TOWARDS A FRAMEWORK FOR MANAGING ELECTRONIC RECORDS IN SCIENTIFIC RESEARCH

BY KALPANA SHANKAR

ABSTRACT: This paper discusses scientific record keeping in the context of current theories of electronic records management. It describes the role of the laboratory notebook and the advent of electronic record keeping in documenting research. This paper also describes weaknesses of existing models of electronic records management with respect to scientific records. Gaps in the understanding of scientific records, organizational culture, and the warrant for scientific record keeping point to a need for developing a framework for evaluating electronic scientific records. The paper concludes with a proposal for further research into developing a framework that would take into account the problems described in the first part of the paper.

Introduction

A universal precept in scientific research is that scientists are dependent upon their written records of experimentation for the success of their work. The new student of science is taught that "proper" record keeping, traditionally done in a bound notebook, is the hallmark of reliability and integrity in research. In fact, the laboratory notebook is usually considered the only legally permissible documentation of research in the assignment of patents and in disputes of intellectual property.[1] A scientist "does science" in a laboratory notebook: it is here that problem framing, experiments, observations, methodology, interpretations, and references occur.

In spite of their long history of use, paper notebooks have their disadvantages. They are difficult to search, index, and manipulate. Recording repetitive experiments and calibrations is tedious. Other media of research, such as video images, are impossible to integrate into the paper notebook. In general, laboratory notebooks contain, in addition to text, graphics, photographs, printouts from instrumentation, samples, and other products of experiments.

To combat these problems and to fulfill other needs, scientists in many corporate and government laboratories have turned to "electronic laboratory notebooks." Such systems have been slow to catch on, but as technology has become more portable and less expensive, and the need to share information across time zones and continents has grown, electronic systems for capturing scientific data have become increasingly popular.[2] Even when scientists don't use formal record-keeping systems, they may keep parts of their laboratory notebooks on desktop computers and file printouts as appropriate.

However, the companies and individuals who have created these systems are generally scientists and systems designers, not records managers or archivists. As a result, these systems have often been optimized (or are being optimized) for the efficient capture of scientific data, without necessarily taking into account long-term archival and records management needs. The fact that neither the scientific nor the archival literature has confronted this problem suggests that the impact of electronic record keeping on science is unclear. A critical assessment of scientific records in the context of current electronic records research is in order.

This paper discusses the role of the laboratory notebook in scientific research and its importance as a record of scientific activity. It places the laboratory notebook in context with other scientific documentation and scientific practice. The paper suggests gaps in the existing electronic records research that need to be addressed if archivists are to help the scientific community take ownership of the electronic records management issues in their own work. Lastly, it proposes a mapping between archival practice and scientific culture that may help archivists and scientists communicate effectively to manage electronic (and paper) records in their purview.

The justification for such an assessment is clear. Scientific records have come under increased scrutiny from the stakeholders of scientific research: the public that pays for it, the scientists that engage in it, and the regulatory environment that governs it. Historians and archivists of contemporary science also lament the lack of documentation of contemporary science and have suggested that a scientific memory crisis is brewing.[3]

A general framework for interpreting scientific records could assist the archivists and records managers of contemporary scientific research in averting such a crisis. However, this paper will show that existing theories of electronic records management are inadequate for describing the records of science. Much of the rest of this paper will discuss scientific record-keeping practice in the context of two current frameworks and examine the shortcomings of those frameworks in assessing electronic scientific records. Based on these shortcomings, the paper outlines a proposal for developing a workable model for archiving scientific records.

Literature Review

It is helpful to start with definitions. What is a "laboratory notebook"? For the purposes of this paper, a "laboratory notebook" may be considered as the daily log of research kept by researchers and other technical staff in performing their scientific duties. Howard Kanare, who has written one of the most exhaustive manuals on the subject, goes so far as to suggest that "the properly kept notebook contains unambiguous statements of 'the truth' as observed by the scientist."[4] Most writers who have discussed proper record keeping concur that certain essential features of the scientific record must exist for the sake of completeness: research notes must be clear, complete, signed by the author, witnessed by another member of the group, and legible. The notes should be dated and kept in date order. Other names for the notebook are "research notes," "logbooks," and "laboratory records." For the purposes of this paper, the laboratory notebook can be electronic, paper, or a hybrid system.

There has been extensive research and writing on the preservation of electronic records in corporate and government environments. The two major theoretical frameworks in this arena are the University of Pittsburgh and the University of British Columbia electronic records projects.[5] While these two projects have taken radically different tacks in examining electronic records, practitioners who have tried implementing one or the other model have been forced to make compromises between theory and practice.

