadata in these notebooks when they are used for recording experiments and organizing project work.

Interestingly the notebooks with more than five authors appear to rely more on Sections than on key-value pairs. There could be a number of reasons for this. For example, as the majority of these notebooks are authored by members of the same group, there may be less of a need for detailed metadata, although several groups mention that they make use of key-value pairs in order to help enable grouping of entries.[12] An alternative reason may be that large numbers of key-values are quite difficult to manage in the current user interface and may actually hinder the rapid location of entries.

Are notebook users adding metadata for their own benefit or for that of readers? The owner of a LabTrove notebook can choose the level of privacy for their notebook. The privacy level of the notebook may indicate if the author intended it to be used for personal record keeping or data dissemination and therefore whether the needs of an audience may have been taken into account.[22] If the author of a notebook intends their work to be read by the public or peers, then including more metadata will enable their work to be found and used more easily. Authors ought therefore to choose to add more metadata in the ELN, although that is no guarantee that it is created to meet the needs of an external audience.[23]

The notebooks included in the survey were divided into three groups based upon their privacy settings:

1. Entirely public (Public)
2. Publicly accessible through an Open ID (Logon)
3. Accessible only with an institutional (university) login (Private)

Notebooks that are only visible to the notebook owner or specifically selected users do not form a part of the survey. The numbers of metadata elements used were compared for each group to determine whether privacy did have an effect on the amount of metadata used. Figure 6 shows that the results are

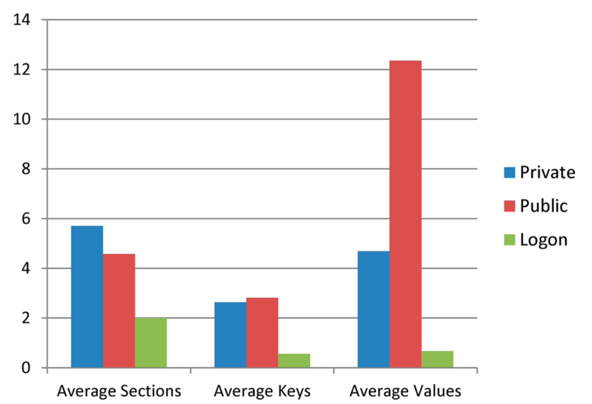


Figure 6. Average number of Section, Key, and Value metadata by the privacy status of the notebook.

somewhat unexpected, with the Logon group of notebooks showing relatively little metadata at all. Looking into this in more depth, the notebooks in the Logon group are not well populated, and those with low metadata numbers are notebooks for discussion by a specific community rather than for formal experiment recording or engagement with the general community. Another notebook is for experiment recording, but it is populated through autoblogging. So the metadata is the same for each entry. These notebooks need to be accessible to users from a variety of different institutions, and this is likely to be the reason why the particular privacy settings have been chosen.

Given that that one of the perceived inhibitors to metadata use is the reluctance to make data public in the first place, it is perhaps surprising to see that the private notebooks have some of the highest average figures for metadata use. However, many of the notebooks that are protected behind institutional logins are intended to be public within their own community.

One of the groups with the largest number of publicly accessible notebooks has expressed anxieties about creating metadata in their notebooks but have recently begun to add more metadata, particularly key-value pairs, to help them organize and group their experiments.[12]

A further study would be needed to determine whether the metadata values contribute to community engagement or help with discovery of experiments and notebooks. LabTrove has a search facility that includes the metadata related to entries, as well as the entry titles, and entry contents. A possible extension to LabTrove in the future could be to provide statistics and data analytics in order to investigate what terms are used for search, how users find a particular notebook or entry, and which entries are popular. From discussions with users, we know more about what the authors of notebooks search for within their own notebooks than we do about how other users use the search facility or what information those users are trying to find. Interviews with users indicate that users primarily use the Navigation menu to locate notebook entries of interest, using all of the different metadata elements that are available including dates, authors, Sections, and key-value pairs. Users also use the search facility, but this is typically to find specific experiment identifiers and can often be many months after the original experiment was recorded. For some of the groups a new researcher may continue work on a set of experiments and therefore take over recording in a notebook for a researcher that has moved on. Researchers in this position have indicated that better metadata would have helped make their job of taking over research easier.

For the moment, the number of authors remains a better indicator of community engagement until statistics on search and the number of unique visitors to a notebook become available.

2.8. LabTrove User Research. The LabTrove development team have undertaken a variety of activities to investigate user behavior and attitudes toward metadata including interviewing both current users and communities interested in adopting LabTrove, usability testing with novice users, and trialling the ELN with students. These activities have provided the opportunity to examine what expectations and understanding users have about metadata and how the design of the ELN might affect their metadata use.

The results of these activities have indicated that although some users are comfortable adding metadata to their notebooks and understand the benefit the metadata provides in helping them to locate information, many others have indicated that they felt it was too difficult to use and did not see the benefit in adding it. A great deal of the anxiety within the user communities about using metadata stems from a fear of having to design their own metadata scheme. The researchers want meaningful examples to follow but feel that taking schemes from other teams would not be very helpful because the scheme would be too specific and not easy to adapt. The alternative of being provided with an example created at a high level would be too generic to be useful. Even though the researchers do not know exactly what metadata they want to use, their reluctance to use schemes from other teams may indicate that they do nevertheless have the ability to recognize a meaningful example when they see one.

Observations from our user studies and the ELN trial with students match the results seen in the metadata survey in that the majority of new users add the minimum amount of metadata required (a single Section). In common with the data from the notebook survey, if adding metadata were not enforced within the notebook, the majority of users would not add it, at least with the current interface design. Students differed in how they approached writing entries in the ELN, but the way they chose to structure their experiments affected how they used metadata. If they chose to break the parts of their experiment into multiple entries, they were more likely to use appropriate metadata to classify their entries by content or activity, whereas those students who created a single entry for an experiment tended to use a catch-all category. Where example Section and Key terms were provided, for example, from exemplar notebooks or template entries, the students were more likely to use metadata, but they did not necessarily select appropriate combinations or contribute their own more relevant terms.

Several of the interviewed users indicated that they were more familiar with tags from using other Web-based tools and felt more comfortable adding metadata as tags. In fact some of the researchers described the metadata that they use as "metadata tags" and wanted to have the metadata visually represented in tag cloud form. There is evidence that social networking is one of the main areas where students encounter and actively use metadata.[24] Some of the users also expressed a desire for different ways of organizing their data in the ELN, a feature that could be facilitated if users were encouraged to provide specific metadata such as the start date of the experiment or sample identifiers. Others viewed metadata as a way of making links to other resources, suggesting that users expect metadata interfaces to behave in a similar manner to familiar social networking applications such as Facebook or Flickr, where information becomes a link to other people or objects that share the tag or name.

Some anxiety about using metadata comes from a fear of using the "wrong" metadata and therefore having to go back and change it, or "messing up" data by creating mismatched metadata between entries. Some of the groups that were reluctant to use metadata in their notebooks for these reasons have more recently begun to add metadata to their notebooks entries and to view metadata as a useful tool for grouping experiments together.[12] Adding metadata retrospectively is not difficult once a suitable scheme has been devised.

3. INVESTIGATING USER-DEFINED METADATA USE ON OTHER PLATFORMS

Although the LabTrove team had some expectations about how metadata might be used and perhaps ought to be used in the ELN, until the survey, we had only a limited knowledge of how metadata was actually being used by the community.[3,11]

The results of the notebook survey provided insights into two main areas. First, that of the utilization of the metadata facilities provided, where it was found that although some groups are comfortable with metadata and are able to design a metadata structure that works effectively, a large percentage of the community adopted a "minimum required" approach, where metadata has only been added because the system requires it. It was also found that metadata use increased with the number of authors for a notebook, but the amount of metadata used was not associated with the privacy status of the notebook. This is not entirely in contrast with studies of other platforms where users can define their own metadata, which suggest that users are often motivated to add user-defined metadata both for social reasons and for personal organization or future retrieval.[25,26]

Second, the survey provided insights into the type of metadata that users were choosing to add to their notebook entries. The results indicate that the majority of the metadata used was high-level rather than specific and is dominated by high-level labels representing the content of the entries that they annotate. The specific metadata describes objects or items such as codes, dates, instruments, materials, topics, and project information. Activities such as techniques, methods, and actions that were used to complete the experiments are relatively poorly represented in the metadata. The vast majority of metadata values are Nouns or Noun-type values, including those that describe activities such as purification, preparation, analysis, lithography, electrophoresis, and filtration, with adjectives and verbs rarely used.

The insights from both areas, utilization and type of metadata, together with feedback from our user community, provide suggestions for changes that might be made to the LabTrove interface to encourage more metadata creation and to improve the quality and usefulness of the metadata that is captured. The results of the survey also raise an important question: Is the metadata used in LabTrove significantly different from other platforms that enable the creation of user-defined metadata? For example, do those other platforms suffer from the same problems of minimum required use, does the number of authors influence the amount of metadata, and are the same kinds of metadata values being used?

Our user research indicated that many researchers have experience with social media and networking platforms such as Facebook and Flickr (section 3.1) that make use of metadata, and presenting familiar interfaces like these might support and encourage the creation of metadata in an ELN. Our examination of other platforms that enable user-defined metadata creation also provided an opportunity to investigate whether particular interface designs might be effective at encouraging user-defined metadata creation. These platforms include Flickr (section 3.1), blogging platforms used by NASA (section 3.2) and other researchers (section 3.3), and myExperiment (section 3.4).

In this section of the paper, we present the results of surveys of user-defined metadata use across different platforms to investigate if the metadata usage seen in LabTrove is unusual and whether the types of metadata used are different across different communities and platforms. We use the results of these investigations to discuss potential enhancements that could be made to improve metadata creation, capture, and quality in LabTrove (section 4).

3.1. Flickr. Flickr is an image sharing social networking site, where users can add tags to describe their photos and videos. The tags are displayed reasonably prominently on the pages displaying the images, together with an invitation to add tags. Interfaces for uploading images on the Web site and also through mobile applications usually have a prominent invitation to add tags, and the user can select from their commonly used and previously used tags. Other Flickr users can also tag the images. This tagging by other members has been used effectively by organizations such as the National Library of Congress and the British Library to increase access to publicly held photography collections and to acquire information and knowledge about the subjects in the photographs from the general public through "The Commons" project.[27] There are also "machine tags"[28] in Flickr that use a special three-part syntax. Machine tags are frequently added by third party Flickr applications to provide information about themselves and geotag photos and for more novel uses such as the Proyecto Agua that uses the machine tags to add metadata about the taxonomic classification of waterborne organisms.[29]

Although it is difficult to gauge what percentage of users use tags in Flickr, of one of the author's (CW) 657 contacts, less than 4% used no tags at all, with an average of 12 tags on each picture. Less than 10% use 20 or more tags.

For the Flickr survey, Flickr Hive Mind,[30] a data-mining tool for the Flickr photography database, was used to acquire tags from the 500 images ranked most interesting that were tagged with both "chemistry" and "experiment".[31] This produced 1381 unique tags with a further 162 removed because they were indecipherable. These tags were then classified by word or phrase type and by topic in the same manner as the LabTrove metadata. Figure 7 shows the breakdown of the Flickr tags by word type and subject. It can be seen that Verb-type usage is slightly higher in Flickr (7%) compared with LabTrove (3%), while usage of the descriptive Adjective-type is much higher in Flickr (14%) compared to LabTrove (2%).

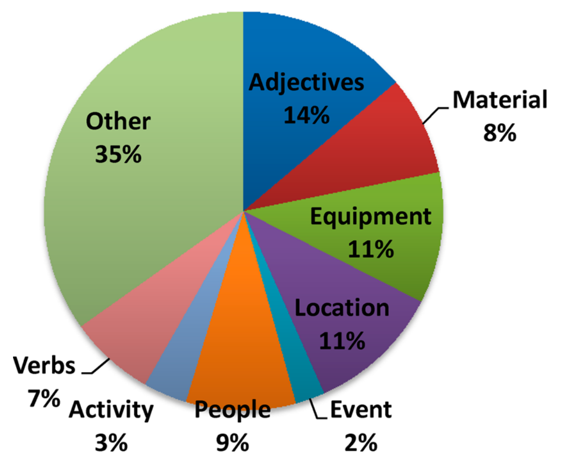


Figure 7. Breakdown of tags on Flickr images tagged with "chemistry" and "experiment" by word type and subject.

According to the tags in Flickr, materials and equipment are clearly important in chemistry experiments, with location and people also common topics of tags. The locations are mostly specific locations, such as specific institutions or towns, but with some more general locations such as "library", "school", and "laboratory". These values are similar to location values used in LabTrove, which include a mixture of specific towns and specific laboratory locations. Metadata about people are more common in Flickr than LabTrove, with mostly general values such as "person", "man", "physicist", and "researcher", but also some specific values such as the names of individuals.

3.2. NASA Blogs. NASA blogs might not seem to have very much in common with chemistry notebooks, but the system that is used by a number of NASA organizations and personnel has some parallels with the LabTrove system. The NASA blogs are based on the WordPress platform.[32] The majority of the NASA blogs use tagging, and it appears that entries are required to have at least one value.

Of the 53 NASA blogs investigated, only 30% of the blogs use three or less metadata items, although nearly 80% of the blogs use the "General" tag. It is unlikely that users are choosing to add the "General" tag deliberately due to the location of the tagging feature far down the interface in the NASA blog system, and therefore, the "General" tag appears to be automatically added to entries when no other tags are included.[32] The inclusion of a default metadata value is normal in WordPress blogs, although it most commonly has a value of "uncategorized". Around 50% of the blogs have multiple authors, including one blog that has 45 authors!

User-defined metadata from the 53 NASA blogs was extracted from tag clouds and category lists to generate 1226 items of metadata. These metadata items were then classified using word or phrase type and subject, as shown in Figure 8.

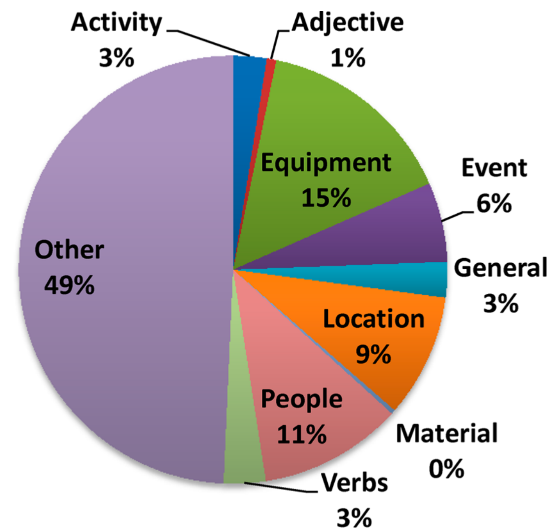


Figure 8. Breakdown of metadata on 53 NASA Blogs by word type and subject.

Full details of the NASA blogs, including names, URLs, extracted metadata, and classifications, can be found in the Supporting Information. The results show that NASA blogs use verb- and adjective-type metadata even less than LabTrove, with the metadata dominated by equipment, such as various rockets and instruments; locations, particularly NASA sites; and people, such as astronauts and NASA leaders, organizations that NASA works with, and the general kinds of people that NASA programs are targeted at, such as women and students.

Figure 9 provides a word cloud of the metadata from 53 NASA blogs, showing clearly how the metadata represents the interests of the NASA organization, and also the frequency of the "General" tag.

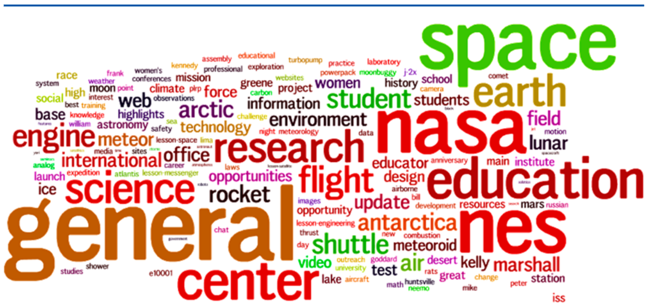


Figure 9. Wordle visualization showing metadata from 53 NASA blogs. Size relates to the relative frequency.

3.3. Chemistry Blogs. The comparison to the NASA blogs is interesting from the point of view of user behavior and word type usage but less relevant in terms of subject classifications of the metadata. In order to address this limitation, metadata from

50 chemistry-related blogs on the WordPress and Blogger platforms were investigated. Both the WordPress and Blogger platforms encourage users to add metadata to their blog entries in the form of tags and categories on WordPress and "labels" on Blogger. Whether the metadata is visible on entries or within the interface in tag clouds or other navigation assistants varies depending on the blog author's settings. WordPress has an interesting feature that encourages the user to add tags to a newly created entry based on the tags that others have used, but this seems to be using only other tags that have been added rather than making suggestions based upon the content of the entry.

The chemistry-related blogs in the study were found through searching Technorati or from recommended chemistry blog compilations. Only blogs that had accessible metadata were chosen for the investigation, with a further 90 blogs excluded from the study because they had little or inaccessible metadata. One of the advantages with the Blogger platform when extracting metadata is that you can get a list of the available labels by adding a query string to the Web address of the blog, whereas extracting metadata from a WordPress blog relies upon a tag cloud or category list being visible in the chosen theme.

