Kyle A. Richardson. Academic Chemists' Use of Laboratory Notebooks and Other Information Management Tools. A Master's Paper for the M.S. in I.S. degree. June, 2009. 60 pages. Advisor: Diane Kelly.

While the role of laboratory notebooks and other scientific information management tools have been studied in the context of corporate research and development, little work has been done to describe similar practices in the academic domain. This study aims to determine how scientific researchers in academia use their laboratory notebooks and other information management tools to aid in their day-to-day research work and if these tools effectively support their collaborative research efforts. An online survey of academic research chemists from four universities in central North Carolina was conducted. Subjects were asked a series of questions to gauge their use of laboratory notebooks, electronic information management tools, and collaboration practices. The response data indicates an evolving trend toward electronic laboratory data and notes; however, the paper-based laboratory notebook remains the primary means of recording experimental data and tracking progress.

Headings:

Science and Technology/Information Management

Chemistry/Laboratory Notebooks

Surveys/Information Tools

ACADEMIC CHEMISTS' USE OF LABORATORY NOTEBOOKS AND OTHER INFORMATION MANAGEMENT TOOLS

by
Kyle A. Richardson

A Master's paper submitted to the faculty
of the School of Information and Library Science
of the University of North Carolina at Chapel Hill
in partial fulfillment of the requirements
for the degree of Master of Science in
Information Science.

Chapel Hill, North Carolina

June 2009

Approved by

____

Diane Kelly

Table of Contents

Introduction

Scientists rely heavily on their laboratory notebooks as the definitive record of their research work. Scientific research is built upon the principle that all results can be replicated given an appropriate description of the experiment, and the laboratory notebook is a vital piece of that documentation. Traditionally, laboratory notebooks have been bound, paper-based artifacts that not only serve as the record of a researcher's work but also as his/her personal research journal (schraefel et al., 2004). In addition to the basic experimental description and results, researchers record insight into their thoughts, ideas, and experiences. These can be extremely important to the overall research process but may or may not have any bearing on the reproducibility of a specific experiment.

With the advancements in technology and scientific instrumentation, scientists are now able to collect huge amounts of data for each experiment they perform. Advanced computing algorithms have aided analysis of these massive data sets, but researchers still suffer from information overload with each new experiment they perform. In addition to the shear amount of data generated, modern scientists have benefited from advancements in technology, specifically personal computers and the Internet, to write and store their published work, communicate and collaborate with fellow researchers, and cross the traditional boundaries of their disciplines. While all of these advances have greatly increased the productivity of scientific research, they have left in their wake a new set of problems.

Scientists now have many disconnected forms and locations where their data, observations, thoughts, and finalized work are maintained and stored. Many have text and document files on their personal computer, raw and analyzed data sets on laboratory computer systems, and a handwritten record that binds it all together in the form of a paper-based laboratory notebook. In recent years, private sector research and development (R&D) has begun to realize the vast amount of knowledge stored in laboratory notebooks that is lost or overlooked due to the inability to effectively search those records (Taylor, 2006). Corporate R&D environments have thus been searching for an electronic laboratory notebook (ELN) solution that would allow their researchers to combine all aspects of their work into a single system that could then be searched by other researchers within the company. They believe this will cut down on redundancy and provide a more accessible record of the company's overall R&D performance. In addition, an ELN system would be designed to integrate into a company's existing R&D workflow, which is nearly entirely computer-based already (Taylor, 2006).

Scientific researchers in academia, however, have remained skeptical of ELN systems and very few have implemented a broad-spectrum electronic information management solution (Butler, 2005). This is most notably due to a fully integrated system's rigidity. Researchers in academia have a lot more freedom than those involved in private sector R&D and do not want to be bound to a one-size-fits-all solution. Most are free to study whatever topics they wish and often workflows vary widely from one research lab to another. While these aspects often support a greater quest for knowledge in areas that the corporate world are uninterested in, it creates a major problem with respect to integrating an entire university community into any standardized electronic information management system. University culture is based on freedom and openness, and this model simply does not support those ideals. Because of the very open nature of the academic environment, there are a number of possibilities for how each laboratory and potentially each researcher have chosen to manage and store all of their information.

Collaborative research is growing in importance in both the academic and private-sector arenas as the demand for scientific advancements and breakthroughs skyrockets. With the advancements in data collection and overall explosion of scientific data, researchers are beginning to develop closer ties with one another. Once scattered and concerned only with their current work, researchers are now actively seeking out colleagues with similar interests to tackle problems too large for a single researcher or laboratory. Collaborative research is becoming more and more commonplace with the advances in communication technologies and the ability to readily access and share data sets over the Internet. Researchers are still concerned with maintaining control over their work but are realizing the importance of multiple thoughts and opinions to continually develop innovative ideas and solutions to today's problems.

The purpose of this study is to determine the various forms the scientific laboratory notebook takes in an academic research environment and whether, in its current form, it is meeting the information management needs of research scientists. This study will also determine if current information management practices are effective in providing support for collaborative research. The specific research questions this study will address are:

1. What role does the laboratory notebook serve in scientific information management practices and what forms does it take?

2. What electronic information management tools, other than the laboratory notebook, are currently being used to support research and collaboration efforts?
3. How effective are researchers' current information management practices and how willing are they to explore alternatives?

Literature Review

In order to develop an understanding of the information management needs of academic researchers in the natural sciences, it is first necessary to gain some insight into the types of research information they currently manage and how they meet their particular information needs. Additionally, since current information management practices revolve around the laboratory notebook, it is important to develop a working knowledge of the various affordances it provides as well as the functions it supports in research work. Beyond the laboratory notebook, scientists use a variety of other, supplemental, information management tools to organize and access their research materials. An overview of the various tools being used will also be presented. Finally, the current state of ELN systems along with the features they provide and needs they meet will be discussed. The importance and challenges of effectively supporting research collaboration will be stressed.

Meeting Information Needs Electronically

In order to understand how scientists choose to manage research information, it is first necessary to understand how they retrieve and store the information required to support their work. Research has shown that the vast majority of scientific information needs are now being met electronically. Over 70% of respondents to a survey on scientific information seeking behavior indicated that they used either citation database searching, such as SciFinder Scholar, or online web searching, such as Google, as their primary means of gathering research information (Hemminger, Lu, Vaughan, & Adams, 2007). Additionally, researchers in the natural sciences, chemists in particular, tend to rely almost solely on peer-reviewed journal articles to support their research and often are interested only in recently published material, generally within the last 5 years (Brown, Blake, Brown, & Tenopir, 2006). Because of this, scientists are able to retrieve nearly all of the research articles of interest to them electronically from either the publisher's website or a digital library and store them electronically on their personal computers. Interestingly enough, although researchers prefer to gather their information electronically, the majority of them still prefer to read materials in printed form (Hemminger et al., 2007). While this behavior is undoubtedly common, it may shed some additional light on the resistance to move to an entirely digital information management system.

Importance of the Laboratory Notebook

As far back as the mid-1990s, private-sector scientists were beginning to understand the need to access and share the information contained in laboratory notebooks. Laboratory notebooks contain a log of all experiments performed by a particular researcher as well as an initial interpretation of their findings (Dessy, 1995). Published journal articles and reports that derive from those findings never include all of the information originally found in the laboratory notebook, and this information can be very important if the work is to be repeated or enhanced by someone else. Additionally, researchers now gather more and more data in digital formats making it harder to integrate with the traditional laboratory notebook. Images and data tables are often printed out and glued into the notebook to maintain records; however, the file names and locations are often not included causing potential retrieval problems when the data is needed for additional analysis or review (Butler, 2005).

Companies and organizations in the private sector have realized that a great deal of the valuable information in laboratory notebooks is being lost (Taylor, 2006). Paper-based notebooks cannot be easily searched and are often difficult to interpret and understand by anyone besides the original author. Additionally, scientists are often performing several experiments in parallel making the chronological progression of the laboratory notebook less useful since a single experiment may skip from page to page as the researcher works on other projects (Dessy, 1995). Laboratory notebooks are also the legal records required as evidence for patent and intellectual property disputes and can cause the organization embarrassment if the writing is eligible or pages are lost or destroyed over the years (Kihlen, 2005). Although these issues may be more commonplace in private sector industry, academic research laboratories and institutes suffer from the same general problems only on a slightly smaller scale.

Information Management Practices

Understanding how researchers currently utilize the laboratory notebook and other supplemental information management tools is extremely important if new tools and solutions are to be developed to support their efforts.

According to a study of industry chemists in the UK, paper-based laboratory notebooks are primarily used to record the measurements taken in the laboratory and observations as the experiment progresses but very little about the actual procedure used to perform the experiment (schraefel et al., 2004). This information is instead recorded in a separate document since the experiment will likely be repeated several times.

Researchers within a discipline or even a particular lab tend to rely on a common knowledge of certain procedures and terms. This results in documentation that is difficult to understand without the community specific knowledge (schraefel et al., 2004). What a researcher records in the laboratory notebook also depends on the researcher's own style. For example, one researcher might record the batch number of a substance used while another may not find that information important (schraefel et al., 2004). Additionally, these notebooks are subjected to highly volatile conditions throughout the course of a normal day as researchers place them wherever they can find room around their workbench and instruments. The potential for damage and destruction of notebook is quite high in this environment, which can cause entire pages of a laboratory notebook or even entire experiments to be lost.

In addition to the laboratory notebook most researchers have at least one, and usually several, other notebooks in which they track different kinds of research related information (Reimer & Douglas, 2004). These can include publication notebooks, group meeting notebooks, travel notebooks, and many others. Additionally, these are generally found in a variety of media formats. For example, a publication notebook could be a Word document that includes notes and an outline for an upcoming publication, while a travel or group meeting notebook could be a simple spiral notebook or notepad that is easily taken to meetings and conferences and contains notes on others work and ideas for future research.

Researchers in the natural sciences seem to generally struggle with how to organize and find their data (Tabard, Mackay, & Eastmond, 2008). In recent years, this problem has escalated due to the huge amounts of data capable of being generated by each experiment. Researchers will generally collect as much data as possible even if they know they will not be analyzing all of it (Birnholtz & Bietz, 2003). The thought is that they can share these raw data sets with other researchers who may be interested in analyzing some portion of it.