One of the first gaps between the scientific and electronic records communities rests in the definition of "records" itself. One of the definitions that Richard Cox of the University of Pittsburgh provides is as follows:

> Recordkeeping systems are information systems which are distinguished by the fact that the information they contain is linked to transactions which they document . . . Recordkeeping systems capture, maintain and access evidence of transactions over time as required by the jurisdiction in which they are implemented and in accordance with common organizational practices . . . Recordkeeping systems support functions of the organizations, and these functions require records of transactions in order to continue daily operations, satisfy administrative and legal requirements, and maintain accountability.[6]

He also indicates his support for Frank Upward's view that

> During the last thirty years there has been a steady shift to emphasizing the importance of data, and this has been at the expense of perceptions about the importance of documents, a term which temporarily was considered to be more applicable to paper records.

It is useful to contrast this definition with the scientists' use of the word "record," where the concept of transaction is vague. Kanare uses the word "record" almost synonymously with the word "data." He writes that the "laboratory *databook* [emphasis added] . . . contains a written record of the researcher's mental and physical activities from experiment and observation, to the ultimate understanding of physical phenomena."[7] Francis Macrina, who cites heavily from Kanare, concurs when he writes that "data contains handscript or affixed typescript which records and reports measurements, observations, calculations, interpretations, and conclusions."[8]

The archival community has been wrestling with the definition of "records" for centuries and continues to examine the applicability of those definitions in the digital environment. In the past, the scientific community has had less reason to come to grips with the definition of "record." However, this is changing as scientists in industrial, government, and academic environments have become concerned with issues of fraud and intellectual property protection. Since so much of the Pittsburgh model is based on its transactional definition of "record"—a definition that is distinctly at odds with scientific understanding (or lack thereof)—one can immediately begin to see that there may be some problems in directly using definitions found in the Pittsburgh model to manage scientific electronic records.

The University of British Columbia research project provides a more abstract and yet, perhaps, more useful definition of "record." The project leaders, Luciana Duranti and Terry Eastwood, rely on the centuries-old practice of diplomatics, a science developed during the seventeenth and eighteenth centuries for understanding the nature of documents and verifying their authenticity.[9] The diplomatics definition they adapted for the electronic environment is that a "record" was an archival document, or "a document produced in the course of practical activity."[10] Clearly, this definition is far more general than the Pittsburgh definition, which is legal and business oriented in its understanding. Although there are few scientists in any scientific discipline who know much about diplomatics, the diplomatics definition of records in electronic form is one with which they would be able to find some affinity.

However, the University of British Columbia's definition of "complete electronic record" would require some rethinking with respect to electronic scientific records. Duranti and Eastwood suggest that, for a record to be considered as having been created, "the entity needs to be created with the intent and the capacity of being communicated."[11]

The problematic phrase is "intent to communicate." The intent to communicate the primary records of science is murky. Kanare writes that the laboratory notebook should be considered a "forum" in which ideas are discussed.[12] Since a "forum" is a public discussion space, Kanare might argue that, indeed, the information recorded in a laboratory notebook, electronic or not, is intended to be communicated. Macrina elaborates on the "intent to communicate" when he writes that "corroboration and verification of scientific results using primary data contained in a laboratory databook may involve individuals other than the primary databook keeper."[13] The laboratory notebook will usually not be shared with others until the creator is ready for publication and, perhaps, not even then. In most environments, the organization (a university, a research laboratory, or a corporation) owns the laboratory notebook. Nevertheless, many scientists consider the notebook "their" space for reflection, observation, and calculation. Duranti and Eastwood argue that successful transmission of a documentary entity is necessary for the entity to be called a record.[14] So when does the "scientific record" become a "record"? The idea of transmission or intent to transmit may need to be refined to accommodate both practice and usage of the scientific meaning of "record."

One possible definition of a scientific record, based on the above discussion, might be "a document created during the course of research activity, written or recorded with the intent to allow others to replicate the activity being documented, outline the process of intellectual and practical actions that make up that activity, and to create documentation for the regulatory administration of that activity." This definition, or one similar to it, would incorporate those ideas that are germane to the archival and scientific professions and insure that the appropriate records are saved for administrative and legal purposes.