User-defined metadata from the 50 chemistry-related blogs were extracted from tag clouds, category lists, and label lists to generate 10,436 items of metadata. These metadata items were then classified using word or phrase type and subject, as shown in Figure 10. Full details of the chemistry-related blogs,

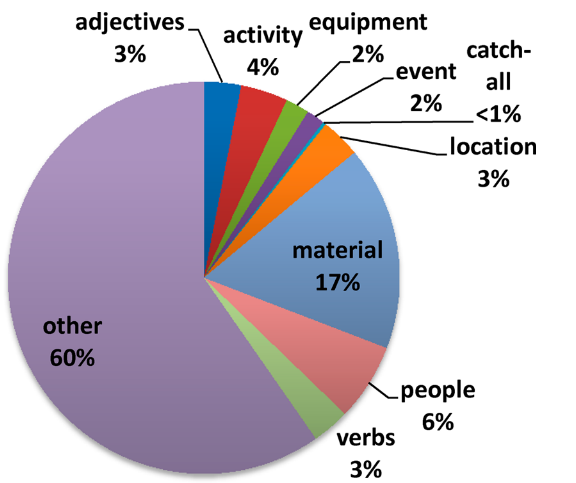


Figure 10. Breakdown of metadata from 50 chemistry-related blogs by word or phrase type and subject.

including names, URLs, extracted metadata, and classifications, can be found in the Supporting Information. The results of the classification are very similar to the NASA blogs in that the number of adjective- and verb-type words or phrases is relatively low, and "things" represent the majority of the metadata, as seen in the other metadata sources. When the metadata is classified by subject, the dominant category comprises values that describe materials such as organisms, chemicals, and drugs. Categories describing equipment, people, and locations have smaller numbers of values. One third of the blogs use a "catch-all" category such as "uncategorized", the default if no metadata is chosen in WordPress, or values like "General" and "Miscellaneous". Figure 11 provides a word cloud of the most common metadata values from the chemistry-related blogs, showing the variety of different topics

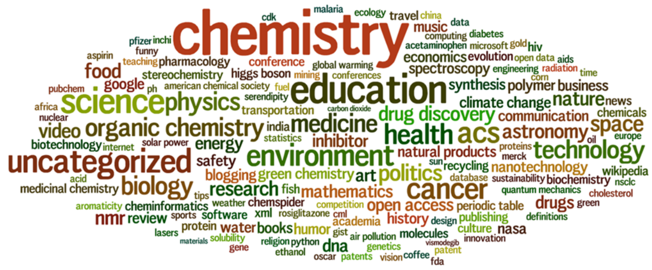


Figure 11. Wordle visualization showing metadata items from 50 chemistry-related blogs. Size relates to the relative frequency.

covered by the blogs. The word cloud also demonstrates how the high-level type metadata becomes more visible and therefore appears more important than the specific terms in this kind of visualization.

3.4. myExperiment. The myExperiment Web site[33] was chosen for surveying metadata because the primary function of the Web site is for researchers to store and share particulars relating to their research methods and experiment workflows.[34] We assumed that these particulars are likely to contain a lot of information about research activities and may therefore be tagged with user-defined metadata that reflects these activities, providing a useful comparison to the experiment data in LabTrove.

The myExperiment Web site enables users to add "tags" to describe their workflows. Users can choose from their previously added tags, which are visible in the interface. An interesting addition to the myExperiment interface is the "This Workflow has no tags!" message if the user has not added any tags to their workflow. The message may be effective at encouraging users to add tags by highlighting the workflows where they have been omitted. A sample of 10% of the most recently added myExperiment workflows suggests this tactic would be successful, with 96% of the workflows including at least one tag.

User-defined metadata from myExperiment was extracted from the "All tags on myExperiment" page[35] to obtain 2349 unique items of metadata. These metadata items were then classified using word or phrase type and subject, as shown in Figure 12. The extracted metadata and classifications can be found in the Supporting Information. The results show that

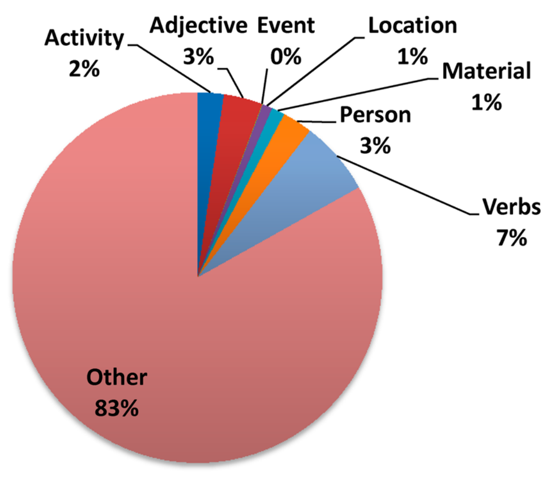


Figure 12. Breakdown of myExperiment tags by word or phrase type and subject.

despite the site being primarily about activities and processes such as experiment workflows and methods, the tags are still dominated by noun-type words and phrases, with similar percentages of verb-type and activity metadata as LabTrove. The metadata in myExperiment is in fact dominated by computer-related subjects, with over one-quarter of the tags comprising abbreviations and acronyms. Materials, equipment, location, and people are relatively insignificant for this community, which is dominated by computer scientists, information professionals, and researchers.

4. ENHANCING METADATA CAPTURE IN LABTROVE

This section provides some general observations and suggestions for how the user interface and other aspects of the metadata provision might be enhanced in LabTrove based on the findings from the initial LabTrove metadata survey, user feedback, and investigations presented in this paper. The range of suggested enhancements to LabTrove consists of the following areas:

- Development of standard schemas
- Use of invitations to add and reuse metadata
- Automation and automatic assistance
- Experiment and other notebook entry profiles
- Visual benefits to adding the metadata
- Educating researchers

It should be noted that there is no implied commitment by the authors to implement these recommendations in future releases of LabTrove. Many of the observations and suggestions in this section will be applicable to other ELN platforms or interfaces designed for recording experiment and research information.

4.1. Observations from Studies. In this section, the high-level observations from the studies are presented together with how they might influence the design of LabTrove and, in particular, the user interface.

4.1.1. Mandatory Metadata, Default Values, and "Blank Canvas" Effect. The clearest result from the LabTrove notebook survey and feedback from user experiences was that the biggest inhibitor to adding useful metadata is the "blank canvas" effect, where the users may be willing to add metadata but do not know where to start. In some cases, users do not even know what metadata is. Providing example metadata in the form of standard schemas would help users overcome some of their difficulty with getting started in a notebook, but the survey results indicate that making the addition of metadata mandatory without such assistance is not helpful at encouraging the creation of meaningful metadata. In LabTrove, a Section is mandatory, but a large percentage of users use a meaningless "catch-all" value for their entries. Default values are added automatically in the NASA logs and the WordPress-authored Chemistry blogs if no user-defined value has been added to the entries. The majority of the NASA blogs surveyed included the default value "General", and many of the Chemistry blogs included the WordPress default value "uncategorized". The presence of a default value may encourage some users to replace it with something more meaningful, but its prevalence in the survey results suggests that this is not very effective. Overcoming the lack of knowledge about what metadata is and how to use it requires appropriate education about data management and metadata facilities.

4.1.2. Different Communities Use Different Terminology. The results presented in this paper show metadata is representative of the community that uses it, and therefore, no one set of terminology is likely to work for all users of an ELN. For example, "Materials" is one of the most commonly used metadata values in the notebook survey, but the terminology (and types) used for the materials included in an experiment does differ significantly depending upon discipline and the type of experiment. For example, "Inputs", "Reagents", "Strain", "Products", "Substance", "Compound", "Molecule type", "Batch", and "Sample" might be appropriate Keys depending on whether the community is involved in biology, chemistry, or drug discovery.[36,37] However, the results do show that commonalities exist at a high level between these different communities, particularly the "label"-type information such as materials, locations, data types, people, events, and equipment, but also specific topics such as "science", "chemistry", "education", "pharmaceuticals", "technology", "environment", and "research". Selecting the correct terminology in the interface is important for ensuring the relevance of the metadata for the user and to encourage subsequent capture. Basic schemas could be developed using these common elements and included in the ELN, while different communities can develop their own extended schemas, which can be shared and imported into the ELN for seeding by example. As mentioned in section 2, templates are a mechanism that can be used to encourage users to enter experimental information with a consistent structure and also using consistent vocabulary for the metadata.

The results of the surveys also indicate that users more readily add certain types of information as metadata. For example, the majority of metadata used describes things and objects, rather than activities, indicating that some information that would be useful is not captured as metadata, such as experiment activities and conditions. Mechanisms could be provided to specifically seek addition of this useful but underrepresented context for the experiment record.

4.1.3. Invitations. The majority of the platforms examined in the surveys included interface designs with invitations to add metadata. For example, in Flickr, the user is invited to "Add a tag" and "Choose from your tags". In WordPress, the user is invited to "Choose from your most used tags". In myExperiment, the user is invited to "Add tags". Other platforms that users may be familiar with also use invitations to encourage the creation of information. For example, in Facebook,[38] the user is encouraged to add people, locations, emotions, activities, feelings, and items into their status updates and photos with questions such as "What's on your mind?", "Where was this taken?", and "Who are you with?". Another example, LinkedIn,[39] also uses questions inviting users to add information about their career history and interests. None of the platforms included in the study have invitations that are as prominent as those seen in Facebook or LinkedIn, but many of the mobile applications for Flickr include more prominent invitations to add tags than the Web interface. Making the invitation to add metadata more inviting and more prominent in the LabTrove interface is likely to increase the amount of metadata that is added to entries.

4.1.4. Making Use of Metadata. In Flickr, any member's tags can be viewed as a list, and tags can also be used in advanced search. In myExperiment, all tags can be viewed as a tag cloud, and selecting multiple tags in a filter can be used to narrow a search for workflows. In the WordPress and Blogger blogs, the user can choose through the use of themes or widgets how their blog is displayed and can choose to enable the metadata to be used as a navigation menu or displayed as a list or tag cloud. It is possible for a user to have their blog configured in such a way on these platforms that the metadata is not visible to an audience. This study is unable to comment on the effectiveness of these approaches, but if we want to encourage users to add metadata, then metadata ought to be prominent and provide useful functionality, such as tag clouds, filtering, and the searching capabilities seen across the platforms. In LabTrove, the metadata added to notebook entries is used to create a navigation menu and is also used to provide advanced search functionality, but given that the capability to add user-defined metadata is a strength of the platform, then it should be possible to make better use of the metadata. As discussed in the following sections, there are strategies we can implement to encourage users to add more metadata and to help users to create "better" metadata. Users can be educated to understand how metadata can be used to help them in their day-to-day work, such as finding previous experiments, characterizing notebook entries, grouping entries, and enabling link creation,[12] as well as why it is important for long-term preservation. Better understanding of the benefits of the metadata together with systems and tools making use of this metadata are beneficial for all users.

4.2. Encouraging Addition of Metadata. Capturing metadata for experiments within the ELN provides the opportunity to make metadata creation and curation easier and simpler.[10] As demonstrated by the LabTrove survey, just providing a mechanism to add metadata is not sufficient. More assistance needs to be provided to users to help them create more metadata but also more consistent and appropriate metadata for their experiments. Schemas and ontologies can be used for providing a valuable starting point and consistent terminology and for interoperability with other systems. Users can be supported through more effective interface designs and automatic assistance for metadata creation, and making use of the metadata through visualizations and enhancing the user interface can help users to get value from adding metadata in their every day work.

4.2.1. Providing a Starting Point with Data Using Schemas. The most effective way of solving the "blank canvas" problem is likely to be developing basic schemas, starting with a single generic schema that can be extended by discipline and then extended further to meet the needs of individual research groups. These schemas could be used in LabTrove to provide users with meaningful options for Sections and key-value pairs that they can select when creating their experiments.

Functionality could be provided in LabTrove to enable users and administrators to create and include basic or more complex schemas into LabTrove. The schema could optionally be available in new notebooks. In the United Kingdom, efforts to harmonize subject classification across the higher education and research sector have resulted in the production of a three-layer classification scheme that could provide a basis for metadata schemas for the ELN.[40] The classification includes different discipline areas, methodologies, instruments, material types, and topics.[41] Where metadata is well used, notebooks include terms that match some of the harmonized subject classifications. Individual communities could develop extended schemas and import these schemas into the LabTrove interface. Schemas could be shared with other groups and adapted as required.

Although there is a plethora of standard schemas and ontologies for science, both generic and domain specific, that could be integrated into LabTrove, there is no single adopted format or model. A benefit of integrating such standards into LabTrove is that they provide the opportunity for the exchange of data as well as merely promoting consistent use of metadata terminology. For example, some recent projects by members of the LabTrove communities have created structured data for sharing between the ELN and different systems based on standard metadata requirements.[12,42]

The problem with the myriad of alternatives is choosing which ones to use. Most of the schemas that have been defined also focus on the structure of data rather than the experimental process as a whole, although notable exceptions exist, for example, ORECHEM,[43,44] CMCS,[9] SEML,[45] CCLRC Scientific Metadata Model,[46] SciPort,[47] and CMO.[48] Other questions to be answered include the following: To what extent the schema should be specialized? How to agree common terminology across disciplines? How much flexibility is permitted within the system to add extensions or customize the metadata to your own use? These questions are beyond the scope of this paper. Milsted et al. have provided more in depth discussion of ontologies and LabTrove.[11]

Formal documentation of schema elements would be valuable to ensure understanding and enable consistency of use. Example notebook entries with metadata included should also be provided with the documentation.

In LabTrove, any schema can take advantage of the two different forms of metadata. Section could be formally used to indicate the generic type of content in the entries and the key-value metadata used for specific information about the experiment. LabTrove has no mechanism for metadata hierarchies, but users could construct Key names to imply a hierarchy, for example, Material_Reagent, Material_Strain, and Material_Product. There are various ways that LabTrove metadata could be used to create taxonomies. Techniques usually applied to tags to create hierarchies could be applied to Keys, while key-value pair data could be extracted to create a hierarchical taxonomy, with the Keys representing one level and their values representing another.[16,49] Although this is a post hoc operation, it could be done on the fly, enabling the hierarchy to be displayed within the interface. Complex ontologies could be created from metadata items by making use of a "metadata operator" and chaining of multiple metadata values.[50] For example, a Key of "Conditions" could have multiple Values that themselves contain individual relationships, such as "Temperature:=200" and "Duration:=24 h", or even sequences of information, with appropriate delimiters defined, for example, "ID:=A1; Temperature:=200; Duration:=24 h; pH:=4.2".

4.2.2. Encouraging Addition of Metadata with Invitations and Profiles. Interface design has an important role in encouraging and supporting users to add their own metadata. User interfaces and application behavior can have a strong influence on the quality of metadata that users create.[6] Providing a visible and easy-to-use mechanism for adding metadata is essential. The benefits and functions of metadata cannot easily be exposed until metadata exists in a notebook, so the most important focus should be on enabling the user to add the metadata as simply as possible.

A very simple change that could be made to LabTrove would be to make the interface for adding metadata more inviting by using "action" words, for example, "Choose a Section", "Add a Key and Value", or "Create a new Key". A variety of techniques could be used to make the metadata creation facility more prominent on the page, but an example that would not require a large redesign would be a Web 2.0 overlay that could be used to ask the user to "Add metadata to describe your experiment or entry" on creation of the entry or after the entry has been saved for the first time.

An alternative to asking the user to type content to be used as metadata is to allow the user to select previously used metadata values, as seen in the interfaces of Flickr, WordPress, Blogger, and myExperiment, or even to select existing content to use as a piece of metadata. The user could select a word or phrase and click a button to add the selected text (or image) as a tag. This selection of existing content to use as metadata is similar to the procedure in the A-Book ELN for Biologists, where the users draw a box around a name, procedure, drawing, or other object in order to label or categorize it for later search.[51]

An effective way to elicit quality metadata from the user is to make use of invitations in the form of questions. If we could ask the user specific questions about their experiment, then we could capture more information. Queries such as "What materials did you use in your experiment?" or "What instruments did you use in your experiment?" are focused on the vocabulary and viewpoint of the user and are therefore more likely to lead to the creation of relevant and meaningful metadata. Metadata that can still be stored using the existing metadata capabilities of LabTrove. Key-value pairs could be used for the storage of questions and answers, for example, using a Key of "Material" or "Instrument" with the associated values extracted from the input field. Multiple values could be captured in this way. Providing answers to question-based invitations would provide more consistency in the information generated and be more usable than the current interface. An "experiment profile" could be developed with standard questions that could be used in the interface to capture information relevant to an experiment. The high-level label-type values that represent a large proportion of the metadata used in the LabTrove community could be used as a starting point for this activity.

The community uses LabTrove for more than creating experiment records, so users could therefore be given the opportunity to select what kind of entry they are creating in their notebook. Each entry type, for example, project information, background literature, a plan, an experiment, or data, could have its own profile, with associated questions predefined for capturing relevant metadata. There is a risk that if no appropriate place is provided to record certain information, then that information will fail to be recorded,[52] so it is important to enable users to be able to define their own values if these are more appropriate. Communities and users should have the option to define their own entry types, such as sample description and instrument configurations that are used by some LabTrove communities, and appropriate properties for any type of entry. The Section could be used to store the information about what kind of entry the user has chosen to create.

In order to help determine what the most effective structures are for templates or profiles and what the best questions to ask in invitations might be, we are currently investigating how different structures and questions can affect what information researchers record for both experiments and metadata.[53,54]

4.2.3. Automatic Assistance. Alternatively, automatic methods could be used to prompt the user to add metadata in the same way that WordPress prompts users to add suggested tags based on the existing tags. The content of entries could be used to "guess" what the content may be about and prompt the user based on the results. For example, commonly used words can be extracted with data mining techniques such as word counts, matching words from a predefined vocabulary, identifying synonyms using a dictionary of terms, or creating metadata based on the format and context of the data files. Other examples include taking metadata from images by creating Keys and Values from the EXIF properties[55] and using a variety of tools to validate and extract the domain-specific metadata from data files.[56] Often data files are in tabular format, such as Excel files, or other structured formats, such as CSV files, where metadata could be extracted by examining the column or row headers. A similar function could be added to identify headers within notebook entries, for example, information formatted as a subheading or table headers.