The actual organization of a scientist's information is often scattered. They are using so many different tools that portions of their research information could be stored in Word documents, e-mail messages, paper-based laboratory notebooks, meeting notebooks, and at times even blogs, wikis, or web pages (Tabard et al., 2008). All of these different sources cannot be easily organized into a single structure and thus retrieval of information is extremely difficult and becomes worse the further back the information was recorded. Often researchers can easily remember what they have been working on over the past week or even month, but beyond that a person's memory is not a dependable retrieval mechanism (Tabard et al., 2008).

While computer use is pervasive in nearly all aspects of scientific research work, nearly 75% of the industry biologists interviewed in a French study continue to use paper-based laboratory notebooks (Tabard et al., 2008). This is due in part to the rigorous scientific training they have undergone. Scientific researchers are taught early on to write clearly and concisely the important aspects of the experiment they are performing and to never delete or edit the information once it has been recorded. Editing information in the laboratory notebook is highly discouraged due to the legal nature of the notebook should patents or intellectual property claims wish to be filed in the future.

Electronic Laboratory Notebooks

Due to the high demand for an integrated digital solution in private sector industry, there has been quite a bit of work done to understand the needs of scientific researchers and to develop novel and even some commercial ELN systems to support those needs. It is important to note that the vast majority of this work concentrates on the corporate environment and provides little insight into the needs specific to academic researchers.

The fundamental role that any successful ELN system should fulfill for private-sector R&D is the ability to maintain accurate, legal records of research in accordance with US patent laws (Myers, 2003). There are a number of challenges that must be addressed here in order to map the paper-based laboratory notebook requirements to an electronic equivalent. Signed notebook pages are necessary to comply with these laws and must be addressed. One solution is to maintain all records on a central server where the data is saved once entered and is unchangeable (schraefel et al., 2004). Another is to implement a public key digital signature, which can only be applied by the author (Myers, 2003). Additionally, the ability to print out physical copies of these documents in the form of standardized reports, which include all the necessary information and look professional is essential (Kihlen, 2005). These documents can then be sent to government agencies if needed as well as be stored on-site as a physical backup of the research work being performed.

ELN systems have been shown to produce higher quality information than traditional, paper-based laboratory notebooks in an industry setting. Researchers recognize that they will not be the only ones using and relying on the information and make a conscious effort to provide lengthier descriptions and more detailed notes regarding their activities

(Kihlen, 2005). They find it beneficial to be able to see what others are working on and how it relates to their own current projects.

A variety of systems, both fully digital as well as digital-paper hybrids, have been developed over the last several years and have identified some of the main features needed by the scientific community. The ability of the system to integrate personal productivity features such as access to e-mail, calendaring, file system documents, and web browser is essential (Myers, 2003; Tabard et al., 2008). Most researchers depend on e-mail as a vital source of scholarly communications from colleagues and collaborators, and the ability to tag messages for a particular project or experiment would avoid the need to re-write information and reduce the transposition errors that might occur. The ability to support advanced search and retrieval of information is also critical. One of the main downsides of the paper-based laboratory notebook is the inability to effectively search those records. Researchers would save time and frustration by being able to search their laboratory notebooks for a specific experiment or concept. Additionally, the ability to provide long-term preservation and archival of research information would be an important feature of any ELN system (Myers, 2003). Researchers need an organized way to provide for long-term storage. Currently this consists mostly of boxes filled with publications, laboratory notebooks, and DVDs of archived raw data files that are not easily retrieved. While these features are all important in order to provide for the scientists needs in an information management system, the most pressing need and sought after feature is the ability to collaborate and share research information.

Scientific research has become an increasingly more collaborative arena. In the past, researchers were hesitant to share data and procedures for fear that another investigator would beat them to a discovery. Today, scientists are generating so much data that they need the help of collaborators to analyze and understand it all. Not only do researchers need an electronic information management system, they need a way to share some or all of the information in that system with others (Birnholtz & Bietz, 2003). They may need to share it with others members of their laboratory or organization or they may want to share it with external collaborators that span across industry and academia as well as around the globe.

Summary

Scientific researchers, specifically those in the natural sciences such as chemistry and biology, rely on their laboratory notebooks to provide the glue that holds all of their research data and materials together. As the Internet and personal computers have become commonplace in society, they have also become a mainstay in research laboratories and an essential part of the research workflow. Scientists are now able to search and retrieve research materials using databases and search engines from the comfort of their offices and laboratories. They are also able to utilize complex instruments and computational algorithms to gather huge amounts of raw data. Once scientists have gathered all of this information, they need to be able to organize it for easier use. This is where many researchers struggle. Their laboratory notes are not adequate to describe all of the different sources of research information they are utilizing. Additionally, multiple experiments and lines of research running in parallel make it more difficult to manage the information using a traditional laboratory notebook.

Electronic laboratory notebooks have become a buzzword in the private sector because of the potential productively increase and streamlining of data and resource management. ELN systems can provide the legal documentation necessary to file patents and often result in higher quality information than that provided in a personal, paper-based laboratory notebook. They can also allow for the integration of all of an organization's information into a central repository of knowledge. This institutional knowledge can then be preserved even as researchers come and go over the years. Additionally, ELN systems can better support the collaborative efforts of scientific researchers today. Many scientists have begun to work in teams on research projects to increase productivity and bring a greater range of knowledge and experience to the problem. The ability to effectively share research information and materials over the Internet to collaborators around the world provides great potential for the advancement of science.

Methods

Academic research chemists' current use of laboratory notebooks and other scientific information management tools were investigated using a survey method. Specifically, an online questionnaire was used to provide an easy interface for subjects to respond without requiring a large time commitment. Surveys are an excellent way to describe the behavior of a large population without direct observation (Babbie, 2007). This allows the investigation of a broad range of subjects from a population that has not been previously described in the literature.

Subjects

While this research could be broadly applied to academic research scientists, this study focuses on researchers in the field of chemistry. Specifically, the target subject population included researchers in the Departments of Chemistry from four major research universities in North Carolina. These included the University of North Carolina at Chapel Hill, North Carolina State University, Duke University, and Wake Forest University. All subjects were over 18 years of age and of varying gender, ethnicity, and race. Additionally, all subjects had completed a bachelor's degree program and were actively conducting chemical research at one of the above mentioned institutions at the time of the study. Completion of a bachelor's degree indicates some level of research experience and a desire to continue conducting research in the field.

Sampling was done by convenience (those who chose to complete the online survey). Subjects were recruited through an e-mail invitation sent to various electronic mailing lists of the departments being studied (Appendix A). The business manager of each department was contacted and asked to forward the survey invitation to all graduate students, postdoctoral researchers, staff scientists, and professors in their department. The e-mail text contained a brief description of the study and an invitation to participate in the study. Halfway through the study, a second reminder e-mail was sent to the department business mangers to be forwarded on to the same mailing lists as the original invitation. Because of the use of various departmental mailing lists, there was no way to determine the exact size of the population and thus it was impossible to accurately compute a response rate.

Care was taken to ensure that all subjects were comfortable participating in the study and that they understood their rights as subjects. Before beginning any data collection, the Institutional Review Board (IRB) was consulted for their approval of this study and all of its proposed procedures. Approval was secured under IRB Study # 09-0721. Subjects were asked to read and acknowledge their understanding of a standard online consent form prior to participating in the study (Appendix B). No identifiable data was collected and all subjects' responses were completely anonymous. The Qualtrics system, which was used to provide the online survey to subjects, does collect the IP address of the computer used to complete and survey, but these were destroyed prior to data analysis to ensure subject anonymity. The researcher was the only one with access to the survey data and it was stored on a secure server throughout the analysis period. The data will be kept as a record of this study after its completion, but, as it is completely anonymous, there is no risk of subjects ever being identified from it and it is highly unlikely that subjects' identities can be deduced from their survey responses.

Subjects will not directly benefit from this study in any way; however, this study does hope to benefit the overall scientific community by providing a descriptive analysis of the current usage of laboratory notebooks and other tools to manage scientific information in an academic environment. This information could then potentially be used to suggest additional tools and strategies to scientific researchers as well as to develop new and innovative electronic solutions to scientific information management.

The Survey Instrument

The questionnaire itself is divided into four sections, one to collect demographic information about subjects and three others that each focus on one of the research questions presented above. The first section contains basic demographic questions so that responses can later be correlated and interesting trends discovered. Questions identify the job position of the subject, how long they have been a full-time researcher, and to which traditional division of chemistry their work relates. The second section focuses on subjects' current use of their laboratory notebooks. Questions were developed to gauge how effective a subject's laboratory notebook currently is and what format(s) it takes. These include how often they record experiments, how often they look back at their notes, and how confident they are that their notes are sufficiently detailed. The third section focuses on the electronic information used by subjects in their research. Questions were developed to gauge the amount of data being generated in electronic formats, what formats they take, and how easy it is to retrieve them. The fourth and final section focuses on the subject's collaborations with other researchers and the tools currently being used to support this. Questions include how frequently data and information sharing among collaborators occurs and how the information is exchanged. A set of 5-point Likert-type items was developed to evaluate effectiveness, confidence, and other similar constructs. The full survey questionnaire can be found in Appendix C.

Procedures

Once IRB approval was secured, the online survey was "launched" via the Qualtrics system on April 22, 2009. The survey remained open approximately two and a half weeks and was officially closed on May 11, 2009. The business manager for each of the departments listed above was contacted and asked to distribute the initial survey invitation to all members of their department meeting the criteria outlined above. They were also asked to send a reminder message to these same individuals approximately one week later. There was no guarantee that the business managers forwarded the e-mail invitation and reminder on to their departments; however, responses were received shortly after the messages were sent. One of the business managers requested more information regarding the study, which was provided, but ultimately elected not to have their department participate.

Subjects were free to complete the survey in any location they felt comfortable, but were reminded to select a location that provided them with an appropriate level of privacy. Each subject's participation in the study began when they clicked on the link to the online survey in either the initial invitation or reminder message. They were then directed to the Qualtrics online system where they were presented with a short description of the study and the online consent form (Appendix B). Subjects then indicated their consent to be a part of the study by clicking the "next" button to begin the actual questionnaire. Each subject took varying amounts of time to fully complete the questionnaire, but it took an average of between 5 and 10 minutes for subjects to complete. Subjects were free to skip any question they did not feel comfortable answering and were able to leave the survey at any time. Once subjects completed the survey they were thanked for their time and participation and were not contacted further.

Once the survey period ended, the survey was “closed” via the Qualtrics system to prevent further responses. The response data was then exported and analyzed with Microsoft Excel and SPSS. All IP addresses, collected by the Qualtrics system, were deleted immediately upon export to ensure all data was completely anonymous. All questions are multiple-choice and represent either nominal or ordinal data. Each question was coded and basic descriptive statistics such as the range of data, the mean, and the mode are reported below.