So far, we have shown that *what* is created is problematic when one discusses scientific records. What may be even murkier is *why* such records must be created and maintained. The project leaders of the Pittsburgh group argue that one of the most persuasive means of implementing the functional requirements they have outlined for electronic records is by appealing to authoritative sources that are trusted and respected by records creators. The researchers in the University of Pittsburgh project have used the concept of "literary warrant" to encompass these ideas. Literary warrant involves the documentation of formal mandates for record keeping that are the onus of the records creators and managers. These can include laws, industry regulations, best practices, standard operating procedures, and corporate rules. Wendy Duff suggests that compiling "literary warrant" can aid in implementing effective records management[15] and, to this end, the Pittsburgh staff hunted through legal, auditing, professional best practices, and information technology literature to collect passages directly in support of the functional requirements.[16]

However, literary warrant is limited by geography and scientific records creators are not. Laws that pertain in one country or state or institution with respect to records creation and retention will probably not pertain in the next. This is particularly problematic for contemporary scientific endeavors. For example, Joan Warnow-Blewett writes that in the discipline of high-energy physics, "members of a single collaboration may . . . come from a score of institutions based in several countries."[17] This organizational structure is not unique to high-energy physics: multinational collaborations are common in many scientific disciplines. Professional best practices suffer the same fate as legal warrant, since members of the same profession working in different countries may not be bound by the same professional codes. "Best practices" will also mean less in a multidisciplinary collaboration, another common organizational framework in research science. The best practices of the physicist will not be the same as those of the biologist working with that physicist on a research project. "Best practices" may also have different meanings for the industrial chemist in partnership with an academic chemist, another common research framework. And in newly established fields of scientific research, best practices may not even have been formally established.

The term "auditing" also has different meaning in the scientific realm than it does in the activities for which Pittsburgh has created its model of literary warrant. Scientific regulators tend to speak more of "data audit"[18] than of financial or legal audit. Although the term can have multiple meanings in the regulatory context, Adil Shamoo suggests that one definition of "data audit" is

> . . . the systematic process by which objective evidence is obtained and evaluated as to assertions about research data and their value to determine the degree of correspondence between those assertions and established or predetermined criteria which can then be communicated to interested parties.[19]

In other words, the purpose of data audit is to insure that the assertions that are made about the results of research actually are derivable from the records of that research. While the concept of data audit has been promulgated for a number of years[20], its use has been limited to the corporate arena and a few government agencies. While industry usually requires that a peer of the researcher witness the laboratory notebook on a regular basis, there are no other "universal" principles of data audit that apply equally to all areas of scientific research, hence, the difficulty of using "data audit" to compile warrant. Moreover, many scientists have been strongly suspicious of data audit and have fought against it on the grounds of its intrusiveness and the increased paperwork that formal data audit generates.[21] In short, "auditing" may not be a term with which archivists will gain authority in the scientific community.

Deriving literary warrant for scientific records is a project in and of itself and clearly was not in the scope of the Pittsburgh project. In spite of the difficulties inherent in compiling literary warrant for scientific record keeping, there still are universal precepts that can be used to answer the question of "why." The strongest and most universal warrant for record keeping in scientific research is a form of "best practice": many scientists would suggest that the "best practice" of scientific research is scientific method.[22] Science derives its reliability from the ability to be replicated: if a given experiment can produce the same results each time the method is followed, then the experiment is considered reliable. Within the context of the scientific method, the ethical scientist must maintain reliable records of his or her work so that others may replicate it. Literary warrant for electronic record keeping, however, will need to be compiled and documented based on the exigencies of specific research environments.

While reliability of the records and the record-keeping process is one of the hallmarks of scientific ethics, "ethics" is not often discussed in the electronic records literature. Ethics investigations and congressional hearings have often focused on the inadequacies of scientific records. Kenneth Freedland and Robert Carney suggest that the ethics of scientific records management are bound up in quality control and, thus, mandate accurate record keeping. While the term "best practices" is not common in the scientific literature, "accountability" is. Scientific research is funded by institutions and heavily supported by government agencies or corporate entities; therefore, records creators are accountable to these bodies. The warrant for scientific record keeping lies in the ethics regulations that govern scientific research, but as yet these regulations and ethics have not found a home in the Pittsburgh model of literary warrant. This is not to suggest that the Pittsburgh model of compiling warrant would be inapplicable, just that, as yet, there has been no systematic progress on compiling scientific warrant that could serve industrial, government, and academic science alike.