Another example is to compare the notebook text to a dictionary of chemical terms while the user is writing the entry or at the point they save the entry. The user could then be prompted to add the chemical identifier as metadata or better still create the metadata automatically and then ask the user to verify it on saving the entry. The dictionaries used should contain any words relevant to the discipline, for example, equipment, instruments, or procedures. Individual users or teams could create their own custom dictionaries. These dictionaries could also contain mappings to URLs, so that all occurrences of the words could link to additional resources, for example, instructions and specifications for instruments or catalogues for chemicals.

Some work has already been done in LabTrove to perform automatic searching and matching of chemical names to create links to the relevant structures in the ChemSpider database.[57] This work could be extended to automatically create metadata for the entries based on the identified compounds.[58] For example, a key-value pair of "Material"-*Name of compound* could be created with the information derived from the data mining activity.

4.2.4. Visualizing Metadata. It is important to provide some immediate visual benefit from adding the metadata. For example, tag clouds could provide a very visual way to not only access the metadata and content beneath them but also to provide information about the use of the notebook. Not only tag-type metadata can be visualized using tag clouds, key-value metadata can be visualized by using the Key as the basis for the title and topic of the tag cloud. For example, a set of values for the Key "Materials" could be used to create a "Materials" tag cloud, while a "Topics" Key could create a "Topics" tag cloud. This information can be reinforcing for the authors of the notebook entries, encouraging them to add more categories and to associate entries with the appropriate metadata values. Tag clouds are also extremely valuable for helping others viewing the notebook to understand what it is about. The metadata can also be used to assist with filtering, grouping, and searching for information. Making these functions more visible, more powerful, and most importantly useful may increase the amount of metadata provided. Utilizing the interest that students have in social networking and the understanding that they have about metadata in that context may also provide alternative approaches for developing interfaces for metadata in an ELN for collaboration and open science.[24]

4.3. Educating Researchers. Our surveys show that Laboratory scientists need education, training, and encouragement to use metadata to curate their experiments and data.[16]

There are two different aspects of educating researchers that can be considered: (1) education about the metadata provisions available within LabTrove and (2) the vitally important need to provide researchers with education in the techniques and importance of metadata and data management in general.

4.3.1. LabTrove Education. Education is never a substitute for a usable interface, but education and training materials are still an important part of the user experience for software. The team have focused on significant improvements to the metadata documentation for LabTrove based on feedback from the LabTrove community. There is also the potential to add video tutorials and case study-style documentation to describe and demonstrate the use of the metadata provisions in LabTrove, appropriate workflows for adding metadata, and designing schemas for research groups. Courses for LabTrove could also be devised for students and researchers or be included as part of other laboratory and data management education. Exemplar notebooks that show best practices for metadata use and data management could also be made publicly available. Another possibility is to create "metadata checklists" that encompass recommended metadata items and a description for each item.[59] A different checklist could be created for different notebook uses and for different domains and could be made available to the wider research community.

4.3.2. Data Management Education. Despite the almost universal and regular use of technology by undergraduates,[60] there is evidence to show that the expectations that we have that researchers from the "Google Generation" will have high levels of skills and information literacy are unrealistic.[61] This observation is true even for researchers in the physical sciences, who are perceived to be more technical.[62] Research suggests that relatively few students actually know what metadata is or how to use it,[24] and knowledge of data management and curation is uncommon among researchers across all disciplines.[63−65] This suggests that one possible strategy for improving metadata use is to target users directly through data management education programs.[24,66−69]

At the University of Southampton, the majority of students receive some training in information retrieval at the beginning of their studies, usually delivered by library staff. These courses help them to develop skills at finding information from different sources, using filtering and Keywords, but further training is necessary to help them to make their own data "findable" in the future. As suggested by Swan and Brown,[66] there is a clear need to provide data management education to researchers at a postgraduate level, and it would be advantageous for undergraduates. Other studies have advocated the importance of appropriate data management training as early as possible in a researcher's career,[67] and we have previously advocated that instruction in data management should begin in school.[16] Such education should aid the student's understanding of organizing, classifying, and adding metadata to their own work, for their own benefit and for others, such as supervisors, the community, and the public.

The adoption of data management planning and education initiatives, such as the JISC Research data management (JISCMRD) projects such as the RDMP and RDMTrain projects, will contribute to improving the quality of metadata by providing examples of good practice and by developing data management training for researchers.[68,69] Education that is tailored to researchers and provides discipline and tool-specific training may be of more value to researchers than generic data management education.[67] The IDMB Archaeology Case Study provides an example of discipline-specific, practice-led, researcher training, including the introduction of a metadata model for use in a repository tool.[67]

Metadata use can be encouraged and facilitated by librarians and information specialists, who have data management expertise. Ideally, each research team would include a data management expert to provide support and advice and to help ensure best practice. If not, then an expert from outside could be brought in to provide assistance with the schema generation process and data management training[66] and maybe even "go native" to help.[16] Managers of projects could also take an active role by getting involved in the generation of schemas and using metadata quality indicators for assessing projects and data sets.[70]

5. CONCLUSIONS

The flexible metadata framework in LabTrove enables users to define metadata that is most appropriate for their work and experiments. The results of the LabTrove notebook survey demonstrate that simply providing the facility to add metadata is not enough to ensure that metadata is added to a notebook, let alone ensuring high quality metadata is added. For metadata to be useful it first has to be present, but enforcing metadata generation is of no benefit if it is low quality, inconsistent, or irrelevant. The LabTrove survey and the results of user research with the LabTrove community have indicated that the major inhibitors of metadata use in an ELN are the following:

- A lack of a defined metadata schema for a notebook or the "blank-canvas effect"
- A lack of knowledge about metadata
- The effort involved in the creation of metadata
- A lack of visibility and perceived benefits of metadata

The aim of the investigations into other platforms that enable users to create their own metadata was to determine whether the patterns of metadata usage by the LabTrove community were unusual and whether their interfaces might suggest enhancements that could be made to improve metadata capture within LabTrove.

The use of metadata in LabTrove may be lower than seen on the other platforms investigated, although it is difficult to tell because of the social nature of these platforms. The most comparable platforms to LabTrove in this regard are the NASA blogs, which shows a higher overall use of metadata but also show the same high use of a meaningless "catch-all" value.

The investigations indicated that the types of metadata that are added do vary across the different platforms and communities, but there are also many commonalities. In terms of word type, LabTrove is similar in verb and adjective use to both the NASA blogs and the Chemistry blogs. But myExperiment and Flickr use a greater proportion of verbs, and Flickr alone uses a significantly higher number of adjectives. LabTrove is notable for having a high percentage of activities described using noun-type words compared to the other platforms. This nominalization may reflect the research nature of LabTrove compared to the other platforms.[71] The subjects described in the metadata have commonalities across the platforms, such as materials, equipment, people, and locations. Each platform shows differences in the dominant subject types that are relevant to that platform. For example, in the NASA blogs, the dominant subject is equipment, describing entries about NASA-developed technologies such as rockets and satellite instruments, while myExperiment is dominated by computer-related subjects describing software or data formats used in the experiment workflows. High-level labels dominate LabTrove Section values and Key terms, a pattern that is different from the metadata seen on the other platforms, although high-level terms such as "chemistry" and "education" are seen across all the platforms. Materials dominate the Value terms in LabTrove, and this term also makes up a high proportion of the Section terms and Key terms used, which is in common with the metadata from the Chemistry blogs investigated.

The metadata annotating digital documents must be present, of high quality, and consistent if it is to be useful. Given that a significant percentage of LabTrove users appear to have adopted a "minimum required" approach to metadata, working to encourage and support the creation of metadata is essential. Three different approaches could be used, either independently or in combination. The first approach is through the development of basic and community-appropriate metadata schemas and vocabularies, which can be used to seed LabTrove with metadata examples. Seeded metadata will help to mitigate the "blank canvas effect", by providing a starting point that users and teams can build upon, and can also improve quality by nurturing the use of consistent terms and abbreviations. The second approach is to apply data mining and other techniques to automate more capture of metadata by extracting key terms or guessing context by analyzing the content of the notebook entries. The final approach, and perhaps the most valuable in terms of encouraging the capture of user-defined metadata, is to enhance the user interface. Users can be encouraged to add metadata to their entries through the use of invitations prompting them to create metadata, reuse metadata, and provide information about specific elements of their experiment. More use can be made of the captured metadata within the interface, which in turn helps users to see the benefits of adding it and encouraging further metadata creation.

Our findings from these surveys have already had an influence on our ELN development activities in Southampton University and are likely to inform our ELN designs in the future.

■ ASSOCIATED CONTENT

Ⓢ Supporting Information

Numerical breakdowns and raw data of the results of the metadata surveys. Because many of the LabTrove notebooks included in the survey were private and not accessible to the public, the "raw" metadata from the LabTrove survey is not included. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail cerys.willoughby@soton.ac.uk.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The research reported in this paper has been made possible by funding from the RCUK e-Science programme (EPSRC Grant GR/R67729, EP/C008863, EP/E502997, EP/G026238, BBSRC BB/D00652X), HEFCE and JISC Data Management Programme, and University Modernisation (UMF), and most recently the RCUK Digital Economy Theme as part of the IT as a Utility Network+ funding (EPSRC EP/K003569).

■ REFERENCES

(1) Boulton, G.; Campbell, P.; Collins, B.; Elias, P.; Hall, W.; Laurie, G.; O'Neill, O.; Rawlins, M.; Thornton, J.; Vallance P.; Walport, M. *Science as an Open Enterprise*, 2012. http://royalsociety.org/uploadedFiles/Royal_Society_Content/policy/projects/sape/2012-06-20-SAOE.pdf (accessed July 11, 2013).

(2) Ball, A. Metadata for Data Citation and Discovery. In *Disseminate, Discover: Metadata for Effective Data Citation*; DataCite, British Library, JISC Workshop, Opus: University of Bath Online Publication Store, 2012. http://opus.bath.ac.uk/30505/ (accessed July 11, 2013).

(3) Bird, C.; Willoughby, C.; Frey, J. Laboratory notebooks in the digital era: Record keeping in chemical and other science laboratories. *Chem. Soc. Rev.* **2013**, *42*, 8157−8175 DOI: 10.1039/c3cs60122f.

(4) Borgman, C. L. *Scholarship in the Digital Age: Information, Infrastructure, and the Internet*. MIT Press: Cambridge, MA, 2007; ISBN: 9780262514903.

(5) Borgman, C. L. Data, disciplines, and scholarly publishing. *Learn. Publ.* **2008**, *21*, 29−38.

(6) Crystal, A.; Greenberg, J. Usability of metadata creation application for resource authors. *Libr. Inf. Sci. Res.* **2005**, *27*, 177−189.

(7) Currier, S.; Barton, J.; O'Beirne, R.; Ryan, B. Quality assurance for digital learning object repositories: Issues for the metadata creation process. *ALT-J.* **2004**, *12*, 5−20.

(8) Ryan, B.; Walmsley, S. Implementing Metadata Collection: A Project's Problems and Solutions. *Learning Technol.* **2003**, *5*, Article 4. http://www.ieeetclt.org/issues/january2003/index.html#3 (accessed October 26, 2014).

(9) Pancerella, C.; Hewson, J.; Koegler, W.; Leahy, D.; Rahn, L.; Yang, C.; Myers, J. D.; Didier, B.; McCoy, R.; Schuchardt, K.; Stephan, E.; Windus, T.; Amin, K.; Bittner, S.; Lansing, C.; Minkoff, M.; Nijsure, S.; Laszewski, G.; Pinzon, R.; Ruscic, B.; Wagner, A.; Wang, B.; Pitz, W.; Ho, Y.; Montoya, D.; Xu, L.; Allison, T.; Green, W.; Frenklach, M. Metadata in the Collaboratory for Multi-Scale Chemical Science. In *DCMI '03 Proceedings of the 2003 International Conference on Dublin Core and Metadata Applications: Supporting Communities of Discourse and Practice – Metadata Research & Applications*, September 28−October 2, 2003, Seattle, WA.http://dcpapers.dublincore.org/pubs/article/view/740 (accessed July 11, **2013**).

(10) Frey, J. Curation of Laboratory Experimental Data as Part of the Overall Data Lifecycle *Int. J. Digital Curation*, **2008**. *3*, 44−62. DOI: 10.2218/ijdc.v3i1. http://ijdc.net/index.php/ijdc/article/view/62/41 (accessed July 11, 2013).

(11) Milsted, A.; Hale, J.; Frey, J.; Neylon, C. LabTrove: A lightweight, Web-based, laboratory "blog" as a route towards a marked up record of work in a bioscience research laboratory. *PLoS One* **2013**, *8*, e67460 DOI: 10.1371/journal.pone.0067460.

(12) Frey, J. G.; Coles, S.; Bird, C. L.; Brocklesby, W. S.; Badiola, K. A.; Williamson, A. E.; Cronshaw, J. R.; Robins, M.; Roberston, M. N.; Todd, M. H.; Chapman, R. T.; Fisher, A.; Grossel, M.; Casson, J.; Hibbert, D. B.; Gloria, D.; Mapp, L. K.; Matthews, B.; Milsted, A.; Minns, R. S.; Mueller, K. T.; Parkinson, T.; Quinnell, R.; Robinson, J. S.; Springate, E.; Tizzard, G. J.; Willoughby, C.; Yang, E.; Ylioja, P. M.; Knight, N. J.; Marazzi, L. Experiences with a researcher-centric ELN. *Chem. Sci.* **2014**, DOI: 10.1039/C4SC02128B.

(13) Gilliland, A. J. Setting the Stage. In *Introduction to Metadata: Pathways to Digital Information*, version 3.0; Baca, M., Ed.; Getty Information Institute: Los Angeles, 2008. http://www.getty.edu/research/publications/electronic_publications/intrometadata/setting.html (accessed October 17, 2014).

(14) Greenberg, J. A quantitative categorical analysis of metadata elements in image-applicable metadata schemas. *J. Assoc. Inf. Sci. Technol.* **2001**, *52*, 917−924 DOI: 10.1002/asi.1170.

(15) Zeng, M. L.; Qin, J. *Metadata*; Neal-Schuman Publishers: New York, 2008; ISBN: 9781555706357.