Advantages and Disadvantages

This study uses a survey method to reach as many members of a large population as possible. However, for the purposes of making this study more realistic for a single researcher with limited time and resources, the population for this study was narrowed to chemists associated with one of four universities in the central North Carolina area. While this study undoubtedly generated interesting data, the data is not sufficiently exhaustive to generalize the study findings to all academic research chemists. Thus, the results of this study will be reported solely to describe this particular subset of the overall scientific community and suggestions will be made to enhance and expand the original study to make it more exhaustive.

Because this study uses a survey method with a static questionnaire consisting of only multiple-choice questions, it should generally be reliable and reproducible. This may be slightly skewed depending on the sample being studied, but, as long as the same instrument is used, another researcher should obtain similar results and be able to come fairly close to reproducing this work. This, of course, is never a certainty because studying human subjects brings additionally complexity as their unique past experiences and moods during the actual study will affect the results.

The validity of the survey instrument has not been established and is based solely on the researcher's knowledge gained from the literature and personal experience as a research chemist and information professional. While it is possible that in the future this instrument, or some variant of it, will undergo and pass validity testing, this process is not part of the work described here. Care was taken when developing the survey questionnaire to ensure that each research question posed in this study was adequately represented, but there is no way to establish its ability to generate valid answers to those questions without extensive testing.

Results

A total of 57 responses to the online survey were received over the two and a half week period it was available. The survey results are presented below based on the four sections of the survey instrument, which also correspond to the research questions motivating this study. Additional analysis of the survey data is also provided.

Demographics

All survey respondents had attained a Bachelor's Degree (60.0%), a Master's Degree (20.0%), or a Doctoral Degree (20.0%). Recruitment was conducted to ensure that a minimum of a bachelor's degree had been attained for participation in the study in order to ensure at least some laboratory research experience.

From the four universities contacted for participation, responses were only received from two, North Carolina State University and the University of North Carolina at Chapel Hill. Researchers affiliated with the University of North Carolina at Chapel Hill provided the majority of the responses (74.5%).

The vast majority of responses came from graduate students (87.3%) with only a few responses from post doctoral researchers, staff scientists, and professors/principal investigators. Consequently, the majority of respondents also indicated that they had 1 to 5 years of full-time research experience (72.7%), which is consistent with a 4 to 5 year graduate program.

All of the traditional divisions of chemistry were represented in the response set; however, more respondents indicated Analytical (34.5%) and Inorganic (27.3%) than any other division. Additionally, a number of respondents indicated Other (21.7%), which generally corresponded to Polymer and Materials Chemistry or a combination of Biological Chemistry with other areas such and Organic or Inorganic Chemistry.

A detailed breakdown of the survey's demographic questions and responses can be found in Table 1.

Question/Response	Frequency	Percent
What is the highest degree you have obtained to date?		
Bachelor's Degree	33	60.0%
Master's Degree	11	20.0%
Doctoral Degree	11	20.0%
With what institution are you affiliated?		
North Carolina State University	14	25.5%
The University of North Carolina at Chapel Hill	41	74.5%
What is your current position?		
Graduate Student	48	87.3%
Postdoctoral Researcher	3	5.5%
Professor/Principle Investigator	3	5.5%
Staff Scientist	1	1.8%
How long have you been a full-time researcher?		
Less than a year	5	9.1%
1 to 5 years	40	72.7%
5 to 10 years	9	16.4%
More than 10 years	1	1.8%
What traditional chemistry division most closely relates to your research work?		
Analytical	19	34.5%
Biological	5	9.1%
Inorganic	15	27.3%
Organic	5	9.1%
Physical	4	7.3%
Other	7	12.7%

Table 1: Summary of Demographic Questions and Responses

The Lab Notebook

Overwhelmingly, respondents indicated that their primary laboratory notebook was hand-written (90.7%). Additionally, 66.7% of respondents indicated they were either satisfied or very satisfied with their current lab notebook. The mean response was 3.69±0.64 corresponding to between neutral and satisfied. A detailed summary of these responses can be found in Table 2.

Question/Response	Frequency	Percent
In what format is your primary lab notebook?		
Electronic	4	7.4%
Hand-written	49	90.7%
Other	1	1.9%
How satisfied are you with your current lab notebook?		
Very Dissatisfied (1)	0	0.0%
Dissatisfied (2)	2	3.7%
Neutral (3)	16	29.6%
Satisfied (4)	33	61.1%
Very Satisfied (5)	3	5.6%

Table 2: Summary of Lab Notebook Format and Satisfaction

Respondents indicated that they recorded laboratory notes primarily in their lab notebooks (89.5%); however, many also use Word documents, Excel spreadsheets, legal pads, and a few use commercial electronic notebook systems. Interestingly, no one indicated the use of blog entries or wiki pages, online collaborative alternatives to these standard documents. A breakdown of these responses is shown in Figure 1.

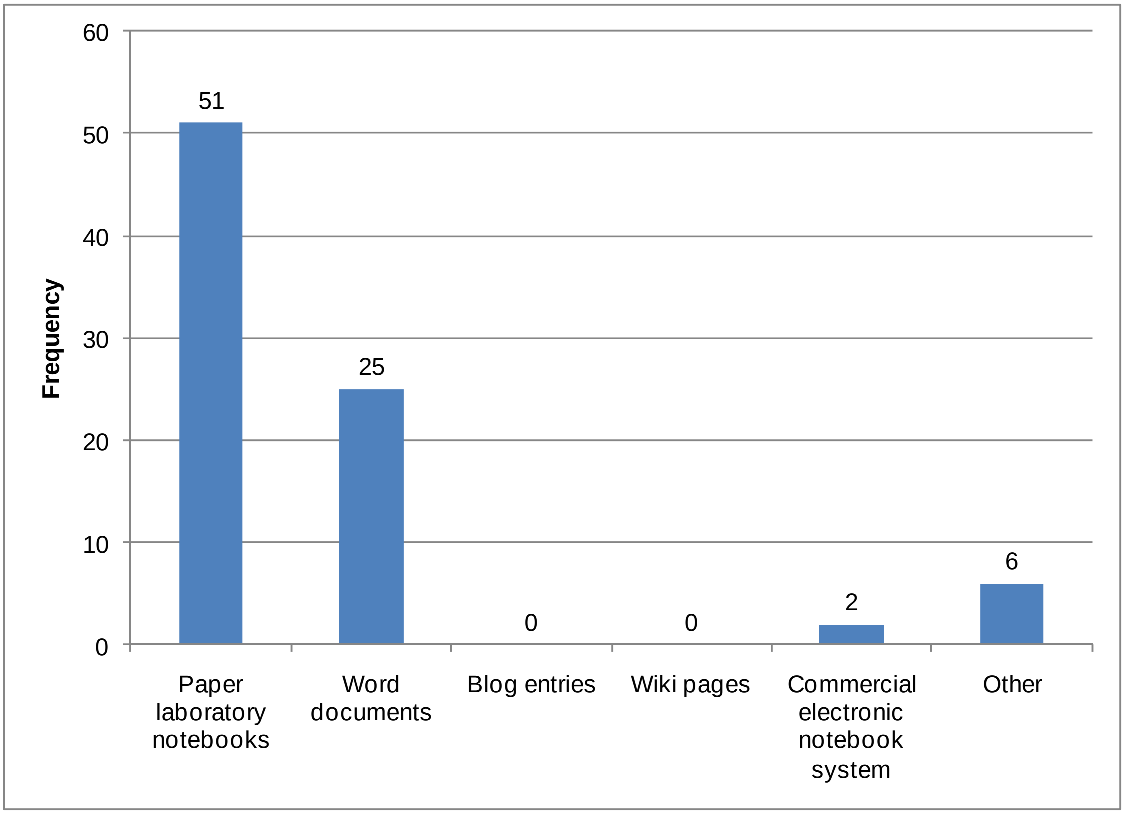

Figure 1: Methods of Recording Laboratory Notes

The majority of respondents (75.9%) indicated that they recorded experiments in their lab notebooks either the same day or within two days of performing an experiment. Only 13.0% indicated that experiments were recorded either within two weeks or over two weeks.

Over half of the respondents (53.7%) indicated that they used a level of detail that was either detailed or very detailed when recording experiments in their lab notebooks. However, the mean response was 3.46±0.88 corresponding to only slightly more detailed than average. Additionally, 70.4% of respondents indicated that they were either confident or very confident that someone else could accurately reproduce an experiment using their notes. The mean response in this case was 2.85±0.76 corresponding to between somewhat confident and confident.

A detailed summary of the survey's laboratory notebook behavior questions can be found in Table 3.

Question/Response	Frequency	Percent
On average, how promptly do you record experiments in your lab notebook?		
Same day	29	53.7%
Within two days	12	22.2%
Within a week	6	11.1%
Within two weeks	4	7.4%
Over two weeks	3	5.6%
In general, what level of detail would you say you use in recording experiments in your lab notebook?		
Very Brief (1)	0	0.0%
Brief (2)	9	16.7%
Average (3)	16	29.6%
Detailed (4)	24	44.4%
Very Detailed (5)	5	9.3%
How confident are you that someone else could use your notes to reproduce an experiment?		
Not Confident (1)	2	3.7%
Somewhat Confident (2)	14	25.9%
Confident (3)	28	51.9%
Very Confident (4)	10	18.5%

Table 3: Summary of Laboratory Notebook Behaviors

Figure 2 illustrates the frequency respondents indicated they look back at their own lab notes that are more than 2 months old. The most prevalent response was 2-3 times a month and accounted for 40.7% of all responses. Additionally, 31.5% of respondents indicated they utilized lab notes older than 2 months only once a month or less.