Much of the need for good record keeping in science is also being driven by recent trends in intellectual property. The assignment of a patent, for example, is based on "first to invent," not "first to file," at least in the United States. The need and desire to patent has become more widespread; even genes and vegetables can be patented. The record-keeping requirements outlined by the United States Patent and Trademark Office binds those who wish to file patents in the United States. In a world where science has become a lucrative commodity, the rewards of keeping good laboratory records can be great. Archivists may find that possible financial reward for functional and complete records is the strongest warrant of all. More likely, however, they may find that fear is an even stronger motivator: highly publicized cases of fraud and misconduct in scientific research have many scientists examining their own record-keeping practices in a critical light.[23]

If there are difficulties answering the "What?" and "Why?" of scientific records, the next question we ask of the literature could be "Who?": Who are the records creators and how does their organizational culture fit into electronic record keeping? Warnow-Blewett, in her documentary analysis of high-energy physics collaborations, found that many of the daily activities were being performed by staff, students, and junior-level scientists.[24] This is confirmed by Helen Samuels who writes, "For collaborative research of all kinds and the majority of scientific and technological efforts, project leaders work as part of a multi-layered team comprising researchers, administrators, and technical assistants . . . . The research staff is made up of graduate students, postdoctoral fellows, and research assistants . . . ."[25]

While identifying records creators may be fairly straightforward, understanding the organizational culture of research may not. It is necessary to understand competing models of electronic records use in order to appreciate the difficulty. For some time, there have been competing models of electronic records use. Duranti argues that electronic records have a "life cycle" comparable to traditional records: records are most heavily used around the time they are created and then drop off in use. At this point, they should move into the archival bond, which Duranti defines as "the originary, determined, and net of relationships that each record has with the records belonging in the same aggregation."[26] While the research science community would probably agree with Duranti's life cycle in theory (scientists tend to use their records most intensively after creation, with use dropping off with time), her emphasis on the archivist as gatekeeper would probably not be in concert with the community's need for self-regulation of its records.

Duranti's view is in direct opposition to the Australian model, which argues that electronic records may follow more of a "continuum" pattern, with cycles of use and nonuse by the creators.[27] Upward suggests that the continuum model creates some new principles for the age of electronic records: an expanded definition of "record," systematization of requirements for record keeping (related to the Pittsburgh Functional Requirements approach), and most importantly for research science records, strong emphasis on the integration of records management into the social processes of the research organization.[28] The post-custodial model argues for the records manager to mediate the local management of records at the site of creation. This approach may be effective in government agencies where knowledge of records management is localized, but the applicability of this approach in a community without records management training has yet to be tested.

In short, scientific records do tend to be used intensively at the time of creation, then drop off in use, and in this regard the life cycle concept would seem most applicable. Cycles of use and nonuse are discipline driven and perhaps best understood by the community itself. But moving records physically into an archives probably will not work in some scientific arenas (such as academia) for cultural reasons. In this regard, the continuum approach, with its emphasis on self-governance and education for records creators, is likely to be more applicable in a scientific community that has traditionally been quite protective of its records. So, there does not appear to be a pat solution to the issues of custody or other electronic records issues, either in the scientific arena or in any other. It would seem, rightly, that scientific practice and scientific records share some aspects of both the life cycle model and the continuum model, at least as articulated by their respective adherents. Perhaps a combination of custodial and post-custodial approaches, mediated by archivists who are sensitive to the evidentiary, regulatory, and long-term cultural import of scientific records, would work best in the research environment. There is a great deal of exploratory work to be done in this arena.

Scientists certainly would not find themselves in concert with Duranti's custodial model of electronic records appraisal, where the authenticity of records lies in their transfer to physical archives.[29] Many scientists are skeptical of "outsiders" (i.e., non-scientists) who claim to know the value of the records created by "insiders."[30] Although many institutions have formal procedures for archiving laboratory notebooks and centralizing their custody, at most noncorporate institutions these procedures are seldom followed. Moreover, there appear to be very few institutions that have established any procedures for archiving electronic records of science. As noted above, fears of the potential loss of intellectual property may also preclude scientists from trusting the records managers and archivists who wish to implement electronic records management programs. Very few scientists would agree with Duranti's assertion that the final disposition of records to the archives is what confers their authenticity. To the scientific community, the authenticity of records lies in scientific method, not in their archival status. Macrina sums up this position thus:

> Authentic data represent the true results of work and observations. When data deviate from this standard because of carelessness, self-deception, or deliberate misrepresentation, they lose their authenticity.[31]

Tackling these gaps of current practice in electronic records management when applied to scientific records constitutes only a preliminary discussion of the problems. The literature on electronic records has not focused on scientific records and, in turn, the users and creators of electronic laboratory notebooks or other record-keeping systems have not been exposed to the relevant literature on electronic records management. In the next section, I will suggest that a general framework for scientific records management can be derived from an understanding of these deficiencies.