(16) Bird, C. L.; Willoughby, C.; Coles, S. J.; Frey, J. G. Data curation issues in the chemical sciences. *Inf. Stand. Q.* **2013**, *25* (3), 4−12 http://dx.doi.org/10.3789/isqv25no3.2013.02.
(17) Labtrove troves used in the study: http://altc.ourexperiment.org/, http://biolab.isis.rl.ac.uk, http://blogs.chem.soton.ac.uk/, http://www.ourexperiment.org/, http://xray.orc.soton.ac.uk/, https://www.enotebook.science.unsw.edu.au/, and http://mueller.ourexperiment.org.
(18) Attribute−Value Pair. Wikipedia.http://en.wikipedia.org/wiki/Key_Value_Pair (accessed May 14, 2014).
(19) Eynard, D.; Mazzola, L.; Dattolo, A. Exploiting tag similarities to discover synonyms and homonyms in folksonomies. *Softw: Pract. Exper.* **2012**, *12*, 1437−1457 DOI: 10.1002/spe.2150.
(20) Rahm, E.; Hai Do, H. Data cleaning: Problems and current approaches. *IEEE Data Eng. Bull.* **2000**, *23* (4), 3−13.
(21) IBM. Generating a Taxonomy for Documents from Tag Data. U.S. Patent US 8346776 B2, 2013. http://www.google.com/patents/US8346776 (accessed July 31, 2013).
(22) Kjellberg, S. I am a blogging researcher: Motivations for blogging in a scholarly context. *First Monday* **2010**, *15*, Article 8. http://firstmonday.org/ojs/index.php/fm/article/view/2962/2580 (accessed July 31, 2013).
(23) Puschmann, C.; Mahrt, M. Scholarly Blogging: A New Form of Publishing or Science Journalism 2.0? In *Science and the Internet*; Tokar, A., Beurskens, M., Keuneke, S., Mahrt, M., Peters, I., Puschmann, C., Weller, K., van Treeck, T., Eds.; Düsseldorf University Press: Düsseldorf, 2012; pp 171−182.
(24) Mitchell, E. T. Metadata Literacy: An Analysis Of Metadata Awareness in College Students. Ph.D. Thesis, School of Information and Library Science, University of North Carolina at Chapel Hill, 2010. http://dc.lib.unc.edu/cdm/ref/collection/etd/id/2910 (accessed July 31, 2013).
(25) Ames, M.; Naaman, M. Why We Tag: Motivations for Annotation in Mobile and Online Media. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*; Rosson, M. B., Gilmore, D., Eds.; Association for Computing Machinery: San Jose, CA, 2007; pp 971−980.
(26) Hammond, T.; Hannay, T.; Lund, B.; Scott, J. Social Bookmarking Tools (I): A General Review. *D-Lib Mag.* **2005**, *11*, article 1. http://www.dlib.org/dlib/april05/hammond/04hammond.html (accessed October 16, 2014).
(27) Flickr: The Commons. https://www.flickr.com/commons (accessed April 7, 2014).
(28) Discussing Machine Tags in Flickr API. https://www.flickr.com/groups/api/discuss/72157594497877875/ (accessed April 7, 2014).
(29) Proyecto Agua. Water Project on Flickr. https://www.flickr.com/people/microagua/ (accessed April 7, 2014).
(30) Flickr Hive Mind. http://flickrhivemind.net (accessed April 7, 2014).
(31) Most Interesting Pictures Tagged with Chemistry and Experiment from Flickr Hive Mind. http://flickrhivemind.net/Tags/chemistry,experiment/Interesting (accessed January 7, 2014).
(32) Yembrick, J. Personal communication, April 3, 2014.
(33) myExperiment. http://www.myexperiment.org (accessed April 7, 2014).
(34) De Roure, D.; Goble, C.; Stevens, R. *The Design and Realisation of the myExperiment Virtual Research Environment for Social Sharing of Workflows*; Future Generation Computer Systems 25; University of Southampton: Southampton, U.K., 2009.
(35) All tags on myExperiment http://www.myexperiment.org/tags/ (accessed April 7, 2014).
(36) Sander, T.; Freyss, J.; von Korff, M.; Renée Reich, J.; Rufener, C. OSIRIS, an entirely in-house developed drug discovery informatics system. *J. Chem. Inf. Model.* **2009**, *49*, 232−246.
(37) Thibault, J. C.; Facelli, J. C.; Cheatham, T. E. iBIOMES: Managing and sharing biomolecular simulation data in a distributed environment. *J. Chem. Inf. Model* **2013**, *53*, 726−736 DOI: 10.1021/ci300524j.
(38) Facebook. https://www.facebook.com (accessed May 23, 2014).
(39) LinkedIn. https://www.linkedin.com (accessed May 23, 2014).
(40) Summary of Peer Review Harmonisation Activities. Research Councils UK. http://www.rcuk.ac.uk/research/efficiency/researchadmin/harmonisation/ (accessed July 31, 2013).
(41) RCUK Research Classifications (data file). http://www.rcuk.ac.uk/research/Efficiency/Pages/harmonisation.aspx (accessed November 2014).
(42) Coles, S. J.; Frey, J. G.; Bird, C. L.; Whitby, R. J.; Day, A. E. First steps towards semantic descriptions of electronic laboratory notebook records. *J. Cheminf.* **2013**, *5:52*, 1−10 DOI: 10.1186/1758-2946-5-52.
(43) Borkum, M.; Frey, J. G.; Coles, S. oreChem: Planning and Enacting Chemistry on the Semantic Web. Microsoft Research eScience Workshop 2010, Berkeley, CA, October 11−13, 2010. http://www.slideshare.net/mark.borkum/orechem-planning-and-enacting-chemistry-on-the-semantic-web (accessed October 22, 2014).
(44) Lagoze, C. The oreChem Project: Integrating Chemistry Scholarship with the Semantic Web. In *Proceedings of the WebSci'09: Society On-Line*, Athens, Greece, 2009.
(45) Kaestle, G.; Shek, E. C.; Dao, S. K. Sharing Experiences from Scientific Experiments. In *Proceedings of the 11th International Conference on Scientific on Scientific and Statistical Database Management*, Cleveland, OH, July 28−30, 1999, pp 168−177.
(46) Sufi, S.; Matthews, B.; Kleese van Dam, K. An interdisciplinary model for the representation of scientific studies and associated data holdings. UK e-Science All Hands Meeting. Nottingham, U.K., September 2−4, 2003, pp 103−110. http://www.nesc.ac.uk/events/ahm2003/AHMCD/pdf/020.pdf (accessed November 2104).
(47) Wang, F.; Liu, P.; Pearson, J.; Azar, F.; Madlmayr, G. Experiment Management with Metadata-based Integration for Collaborative Scientific Research. ICDE'06 Proceedings of the 22nd International Conference on Data Engineering, IEEE Computer Society, Atlanta, GA, April 3−7, 2006, p 96. DOI: 10.1109/ICDE.2006.65.
(48) Chemical Methods Ontology (CMO). Royal Society of Chemistry. http://www.rsc.org/ontologies/CMO/index.asp (accessed October 22, 2014).
(49) Heymann, P.; Garcia-Molina, H. *Collaborative Creation of Communal Hierarchical Taxonomies in Social Tagging Systems*; Technical Report 2006-10; Stanford University: Stanford, CA, 2006. http://ilpubs.stanford.edu:8090/775/1/2006-10.pdf (accessed July 31, 2013).
(50) Robson, B. Clinical and pharmacogenomic data mining: 4. The FANO program and command set as an example of tools for biomedical discovery and evidence-based medicine. *J. Proteome Res.* **2008**, *7*, 3922−3947 DOI: 10.1021/pr800204f.
(51) Mackay, W. E.; Pothier, G. The A-Book: An augmented laboratory notebook for biologists. *ERCIM News* **2001**, *46*, 52−53.
(52) Swinglehurst, D.; Greenhalgh, T.; Roberts, C. Computer templates in chronic disease management: Ethnographic case study in general practice. *BMJ Open* **2012**, *2*, e001754 DOI: 10.1136/bmjopen-2012-001754.
(53) Willoughby, C.; Bird, C.; Frey, J. User-defined metadata: Using cues and changing perspectives. 2014, submitted.
(54) Willoughby, C.; Frey, J. Effects of using structured templates for recording chemistry experiments. 2014, in preparation.
(55) Exchangeable Image File Format. Wikipedia. http://en.wikipedia.org/wiki/Exchangeable_image_file_format (accessed May 23, 2014).
(56) Downing, J.; Murray-Rust, P.; Tonge, A. P.; Morgan, P.; Rzepa, H. S.; Cotterill, F.; Harvey, M. J. SPECTRa: The deposition and validation of primary chemistry research data in digital repositories. *J. Chem. Inf. Model.* **2008**, *48*, 1571−1581 DOI: 10.1021/ci7004737.
(57) Chemspider. http://www.chemspider.com/About.aspx? (accessed July 29, 2013).
(58) Day, A. Search and Insert from ChemSpider in LabTrove and Other Websites with TinyMCE Editors, 2013. http://www.chemspider.com/blog/search-and-add-from-chemspiderdirectly-from-

labtrove-and-other-blog-based-websites-which-use-tinymce.html (accessed July 11, 2013).
(59) Higdon, R.; Stewart, E.; Stanberry, L.; Haynes, W.; Choiniere, J.; Montague, E.; Kolker, E. MOPED enables discoveries through consistently processed proteomics data. *J. Proteome Res.* **2014**, *13*, 107−113 DOI: 10.1021/pr400884c.
(60) Smith, S. D.; Salaway, G.; Caruso, J. B. The ECAR Study of Undergraduate Students and Information Technology. EDUCAUSE Center for Applied Research, 2009. http://net.educause.edu/ir/library/pdf/ers0906/rs/ers0906w.pdf (accessed July 31, 2013).
(61) Rowlands, I.; Nicholas, D.; Williams, P.; Huntington, P.; Fieldhouse, M.; Gunter, B.; Withey, R.; Jamali, H. R.; Dobrowolski, T.; Tenopir, C. The Google generation: The information behaviour of the researcher of the future. *Aslib Proc.* **2008**, *60*, 290−310.
(62) Goodger, J.; Worthington, W. Research Data Management Training for the Whole Project Lifecycle. In *Physics & Astronomy Research (RDMTPA) JISC Final Report*. Research Data Toolkit, University of Hertfordshire, 2013. http://research-data-toolkit.herts.ac.uk/document/rdmtpa-final-report/ (accessed July 11, 2013).
(63) Whitton, M.; Takeda, K. Data Management Questionnaire Results: IDMB Project. University of Southampton: Southampton, GB, 2011. http://eprints.soton.ac.uk/196243/1/IDMB_Survey_Report.pdf (accessed July 11, 2013).
(64) Whitton, M. Institutional Data Management Blueprint (IDMB) Questionnaire Quantitative Data. University of Southampton: Southampton, GB, 2011. http://eprints.soton.ac.uk/195959/ (accessed July 11, 2013).
(65) Ward, C.; Freiman, L.; Jones, S.; Molloy, L.; Snow, K. Making Sense: Talking Data Management with Researchers. *Int. J. Digital Curation*. **2011**, *6*, 265−273. http://www.ijdc.net/index.php/ijdc/article/view/197 (accessed July 11, 2013).
(66) Swan, A.; Brown, S. Skills, Role & Career Structure of Data Scientists & Curators: Assessment of Current Practice & Future Needs; Report to the JISC; 2008. http://www.jisc.ac.uk/media/documents/programmes/digitalrepositories/dataskillscareersfinalreport.pdf (accessed July 11, 2013).
(67) Earl, G.; White, W.; Wake, P. *IDMB Archaeology Case Study: Summary*; University of Southampton: Southampton, GB, 2011. http://eprints.soton.ac.uk/id/eprint/196237 (accessed July 11, 2013).
(68) Managing Research Data Programme 2011-13. JISC. http://www.jisc.ac.uk/whatwedo/programmes/di_researchmanagement/managingresearchdata.aspx (accessed July 11, 2013).
(69) Hodson, S. Meeting the Research Data Challenge JISC Briefing Paper, 2009. http://www.jisc.ac.uk/publications/briefingpapers/2009/bpresearchdatachallenge.aspx#downloads (accessed July 11, 2003).
(70) Bruce, T. R.; Hillmann, D. Metadata Quality in a Linked Data Context. *VoxPopuLII*, 2013. http://blog.law.cornell.edu/voxpop/2013/01/24/metadata-quality-in-a-linked-data-context/ (accessed July 31, 2013).
(71) Biber, D.; Gray, B. Nominalizing the Verb Phrase in Academic Science Writing. In *The Verb Phrase in English*, 1st ed.; Aarts, B.; Close, J.; Leech, G.; Wallis, S., Eds.; Cambridge University Press: Cambridge, 2013; pp 99−132. http://dx.doi.org/10.1017/CBO9781139060998.006 (accessed November 2014).

User-Defined Metadata: Using Cues and Changing Perspectives

Willoughby, C., Bird, C.L., & Frey, J.G. User-Defined Metadata: Using Cues and Changing Perspectives. International Journal of Digital Curation. 10(1), 18-47, as described in **Chapter 7: Generating metadata by using cues and changing perspectives**.

IJDC | *Peer-Reviewed Paper*

User-Defined Metadata: Using Cues and Changing Perspectives

Cerys Willoughby
University of Southampton

Colin Bird
University of Southampton

Jeremy Frey
University of Southampton

Abstract

User-defined metadata is useful for curating and helping to provide context for experiment records, but our previous investigations have demonstrated that simply providing the facility to add metadata is not enough to ensure that metadata is added, let alone to ensure that the metadata is of high quality. For metadata to be useful it first has to be present, but enforcing metadata generation is of no benefit if it is low quality, inconsistent, or irrelevant. Researchers need support. One strategy to encourage more effective metadata creation is to design user interfaces that invite users to add metadata by asking them questions. If we ask users specific questions about their experiments and other activities then we could capture more relevant or useful metadata, although there is a risk that asking the wrong questions may lead to loss of valuable metadata terms or the creation of irrelevant material. In this paper we report on a study to investigate how different questions could be used to generate metadata by eliciting information in three different conditions: free recall, changing perspective by thinking about search terms to help someone else, and providing cues by using a set of topic-based questions. We also investigate how responses varied with different information types. The results of the study show that different terms are created under the different conditions, as expected. The use of cues generates the highest numbers of terms and the most diverse range, including elements that are not captured in other conditions. However, important themes generated in other conditions are not produced because the cues to create them are missing. The study also generated a number of unexpected findings, including responses describing information that is not in the original material: personal opinions and experiences, and comments about the information text itself. These unexpected responses have both positive and negative consequences for the generation of metadata and the curation of scientific records. The results of studies using these techniques to capture metadata for chemistry experiments are also discussed.

Received 13 October 2014 | *Revised* 5 January 2015 | *Accepted* 10 February 2015

Correspondence should be addressed to Cerys Willoughby, School of Chemistry, University of Southampton, Southampton, SO17 1BJ. Email: cerys.willoughby@soton.ac.uk

An earlier version of this paper was presented at the 10th International Digital Curation Conference.

The *International Journal of Digital Curation* is an international journal committed to scholarly excellence and dedicated to the advancement of digital curation across a wide range of sectors. The IJDC is published by the University of Edinburgh on behalf of the Digital Curation Centre. ISSN: 1746-8256. URL: http://www.ijdc.net/

http://dx.doi.org/10.2218/ijdc.v10i1.343
DOI: 10.2218/ijdc.v10i1.343

Introduction

Scientific records include not only data but also details of ideas, plans, methods, observations, results and analyses. Science depends on these records being preserved and maintained, and therefore adequately curated in such a way that the information is both searchable and reusable (Bird, Willoughby, Coles and Frey, 2013a). Metadata is essential for efficient access, but also for search and retrieval, reuse, and providing context for the data (Gilliland, 2008; Zeng and Qin, 2008; Bird, Willoughby and Frey, 2013b). Effective metadata for scientific records can have value outside of archive systems to both help the original author of the record to locate and reuse their own materials, and to enable public or collaborator access to research data as soon as it has been produced. One of the most important elements that metadata provides for experiments is the context, without which digitally captured data in particular becomes meaningless (Borgman, 2008; Frey, 2008). The lack of formal curation as part of managing data in chemistry can lead to the loss of data that would have been useful if it had been preserved, and reinterpreting paper archives is difficult (Downing, et al., 2008).

Although some metadata can be created automatically, some tasks require ‘human-generated’ or ‘user-defined’ metadata to capture the full context (Currier, et al., 2004; Greenberg and Robertson, 2002). The humans required to create the metadata may or may not be subject specialists or information professionals, and may or may not be the original creator of an information object (Gilliland, 2008). Although there are potential roles for librarians and information specialists to assist with curating scientific records, particularly later in the process, the researchers themselves are best placed to provide the context for the experiment and therefore generate the experiment metadata (Bird, Willoughby, Coles and Frey, 2013a; Frey, 2008).

Curation should be a concern to the researcher, and not be seen as something that happens later, for example as part of the publication process. However, previous studies of researcher behaviour and attitudes have identified the so-called ‘burden of curation’ that leads to difficulties encouraging researchers to generate appropriate metadata for their data (Borgman, 2007, 2008; Crystal and Greenberg, 2005; Ryan and Walmsley, 2003; Frey, 2008). Content creators see metadata creation as extra work, owing to limited understanding of the rationale and value of adding metadata, combined with a lack of incentives for creating it (Greenberg, 2004; Currier et al., 2004). On top of a shortage of rewards for data management comes a reluctance to share research data in the first place, particularly by chemists whose data has high value to industrial funders (Borgman, 2010).

The expanding use of electronic laboratory notebooks (ELNs) provides an opportunity to assist the creators of scientific records by designing curation into the experiment process and encouraging ‘curation at source’, making metadata creation more effective, efficient, and less error-prone (Frey, 2008). However, just providing the capability to add metadata within experiment records is not sufficient to ensure that researchers will add it or that it will be useful for curation. In our own studies we have investigated the effectiveness of metadata use within an ELN developed at the University of Southampton, which enables users to add their own user-defined metadata to their scientific records[1]. Our findings indicated that metadata is not effectively used

1 LabTrove: http://www.labtrove.org

and a large percentage of users have adopted a 'minimum required approach' where metadata has only been added to notebook entries because it is required by the system (Willoughby, et al., 2014). We also found that researchers do not necessarily understand what metadata is and how it can benefit them. Users may be willing to add metadata but they do not know where to start. Some of our users also expressed anxieties about using the "wrong" metadata and the effort involved in fixing problems later. Our previous study also indicated that different types of words were used for metadata depending on the information type – text or image – being annotated; in particular adjectives were more common on photographs than on text-based materials, and verbs were relatively low across all information types. The lack of certain word types may indicate that certain important information, such as techniques and processes used in experiments, is not represented in the metadata.

Researchers are not experts in the complex task of metadata creation and find it even more difficult to create metadata that will be useful for others (Borgman, 2008). A variety of strategies could be adopted to help with these difficulties, including better data management education for researchers, the use of predefined taxonomies, and using data mining to automatically extract metadata. Another strategy is to create well-designed systems that support researchers with creating metadata (Crystal and Greenberg, 2005; Greenberg et al., 2003). One approach is to design interfaces that make use of question-based invitations prompting researchers to add information about specific elements of their experiment.

Using Cues and Changing Perspective

Previous experiences and knowledge influence what information we remember, but other factors can influence what we remember or choose to recall when asked (Tourangeau et al., 2000; Jabine et al., 1984). The conditions under which subjects are asked to recall information affects what they remember and how they present the information. Cues can act as reminders, but may also shape the information that is recalled (Tourangeau et al., 2014; Higham and Tam, 2005; Marian and Neisser, 2000). Subjects can also be prompted to recall different information when asked to recall something from a different perspective (Dudukovic et al., 2004; Tversky and Marsh, 2000). The findings from such studies have been used to enhance methods for information collection, including witness interviews (Memon et al., 2010) and designs of surveys and questionnaires (Schwarz, 2007; Tourangeau et al., 2000; Jobe and Mingay, 1989). That cues might be useful for aiding in metadata creation has been noted previously by Crystal and Greenberg (2005).

Asking researchers specific questions with cues that prompt them to recall particular information of interest, or encouraging them to change their perspective, should change what information they provide compared to free recall, in which they choose the first information that comes to mind. What is not clear is exactly what these different responses might be and whether they will be beneficial for the generation of metadata. More relevant or useful metadata may be captured, but there is a risk that asking the wrong questions may lead to loss of valuable metadata terms or the creation of irrelevant material. In this paper we report on a study to investigate whether information that could be usefully used as metadata is generated when responses are elicited under the following three conditions:

- **Free recall:** The user is asked to write up to ten words or phrases that come to mind when they look at the photograph or read the text in a questionnaire.

- **Audience (change perspective):** The user is asked to write up to ten words or phrases that come to mind for use in a search engine when they imagine helping someone else to find a similar picture or piece of text on the internet.

- **Profile (cued recall):** The user is asked to provide word or short phrase answers to a number of questions to elicit specific information about the photograph or text, including locations, people, equipment, activities or actions, other objects, and any other words or phrases of their choice to describe the material type.

These conditions have been chosen to investigate whether these mechanisms are suitable for generating user-defined metadata for facilitating discovery (*audience*) by asking the participants to identify words that would be useful for performing a search for similar information, and for capturing context (*profile*) by asking the participants to describe elements that are important for a scientific record. The questions in the *profile* condition are derived from user-defined metadata we have observed to be commonly provided when describing chemistry experiments (Willoughby, et al., 2014).