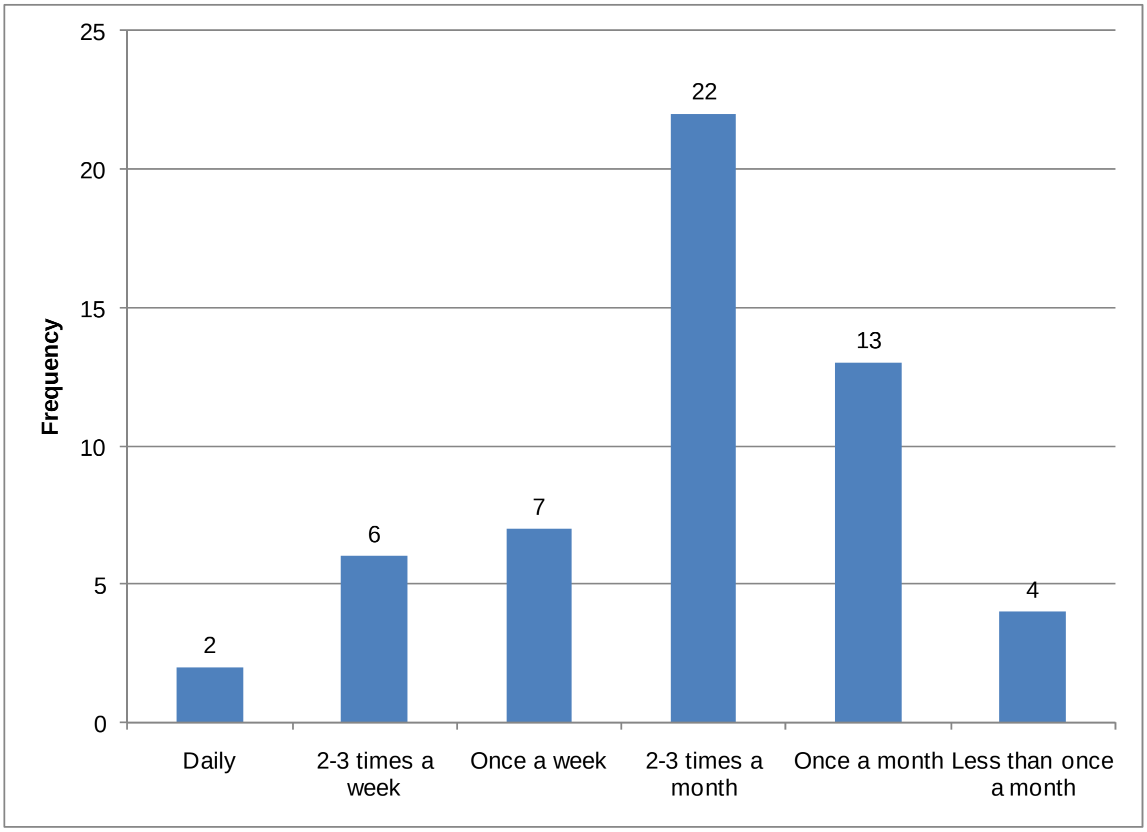

Figure 2: Frequency Respondents Look at Lab Notes More than 2 Months Old

A majority (70.4%) of respondents indicated it was either easy or very easy to find the information they are looking for when they look back at these notes. Another 25.9% indicated they were neutral and only 3.7% indicated that this process was difficult for them. The mean response to this item was 2.19±0.73 corresponding to slightly more neutral than easy.

Figure 3 illustrates the frequency respondents indicated they need to look at another researcher's laboratory notes or data. Over half (61.1%) indicated they needed to utilize another researcher's notes either less than once a month or never. Another 33.4% responded 2-3 times a month or once a month, and only 5.6% responded once a week or more.

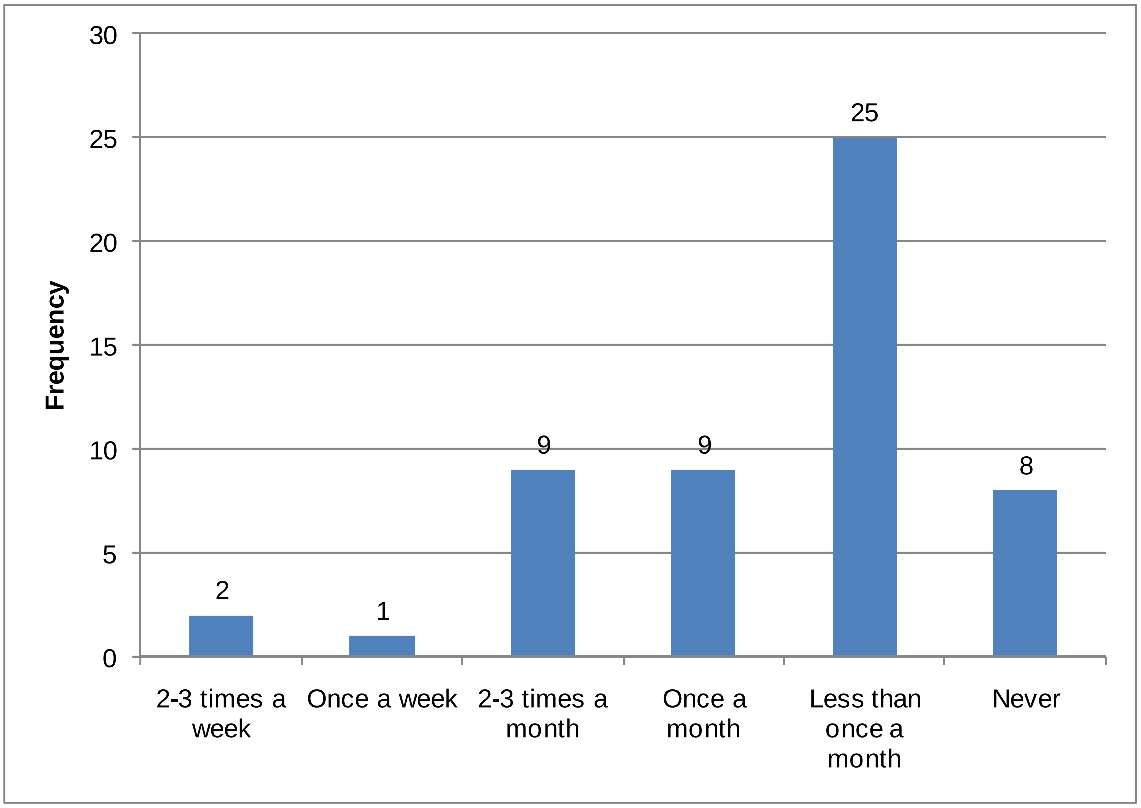

Figure 3: Frequency Respondents Look at another Researcher's Lab Notes or Data

Of those who indicated that they utilize another researcher's notes or data, more than half (57.7%) indicated they were neutral on the difficulty of finding the information they needed in those notes. Another 30.8% responded that it was either difficult or very difficult to retrieve this information, and only 11.5% or 6 respondents indicated this was either an easy or very easy task. The mean response in this case was 2.77±0.78 corresponding to slightly more difficult than neutral.

A detailed summary of the survey's questions relating to laboratory notebook effectiveness can be found in Table 4.

Question/Response	Frequency	Percent
How often do you look back at your own lab notes that are more than 2 months old?		
Daily	2	3.7%
2-3 times a week	6	11.1%
Once a week	7	13.0%
2-3 times a month	22	40.7%
Once a month	13	24.1%
Less than once a month	4	7.4%
When you do, how difficult is it to find the information you are looking for?		
Very Easy (1)	8	14.8%
Easy (2)	30	55.6%
Neutral (3)	14	25.9%
Difficult (4)	2	3.7%
Very Difficult (5)	0	0.0%
How often do you need to look at another researcher's laboratory notes or data?		
2-3 times a week	2	3.7%
Once a week	1	1.9%
2-3 times a month	9	16.7%
Once a month	9	16.7%
Less than once a month	25	46.3%
Never	8	14.8%
When you do, how easy is it to find the information you are looking for?		
Very Difficult (1)	3	5.8%
Difficult (2)	13	25.0%
Neutral (3)	30	57.7%
Easy (4)	5	9.6%
Very Easy (5)	1	1.9%

Table 4: Summary of Laboratory Notebook Effectiveness

Electronic Information

Generally, respondents indicated that their research information takes on many different forms, both printed and electronic. The most frequent responses were spreadsheets and electronic copies of journal articles, manuscripts, etc. each chosen by 46 respondents; however, all of the responses were chosen by a minimum of 31 of the survey respondents. The exception was a response of other, indicated by only 2 respondents, who also indicated using verbal communication and theses and dissertations as valuable sources of research information. Table 5 shows a complete listing of responses and their corresponding frequencies.

Response	Frequency
E-mail messages	34
Spreadsheets	46
Word documents	38
Printed photos or images	31
Digital or scanned images	34
Printed journal articles, manuscripts, etc.	44
Electronic copies of journal articles, manuscripts, etc.	46
Laboratory experiment results (electronic)	42
Other	2

Table 5: Forms of Research Information

Figure 4 illustrates that the majority of respondents collect more than 80% of both their raw and analyzed data in electronic formats. More than half (59.6% and 63.5% respectively) of those surveyed indicated more than 80% of their data is collected in an electronic format. Only 11.5% and 9.6% respectively indicated less than 20% of their data is collected in an electronic format.