Proposal and Methodology

There exists a need for a framework of electronic scientific records appraisal that accommodates the needs, assumptions, and practices of scientific research.[32] Some of these issues were discussed above and can be summarized as follows:

- Any framework that is derived must work for both paper-based and electronic records, since most scientists still use a hybrid record-keeping system.
- The archival definition of "records" needs to be reassessed and perhaps expanded in light of the use of the term "records" in the scientific community.
- The warrant for record keeping should include accountability, ethics, defense of intellectual property, and quality control as these concepts apply in the conduct of science.
- The organizational culture of scientific practice needs to become cognizant of record keeping as understood by archivists and records managers.

- The scientific community may need an alternative to the archival custodial model of records management. In general, the scientific community is suspicious of external management of its records.
- The scientific meaning of authenticity and reliability of records needs to be incorporated into any framework for electronic record keeping.

Much of the literature of electronic records management is not easily understood by the creators of the records. Therefore, a bridge is needed to span the gap between the archival and scientific communities. One possible framework is journalistic: the six questions of the reporter can provide an easily understood translation of difficult archival concepts. This mapping could make implementation of electronic records projects easier and more acceptable in the scientific community. These journalistic questions can be summarized thus:

- Why is the long-term management of scientific records important?
- Who is creating scientific records and who is using them?
- What records are they creating?
- When are records being created and when should they move into the archival bond?
- Where are records being created and where does the responsibility for their management lie?
- How are records stored?

At first glance, these questions are nothing new for the practicing archivist. Archivists are trained to consider these questions of context whenever they examine any record-keeping system, whether paper or electronic. However, these questions are new to the majority of the scientific community, many of whom may not be familiar with a formal records management program at all. What follows is a translation of these questions of archival processes and culture into the vernacular of contemporary scientific discourse. The culture of contemporary science, particularly since World War II, has changed dramatically from the work of one individual scientist in one laboratory working on one project. Instead, large-scale multi-institutional collaborations, with records in electronic and paper form distributed in numerous bureaucratic and scientific environments, are the norm. Making electronic records management work in the complex, distributed, highly regulated, and competitive arena of contemporary science may require understanding of scientific culture, vocabulary, and practice.

Why are records important?

The first step in undertaking an electronic records management program is a consciousness-raising step: participants in the function or organization being studied must be made aware that they are generating records and that those records need to be complete and accurate. In scientific research, this amounts to dated, authenticated research notes (whether electronic or paper) and related documentation. Even beginning scientific researchers need to be inculcated with the message that their records are important. Warrant stems from three principal sources. First, the ethics and practice of research demand accurate records. This argument can be bolstered by professional guidelines (where they exist) and the pedagogical literature. Second, research organizations and funding agencies have their own policies that must be met. These policies tend to center on protection of intellectual property rights, protection of human and animal subjects, and hazardous material use. Third, warrant can be compiled from the intellectual property sector. Patent law and other legal documents pertaining to the assignment of data ownership come into play. If research is undertaken in a corporate environment, the company will have another set of guidelines pertaining to records (usually drawn up as a result of the other warrants). These guidelines will insure that records will be compliant with laws and complete for the sake of research needs and ethics.

Who is creating records and who is using them?

As noted earlier, contemporary scientific practice, particularly in the academic arena, often employs students, faculty, and staff, with the work distributed over numerous offices and laboratories in both the private and public sectors. For example, university records managers may need to interview a number of scientists, staff, students, and administrators to understand the universe of electronic records that are being created and their relationship with the paper laboratory notebooks that are generally still the primary scientific record. Even if the archivist cannot collect all of the records, the awareness that the records are distributed in nature will help place the records that are collected in a broader context of institutional records. As for the needs of both primary users of scientific records (scientists and regulators) and the secondary users (archivists, historians, and sociologists of science), this is a subject that deserves further analysis. Almost no research has been done on the subsequent, post-publication uses to which scientists and others put their laboratory notebooks and related records.

What records are they creating?

Once the records creators are identified, the archivist must understand what records are being created. A broad understanding of the term "record" must be in force. Here, some observation and understanding of the research being undertaken will help. Raw and derived data, research notes, observations, communications about the project itself, grant reports, logbooks, instrumentation recordings, images, and sound recordings may all be part of the laboratory notebook. The importance of the preservation of these individual components will depend upon the discipline. The subject of the project, the participants, the dates, and the format of data, should help to place records in context.