Records in ELNs can be both structured, through the use of templates, or unstructured. For this study information types representing photographs, structured text-based material, and unstructured text-based material were included to investigate whether the information type had any effect on the responses produced in the different conditions. In our previous study we found that different types of words were used for metadata depending on the information type being annotated; in particular adjectives were more common on photographs than on text-based materials, and verbs were relatively rare across all information types. The lack of certain word types may indicate that certain important information is not being represented in the metadata, in particular the techniques and processes used in an experiment. For this reason we wanted to investigate whether the different conditions would result in a different pattern of word use. It was anticipated that more adjectives would be seen in the photograph, compared to the text-based materials, based on our previous findings, and that more verbs would be seen in the *profile* condition because a specific question would be asked about the activities present in the material.

Methods

Our study used an online questionnaire to present information objects of different types: photographs, structured text, and unstructured text. Each of the photographs contains people engaged in activities, the structured texts are cookbook-style recipes, and the unstructured texts each describe a different procedure, such as fixing a puncture, to reflect the active nature of experiments. Participants were asked to complete a series of questionnaires in sequence, each containing one of the materials and associated cues (see Table 1, Appendix 1 and Appendix 2). Each participant was randomly directed to one sequence, enabling a comparison to be made with different cues to be used for the same materials.

Table 1. Information types and recall conditions used in the survey.

Information Object	Information Type	Condition
Plenary meeting	Photograph (1)	Free recall
Goan Chicken Recipe	Structured (1)	Free recall
Fixing a puncture	Unstructured (1)	Free recall
Indian band	Photograph (2)	Change perspective or Cued recall
Victoria Sponge cake recipe	Structured (2)	Change perspective or Cued recall
Changing oil in a car	Unstructured (2)	Change perspective or Cued recall
Business conference	Photograph (3)	Change perspective or Cued recall
Beef bourguignon recipe	Structured (3)	Change perspective or Cued recall
Changing guitar strings	Unstructured (3)	Change perspective or Cued recall

Each response was classified by word or phrase type into Noun-type, Verb-type, and Adjective-type. Single words are classified using their dictionary class. Where possible, phrases are classified using the appropriate type, for example ‘Indian classical’ is classified as an adjective-type whilst ‘Indian instruments’ is classified as noun-type. Longer phrases were sometimes given as responses, and these are classified as ‘statements’ or ‘questions’ depending on the content. Each response was also categorised into one of seven categories (People, Activity, Location, Materials, Event, Equipment or Other) to enable the distribution of topics to be examined. These categories were based upon common topics used for metadata in our previous study (Willoughby et al., 2014). Themes were also identified using qualitative analysis for each information object. The themes in the *profile condition* are derived from the question categories (Location, People, Equipment etc.) but the responses to each question are also analysed for sub-themes. For example, ‘appearance’, ‘behaviour’, ‘feelings’, or ‘roles’ might be themes describing responses when subjects are asked what words they would use to describe People for a particular information object.

Results

The first notable difference between the different recall conditions are the number of phrases that each condition generated. The *profile condition*, with an average of 16.8 words and phrases per participant, produced more than three times as many as the *audience condition*, at only 5.1 per participant, with an average of 6.3 produced in the *free recall condition*. In the *free recall condition* the average number of words and phrases did not vary greatly by information type, but in both the *audience* and *profile* conditions there was a slightly higher average for the photograph compared to both text types, suggesting it is easier to generate metadata terms for photographs.

Word Type

Asking the subjects for words that come to mind when they look at information is similar to asking a user to 'tag' their data and we therefore expected to see a similar distribution of word-types in the *free recall condition* as seen in our previous study. A similar pattern is seen with high numbers of nouns and a relatively small number of verbs, but the number of adjectives is higher. The *audience* and *profile* conditions are also dominated by noun-type words and phrases. The *audience condition* in general has fewer verbs and adjectives than the *profile condition*, particularly for the photograph and structured text. The *profile condition* has the most consistent distribution of word types across both the images and text, and in general has the highest numbers of verbs. The photographs have the highest number of adjectives across all conditions, except for Instructions 2 in the *profile condition,* where many adjectives are used to describe the oil change, e.g. "dirty", "dangerous", and "difficult". This information object also had fewer verbs in both the *profile* and *audience* conditions, in part because most subjects used the term "oil change" to describe the activity rather than "changing oil". Statements were only produced in the *free recall condition* for the text, particularly in the unstructured text where 75% of the statements describe properties or comments about the material itself. The full breakdown of word-types for each condition and information object can be seen in Figure 1.

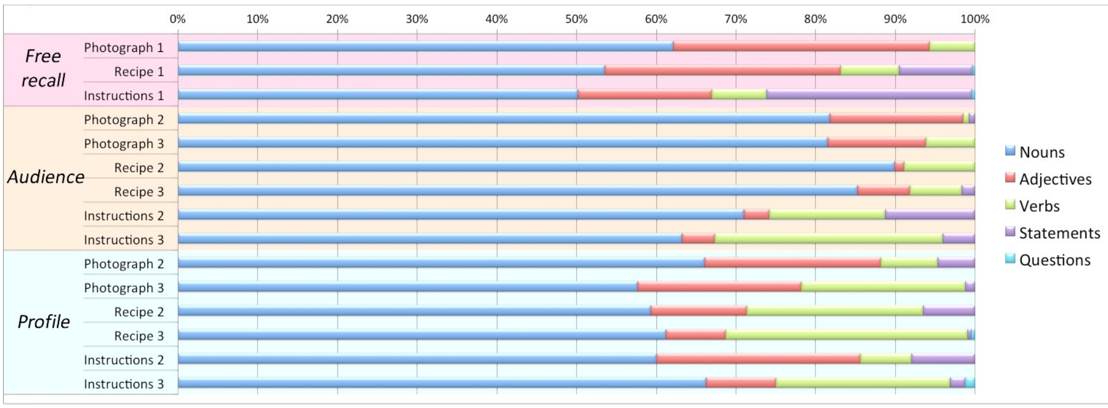


Figure 1. Word types for each condition and information object.

Topics

Each response was categorised according to topic into one of the following categories: People, Activity, Location, Materials, Event, Equipment or Other. The topic classification exercise demonstrated that each information type generates a different pattern of topics based upon the content of the material in the *free recall* and *audience* conditions. People are important aspect of photographs, with the distribution of other topics depending on the content. For example, terms describing the instruments – classified as Equipment – are common for the Indian band photograph and terms describing the venue and furnishings – classified as Location – are common for the business conference photograph. The recipes are dominated by terms classified as Materials describing the ingredients and food. In the *free recall condition* the unstructured text is dominated by terms classified as Other describing the text itself, for example, "British English", "complex", and "diagrams needed!!!". In the *audience* *condition*, the unstructured text is dominated by terms classified as Activity describing the activities in the text. In contrast, the *profile condition* provides a more consistent balance of topics across all the information types, as was expected as a result of asking directly for responses matching these topics. However, the topics that have not had cues provided for them, Events and Materials, are significantly under represented in these responses. This is particularly noticeable for the recipes, where very few responses in the *profile condition* relate to the ingredients in the recipe. The full breakdown of topics for each condition and information object can be seen in Figure 2.

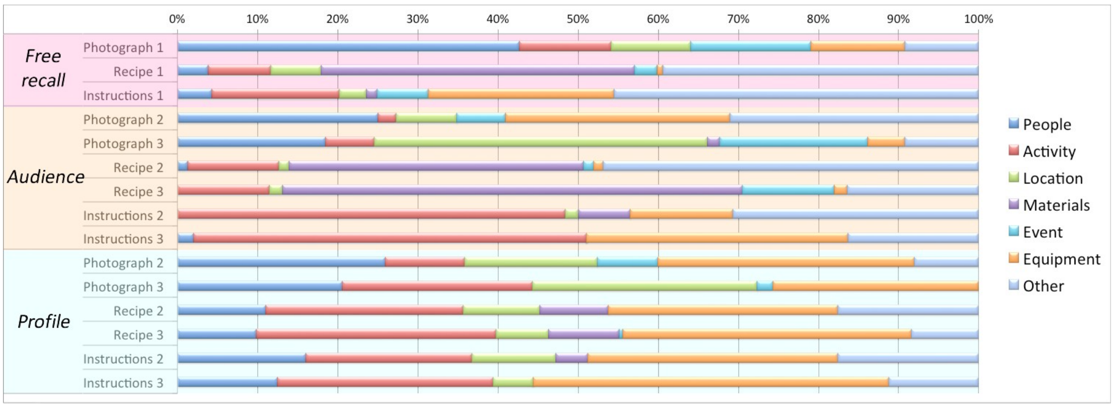


Figure 2. Topics for each condition and information object.

Themes

The responses for each information object were analysed to identify themes. To a large extent the themes match the content of the materials themselves and are similar for each information type across the *audience* and *free recall* conditions. In the *profile condition* the majority of terms relate to the questions that were asked and therefore represent the major themes in this condition. Table 2 shows the major themes for each information object in the *free recall* and *audience* conditions.

Table 2. Information types and recall conditions used in the survey.

Condition	Information Object	Major Themes	
Free recall	Photograph 1	Function of the event	Equipment
		People: activities, attitudes, appearance	
	Recipe 1	Food/Ingredients	Instructions / Recipe
		Taste	Personal
	Instruction 1	Fixing the puncture-activities	Instructions
		Fixing the puncture-equipment	Personal

Condition	Information Object	Major Themes	
Audience (change perspective)	Photograph 2	Function of the event People	Instruments
	Photograph 3	Function of the event Location	People
	Recipe 2	Recipe Activities Food/Ingredients	Event Search suggestions
	Recipe 3	Recipe Ingredients	Event Origin of dish
	Instructions 2	Activity (high-level and steps) Material type (instructions)	Search suggestions
	Instructions 3	Equipment Material type (instructions)	Activity (high-level)

Compared to the other conditions, more terms are used, more objects are identified, and a greater diversity of terms is provided in the *profile condition*. This greater diversity of terms included more descriptive terms, including personal opinions or guesses about aspects of the content. The *profile condition* also includes terms that describe information that is not visible or mentioned in the material, for example, equipment or locations that are not mentioned in the text. These external elements and guesses are common responses for the People and Location questions for the structured and unstructured text because these are elements that are not described in the text. Although some subjects indicated confusion about the question, more than 75% provided responses for these questions that included personal terms describing people and environments related to but not present in the content. The *profile condition* is missing significant themes seen in the other conditions, such as the event for the photographs and the ingredients in the recipes, as mentioned before.

Personal opinions and guess responses, such as guesses about the type of music, origin of the band, location and roles of the people in the photographs, are seen in the *audience condition* but they are less common. Uniquely, the responses in the *audience condition* include ‘search suggestions’, such as “good food (website)”, “halfords car manual”, and “site:youtube.com how to change oil on car”. Personal opinions and external elements are even more common in the *free recall condition*, making up a large percentage of the responses. For example, in this condition subjects provided responses on the perceived feelings and activities of people in the photograph; the taste, smell, and appeal of the food in the recipe; feelings about the activity; suggested improvements; and personal experiences or an imagined situation (e.g. “glass”, “rain”, “dirty hands”, “teaching my child to repair”) for the instructions.

See Appendix 3 and Appendix 4 for visualisations of the *free recall*, *audience* and *profile* results.

 doi:10.2218/ijdc.v10i1.343

Discussion

We expected to see different responses generated for the different information types – text or image – when we presented subjects with cues or asked them to change perspective compared to using *free recall*. We also expected that different word types would be generated if the information type was text-based or a photograph. The results show that both of these expectations were correct, although the results of the word type analysis were more complex than originally anticipated. The request for "what words or phrases come to mind for use in a search engine?" in the *audience condition* generates terms that would be useful for searching for similar material, although the terms most commonly used could have been derived from a suitable title for the materials. The *profile condition* generated significantly more terms with a greater breadth and diversity for the topics requested than seen in the other conditions. On the negative side, however, important context about the content was missing because the relevant cues were not provided, such as the missing ingredients in the recipes. This matches findings from other studies where the inflexibility of templates or questionnaires restricts what an author enters (McGlade et al., 2012) and information that is not specifically asked for is not recorded (Swinglehurst et al., 2012).

To understand whether the responses in the *profile condition* are useful for curation, we assessed whether the terms produced are relevant to each question category. In general, the most relevant responses were for the Equipment, Activity and Other objects categories. The People and Location topic sections had the least relevant responses because of the confusion around these categories, but the vast majority of responses were in fact relevant, although some slightly unexpected terms for Location were produced, such as kitchen objects and parts of a guitar or car.

An unexpected finding during the analysis of the results of the study was that many of the responses described information not included in the photograph or text. For example, in the *free recall condition*, the terms "dinner" and "wifi" are used to describe the photograph, "naan bread" and "lime" are not in the recipe instructions, and "stuck at roadside" and "glass" are not mentioned in the unstructured text. Other responses include terms that describe personal opinions or personal experience, for example, in the *free recall condition*, terms such as "bored" and "disinterested" describe feelings of subjects in the photograph; "delicious", "too hot" and "my child won't eat this" describe opinions about the recipe, and "tedious" and "I couldn't do this" describe opinions about the activity in the unstructured text. Examples of external elements and personal responses are seen in all of the conditions, but are highest in the *free recall* and *profile* conditions. The *profile condition* may promote the inclusion of information not mentioned in text because questions are asked about elements not mentioned, leading to guesswork by the participants, and providing cues prompts recall of previous experiences. More external elements are included for the photographs in the *audience condition* suggesting that participants are willing to make guesses about locations, events, activities, objects and people that they can see, or perhaps the visual content of the photographs itself cues more personal recall. Despite the inclusion of external elements, terms have been produced in all the conditions that could be valuable for curation, and it could be considered that the unexpected terms provide useful additional context.

Another unexpected finding was that many of the responses described properties, formatting, or opinions about the information objects themselves, for example "verbose", "familiar format" and "needs bullets". In general, the number of responses describing the information objects increased with the increasing complexity of the material. Unstructured materials generated the most terms, particularly in the *free recall condition*, with nearly three times as many responses about the information itself when compared to the structured text. Responses included comments about the structure, complexity and language of the text, as well as personal opinions about improving both format and content (e.g. “methodical”, “well-laid out”, “wrong order”). A large number of responses disagree with the instructions and are contradictory, for example “nobody patches things anymore” and “patches are fine”.

In order to see how the results relate to what terms an expert would use to classify the different information sources, an experienced librarian was asked to provide an expert opinion about the terms they would use if they were asked to curate the information objects used in the study. Usual practice for the librarian would be to select tags/categories from a prescribed list (effectively a controlled vocabulary). Appendix 5 shows the descriptive categories or label assigned by the librarian to each of the information types in the study compared to the results from the different cues in the study. The results show that none of the different capture conditions closely matches all the terms that the expert would use, suggesting that including a specialist to add metadata at a later stage would be very beneficial for consistency in curation, although the metadata generated by the non-specialists using these methods is much richer in contextual detail.

Capturing Experiment Metadata in Practice

The results of the study suggested that using these different approaches to gather metadata was worth further investigation. Additional research has included using these different methods to capture experiment metadata in two studies with chemistry researchers. One study included 60 academics, divided into small groups of staff and research students, completing an experiment using Alka-Seltzer to power Lego cars. Each team was asked to record their experiments on one of three different templates, each employing a different way to capture metadata: ‘free recall’, ‘cued recall’, and ‘keywords’ based on the conditions in this study. The results of the free recall were similar to those seen reported in this study, with 62% nouns, 15% adjectives and the remainder equal numbers of questions and statements. Themes included components of the experiment (e.g. “Lego”, “car”, “aerodynamics”), conditions (e.g. “weight of car”, “how much water”, “shape and design”), results (e.g. “Passenger survived in second go”, “We are happy with the final result of 3.2 meter”), and personal elements (e.g. “childhood memories”, “fireworks”, “fun”, “exciting”).

The cued recall asked questions about participants in the experiment, chemicals and materials, instruments or equipment, location, activities and techniques, and any other things that it might be useful to know. All of the groups provided useful and relevant responses to the majority of questions, including the other information question, where 40% responded and provided useful information. Naming of the chemicals used was variable, with a mixture of specific chemical names, formulae and common names, depending on the team. One team chose “H2O + aspirin”, as though representing the reaction. There was some confusion between Materials and Equipment, with groups including the same terms as each other but in different sections. An unexpected response from all the teams was including equipment they used for recording the experiment. Location had different levels of specificity, from a precise location within a room and full address, to only a room name. One group provided a useful description of the environment: “Regular wooden floor, relatively smooth”. Responses to the request for ‘keywords’ were more restricted and produced similar results to the *audience condition* in the main study, with a small number of responses like the title, main components (e.g. “Lego”, “rocket”), and terms related to the Alka-Seltzer reaction.

A more formal study involved 20 undergraduate students in an Organic Chemistry Summer School, who were asked to create a write-up of their experiments after the event using different templates. Each template included a free recall section, where the students were asked to write down up to ten things that came to mind when they thought about the experiment, and one of the templates included ‘profile-style’ questions asking about chemicals or materials, equipment or instruments, location, activities or techniques, and other useful information that they had used in their experiment. 361 free recall responses were collected from three chemistry experiments. 67% of the responses were statements rather than words or phrases. The responses were consistent for each individual, with some choosing mostly personal responses, and others providing a mixture of chemicals, equipment and technique. Overall activities, including techniques, procedures and conditions of the experiment, were the dominant theme of the responses at 40%; followed by materials and details about the reaction at 25%; personal comments, such as the students’ feelings about the experiment, what they learnt and personal worries or problems accounted for 14% of the responses. Less common responses for the free recall were Equipment (9%), Safety (6%), and results of the experiment (6%). Only one response mentioned the location of the experiment.