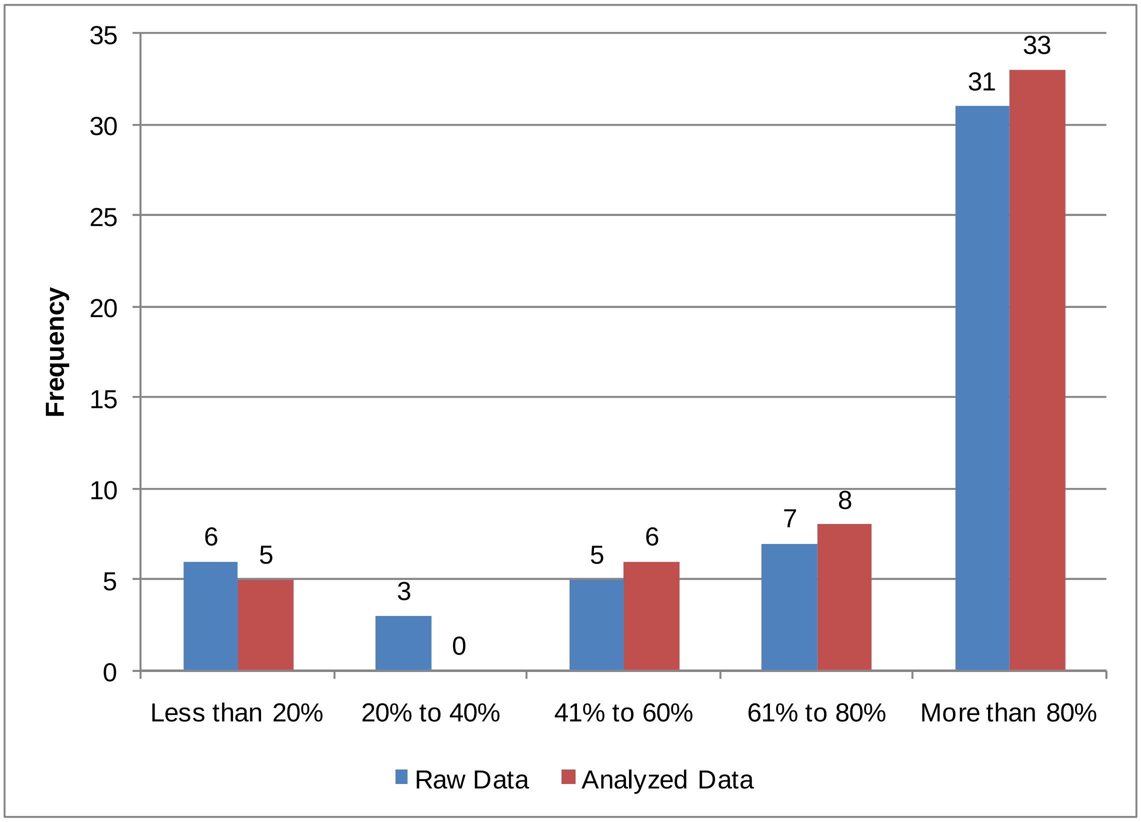

Figure 4: Percent of Data Collected in Electronic Formats

Respondents were asked to indicate all of the locations where they store electronic research data. Table 6 shows that primarily respondents store their data on their own personal computers or locally on lab owned machines. Other locations indicated by respondents included USB flash drives, external hard drives, and other portable storage devices.

Response	Frequency
Personal computer	41
Lab owned computer	35
Lab owned file server	18
University provided file storage	20
Other	9

Table 6: Electronic Data Storage Locations

Table 7 summarizes how easy respondents feel it is to search for and find their electronic research data from previous experiments. The mean response was 2.10±0.78

corresponding to a response of easy. The majority (72.6%) indicated that it was either easy or very easy to retrieve this information while only 3.9% indicated it was difficult.

Response	Frequency	Percent
Very Easy (1)	11	21.6%
Easy (2)	26	51.0%
Neutral (3)	12	23.5%
Difficult (4)	2	3.9%
Very Difficult (5)	0	0.0%

Table 7: Ease of Retrieving Electronic Research Data from Past Experiments

Figure 5 illustrates the likelihood respondents would be willing to move to an entirely electronic lab notebook system given the appropriate tools and training. The mean response was 3.15±1.24, which generally corresponds to an undecided response; however, due to the large standard deviation, it is not possible to establish a general trend from the response data. For example, nearly half (44.2%) indicated they were either unlikely or very unlikely to move to an entirely electronic system, and 34.6% indicated they were either likely or very likely to do so.

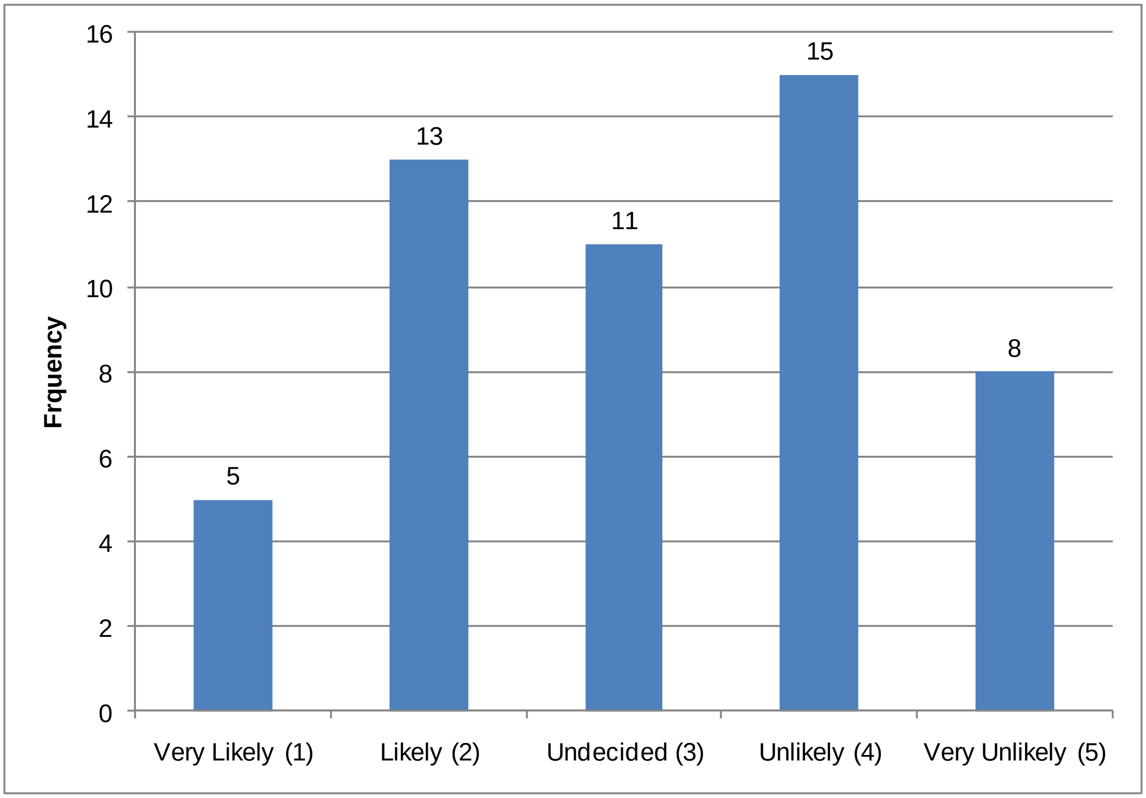

Figure 5: Likelihood of Moving to an Entirely Electronic Lab Notebook System

Collaboration

Respondents were asked to indicate the frequency with which they collaborate and/or share data with other researchers in their own lab as well as with researchers outside of their lab. Figure 6 illustrates their responses and shows that, in general, more collaboration occurs among researchers within a lab than externally. This is expected given members of a lab are generally working on similar projects and report to the same principle investigator.

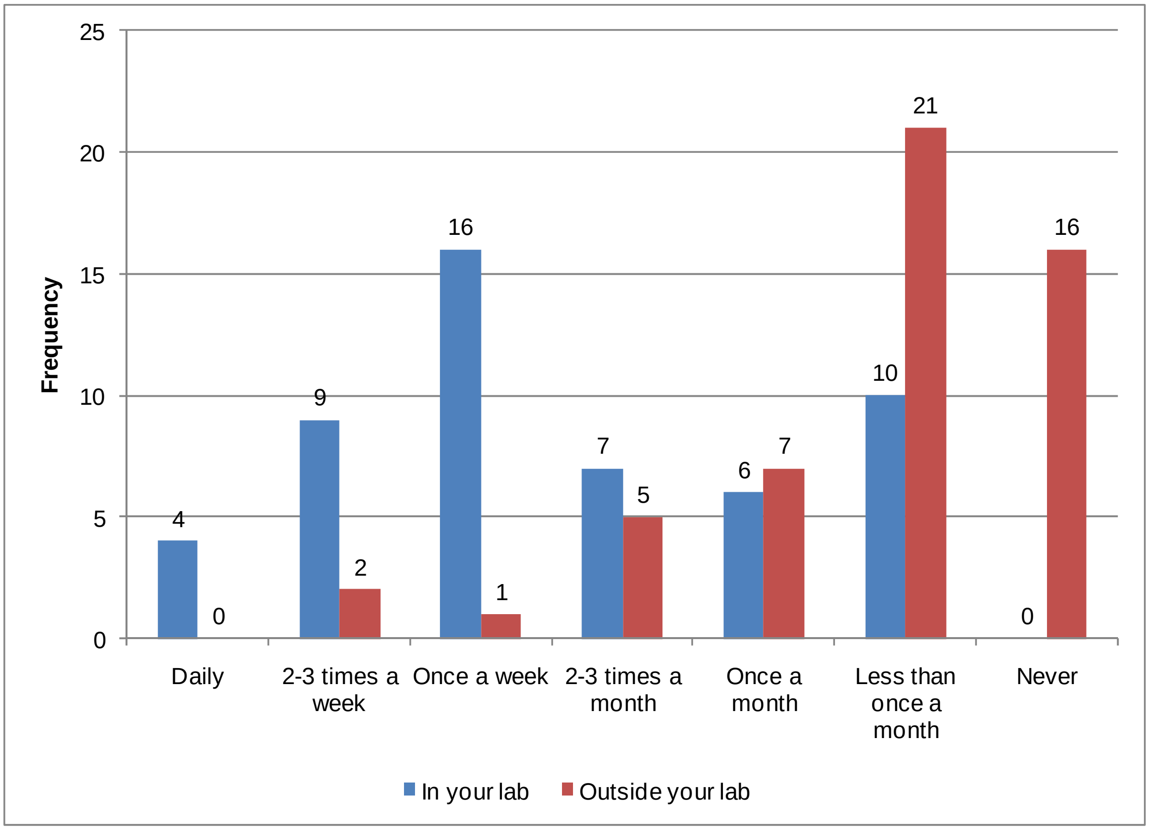

Figure 6: Frequency of Collaboration with Other Researchers

Figure 7 summarizes the various tools respondents indicated they used to support their collaborative and data sharing efforts with other researchers both inside and outside of their laboratory. In general, e-mail is the most prevalent method for sharing data with outside researchers, and e-mail and portable storage devices are the most prevalent for sharing data internally. Shared file storage, presumably provided by either the lab itself or the institution, was a close second for internal collaboration. Only one respondent indicated the use of an online document management system and no one indicated the use of blogs or wikis for collaborating either within or outside of their laboratory. This is especially interesting given the express purpose of these online tools is to support collaborative work.