When are records being created and how long do they need to be maintained before they move into the archival bond?

Archivists can look to documentation and functional analysis studies of archiving the records of science to understand the process of records creation and maintenance needs. In traditional records management, records managers examine organizational hierarchies to determine from which levels and individuals records need to be managed. By contrast, Samuels and others have suggested that the complexity and "flattening" of contemporary organizations require new approaches to documentation. Samuels used the term "functional analysis" to describe this approach. Instead of using the organizational chart alone, an archivist or records manager, using functional analysis, approaches the task of records management by focusing on the functions performed within and by the organization. He or she then assesses the documentation produced as by-products of that function.[33] This approach may be particularly successful in scientific enterprises and may help in furthering understanding of the evidentiary needs that scientific records must fulfill.

Where are records being created?

Most research will probably generate a combination of paper and electronic records. Proprietary software for recording laboratory data, Intranets, World Wide Web pages, instrumentation recordings, E-mail, word processors, spreadsheets, databases, and videoconferencing systems can all generate electronic records. Systems for managing digital records should be integrated with the paper laboratory notebooks and research notes that are still in use by the vast majority of scientific research projects. In many cases, the software that generated data or recorded it should also be kept, since most electronic records would be unreadable without the software.

On a broader level, *where* records are created can cross time zones and continents. The archivist must contend with the trans-national nature of collaborations. Scientific collaborations also cross the boundary between the private and public spheres, making the responsibility for managing the records murky. Clarifying the responsibilities and scope of the archivist's role is a delicate and difficult process in the case of these collaborations.

How are records stored?

The archivist must then determine how records are to be retained. Policies generally exist for the disposition of paper laboratory notebooks, but the archivist must examine whether this policy is effective for electronic records and hybrid systems. Whatever formats are chosen for records retention, archived records will need to be eye readable or easily made so. How long records are to be retained will have an impact on how records are maintained, since different media have different shelf lives.

Lastly, the archivist must contend with where records are to be retained. Answering this question will depend upon policy and pragmatism. Different projects have different access needs and, thus, may need to maintain custody for different lengths of time. In some research organizations, there may be policies for archiving laboratory notebooks, but in practice these policies may never be implemented. Understanding the organizational culture of the institution at which research is being performed, the demands that are made upon research materials, and the data ownership policies that pertain to the records in question will help clarify where and how records are to be kept.

Clearly, this framework needs extensive research and testing to see if it provides a good map between the archival concepts and the scientific vernacular. However, it could be a big step forward for managing electronic records in research science and, most importantly, for gaining the enthusiastic acceptance of scientists for records management. One of the advantages of this framework is flexibility. The principles outlined can be applied to any scientific discipline. The model is also scalable: the same general steps could be followed to appraise the records of a lone researcher or a transnational collaboration. There are no pieces that pertain exclusively to electronic records. Electronic record-keeping systems are not universal and not standardized: any mechanism for managing scientific records must work across a variety of software and hardware platforms and permit the integration of paper records. If each of the steps is documented appropriately, the records will reside in a rich contextual environment.

Most importantly, the framework provides a way to relate arcane concepts to the nonarchivally trained (and suspicious) creator of records. Much of the success of records appraisal rests on the acceptance of records management by the records creators. Scientists have traditionally been wary of permitting their records to fall into the hands of the nonscientifically trained and are increasingly afraid of loss of ownership. Improving communication between archivists and records managers and the scientific community will help meet the needs of both professional communities.

Conclusion

This paper began with a discussion of the primary record of scientific activity, the laboratory notebook, and the problems for electronic records management that all electronic scientific records can engender. It examines why current thinking on electronic records is inadequate for the appraisal and disposition of scientific records. The problems begin with divergent uses of the word "record" in the archival and scientific communities, but do not end there. Much of the research that has been done on the management of electronic records has been for business and government records. There needs to be more research into the management of electronic records for the scientific community, since federal regulations governing scientific research often mandate extensive retention periods.

The laboratory notebook in paper or electronic form is contested ground. Because of its multiple roles as a stage for scientific inquiry, intellectual property, regulatory document, and historical document, there exist numerous professional communities that have a vested interest in the long-term preservation of laboratory notebooks. Increasingly, as scientists and organizations move towards formal electronic laboratory notebooks, the scientific, archival, and regulatory communities will need to communicate to make sure that such notebooks serve the purposes of all in ways that are as intuitive and long lasting as the paper laboratory notebook and as sophisticated (if not more so) as the desktop computer. Even when formal electronic laboratory notebook systems are not being used, scientists and scientific support staff may still be generating electronic "laboratory notebooks" in a more ad hoc fashion. Eliciting the whereabouts and importance of these ad hoc electronic records may require yet more communication between archivist and scientist.