The results of the cued recall were similar to the Lego cars study. All students provided relevant answers for all the profile questions except for ‘Other useful information’, where 80% provided a response, but the majority were only ideas of what could be recorded and not useful information. The ideas mirrored exactly the examples provided in the template: sample identifiers, safety information, and settings for the analysis. The useful responses included specific safety information, strongly suggesting the examples themselves acted as a cue (Crystal and Greenberg, 2005).

Terms used for the responses are consistent between subjects for all of the specific profile questions. The majority of words and phrases produced were noun-type, but the activities and techniques question generated a mixture of nouns and verbs, including nominalized activities such as “extraction”, “rotary evaporation” and “vacuum desiccation”. The materials question generated a range of chemicals with a mixture of chemical names, common names and molecular formula, illustrating the problems of consistency with the myriad different ways of representing chemical structures. Data mining could be used to extract chemical names from the text of the student’s reports, for example using ChemSpider (Day, 2013). However, our results suggest that many names may be missed or incorrectly identified because of unrecognised abbreviations, typos and ambiguous values. Metadata collected by asking what materials were involved in the experiment would actually be more accurate. Tools such as ChemSpider could be used to convert responses using the InChi standard or the SMILES specification for the purposes of metadata use and linking. In the reports, only 50% of the students included any mention of equipment at all, whereas the profile question generated an average of 5.8 responses, again indicating that data mining would have limited value. For Location the majority provided the specific laboratory, with some including bench location and university address, while others used less specific terms such as “in labs” and “fume cupboard”. Giving an example Location would help generate consistent responses.

Conclusions

Different metadata terms are useful for the purposes of search and providing context (Greenberg, 2001). Using free recall, cues, and changing perspective are ways of generating terms for a specific purpose, such as search terms for a particular audience, or to produce terms that are valuable for particular areas of the scientific record, such as describing activities, equipment or location that might not otherwise be captured. The findings of the studies demonstrate that these techniques have both positive and negative consequences. On the one hand, valuable personal knowledge and insights can be captured to add context, but on the other irrelevant and possibly misleading information may have to be removed as part of a later process. Ensuring that the correct questions are identified is also important to ensure that all the relevant information is captured. When applied to actual scientific experiments, cues are useful for capturing context for the experiment, without generating personal elements that are perhaps less useful, as produced by free recall. Asking for search terms or 'keywords', together with input from an information specialist, may be helpful for providing high level metadata that is useful for search purposes.

Acknowledgements

The authors are very grateful to Fiona Bell for conducting the formal classification exercise reported above. The research reported in this paper has been made possible by funding from the RCUK e-Science programme (EPSRC grant GR/R67729, EP/C008863, EP/E502997, EP/G026238, BBSRC BB/D00652X), the HEFCE and JISC Data Management Programme, the University Modernisation (UMF), and most recently the RCUK Digital Economy Theme as part of the IT as a Utility Network+ funding (EPSRC EP/K003569).

References

Bird, C.L., Willoughby, C., Coles, S. J., & Frey, J.G. (2013a). Data curation issues in the chemical sciences. *Information Standards Quarterly, 25*(3), 4-12. doi:10.3789/isqv25no3.2013.02

Bird, C.L., Willoughby, C., & Frey, J.G. (2013b). Laboratory notebooks in the digital era: the role of ELNs in record keeping for chemistry and other sciences. *Chemical Society Reviews*, *42*(20), 8157–8175. doi:10.1039/c3cs60122f

Borgman, C.L. (2007). *Scholarship in the digital age: Information, infrastructure, and the internet*. Cambridge, MA, MIT Press.

Borgman, C.L. (2008). Data, disciplines, and scholarly publishing. *Learned Publishing, 21*(1), 29–38. doi:10.1087/095315108X254476

Borgman, C.L. (2010). *Research data: Who will share what, with whom, when, and why?* Paper presented at the China-North America Library Conference, Beijing. Retrieved from http://works.bepress.com/borgman/238

Crystal, A., & Greenberg, J. (2005). Usability of a metadata creation application for resource authors. *Library and Information Science Research, 27*(2), 177–189. doi:10.1016/j.lisr.2005.01.012

Currier, S., Barton, J., O'Beirne, R., & Ryan, B. (2004). Quality assurance for digital learning object repositories: Issues for the metadata creation process, *ALT-J, Research in Learning Technology, 12*(1), 5–20. doi:10.1080/0968776042000211494

Day, A. (2013). Search and insert from ChemSpider in LabTrove and other websites with TinyMCE editors. Retrieved from http://www.chemspider.com/blog/search-and-add-from-chemspiderdirectly-from-labtrove-and-other-blog-based-websites-which-use-tinymce.html

Downing, J., Murray-Rust, P., Tonge, A.P., Morgan, P., Rzepa, H.S., Cotterill, F., & Harvey, M.J. (2008). SPECTRa: The deposition and validation of primary chemistry research data in digital repositories. *Journal of Chemical Information and Modeling, 48*(8), 1571–1581. doi:10.1021/ci7004737

Dudukovic, N.M., Marsh, E.J., & Tversky, B. (2004). Telling a story or telling it straight: the effects of entertaining versus accurate retellings on memory. *Applied Cognitive Psychology, 18*, 125–143. doi:10.1002/acp.953

Frey, J.G. (2008). Curation of laboratory experimental data as part of the overall data lifecycle. *International Journal of Digital Curation, 3*(1), 44–62. doi:10.2218/ijdc.v3i1.41

Gilliland, A.J. (2008). Setting the stage. In Murtha Baca (Ed.), *Introduction to Metadata: Pathways to Digital Information, Online Edition, Version 3.0*. Los Angeles, CA: Getty Information Institute. http://www.getty.edu/research/publications/electronic_publications/intrometadata/setting.html

Greenberg, J. (2001). A quantitative categorical analysis of metadata elements in image-applicable metadata schemas. *Journal of the American Society for Information Science and Technology, 52*(11), 917–924. doi:10.1002/asi.1170

Greenberg, J. (2004). Metadata extraction and harvesting. *Journal of Internet Cataloging, 6*(4), 59–82. doi:10.1300/J141v06n04_05

Greenberg, J., & Robertson, W.D. (2002). Semantic web construction: An inquiry of authors' views on collaborative metadata generation. Paper presented at the International Conference on Dublin Core and Metadata for e-Communities. Italy. Retrieved from http://dcpapers.dublincore.org/pubs/article/view/693

Greenberg, J., Crystal, A., Robertson, W., & Leadem, E. (2003). Iterative design of metadata creation tools for resource authors. Paper presented at the International Conference on Dublin Core and Metadata Applications. North America. Retrieved from http://dcpapers.dublincore.org/pubs/article/view/733

Higham, P.A. & Tam, H. (2005). Generation failure: Estimating metacognition in cued recall. *Journal of Memory and Language 52*, 595–617. doi:10.1016/j.jml.2005.01.015

Jabine, T.B., Straf, M.L., Tanur, J.M., & Tourangeau, R. (Eds.). (1984). *Cognitive aspects of survey methodology: Building a bridge between disciplines*. Washington, DC: National Academy Press.

Jobe, J., & Mingay, D. (1989) .Cognitive research improves questionnaires. *American Journal of Public Health, 79*(8), 1053–1055. doi:10.2105/AJPH.79.8.1053

Marian, V. & Neisser, U. (2000). Language-dependent recall of autobiographical memories. *Journal of Experimental Psychology General, 129*(3), 361–368. doi:10.1037/0096-3445.129.3.361

McGlade, K., Cargo, C., Fogarty, D., Boohan, M., & McMullin. M. (2012). Handwritten undergraduate case reports. *The Clinical Teacher, 9*(2) 112–118. doi:10.1111/j.1743-498X.2011.00494.x

Memon, A., Meissner, C.A., & Fraser, J. (2010). The cognitive interview: A meta-analytic review and study space analysis of the past 25 years. *Psychology, Public Policy, and Law, 16*(4), 340–372. doi:10.1037/a0020518

Milsted, A., Hale, J., Frey, J., & 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

Ryan, B., & Walmsley, S. (2003). Implementing metadata collection: A project's problems and solutions. *Learning technology, 5*(1). Retrieved from http://www.ieeetclt.org/issues/january2003/index.html#3

Schwarz, N. (2007). Cognitive aspects of survey methodology. *Applied Cognitive Psychology, 21*(2), 277–287. doi:10.1002/acp.1340

Swinglehurst, D., Greenhalgh, T., & Roberts, C. (2012). Computer templates in chronic disease management : Ethnographic case study in general practice. *BMJ Open, 2*(6), e001754. doi:10.1136/bmjopen-2012-001754

Tourangeau, R., Conrad, F.G., Couper, M.P., & Ye, C. (2014). The effects of providing examples in survey questions. *Public Opinion Quarterly, 78*(1), 100–125. doi:10.1093/poq/nft083

Tourangeau, R., Rips, L.J., & Rasinski, K. (2000). *The psychology of survey response*. Cambridge University Press.

Tversky, B., & Marsh, E.J. (2000). Biased retellings of events yield biased memories *Cognitive psychology, 40*(1), 1–38. doi:10.1006/cogp.1999.0720

Willoughby, C., Bird, C.L., Coles, S.J., & Frey, J.G. (2014). Creating context for the experiment record. User-defined metadata: Investigations into metadata usage in the LabTrove ELN. *Journal of Chemical Information and Modeling, 54*(12), 3268–3283. doi:10.1021/ci500469f

Zeng, M.L., & Qin, J. (2008). *Metadata.* New York: Neal-Schuman.

List of other paper contributions

Frey, J. G., Coles, S. J., Bird, C. L., & Willoughby, C. (2015). Collection, Curation, Citation at Source: Publication@Source 10 years on. *International Journal of Digital Curation* (Accepted).

Frey, J. G., Coles, S., Bird, C. L., Brocklesby, W. S., Badiola, K. A., Williamson, A. E., Cronshaw, J.R., Robins, M., Roberston, M.N., Todd, M.H., Chapman, R.T., Fisher, A., Grossel, M., Casson, J., Hibbert, D.B., Gloria, D., Mapp, L.K., Matthews, B., Milsted, A., Minns, R.S., Mueller, K.T., Parkinson, T., Quinnell, R., Robinson, J.S., Springate, E., Tizzard, G.J., Willoughby, C., Yang, E., Ylioja, P.M., Knight, N.J., Marazzi, L. (2015) Experiences with a researcher-centric ELN. *Chemical Science*. **6**, 1614-1629. doi:10.1039/C4SC02128B

Bird, C.L., Willoughby, C., Coles, S.J., & Frey, J.G. (2013) Data Curation Issues in the Chemical Sciences. *Information Standards Quarterly*, Fall 2013, 25(3): 4-12. http://dx.doi.org/10.3789/isqv25no3.2013.02

Bird, C.L., Willoughby, C. & Frey, J.G. (2013) Laboratory notebooks in the digital era: the role of ELNs in record keeping for chemistry and other sciences. *Chemical Society Reviews*. 42, 8157-8175 DOI: 10.1039/C3CS60122F

Frey, J.G., Bird, C.L., Willoughby, C. (2013) Smart Meetings: Experimenting with Space. In *Smart Innovation, Systems and Technologies, Vol. 18: Innovation through Knowledge Transfer 2012*, Howlett, R.J.; Gabrys, B.; Musial-Gabrys, K.; Roach, J. (Eds.), Springer

Frey, J.G., Bird, C.L., Willoughby, C. (2013) Human Aspects of Smart Spaces for Knowledge Transfer. In *Smart Innovation, Systems and Technologies, Vol. 18: Innovation through Knowledge Transfer 2012*, Howlett, R.J.; Gabrys, B.; Musial-Gabrys, K.; Roach, J. (Eds.), Springer

Frey, J.G., Bird, C.L., Willoughby, C. (2013) Smart Meeting Spaces for Knowledge Transfer. In *Smart Innovation, Systems and Technologies, Vol. 18: Innovation through Knowledge Transfer 2012*, Howlett, R.J.; Gabrys, B.; Musial-Gabrys, K.; Roach, J. (Eds.), Springer

Appendices

The appendices for this thesis can be found as digital files on the attached CD in the folders described below.

Appendix A Using structured templates

A.1 Study 1 – Organic Chemistry Summer School – Paper-based

Files related to Study 1, as detailed in Chapter 5: Asking the right questions: Effects of Using Structured Templates for Recalling Chemistry Experiments.

A.1.1 NoTemplateCondition.pdf

Paper-based questionnaire for the No Template condition used in Study 1, as described in 5.3.

A.1.2 TemplateCondition.pdf

Paper-based questionnaire for the Template condition used in Study 1, as described in section 5.3.

A.1.3 OCSS1 - Results.xls

Excel file with results from the different conditions produced by Study 1, as described in 5.3.2.

A.2 Study 2 – Organic Chemistry Summer School – Computer-based

Files related to Study 2, as detailed in Chapter 5: Asking the right questions: Effects of Using Structured Templates for Recalling Chemistry Experiments.

A.2.1 NoTemplateCondition.pdf

Computer-based questionnaire for the No Template condition used in Study 2, as described in 5.4.

A.2.2 TitlesTemplateCondition.pdf

Computer-based questionnaire for the Titles Template condition used in Study 2, as described in 5.4.

A.2.3 ProfileTemplateCondition.pdf

Computer-based questionnaire for the Template condition used in Study 2, as described in 5.4.

A.2.4 OCSS2 - Results.xls

Excel file with results from the different conditions produced by Study 2, as described in 5.4.2.

A.3 Study 3 – Lego Rocket cars

Files related to Study 3, as detailed in Chapter 5: Asking the right questions: Effects of Using Structured Templates for Recalling Chemistry Experiments.

A.3.1 NoTemplateCondition.pdf

Computer-based questionnaire for the No Template condition used in Study 3, as described in 5.5.

A.3.2 TitlesTemplateCondition.pdf

Computer-based questionnaire for the Titles Template condition used in Study 3, as described in 5.5.

A.3.3 ProfileTemplateCondition.pdf

Computer-based questionnaire for the Template condition used in Study 3, as described in 5.5.

A.3.4 LegoRocketCarResults.xls

Excel file with results from the different conditions produced by Study 3, as described in 5.5.2.

Appendix B Bath User Studies

B.1 Bath User Studies.pdf

Participant instructions for the two user studies/experiments run at Bath University with the LabTrove ELN, as described in 6.2.

Appendix C Online Metadata Study

C.1 Materials used.pdf

The photographs, structured, and unstructured information objects used in the questionnaires for the online metadata study, as described in 7.3.

C.2 Questions used.pdf

The questions used in the questionnaires for each of the conditions for the online metadata study, as described in 7.3.

C.3 Visualisation of results.pdf

Visualisations of the results from the online metadata study for Free recall, and comparisons of the Audience and Profile conditions, as described in 7.3.2.

C.4 Formal classifications.pdf

Formal classifications of the information objects described in C.1 by an information specialist, as described in 7.3.

C.5 Online_Metadata_Results.xls

Excel file containing analysis of the responses and a summary of the results from the online metadata study, as described in 7.3.2.

Appendix D Notelus Code

D.1 Notelus.zip

Code for the Notelus ELN iOS application, as described in Chapter 9: Notelus: An Exemplar Mobile ELN.

D.2 Plan Buddy for Notelus.zip

Code for the Plan Buddy for Notelus iOS application for generating XML plans for use with Notelus, as described in Chapter 9: Notelus: An Exemplar Mobile ELN.

List of References

Allen, B.P., and Lindsay, D.S. (1998). Amalgamations of memories: intrusion of information from one event into reports of another *Applied Cognitive Psychology 12*, 277-285.

Ames, M. & Naaman, M. (2007). Why We Tag: Motivations for Annotation in Mobile and Online Media. *In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*; Rosson, M. B., Gilmore, D., Eds.; Association for Computing Machinery: San Jose, CA. pp 971−980.

Amorim, R. C., Castro, J. A., Rocha, J., & Ribeiro, C. (2014). LabTablet : semantic metadata collection on a multi-domain laboratory notebook. In S. Closs, R. Studer, E. Garoufallou, & M. A. Sicilia (Eds.), *Metadata and Semantics Research*: 8th Research Conference, MTSR 2014, Karlsruhe, Germany, November 27-29, 2014. Proceedings (pp. 193–205). Switzerland: Springer International Publishing.

Ancona, M., Dodero, G., & Gianuzzi, V. (1999). RAMSES: a mobile computing system for field archaeology. In (H.-W. Gellersen, Ed.) *Handheld and Ubiquitous Computing*. (Lecture Notes in Computer Science. 1707, pp. 222–233). Springer Berlin Heidelberg.

Anderson, R. C. and Pichert, J.W. (1978). Recall of previously unrecallable information following a shift in perspective *Journal of Verbal Learning and Verbal Behavior 17*, 1-12.

Anderson, S.J. and Conway, M.A. (1993). Investigating the structure of autobiographical memories *Journal of Experimental Psychology: Learning, Memory, and Cognition 19*(5), 1178-1196.

Arnstein, L., Franza, R., Borriello, G., & Consolvo, S. (2002). Labscape: a smart environment for the cell biology laboratory. *IEEE Pervasive Computing*, *1*(3), 13–21. doi:10.1109/MPRV.2002.1037717

Badiola, K.A., Bird, C.L., Brocklesby, W.S., Casson, J., Chapman, R.T. Coles, S.J., Cronshaw, J.R., Fisher, A., Frey, J.G., Gloria, D., Grossel, M.C., Hibbert, D.B., Knight, N., Mapp, L.K., Marazzi, L., Matthews, B., Milsted, A., Minns, R.S., Mueller, K.T., Murphy, K., Parkinson, T., Quinnell, R., Robinson, J.S., Robertson, M.N., Robins, M., Springate, E., Tizzard, G., Todd, M.H., Williamson, A.E., Willoughby, C., Yang, E., & Ylioja, P.M. (2015). Experiences with a researcher-centric ELN. *Chemical Science*. doi:10.1039/C4SC02128B

Barsalou, L.W. and Sewell, D. (1985). Contrasting the representation of scripts and categories. *Journal of Memory and Language. 24*, 646-665.