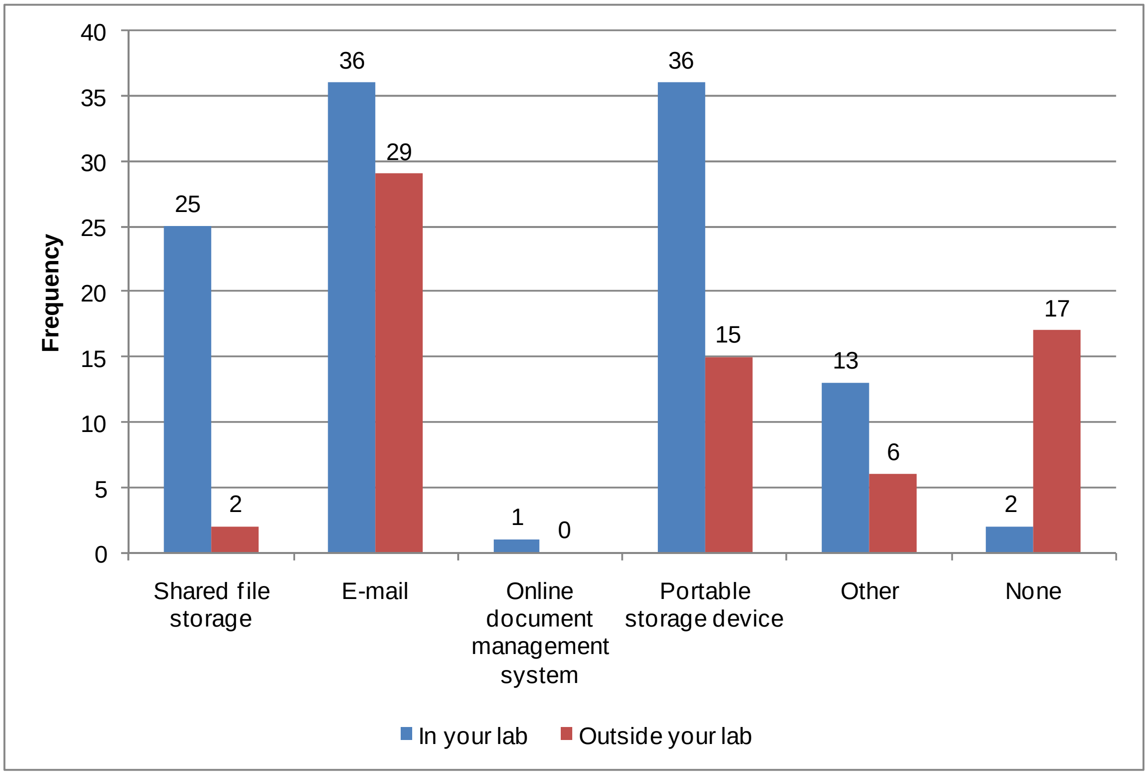

Figure 7: Tools Used to Share Research Information with Collaborators

Table 8 summarizes how effective respondents feel their current collaboration and data sharing tools are in supporting their needs. The mean response was 1.96±0.77 generally corresponding to a response of somewhat effective. Over three-quarters (80.7%) of respondents indicated they feel their current tools are either somewhat effective or very effective. Only 3.8% indicated a response of somewhat ineffective, and no one indicated a response of very ineffective.

Response	Frequency	Percent
Very Effective (1)	14	26.9%
Somewhat Effective (2)	28	53.8%
Neutral (3)	8	15.4%
Somewhat Ineffective (4)	2	3.8%
Very Ineffective (5)	0	0.0%

Table 8: Effectiveness of Current Collaboration Tools

Table 9 summarizes how open respondents are to exploring new electronic collaboration tools. The mean response was 2.29±0.86 corresponding to a response of slightly more neutral than somewhat open. Most (66.6%) respondents indicated they were either somewhat open or very open to exploring new technologies, but 7.9% indicated they were either somewhat against or very against exploring new tools. It is also important to note that approximately one-quarter (25.5%) of respondents were on the fence regarding this issue providing a response of neutral.

Response	Frequency	Percent
Very Open (1)	7	13.7%
Somewhat Open (2)	27	52.9%
Neutral (3)	13	25.5%
Somewhat Against (4)	3	5.9%
Very Against (5)	1	2.0%

Table 9: Openness to Exploring New Electronic Collaboration Tools

Respondents were also asked what concerns would prevent them from considering alternative collaboration tools. Figure 8 summarizes their responses.

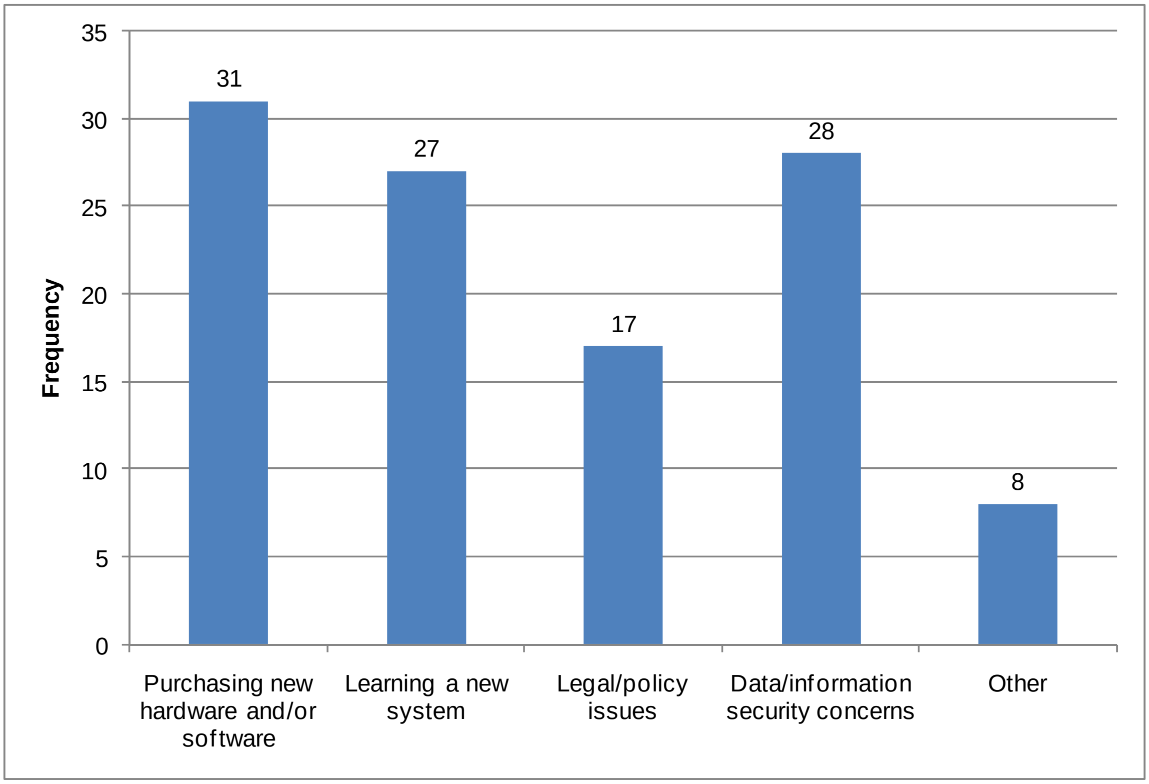

Figure 8: Concerns Preventing Consideration of Alternative Collaboration Solutions

Generally, a wide variety of concerns were expressed, but none of them stood out as overwhelming the others. Additional concerns respondents had were the ease of use of alternative tools, the longevity and reliability of the software and systems, and the time and effort required to transition from their current tools into a new system.

Analysis

Simple Chi-Square analyses were performed in order to ascertain if there were any significant differences amongst respondents' demographics and several key survey questions. The division of chemistry, length of time as a researcher, affiliated institution, and current position of each respondent were compared to their satisfaction with their current lab notebook, willingness to move to an entirely electronic lab notebook, effectiveness of currently used collaboration tools, and openness to exploring new electronic collaboration tools.

Table 10 summarizes the results of the Chi-Square analyses performed on the survey data.

Variable	Pearson Chi-Square	Degrees of Freedom	p-value
Traditional Chemistry Division			
Satisfaction with Current Lab Notebook	16.413	15	0.355
Willingness to Move to an Entirely Electronic Lab Notebook	27.208	20	0.130
Effectiveness of Current Collaboration Tools	13.81	15	0.540
Openness to Explore New Electronic Collaboration Tools	17.231	20	0.638
Length of Time as Researcher			
Satisfaction with Current Lab Notebook	6.402	9	0.699
Willingness to Move to an Entirely Electronic Lab Notebook	19.854	8	**0.011**
Effectiveness of Current Collaboration Tools	5.497	6	0.482
Openness to Explore New Electronic Collaboration Tools	19.031	8	**0.015**
Affiliated Institution			
Satisfaction with Current Lab Notebook	2.915	3	0.405
Willingness to Move to an Entirely Electronic Lab Notebook	11.952	4	**0.018**
Effectiveness of Current Collaboration Tools	1.264	3	0.738
Openness to Explore New Electronic Collaboration Tools	5.467	4	0.243
Current Position			
Satisfaction with Current Lab Notebook	12.084	9	0.209
Willingness to Move to an Entirely Electronic Lab Notebook	17.288	12	0.139
Effectiveness of Current Collaboration Tools	2.543	9	0.980
Openness to Explore New Electronic Collaboration Tools	34.485	12	**0.001**

Table 10: Summary of Chi-Square Analyses (significant p-values are highlighted)

The results show significant differences at the 0.05 level for four of the analyses performed. This indicates differences in the distribution of responses based on demographic data and is not indicative of differences in the mean responses observed. First, there is a significant difference between both a respondent's willingness to move to an electronic notebook system and their openness to explore new collaboration tools with their length of time as a researcher. It is important to note, however, that the majority of respondents (39 of 52) indicated the same length of time as a researcher (1-5 years). Given such little data in the other categories, it is difficult to generalize this trend without further information. A significant difference between a respondent's openness to explore new collaboration tools with their current position was also observed. Similarly, the majority of respondents (47 of 54) indicated the same current position (Graduate Student). Additionally, a significant difference was found between a respondent's willingness to move to an electronic notebook with their affiliated institution. Again, the majority of respondents (38 of 51) indicated the same institution (University of North Carolina at Chapel Hill). Further work is necessary to verify the trends observed and to determine why these factors are important. It is interesting to note that no significant differences were detected for any of the comparisons with a researchers' division of chemistry. Given the natural differences in division research work, certain differences were expected to fall along these boundaries.

Discussion

The results of this study will be examined and discussed in detail to provide insight into the original research questions posed as well as to highlight potential implications of this work and suggestions for further research.