This framework does not address the integration of formal electronic laboratory notebooks with other information systems such as data analysis packages, file management systems, electronic document management systems, and imagebases. Such integration adds more complexity to the electronic records management issues and can make long-term preservation of records even more difficult.

Ultimately, the success or failure of electronic records management in science lies in the ability to bridge a cultural and language gap. In short, archivists must learn to speak to scientists and scientists must be made aware of the work of the archivists. However, many issues are raised about the role of the professional records manager and archivist in the management of contemporary records of science, as well as the responsibilities of the records creators.

Both communities must be made aware that they are documenting the cultural heritage of contemporary science as well as evidentiary records. The archivist of electronic scientific records must consider context as well as content (an issue that archivists are professionally trained to consider), a point that may be lost upon the scientific community. However, the scientist is much more familiar with the environment in which those records are created, and this environment influences policy.

Some authors have suggested that scientific record keeping has changed with the advent of digital media, but to date no one has assessed the extent of these changes.[34] There is interest in the evolving practice of electronic record keeping, but the gap between the scientific community and the archival and records management communities has resulted in a lag in research into scientific electronic records management. Addressing these problems would have an impact on many communities, not the least of which are scientists and records managers. Although the primary beneficiaries would be scientists and records managers, the level of coherent documentation that will be created will prove useful to historians and sociologists of contemporary science. I suggest that a multiplicity of methodological approaches, archival and sociological, will prove to be the most powerful approach for creating a useful and usable framework for scientific records management in the digital age.

ABOUT THE AUTHOR: Kalpana Shankar is a doctoral student working with Dr. Anne Gilliland-Swetland in the Department of Information Studies at the University of California, Los Angeles. She has a bachelor of arts degree in molecular biology from Princeton University and graduate training in biophysics from the University of Pennsylvania. She has worked in software development, indexing and thesaurus construction, and database design and systems analysis. She also worked as a graduate fellow at CENSA, the Collaborative Electronic Notebook Systems Association, in Woburn, Massachusetts, and is currently a research assistant on InterPARES (International Project on the Preservation of Authentic Records in Electronic Systems).