Beagrie, N. (2006). Digital curation for science, digital libraries, and individuals. *International Journal of Digital Curation. 1*(1), 3–16.

Bell, G., & Gemmell, J. (2007). A digital life. *Scientific American. 296*(3), 58–65.

Bergman, E.T., and Roediger, H.L. (1999). Can Bartlett's repeated reproduction experiments be replicated? *Memory & Cognition. 27*(6), 937-947.

Berry, C. J., Henson, R. N. A., & Shanks, D. R. (2006). On the relationship between repetition priming and recognition memory: Insights from a computational model. *Journal of Memory and Language. 55*, 515–533.

Bird, C. L., & Frey, J. G. (2013). Chemical information matters: an e-Research perspective on information and data sharing in the chemical sciences. *Chemical Society Reviews*, 42(16), 6754–76. doi:10.1039/c3cs60050e

Bird, C. Willoughby, C., and Frey, J. (2013a). Laboratory notebooks in the digital era: Record keeping in chemical and other science laboratories *Chemical Society Reviews. 42*, 8157-8175. DOI: 10.1039/c3cs60122f

Bird, C. L., Willoughby, C., Coles, S. J., & Frey, J. G. (2013b). Data Curation Issues in the Chemical Sciences. *Information Standards Quarterly. 25*(3), 4–12. doi:10.3789/isqv25no3.2013.02

Bly, S. (1998). Special section on collaboratories. *Interactions*, 5(3), 31. doi:10.1145/275269.275273

Borgman, C.L. (2007). *Scholarship in the Digital Age: Information, Infrastructure, and the Internet*. Cambridge, MA, MIT Press.

Borgman, C.L. (2008). Data, disciplines, and scholarly publishing. *Learned. Publishing. 21*(1), 29-38.

Borgman, C. L. (2010). Research Data: Who will share what, with whom, when, and why? *The RatSWD Working Paper Series. 161*, 1–21. Retrieved from http://core.kmi.open.ac.uk/download/pdf/6530014.pdf

Borgman, C. L. (2012). The conundrum of sharing research data. *Journal of the American Society for Information Science and Technology*. *63*(6), 1059–1078. doi:10.1002/asi.22634

Borman, S. (1994). Electronic Laboratory Notebooks may revolutionize research record keeping. *Chemical & Engineering News*. *72*(21), 10–20.

Boulton, G., Campbell, P., Collins, B., Elias, P., Hall, W., Laurie, G., O'Neill, O., Rawlins, M., Thornton, J., Vallance P., and Walport, M. (2012) *Science as an open enterprise*. The Royal Society Science Policy Centre report. 02/12. Retrieved from: http://royalsociety.org/uploadedFiles/Royal_Society_Content/policy/projects/sape/2012-06-20-SAOE.pdf (Accessed 11 July 2013).

Bower, G.H., Black, J.B. and Turner, T.J. (1979). Scripts in memory for text. *Cognitive Psychology 11*, 177-220.

Brandl, P., Richter, C., & Haller, M. (2010). NiCEBook - supporting natural note taking. In *Proceedings of the 28th international conference on Human factors in computing systems - CHI '10* (pp. 599–608). New York, New York, USA: ACM Press. doi:10.1145/1753326.1753417

Bransford, J, and Johnson, M. (1972). Contextual prerequisites for understanding: Some investigations of comprehension and recall. *Journal of Verbal Learning and Verbal Behavior. 11*, 717-726.

Brewer, W. F. and Treyens, J.C. (1981). Role of schemata in memory for places *Cognitive Psychology. 13*, 207-230.

Bromme, R., Rambow, R., & Nückles, M. (2001). Expertise and estimating what other people know: the influence of professional experience and type of knowledge. *Journal of Experimental Psychology*. Applied, 7, 317–330. doi:10.1037/1076-898X.7.4.317

Bruce, S. (2002). A Look at the State of Electronic Lab Notebook Technology. *Scientific Computing*. Retrieved from http://www.scientificcomputing.com/articles/2002/12/look-state-electronic-lab-notebook-technology

Cardenas, M. (2014). Electronic Laboratory Notebooks versus Paper Laboratory Notebooks: A Comparison of Undergraduate Experimental Engineering Laboratory Submissions. In L. Gossage, P. Marayong, & M. Dyrud (Eds.), *Proceedings of the 2014 American Society for Engineering Education Zone IV Conference* (pp. 96–107). Long Beach, California. Retrieved from http://works.bepress.com/mary_cardenas/14/

Carpenter, S. (2012). Must a Paper Trail Be Paper? *Science Careers*. http://sciencecareers.sciencemag.org/career_magazine/previous_issues/articles/2012_09_14/caredit.a1200103 (Accessed 9th December 2014)

George Chin, J., & Lansing, C. S. (2004). Capturing and supporting contexts for scientific data sharing via the biological sciences collaboratory. *CSCW '04: Proceedings of the 2004 ACM Conference on Computer Supported Cooperative Work*. *6*(3), 409–418. doi:10.1145/1031607.1031677

Chin, G., Myers, J., & Hoyt, D. (2002). Social Networks in the Virtual Science Laboratory. *Communications of the ACM*. *45*(8), 87–92.

Cianciarulo, D., & Guerra, F. (2007). Digital recording in archaeological excavation using Tablet PC. *Proceedings XXI CIPA International Symposium*. 1–6.

Coleman, A. M. (2009). *Dictionary of psychology*. Oxford University Press, Oxford.

Coles, S.J., Kilpin, K.J., Willoughby, C., & Frey, J.G. (2015). Dial a Molecule iLabber trial. (In preparation, April 2015).

Coles, S. J., Frey, J. G., Bird, C. L., Whitby, R. J., & Day, A. E. (2013). First steps towards semantic descriptions of electronic laboratory notebook records. *Journal of Cheminformatics*. *5*:*52*. http://www.jcheminf.com/content/5/1/52

Collins, A.M. and Loftus, E.F. (1975). A spreading-activation theory of semantic processing. *Psychological Review 82*(6), 407-428.

Conway, M. A. (2009). Episodic memories. *Neuropsychologia*. *47*, 2305–2313. doi:10.1016/j.neuropsychologia.2009.02.003

Conway, M.A., and Pleydell-Pearce, C.W. (2000). The construction of autobiographical memories in the self-memory system. *Psychological Review, 107*(2), 261-288.

Cooke, R., & Schraefel, m. c. (2004). Signature Flip and Clip: Virtually Flipping and Dog Earing Pages in a Digital Lab Book. In *User Interface Software Technology (UIST)* (pp. 1–4). Santa Fe, New Mexico. Retrieved from http://eprints.soton.ac.uk/259251/

Coppin, P. & Hockema, S. (2009). Learning from the information workspace of an information professional with dyslexia and ADHD. In *2009 IEEE Toronto International Conference Science and Technology for Humanity* (TIC-STH) (pp. 801–807). IEEE. doi:10.1109/TIC-STH.2009.5444390

Cosentino, S., Chute, D., Libon, D., Moore, P., & Grossman, M. (2006). How does the brain support script comprehension? A study of executive processes and semantic knowledge in dementia *Neuropsychology 20*(3), 307-18.

Committee on Science, Engineering, and Public Policy, (COSEPUP) (2009). *On Being a Scientist: A Guide to Responsible Conduct in Research: Third Edition*. The National Academies Press. http://www.nap.edu/catalog/12192/on-being-a-scientist-a-guide-to-responsible-conduct-in.

Crestani, F. (1997). Application of Spreading Activation Techniques in Information Retrieval. *Artificial Intelligence Review. 11*(6) 453-482

Crystal, A. & Greenberg, J. (2005). Usability of metadata creation application for resource authors. *Library & Information Science Research*. 27(2), 177-189.

Currier, S., Barton, J., O’Beirne, R., Ryan, B. (2004). Quality assurance for digital learning object repositories: issues for the metadata creation process. *ALT-J, Research in Learning Technology.* 12(1), pp. 5-20.

Dahlstrom, E. & Bichsel, J. (2014). *ECAR Study of Undergraduate Students and Information Technology, 2014*. EDUCAUSE Center for Applied Research, Louisville, CO. Retrieved from http://www.educause.edu/library/resources/2014-student-and-faculty-technology-research-studies

Dando, C.J., Wilcock, R., Milne, R., & Henry, L. (2009). A modified cognitive interview procedure for frontline police investigators. *Applied Cognitive Psychology. 23*, 698-716.

Day, A. (2013). *Search and Insert from ChemSpider in LabTrove and Other Websites with TinyMCE Editors*. Retrieved from http://www.chemspider.com/blog/search -and-add-from-chemspiderdirectly-from-labtrove-and-other-blog-based-websites -which-use-tinymce.html (accessed July 11, 2013).

De Roure, D. & Frey, J. (2007). Three Perspectives on Collaborative Knowledge Acquisition in e-Science. In, *Workshop on Semantic Web for Collaborative Knowledge Acquisition (SWeCKa) 2007, Hyderabad, India*

De Roure, D., Goble, C., & Stevens, R. (2009). The Design and Realisation of the myExperiment Virtual Research Environment for Social Sharing of Workflows. *Future Generation Computer Systems. 25*, 561-567. doi:10.1016/j.future.2008.06.010

Desimone, L. M., & Le Floch, K. C. (2004). Are we asking the right questions? Using cognitive interviews to improve surveys in education research. *Educational Evaluation and Policy Analysis. 26*(1), 1–22. doi:10.3102/01623737026001001

Downing, J., Murray-Rust, P., Tonge, A. P., Morgan, P., Rzepa, H. S., Cotterill, F., Day, N., & Harvey, M. J. (2008). SPECTRa: the deposition and validation of primary chemistry research data in digital repositories. *Journal of Chemical Information and Modeling*. 48(8), 1571–81. doi:10.1021/ci7004737

Drake, D. J. (2007). ELN implementation challenges. *Drug Discovery Today*. *12*(15-16), 647–9. doi:10.1016/j.drudis.2007.06.010

Echtler, F., Häussler, M., & Klinker, G. (2010). BioTISCH. In *Proceedings of the 28th of the international conference extended abstracts on Human factors in computing systems - CHI EA '10* (pp. 3439–3444). New York, New York, USA: ACM Press. doi:10.1145/1753846.1753998

Eisenberg, A. (1982). "Keeping a Laboratory Notebook." *Journal of Chemical Education, 59*:12, 1045–1046.

Elliott, M.H. (2004). It's not about the paper. *Scientific Computing & Instrumentation.* http://www.scientificcomputing.com/articles/2004/08/its-not-about-paper?

Elliott, M. H. (2012). Tablets and ELN: A Honeymoon. *Scientific Computing. 29*, 9–12.

Fakas, G. J., Nguyen, A. V., & Gillet, D. (2005). *Computer Supported Cooperative Work (CSCW)*, *14*, 189–216.

Fear, K. (2011). "You made it, you take care of it": Data Management as Personal Information Management. *International Journal of Digital Curation*, *6*(2), 53–77.

Fisher, R.P., Geiselman, R.E., and Amador, M. (1989). Field test of the Cognitive Interview: enhancing the recollection of actual victims and witnesses of crime *Journal of Applied Psychology 74*(5), 722-727.

Fountain, A., Hall, W., Heath, I., & Davis, H. (1990). Microcosm: An Open Model for Hypermedia with Dynamic Linking. Pages 298–311 of: Rizk, A., Streitz, N., & Andre, J. (eds), *Hypertext: Concepts, Systems and Applications, Proceedings of the Hypertext '90 Conference*, INRIA, France. ACM Press.

Frey, J. (2008). Curation of Laboratory Experimental Data as Part of the Overall Data Lifecycle *The International Journal of Digital Curation. 1*(3). DOI: 10.2218/ijdc.v3i1. http://ijdc.net/index.php/ijdc/article/view/62/41 (Accessed 11 July 2013).

Frey, J. G., Coles, S. J., Bird, C. L., & Willoughby, C. (2015). Collection, Curation, Citation at Source: Publication@Source 10 years on. *International Journal of Digital Curation* (Accepted).

Frey, J.G., Hughes, G.V., Mill, H.R., schraefel, mc, Smith, G.M. and De Roure, D. (2004). Less is More: Lightweight Ontologies and User Interfaces for Smart Labs. *Proceedings of the UK e- Science All Hands Meeting, Nottingham, pp 8, EPSRC 2004, ISBN 1-90 4425-21-6.*

Funnell, E. (2001). Evidence for scripts in semantic dementia: Implications for theories of semantic memory. *Cognitive neuropsychology. 18*(4), 323-41.

Geiselman, R.E., Fisher, R.P., MacKinnon, D.P., and Holland, H.L. (1985). Eyewitness memory enhancement in the police interview: cognitive retrieval mnemonics versus hypnosis. *The Journal of applied psychology. 70*, 401-412.

Gemmell, J., Bell, G., Lueder, R., & Francisco, S. (2006). MyLifeBits : A Personal Database for Everything. *Communications of the ACM*. 49(1) 88-95.

Giles, J. (2012). Going paperless: The digital lab. *Nature*, *481*(7382), 430–1. doi:10.1038/481430a

Gilliland, A.J. (2008). Setting the stage. In Murtha Baca (Ed.), *Introduction to Metadata: Pathways to Digital Information, Online Edition, Version 3.0.* Los Angeles, CA: Getty Information Institute. http://www.getty.edu/research/publications/electronic_publications/intrometadata/setting.html

Goddard, N. H., Macneil, R., & Ritchie, J. (2009). eCAT: Online electronic lab notebook for scientific research. *Automated Experimentation, 1*(1), 4. doi:10.1186/1759-4499-1-4

Good Laboratory Practice (GLP). Second Edition (2009). World Health Organization ISBN-13: 9789241547550

Goodger, J. & Worthington, W. (2013). Research Data Management Training for the Whole Project Lifecycle. In *Physics & Astronomy Research (RDMTPA) JISC Final Report. Research Data Toolkit,* University of Hertfordshire. http://research-data-toolkit.herts.ac. uk/document/rdmtpa-final-report/ (accessed July 11, 2013).

Greenberg, J. (2004). Metadata Extraction and Harvesting. *Journal of Internet Cataloging*. 6(4), 59-82. doi:10.1300/J141v06n04_05

Greenberg, J., & Robertson, W. D. (2002). Semantic Web Construction: An Inquiry of Authors' Views on Collaborative Metadata Generation. *Proceedings of the International Conference on Dublin Core and Metadata for e-Communities 2002* (pp. 45-52).

Hammond, T., Hannay, T., Lund, B., & Scott,J. (2005) Social Bookmarking Tools (I): A General Review. *D-Lib Mag*. 11, article 1.

http://www.dlib.org/dlib/april05/hammond/04hammond. html (accessed October 16, 2014).

Harding, L. (2005). History of modern man unravels as German scholar is exposed as fraud. *The Guardian, 19 February 2005.* Guardian News and Media: London, UK. http://www.guardian.co.uk/science/2005/ feb/19/science.sciencenews (accessed on 30 January 2015)

Hayes, M. (2012). *Electronic Lab Notebooks* http://postdocexperience.scienceblog.com/2012/11/19/electronic-lab-notebooks/ (Accessed 9th December 2014)

Hice, R. (2009) Roadmap to a Clear Definition of ELN. *Scientific Computing.* http://www.scientificcomputing.com/blogs/2009/05/roadmap-clear-definition-eln

Hodges, J. & Patterson, K. (2007). Semantic dementia: a unique clinicopathological syndrome *Lancet neurology 6*, 1004-1014.

Higdon, R.; Stewart, E.; Stanberry, L.; Haynes, W.; Choiniere, J.; Montague, E.; Kolker, E. (2014). MOPED enables discoveries through consistently processed proteomics data. *Journal of Proteome Research. 13*, 107−113 DOI: 10.1021/pr400884c.

Hodges, S., Williams, L., & Berry, E. (2006). SenseCam: A retrospective memory aid. In P. Dourish & A. Friday (Eds.), *UbiComp 2006*, LNCS 4206 (pp. 177–193). Springer-Verlag Berlin Heidelberg. Retrieved from http://link.springer.com/chapter/10.1007/11853565_11

Hoffmann, C. (2013). Processes on Paper: Writing Procedures as Non-Material Research Devices. *Science in Context*, *26*(02), 279–303. doi:10.1017/S0269889713000069

Holleman, B. C., & Murre, J. M. J. (2008). Getting from Neuron to Checkmark: Models and Methods in Cognitive Survey Research. *Applied Cognitive Psychology, 22*(5) 709–732.

Holmes, F. L. (1987). Scientific Writing and Scientific Discovery. *Isis, 78*(2), 220–235.

Holmes, F. L. (2004). *Investigative pathways: Patterns and stages in the careers of experimental scientists*. New Haven, CT: Yale University Press.

Hughes, H. Mills, D. De Roure, J. G. Frey, L. Moreau, M. C. Schraefel, G. Smith, & E. Zaluska, (2004). *Organic & Biomolecular Chemistry, 2,* 3284–93.

Ince, D. (2011). The Duke University scandal – what can be done? *Significance*, *8,* 113–115.

Jabine, T.B., Straf, M.L., Tanur, J.M., & Tourangeau, R. (Eds.). (1984). *Cognitive aspects of survey methodology: Building a bridge between disciplines*. Washington, DC: National Academy Press.

Jewell, M., Costanza, E., & Frankland, T. (2012). The Xeros Data Model: Tracking Interpretations of Archaeological Finds*. In Provenance and Annotation of Data and Processes: Lecture Notes in Computer Science Volume 7525*, (pp. 139-151). Springer Berlin Heidelberg.