Role of the Laboratory Notebook

Paper-based laboratory notebooks remain a researcher's primary means for recording their data and observations. Over 90% of respondents indicated they used a paper laboratory notebook when recording experimental notes. While the paper-based lab notebook is not dying off, additional electronic means of recording information are beginning to take hold. For example, 46% of respondents indicated they use Word documents to record laboratory notes. This is especially significant given the amount of data being collected in an electronic format. Over 60% of respondents indicated that more than 80% of their raw and analyzed data are collected in an electronic format. The ability to integrate this data with their laboratory notes for a given experiment would provide a significant improvement in the organization of their research materials.

Respondents indicated that they were generally confident in their own abilities to record experimental notes with sufficient detail; however, they were less confident that they would be able to interpret another researcher's notes. It seems vitally important to the research enterprise that these records be easily understood by more than the researcher recording them. Perhaps the expansion of collaborative research work would build additional confidence in researchers' abilities to understand the notes of others.

Similarly, respondents indicated it was generally easy for them to go back and retrieve laboratory notes they had previously taken but that it was far more difficult to do the same with another researcher's notes. Again, the ability to share information among researchers, both past and present, is extremely important to research productivity. An electronic laboratory notebook may be easier to search but will not improve its overall content. If retrieval is the underlying problem, an online system would be able to effectively index all researchers' laboratory notes with hopes of preventing overlapping work; however, if the underlying problem is the actual detail of the notes, researcher's will have to make a concerted effort to improve what they record. It is also important to mention that the organization of the laboratory notes may prove to be an obstacle in effective retrieval. This seems especially likely in this case given researchers seem to easily find what they need in their own organizational system but have difficultly when trying to search through another researcher's system. An electronic system could provide additional organizational structure which could alleviate this issue; however, any standardized organizational system among researchers would be able to accomplish the same goal.

In general, the laboratory notebook remains the most important tool of a researcher. It is the one resource that binds all of his/her work together and provides meaning to the experimental data collected. As the laboratory notebook evolves to meet the growing needs of researchers, an increase will be seen in the number of electronic tools being used to record and retrieve laboratory notes as well as the eventual integration into an entirely electronic laboratory. The corporate R&D community is certainly ahead of the game in this respect, but academia will catch up as they begin to see the need for increased organization of their data and notes to provide more effective resources for future researchers.

Current Information Management Tools

Besides the traditional laboratory notebook, respondents indicated the use of a wide variety of other resources to support their work. Journal articles, in both printed and electronic form, topped the list followed by spreadsheets and e-mail. Over 30% of respondents also indicated the use of both printed and digital images. With so many varying forms of information, it seems that it would be difficult for researchers to compile a high-level view of their data without considerable effort. Interestingly, respondents generally indicated that it was easy for them to search for and find electronic data files and information from previous experiments. Perhaps this implies an underlying organization of the information by each researcher. If this is the case, the problem remains that these organizational structures are not consistent making it fairly difficult for outsiders to find the information they need.

As mentioned above, a large portion of the data collected by researchers is already in electronic form; however, there does not seem to be any consistent location for the storage of this information. Nearly 80% of respondents indicated they store their data on their own personal computer. Although the survey did not ask whether this data is backed up, it is unlikely that everyone backs up their personal computers. This could potentially lead to data security and integrity problems not to mention what happens to the data when a researcher leaves. It is vital to the growth and productivity of academic research that a record of the data collected in a laboratory remains with that laboratory. Over 60% of respondents did indicate the use of laboratory owned machines to store their data, which is certainly a step in the right direction. Less the 40% indicated they used laboratory or university provided file servers for storage. This would be the safest and most reliable mechanism for the storage and archival of scientific data and should be encouraged as researchers explore new information management practices. Interestingly, nearly 20% of respondents also indicated the use of portable storage devices. Hopefully these are used as a secondary backup only, but either way these devices should be discouraged in the future as they could potentially violate the privacy and security of the data if they are ever lost or stolen.

Generally speaking, although the use of electronic data is prevalent throughout the scientific research community, researchers are skeptical of moving to an entirely electronic organizational system. The results show that more than half of respondents were either undecided or unlikely to move to such a system given the appropriate tools. Scientists still do not trust the longevity and reliability of electronic storage and will need to be introduced to its advantages over time. No system is perfect and hard drives will crash from time to time, but the use of redundant file servers as opposed to single personal or laboratory machines will reduce if not eliminate the risk of data loss.

Effectiveness of Current Practices

Overall, respondents indicated they were satisfied with their current information management practices and felt that these practices were effective in meeting their research needs. Nearly 70% of those surveyed indicated they were satisfied with their current laboratory notebook. More than 80% indicated they felt their current tools for sharing data with collaborators are effective. It seems that researchers are generally happy with their current toolkit and are unlikely to make any major changes to their current practices until they feel otherwise. Many did express openness to exploring new electronic tools to support their work; however, they have many concerns about moving away from their current tools. Their concerns generally stem from monetary issues such as needing to purchase new hardware and software or additional equipment and training as well as from data security concerns such as privacy and longevity. It is interesting that scientists, individuals always on the front lines of discovery and innovation, are so hesitant to adopt new technologies for the organization and storage of their information.

Implications

There is a great deal known about the information and data management practices and needs for corporate R&D environments, but very little is known about the academic environment. This is presumably due to the vast interest corporations have placed in understanding and improving their practices in recent years; however, this is also partly due to the wide array of differences between not only the two environments but between individual researchers and laboratories in academia. Researchers in academia have very specific needs due to their particular areas of interest within science and even their own discipline. This makes it difficult to provide a standard solution that will effectively support the varying needs of so many different researchers. By discovering what academic researchers are currently using to manage their data, it will be possible to suggest additional tools that may be useful to them and develop a better understanding of their needs in general.

The results of this research have the potential to influence the future development of ELN systems for universities and to drive scientists toward a more organized digital environment for storing and archiving their results. Archivists, university administrators, ELN software developers, and especially the scientists themselves will be interested in the results of this work. By understanding what academic scientists find useful, the university community can learn from each other how best to manage the massive amounts of data being generated by the scientific community in the modern age.

Additionally, this research has implications across several academic disciplines including information science, computer science, and the natural sciences. It will contribute to understanding the information organization and management needs of academic scientists. It will provide a profile of the current state of technologies being used to support scientific research and collaborative efforts in academia, and it will establish a desire among scientists to find and implement technologies which will better support their data management needs.

Future Work

Analysis of the survey results shows that the majority of respondents were graduate students, with 1-5 years of full-time research experience, working at the University of North Carolina at Chapel Hill. These results provide an interesting look at this population of subjects but do not provide generalizable data from which conclusions can be draw for all academic chemists. Future work should focus on obtaining a larger, more exhaustive sample. For example, using the American Chemical Society mailing lists as a sampling frame would provide a wide range of subjects from all areas of chemistry and all experience levels. This tactic was considered for this study, but was simply not feasible given the time and resources to complete the work. Additionally, the survey instrument should be revised to incorporate trends found in this study. For example, no one indicated the use of Web 2.0 collaboration technologies such as wikis or blogs. These items should thus be deleted and new ones added to incorporate more technologies specifically geared toward researchers. Finally, future work should concentrate on generalizing this work in hopes of someday applying it to all academic science disciplines. As the lines of disciplines continue to blur it will become increasingly more important that scientists from all fields be able to effectively communicate and share their resources with one another.

Other interesting issues worthy of further research include studying why there are so few external collaborations among academic research chemists. The literature indicates that collaborations are increasing and becoming commonplace in corporate R&D environments. Perhaps this lack of external collaborations parallels to collaborating between corporations in the private-sector. Whatever the case, it is certainly worthy of additional research. It would also be interesting to dig deeper into the area of commercial electronic lab notebooks. Two respondents indicated that they used systems like this and it would be valuable to find out why they switched to such a system, how hard the transition was, and generally how they like using it. This data could then be used to provide reasoning for or against switching to fully integrated ELN systems.

Conclusion

This study provides a description of the laboratory notebook's role in the scientific information management practices of academic research chemists. It also provides insight into the effectiveness of these practices as well as how current information management tools are being leveraged to support their day-to-day work.

A survey method was used to gather data via an online questionnaire, which consisted of 28 closed questions. Questions were developed to gauge how researchers utilize their laboratory notebooks, what additional tools they use for information management, and whether these tools effectively support research and collaboration.

The results indicate that the vast majority of research chemists still utilize paper-based laboratory notebooks as their primary means of record keeping and information management. Although respondents primarily use a paper-based laboratory notebook, many indicated that they also use additional, electronic tools to support their work. The use of Word documents and Excel spreadsheets were most common; however, a couple of individuals also indicated using a commercial ELN system. Surprisingly, collaborations do not appear to be as important in the academic domain as they are in the private-sector. The majority of respondents indicated they collaborate with individuals outside of their lab less than once a month if at all. Overall, the results showed that researchers were generally happy with their current practices and, as suggested by previous work, are reluctant to move to an entirely electronic system.

This work serves as a beginning point for developing a picture of scientific information management in the academic domain. Looking forward, additional studies should expand on this work to fully describe the role of laboratory notebooks and other information management tools in the day-to-day workflows of all scientific researchers. By fully understanding how scientists organize and use their information, advances can be made to ELN technologies that provide the features needed by academic researchers to ensure productive and efficient discovery in the digital age.

References

Babbie, E. (2007). *The practice of social research* (11th ed.) Thomson Wadsworth.

Birnholtz, J. P., & Bietz, M. J. (2003). Data at work: Supporting sharing in science and engineering. *GROUP '03: Proceedings of the 2003 International ACM SIGGROUP Conference on Supporting Group Work,* Sanibel Island, Florida, USA. 339-348.

Brown, C., Blake, C., Brown, E. A. R., & Tenopir, C. (2006). How chemists are really finding and using information in our digital environment. *Proceedings of the American Society for Information Science and Technology, 43*(1), 38.

Butler, D. (2005). Electronic notebooks: A new leaf. *Nature, 436*(7047), 20-21.

Dessy, R. E. (1995). Electronic lab notebooks: A shareable resource. *Analytical Chemistry, 67*(13), 428A.

Hemminger, B. M., Lu, D., Vaughan, K. T. L., & Adams, S. J. (2007). Information seeking behavior of academic scientists. *Journal of the American Society for Information Science and Technology, 58*(14), 2205-2225.

Kihlen, M. (2005). Electronic lab notebooks - do they work in reality? *Drug Discovery Today, 10*(18), 1205.