NOTES

1. Howard M. Kanare, *Writing the Laboratory Notebook* (American Chemical Society, 1985): 2.
2. CENSA, the Collaborative Electronic Notebooks Systems Association, in Woburn, Massachusetts, is dedicated to developing accountable record-keeping systems for corporate research and development. More information can be found at <http://www.censa.org>.
3. Gavan McCarthy and Tim Sherratt, "Mapping Scientific Memory: Understanding the role of record-keeping in scientific practice," *Archives and Manuscripts* 24 (1996): 1.
4. Kanare, 1.
5. Both projects are too extensive to discuss in this paper. As needed, concepts from both will be selected and articulated in this paper. The University of British Columbia project has evolved into a larger project of international scope and participation, the InterPARES (International Research on Permanent Records in Electronic Systems) project. For more information, see <http://www.interpares.org>.
6. Richard Cox, "The Record: Is It Evolving?," *Records & Retrieval Report* 10 (March 1994), retrieved June 8, 2000, from the World Wide Web: <http://www.lis.pitt.edu/~nhprc/Pub15.html>.
7. Kanare, 7.
8. Francis L. Macrina, *Scientific Integrity: An Introductory Text with Cases* (Washington, DC: American Society for Microbiology, 1995): 43.
9. Luciana Duranti and Terry Eastwood, "Protecting Electronic Evidence: A Progress Report on a Research Study and Its Methodology," *Archivi & Computer* 3 (1995): 213.
10. Luciana Duranti, Terry Eastwood, and Heather MacNeil, "The Preservation of the Integrity of Electronic Records" project. *Template 1: What is a Record in the Traditional Environment?* (March 5, 1997), retrieved June 8, 2000, from the World Wide Web: <http://www.slais.ubc.ca/users/duranti/tem1.htm>.
11. Duranti, et al.
12. Kanare, 5.
13. Macrina, 43.
14. Duranti, et al., "Preservation." *Template 5: What Is a Record in the Electronic Environment?* (March 3, 1997), retrieved December 3, 1997, from the World Wide Web: <http://www.slais.ubc.ca/users/duranti/tem5.htm>.
15. Wendy Duff, "Ensuring the Preservation of Reliable Evidence: A Research Project Funded by the NHPRC," *Archivaria* 42 (fall 1996): 37.
16. Kimberly J. Barata, "Functional Requirements for Evidence in Recordkeeping: Further Developments at the University of Pittsburgh," *Bulletin of the American Society for Information Science* 23 (June/July 1997): 14.
17. Joan Warnow-Blewett, "Documenting Postwar Science: The Challenge of Change," in *Recovering Science: Strategies and Models for the Past, Present, and Future Proceedings of a Conference Held at the University of Melbourne November 1992*, ed. Tim Sherratt, Lisa Jooste, and Rosanne Clayton (Australian Science Archives Project, Canberra, 1995), retrieved December 3, 1997, from the World Wide Web: <http://www.asap.unimelb.edu.au/confs/recovering/blewett.htm>.
18. It has been noted that the scientific community tends to use "data" and "record" almost interchangeably.
19. S.E. Loeb and Adil E. Shamoo, "Data Audit: Its Place in Auditing," in *Principles of Research Data Audit*, ed. Adil E. Shamoo (Gordon and Breach Science Publishers, 1989): 19.
20. Kenneth E. Freedland and Robert M. Carney, "Data Management and Accountability in Behavioral and Biomedical Research," *American Psychology* 17 (May 1992): 641.
21. Adil E. Shamoo and Zoltan Annau, "Data Audit – Historical Perspectives," in *Principles of Research Data Audit*, ed. Adil E. Shamoo (Gordon and Breach Science Publishers, 1989): 8.
22. It can be argued that the "universality" of scientific method, however, has been called into question by contemporary critics of science, particularly by postcolonial and feminist scholars. However, examining the impact of these critiques of modern science on the record is outside the scope of this paper.
23. In recent years, the most highly publicized investigation of scientific fraud centered on the work of Theresa Imanishi-Kari, a colleague of David Baltimore, the 1975 Nobel Laureate in Medicine. After more than ten years of federal investigation, in which the Secret Service examined the laboratory notebooks of all concerned, all were exonerated of fraud. For an extensive treatment of the subject,

consult Daniel J. Kevles's recent book. *The Baltimore Case: A Trial of Politics, Science, and Character* (Norton, 1998).

24. Joan Warnow-Blewett, "Commentary," *American Archivist* 57 (winter 1994): 116.
25. Helen W. Samuels, *Varsity Letters: Documenting Modern Colleges and Universities* (Metuchen, NJ and London: Society of American Archivists and The Scarecrow Press, Inc., 1992): 116. The author of this paper notes that the hierarchy extends even further, having been personally involved in university-affiliated research projects as a high school student and later as a university student.
26. Luciana Duranti, "The Preservation of the Integrity of Electronic Records," *Proceedings of the DLM Forum on Electronic Records Research.* Brussels, Belgium, 18–20 December 1996 (European Commission, 1997): 61.
27. Hans Hofman, "Off the Beaten Track: the Archivist Exploring the Outback of Electronic Records," in *Playing for Keeps: the Proceedings of an Electronic Records Management Conference hosted by the Australian Archives*, ed. Stephen Yorke. Canberra, Australia, 8–10 November 1994 (Australian Archives, 1995), <http://www.aa.gov.au/AA_WWW/AA_Publications/P4K/Hofman.htm>. The model that Hofman suggests is adapted from the Australian Information Exchange Steering Committee's (IESC) publication, *The Management of Electronic Documents in the Australian Public Service* (1993).
28. Frank Upward, "Structuring the Records Continuum Part One: Post-custodial principles and properties," *Archives and Manuscripts* 24 (1996): 277.
29. Duranti, *Proceedings*, 60–65.
30. Shamoo and Annau, "Data Audit: Its Place in Auditing," 8.
31. Macrina, "Scientific Record Keeping," in *Scientific Integrity,* 44.
32. Some might argue that left out of this proposal are the issues related to the documentation of associated records that are not directly products of research, but support the research: administration of research and published dissemination of results. The existing models of electronic records management (especially the Pittsburgh project) are more suited to the administration and business of research than documenting the research itself, for one. Furthermore, Haas, et al., and Samuels have already done extensive work on documentation of these research-related functions.
33. Samuels, 5.
34. An interesting study would be to examine the practices of record keeping over the last 30 years and note differences and trends in different disciplines.