Jones, P., & Nemeth, C. (2005b). Cognitive Artifacts in Complex Work. In Y. Cai (Ed.), *Ambient intelligence for scientific discovery: Lecture Notes in Artificial Intelligence Volume 3345.* (pp. 152–183). Springer Berlin Heidelberg.

Katifori, A., Vassilakis, C., & Dix, A. (2010). Ontologies and the Brain: Using Spreading Activation through Ontologies to Support Personal Interaction *Cognitive Systems Research, 11*(1) 25–41.

King, A. (2013). Notebooks go digital. *Chemistry World*, (May 2013). Retrieved from http://www.rsc.org/chemistryworld/2013/05/electronic-lab-notebook-review

Klokmose, C. N., & Zander, P. (2010). Rethinking Laboratory Notebooks. In M. Lewkowicz, P. Hassanaly, V. Wulf, & M. Rohde (Eds.), *Proceedings of COOP 2010, Computer Supported Cooperative Work* (pp. 119–140). London: Springer London.

Kohls, C. & Scheiter, K. (2008). The relation between design patterns and schema theory. In *Proceedings of the 15th Conference on Pattern Languages of Programs (PLoP '08).* ACM, New York, NY, USA.

Lass, B. (2011). Implementing Electronic Lab Notebooks Part 3. *Scientific Computing*. Retrieved from:

http://www.scientificcomputing.com/articles/2011/07/implementing-electronic-lab-notebooks-part-3#.UguhpRZrHN4 [Accessed 13 Aug 2013]

Latour, B., & Woolgar, S. (1986). *Laboratory life: The construction of scientific facts.* Princeton: Princeton University Press.

Levelt, W. J. M., Drenth, P., & Noort, E. (Eds.). (2012). *Flawed science: The fraudulent research practices of social psychologist Diederik Stapel.* Tilburg University. http://hdl.handle.net/11858/00-001M-0000-0010-258A-9

Loftus, E, and Palmer, J. (1974). Reconstruction of automobile destruction: An example of the interaction between language and memory *Journal of Verbal Learning and Verbal Behavior 13*, 585-589.

Luff, P., Heath, C., Norrie, M., Signer, B., & Herdman, P. (2004). Only touching the surface. In *Proceedings of the 2004 ACM conference on Computer supported cooperative work - CSCW '04* (Vol. 6, pp. 523–532). New York, USA: ACM Press. doi:10.1145/1031607.1031695

Mackay, W. E., & Pothier, G. (2001). The A-Book : An Augmented Laboratory Notebook for Biologists. ERCIM News, (46), 52–53.

Mackay, W. E., Pothier, G., Letondal, C., Bøegh, K., & Sørensen, H. E. (2002). The Missing Link : Augmenting Biology Laboratory Notebooks. In *Proceedings of the 15th annual ACM symposium on User interface software and technology* (Vol. 4, pp. 41–50). Paris, France: ACM.

MacNeil, R. (2011). Apps are overtaking web sites: Will that happen in scientific research as well? *The electronic lab notebook blog.* http://elnblog.axiope.com/?p=896 [Accessed 12 August 2013].

MacNeil, R. (2011b). Paper lab book versus electronic lab notebook: a real life comparison. *The electronic lab notebook blog*. Retrieved from http://elnblog.axiope.com/?p=956

Marques, J.F., Canessa, N., Siri, S., Catricalà, E., & Cappa, S. (2008). Conceptual knowledge in the brain: fMRI evidence for a featural organization. *Brain research, 1194*, 90-99.

Marsh, E.J. and Tversky, B. (2004) Spinning the stories of our lives *Applied Cognitive Psychology 18*, 491-503.

Matschke, C., Moskaliuk, J., Bokhorst, F., Schümmer, T., & Cress, U. (2014). Motivational factors of information exchange in social information spaces. *Computers in Human Behavior*, *36*, 549–558. doi:10.1016/j.chb.2014.04.044

McAllister, James W. (2012). Climate Science Controversies and the Demand for Access to Empirical Data. *Philosophy of Science, 79*(5), 871-880.

McGlade, K., Cargo, C., Fogarty, D., Boohan, M. & McMullin. M. (2012). Handwritten Undergraduate Case Reports. *The Clinical Teacher.* 9(2), 112–8.

Mcphillips, T., Song, T., Kolisnik, T., Aulenbach, S.,Belhajjame, K., Bocinsky, K., Cao, Y., & Chirigati, F. (2015). *YesWorkflow: A User-Oriented, Language-Independent Tool for Recovering Workflow Information from Scripts.* A paper presented to IDCC15, London, 9-12 February.

Memon, A., Meissner, C.A., & Fraser, J. (2010). The cognitive interview: a meta-analytic review and study space analysis of the past 25 years *Psychology, Public Policy, and Law 16*, 340 -372.

Milsted, A., Hale, J., Frey, J., & 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

Mitchell, E. T. (2009). *Metadata Literacy: An Analysis Of Metadata Awareness in College Students.* Ph.D. Thesis, School of Information and Library Science, University of North Carolina at Chapel Hill, 2010. http://dc.lib.unc.edu/cdm/ref/collection/etd/id/2910 (ac- cessed July 31, 2013).

Myers, J. D. (2003). Collaborative electronic notebooks as electronic records: Design issues for the secure electronic laboratory notebook (ELN). *SIMULATION SERIES*, *35*(1), 13-22.

Myers, J. D., Mendoza, E. S., & Hoopes, B. L. (2001). A Collaborative Electronic Laboratory Notebook. In M. H. Hamza (Ed.), *Proceedings of the IASTED International Conference, Internet and Multimedia Systems and Applications* (pp. 334–338). Anaheim, CA: IASTED, Acta Press.

Nickerson, R. S. (1999). How we know—and sometimes misjudge—what others know: Imputing one's own knowledge to others. *Psychological Bulletin, 125*(6), 737–759. doi:10.1037/0033-2909.125.6.737

Pancerella, C., Hewson, J., Koegler, W., Leahy, D., Rahn, L., Yang, C., Myers, J.D., Didier, B., McCoy, R., Schuchardt, K., Stephan, E., Windus, T., Amin, K., Bittner, S., Lansing, C., Minkoff, M., Nijsure, S., Laszewski, G., Pinzon, R., Ruscic, B., Wagner, A., Wang, B., Pitz, W., Ho, Y., Montoya, D., Xu, L., Allison, T., Green, W., Frenklach, (2003) M. Metadata in the Collaboratory for Multi-Scale Chemical Science in *DCMI '03 Proceedings of the 2003 International Conference on Dublin Core and Metadata Applications: supporting communities of discourse and practice – metadata research & applications*. http://dcpapers.dublincore.org/pubs/article/view/740. (Accessed 11 July 2013).

Reder, L. M., Park, H., & Kieffaber, P. D. (2009). Memory systems do not divide on consciousness: Reinterpreting memory in terms of activation and binding. *Psychological Bulletin, 135*(1), 23–49. doi:10.1037/a0013974

Reimer, Y. J., Hagedal, M., Wolf, P., & Bahls, B. (2011). Turning the desktop inside-out: Evaluating information access and management through a single interface. *Journal of the American Society for Information Science and Technology, 62*(12), 2327–2346. doi:10.1002/asi.21627

Roediger, H.L. (2008). Relativity of remembering: why the laws of memory vanished *Annual Review of Psychology 59*, 225-54.

Rogers, Y., Sharp, H., & Preece, J. (2011). *Interaction design: beyond human-computer interaction*. John Wiley & Sons.

Rowlands, I.; Nicholas, D.; Williams, P.; Huntington, P.; Fieldhouse, M.; Gunter, B.; Withey, R.; Jamali, H. R.; Dobrowolski, T.; Tenopir, C. (2008). The Google generation: The information behaviour of the researcher of the future. *Aslib Proceedings*. *60*, 290–310.

Rudolphi, F., & Goossen, L. J. (2012). Electronic laboratory notebook: the academic point of view. *Journal of Chemical Information and Modeling*, *52*(2), 293–301. doi:10.1021/ci2003895

Ryan, B. & Walmsley, S. (2003). Implementing metadata collection: a project's problems and solutions. *Learning technology. 5*(1), http://lttf.ieee.org/learn_tech/issues/january2003/index.html#3 (Accessed 11 July 2013).

Sam A, de Castro P, Pablo E, Jorge E, Hanwell MD, Peter M-R, Paul S, Jens T, Joe T. (2011). The Quixote project: Collaborative and Open Quantum Chemistry data management in the Internet age. *Journal of Cheminformatics*, 3:38. doi:10.1186/1758-2946-3-38

Scholl, P. M., & Van Laerhoven, K. (2014). Wearable digitization of life science experiments. In *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing Adjunct Publication - UbiComp '14 Adjunct*, 1381–1388. doi:10.1145/2638728.2641719

Schott, B.H., Henson, R.N., Richardson-Klavehn, A., Becker, C., Thoma, V., Heinze, H. & Düzel, E. (2005). Redefining implicit and explicit memory: the functional neuroanatomy of priming, remembering, and control of retrieval *Proceedings of the National Academy of Sciences of the United States of America 102:4*, 1257-62.

Schwarz, N. (2007). Cognitive aspects of survey methodology. *Applied Cognitive Psychology, 21*(2), 277–287. doi:10.1002/acp.1340

Shankar, K. (2004). Remembering and Forgetting: Exploring the Role of the Record in Conducting Scientific Research. In Rayward, B. & Bowden, M. (Eds) *Proceedings of the 2nd Conference of the History and Heritage of Scientific Information Systems*, *ASIS&T Monograph Series*, 118-126.

Shankar, K. (2006). Recordkeeping in the Production of Scientific Knowledge: An Ethnographic Study. *Archival Science*, *4*, 367–382. doi:10.1007/s10502-005-2600-1

Shankar, K. (2007). Order from chaos: The poetics and pragmatics of scientific recordkeeping. *Journal of the American Society for Information Science and Technology, 58*(10), 1457–1466. doi:10.1002/asi.20625

Shankar, K. (2009). Ambiguity and legitimate peripheral participation in the creation of scientific documents. *Journal of Documentation, 65*(1), 151–165. doi:10.1108/00220410910926167

Sharifian, F. (2001). The mnemonic influence of self-cues on narrative recall. *Issues In Educational Research*, *11*(1), 15-24. http://www.iier.org.au/iier11/sharifian.html

Solomon, P. (1997). Discovering information behavior in sense making. II. The Social. *Journal of the American Society for Information Science, 48*(12), 1109–1126.

Spiers, H., Maguire, E.A, and Burgess, Neil. (2001). Hippocampal amnesia *Neurocase: case studies in neuropsychology, neuropsychiatry, and behavioural neurology 7*, 357-382.

Stangor, C. and McMillan, D. (1992). Memory for expectancy-congruent and expectancy-incongruent information: A review of the social and social developmental literatures. *Psychological Bulletin, 111*(1), 42-61.

Swan, A., & Brown, S. (2008). *The skills, role and career structure of data scientists and curators: an assessment of current practice and future needs.* Report to the JISC (pp. 1–32). Retrieved from http://ie-repository.jisc.ac.uk/245/1/DataSkillsReport.doc

Swinglehurst, D., Greenhalgh, T., & Roberts, C. (2012). *Computer Templates in Chronic Disease Management: Ethnographic Case Study in General Practice.* 1–16. doi:10.1136/bmjopen-2012-001754.

Tabard, A., Hincapié Ramos, J. D., & Bardram, J. (2012). The eLabBench in the wild – Supporting Exploration in a Molecular Biology Lab. *In Proceedings of the 2012 ACM annual conference on Human Factors in Computing Systems - CHI '12.* (pp. 3051–3060). Austin, Texas, USA: ACM Press. doi:10.1145/2207676.2208718

Tabard, A., Mackay, W. E., & Eastmond, E. (2008). From individual to collaborative: The Evolution of Prism, a Hybrid Laboratory Notebook. In *Proceedings of the ACM 2008 conference on Computer supported cooperative work - CSCW '08* (pp. 569–578). New York, New York, USA: ACM Press. doi:10.1145/1460563.1460653

Tabard, A., Hincapié-Ramos, J., & Bardram, J. (2012). The eLabBench in the wild: supporting exploration in a molecular biology lab. In *Proceedings of the*

2012 ACM Annual Conference on Human Factors in Computing Systems - CHI '12, (pp. 3051–3060). doi:10.1145/2207676.2208718

Taylor, K. T. (2011). Evolution of Electronic Laboratory Notebooks. In S. Ekins, M. A. Z. Hupcey, & A. J. Williams (Eds) *Collaborative Computational Technologies for Biomedical Research*. (pp. 303–320) John Wiley & Sons, Inc., Hoboken, NJ, USA.

Terrenghi, L., Hilliges, O., & Butz, A. (2007). Kitchen stories: Sharing recipes with the Living Cookbook. *Personal and Ubiquitous Computing, 11*(5), 409–414. doi:10.1007/s00779-006-0079-2

Tourangeau, R., Conrad, F. G., Couper, M. P., & Ye, C. (2014). The Effects of Providing Examples in Survey Questions. *Public Opinion Quarterly, 78*(1), 100–125. doi:10.1093/poq/nft083

Tourangeau, R., Rips, L.J., & Rasinski, K. (2000). *The Psychology Of Survey Response.* Cambridge University Press.

Tuckey, M.R. and Brewer, N. (2003a). How schemas affect eyewitness memory over repeated retrieval attempts. *Applied Cognitive Psychology 17*, 785-800.

Tuckey, M.R. and Brewer, N. (2003b). The influence of schemas, stimulus ambiguity, and interview schedule on eyewitness memory over time *Journal of Experimental Psychology: Applied 9*, 101-118.

Tulving, E. (2002). Episodic memory: from mind to brain *Annual review of psychology 53*, 1-25.

Tversky, B. and Marsh, E.J. (2000). Biased retellings of events yield biased memories *Cognitive psychology 40:1*, 1-38.

Tweney, R. (1991). Faraday's 1822 "Chemical hints" notebook and the semantics of chemical discourse. Bulletin for the History of Chemistry, 11, 51–55. Retrieved from http://www.scs.illinois.edu/~mainzv/HIST/bulletin_open_access/num11/num11 p51-55.pdf

UWM Division of Information Technology (2012). *Electronic Lab Notebook Pilot at the University of Wisconsin-Madison.* Retrieved from http://academictech.doit.wisc.edu/ideas/electronic-lab-notebooks

Van Kleek, M. G., Bernstein, M., Panovich, K., Vargas, G. G., Karger, D. R., & Schraefel, M. (2009). Note to self: Examining Personal Information Keeping in a Lightweight Note-Taking Tool. In *Proceedings of the 27th international conference on Human factors in computing systems – CHI 09* (pp. 687–694). New York, New York, USA: ACM Press.

Wagenaar, W.A. (1986). My Memory: A study of autobiographical memory over six years *Cognitive Psychology 18*, 225-252.

Ward, C.; Freiman, L.; Jones, S.; Molloy, L.; Snow, K. Making Sense: Talking Data Management with Researchers. *International Journal of Digital Curation.* 2011, 6,265−273. http://www.ijdc.net/index.php/ijdc/ article/view/197 (accessed July 11, 2013).

Warwick, C., Fisher, C., Terras, M., Baker, M., Clarke, A., Fulford, M., Grove, M., O'Riordan, E., & Rains, M. (2009). iTrench: A study of user reactions to the use of information technology in field archaeology. *Literary and Linguistic Computing*, 24(2), 211–223.

Weng, Y. H., Sun, F. S., & Grigsby, J. D. (2012). GeoTools: An android phone application in geology. *Computers and Geosciences, 44*, 24–30. doi:10.1016/j.cageo.2012.02.027

Whitton, M. *Institutional Data Management Blueprint (IDMB) Questionnaire Quantitative Data*. University of Southampton: Southampton, GB, 2011. http://eprints.soton.ac.uk/195959/ (accessed July 11, 2013).

Whitton, M.; Takeda, K. *Data Management Questionnaire Results: IDMB Project.* University of Southampton: Southampton, GB, 2011. http://eprints.soton.ac.uk/196243/1/IDMB_Survey_ Report.pdf (accessed July 11, 2013).

Willoughby, C., Bird, C. L., Coles, S. J., & Frey, J. G. (2014). Creating Context for the Experiment Record. User-Defined Metadata: Investigations into Metadata Usage in the LabTrove ELN. *Journal of Chemical Information and Modeling, 54(*12), 3268–83. doi:10.1021/ci500469f

Willoughby, C., Bird, C., & Frey, J. (2015). User-Defined Metadata: Using Cues and Changing Perspectives. *International Journal of Digital Curation, 10*(1), 18–47. doi:10.2218/ijdc.v10i1.343

Wilson, D. F. (2011). *Purposive Variation in Recordkeeping in the Academic Molecular Biology Laboratory.* PhD Thesis. University of Glasgow. Retrieved from http://theses.gla.ac.uk/2482/

Wright, J.M. (2009) Make it better but don't change anything. *Automated Experimentation*, **1**:5. doi:10.1186/1759-4499-1-5

Yeh, R., Liao, C., Klemmer, S., Guimbretière, F., Lee, B., Kakaradov, B., Stamberger, J., & Paepcke, A. (2006). "ButterflyNet: a mobile capture and access system for field biology research." In *Proceedings of the SIGCHI conference on Human Factors in computing systems*, pp. 571-580. ACM, 2006.

Yeo, R. (2014) *Notebooks, English Virtuosi, and Early Modern Science***.** University of Chicago Press.

Zeng, M.L., & Qin, J. (2008). *Metadata*. New York: Neal-Schuman. ISBN: 978-1555706357

Zhao, J., Goble, C., Stevens, R., & Bechhofer, S. (2004). Semantically linking and browsing provenance logs for e-science. (M. Bouzeghoub, C. Goble, V. Kashyap, & S. Spaccapietra, Eds.) *Semantics of a Networked World. Semantics for Grid Databases Volume 3226. (pp158-176).* Springer Berlin Heidelberg.