Myers, J. D. (2003). Collaborative electronic notebooks as electronic records: Design issues for the secure electronic laboratory notebook (ELN). Paper presented at the *Proc. of CTS'03,* Reimer, Y. J., & Douglas, S. A. (2004). Ethnography, scenario-based observational usability study and other reviews inform the design of a web-based E-notebook. *International Journal of Human-Computer Interaction, 17*(3), 403-426.

schraefel, m. c., Hughes, G. V., Mills, H. R., Smith, G., Payne, T. R., & Frey, J. (2004). Breaking the book: Translating the chemistry lab book into a pervasive computing lab environment. *CHI '04: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems,* Vienna, Austria. 25-32.

Tabard, A., Mackay, W. E., & Eastmond, E. (2008). From individual to collaborative: The evolution of prism, a hybrid laboratory notebook. *CSCW '08: Proceedings of the ACM 2008 Conference on Computer Supported Cooperative Work,* San Diego, CA, USA. 569-578.

Taylor, K. T. (2006). The status of electronic laboratory notebooks for chemistry and biology. *Current Opinion in Drug Discovery & Development, 9*(3), 348.

Appendix A: Survey Invitation

To: [Electronic Mailing Lists]
Subject: Invitation to participate in a study of chemists' use of laboratory notebooks

Study Title: Academic Chemists' Use of Laboratory Notebooks and Other Information Management Tools
Principal Investigator: Kyle Richardson (kylerichardson@unc.edu, (919) 259-1034)
Faculty Advisor: Dr. Diane Kelly (dianek@email.unc.edu, (919) 962-8065)
School of Information & Library Science, The University of North Carolina at Chapel Hill

Please consider participating in this study of how chemists in academia use laboratory notebooks and other information management tools. As a graduate student and a former research chemist, I understand your time is valuable and would be grateful for your assistance.

Below is a link to an online survey with 28 multiple-choice questions that should take between 10 and 15 minutes to complete. The survey will be available until May 11, 2009. Your responses will contribute to research work that will be published as a Master's Paper at The University of North Carolina at Chapel Hill.

I hope that you will consider helping me complete my graduate work by participating in this study. Please feel free to contact my faculty advisor or me with any questions you might have. Our contact information can be found at the top of this message.

Thank you for your time and consideration.

[Link to online survey]

Appendix B: Online Consent Form

Academic Chemists' Use of Laboratory Notebooks and Other Information Management Tools
IRB Study # 09-0721

Principal Investigator: Kyle Richardson (kylerichardson@unc.edu, (919) 259-1034)
Faculty Advisor: Dr. Diane Kelly (dianek@email.unc.edu, (919) 962-8065)
School of Information & Library Science, The University of North Carolina at Chapel Hill

Thank you for your interest in this study of the use of laboratory notebooks and other scientific information management tools by research chemists in academia. You have been selected for this study because you are a chemical researcher currently working in academia. Please do not complete this study if you are not an academic research chemist who has completed at least a bachelor's degree.

About this study:

- **What's involved:** The survey consists of 28 multiple-choice questions. It should take approximately 10 to 15 minutes to complete.

- **Risks:** This survey poses no more risk than you experience in normal daily living.

- **Benefits:** You may experience the satisfaction that comes with research and discovery, but you will not benefit otherwise from this study.

- **Your privacy:** By clicking to enter the survey, you are giving permission to use your data in this study. The results of this study will be published in a master's paper, but the paper will not contain information that will identify you. Your data will be anonymous. All the information you provide will be used responsibly and will be protected against release to unauthorized persons.

  Please be sure that you take steps to safeguard your privacy as well. Choose a place that allows you enough privacy to comfortably complete the survey.

- **Protection of survey data:** The Qualtrics system maintains data behind a firewall, and only the owner of the survey, who must provide password and user id, accesses the data. All pieces of data are keyed to that owner identification and cannot be accessed by anyone other than the owner or, by the owner's request, technical assistance staff. Technical assistance staff includes server administrators at Qualtrics who will respond to hardware or software failures, or Teresa Edwards, the UNC administrator for the Qualtrics Software Agreement. Ms. Edwards has completed Human Subjects Research certification at UNC-CH, and will only access survey data at the account owner's request.

- **Payment:** You will receive no payment or compensation for participating in this study.

- **Your rights:** You have the right to ask, and have answered, any questions you may have about this research. If you have questions, or concerns, you should contact the researchers listed at the top of this form. All research on human volunteers is reviewed by a committee that works to protect your rights and welfare. If you have questions or concerns about your rights as a research subject you may contact, anonymously if you wish, the Institutional Review Board at (919) 966-3113 or via e-mail at IRB_subjects@unc.edu.

- **Voluntary participation:** Your decision whether or not to participate in this study is voluntary and will not affect your standing at UNC-CH or any other institution. You may choose not to be in the study or to stop being in the study before it is over at any time. You will not be offered or receive any special consideration if you take part in this research.

If you click on the button below and submit a completed survey, you are indicating your agreement to participate based on reading and understanding this form. If you have any questions, please contact an investigator identified at the top of this form prior to completing the survey.

If you do not wish to participate in this study, please navigate away from this web page.

Based on the information above, I agree to participate in this study by clicking the "next" button below.

Appendix C: Survey Instrument

Section 1: Demographics

1. What is the highest degree you have obtained to date?
   - High School Diploma
   - Bachelor's Degree
   - Master's Degree
   - Doctoral Degree
   - Other Professional Degree

2. With what institution are you affiliated?
   - Duke University
   - North Carolina State University
   - The University of North Carolina at Chapel Hill
   - Wake Forest University
   - Other

3. What is your current position?
   - Graduate Student
   - Postdoctoral Researcher
   - Professor/Principle Investigator
   - Staff Scientist
   - Other (please specify)

4. How long have you been a full-time researcher?
   - Less than a year
   - 1 to 5 years
   - 5 to 10 years
   - More than 10 years

5. What traditional chemistry division most closely relates to your research work?
   - Analytical
   - Biological
   - Inorganic
   - Organic
   - Physical
   - Other (please specify)

Section 2: The Lab Notebook

6. In what format is your primary lab notebook?
   - Electronic
   - Hand-written
   - Other (please specify)

7. How satisfied are you with your current lab notebook?
   - Very Dissatisfied
   - Dissatisfied
   - Neutral
   - Satisfied
   - Very Satisfied

8. In which of the following ways do you record laboratory notes? (Please check all that apply)
   - Paper laboratory notebooks
   - Word documents
   - Blog entries
   - Wiki pages
   - Commercial electronic notebook system
   - Other (please specify)

9. On average, how promptly do you record experiments in your lab notebook?
   - Same day
   - Within two days
   - Within a week
   - Within two weeks
   - Over two weeks

10. In general, what level of detail would you say you use in recording experiments in your lab notebook?
    - Very Brief
    - Brief
    - Average
    - Detailed
    - Very Detailed

11. How confident are you that someone else could use your notes to reproduce an experiment?
    - Not Confident
    - Somewhat Confident
    - Confident
    - Very Confident

12. How often do you look back at your own lab notes that are more than 2 months old?
    - o Daily
    - o 2-3 Times a Week
    - o Once a Week
    - o 2-3 Times a Month
    - o Once a Month
    - o Less than Once a Month
    - o Never

13. When you do, how difficult is it to find the information you are looking for?
    - o Very Easy
    - o Easy
    - o Neutral
    - o Difficult
    - o Very Difficult

14. How often do you need to look at another researcher's laboratory notes or data?
    - o Daily
    - o 2-3 Times a Week
    - o Once a Week
    - o 2-3 Times a Month
    - o Once a Month
    - o Less than Once a Month
    - o Never

15. When you do, how easy is it to find the information you're looking for?
    - o Very Difficult
    - o Difficult
    - o Neutral
    - o Easy
    - o Very Easy

Section 3: Electronic Information

16. What forms of information, other than your lab notebook, contribute to your research? (Please check all that apply)
    - o E-mail messages
    - o Spreadsheets
    - o Word documents
    - o Printed photos or images
    - o Digital or scanned images
    - o Printed journal articles, manuscripts, etc.
    - o Electronic copies of journal articles, manuscripts, etc.
    - o Laboratory experiment results (electronic)
    - o Other (please specify)

17. What percentage of the *raw* data you collect are in an electronic format?
    - Less than 20%
    - 20% to 40%
    - 41% to 60%
    - 61% to 80%
    - More than 80%

18. What percentage of your *analyzed* data are in an electronic format?
    - Less than 20%
    - 20% to 40%
    - 41% to 60%
    - 61% to 80%
    - More than 80%

19. Where do you store your electronic data? (Please check all that apply)
    - Personal computer
    - Lab owned computer
    - Lab owned file server
    - University provided file storage
    - Other (please specify)

20. How easy is it to search for and find electronic research data from your past experiments?
    - Very Easy
    - Easy
    - Neutral
    - Difficult
    - Very Difficult

21. How likely are you, given the appropriate tools, to move to an entirely electronic lab notebook system?
    - Very Likely
    - Likely
    - Undecided
    - Unlikely
    - Very Unlikely

Section 4: Collaboration

22. How often do you collaborate or share data with other researchers *in your lab*?
    - Daily
    - 2-3 Times a Week
    - Once a Week
    - 2-3 Times a Month
    - Once a Month

- Less than Once a Month
- Never

23. What, if any, tools do you use to share research information with other researchers *in your lab*? (Please check all that apply)
    - Shared file storage
    - E-mail
    - Online document management system
    - Wiki
    - Blog
    - Portable storage device (USB drive, External hard drive, etc.)
    - Other (please specify)
    - None

24. How often do you collaborate or share data with other researchers *outside of your lab*?
    - Daily
    - 2-3 Times a Week
    - Once a Week
    - 2-3 Times a Month
    - Once a Month
    - Less than Once a Month
    - Never

25. What, if any, tools do you use to share research information with other researchers *outside of your lab*? (Please check all that apply)
    - Shared file storage
    - E-mail
    - Online document management system
    - Wiki
    - Blog
    - Portable storage device (USB drive, External hard drive, etc.)
    - Other (please specify)
    - None

26. How effective do you feel your current tools are in supporting the sharing of data and information?
    - Very Effective
    - Somewhat Effective
    - Neutral
    - Somewhat Ineffective
    - Very Ineffective

27. How open would you be to exploring new electronic tools for the sharing of research information amongst collaborators?
    - Very open

- Somewhat open
- Neutral
- Somewhat against
- Very against

28. What would prevent you from considering alternative solutions to research information sharing? (Please check all that apply)
    - Purchasing new hardware and/or software
    - Learning a new system
    - Data/information security concerns
    - Other (please specify)