the Beckman Center for the History of Chemistry
AMERICAN CHEMICAL SOCIETY • UNIVERSITY OF PENNSYLVANIA •
AMERICAN INSTITUTE OF CHEMICAL ENGINEERS

The Arnold and Mabel Beckman Center for the History of Chemistry
3401 Walnut Street • Philadelphia, PA • 19104-6228 • USA
Telephone: (215) 898-4896 • Telefax: (215) 898-3327

Director: Arnold Thackray

M. Susan Lindee
5 December 1990

NARRATIVE
OF A PROPOSAL

**DOCUMENTING THE HUMAN GENOME INITIATIVE:
RECORDS RETENTION TO INFORM
SCIENCE POLICY AND ENHANCE HISTORICAL RESEARCH**

Introduction

The Human Genome Initiative offers an unparalleled opportunity to apply innovative archival methods to an important project in contemporary biomolecular sciences. The Arnold and Mabel Beckman Center for the History of Chemistry proposes to develop a documentation strategy that will improve records collection and retention for the genome project, that will help archivists efficiently allocate limited resources, and that will generate materials for future policy analysis, bio-ethical case studies and historical investigations.

The genome project raises questions concerning bioethics, multi-institutional collaboration, the role of technology in the production of knowledge, "big science" in biology, the interaction of instrumentation and conceptualization, and the impact of political priorities on scientific research. Addressing these questions will require the input of the scientific community, the public, and of scholars in the humanities and social sciences. All these groups therefore have a stake in the compilation and preservation of relevant materials.

Identifying and preserving a meaningful record of contemporary sciences has proven to be an extremely complex task. Archivists have begun rethinking methods traditionally used to identify records of enduring research value. Many now promote the need for "documentation strategies" to guide the selection of materials for preservation (Warnow-Blewett, 1987; Haas et al, 1985). Such strategies can focus on particular fields, issues or problems, on particular institutions, or on geographical locations. The American Institute of Physics study of Department of Energy document preservation policies illustrated the value of

efficient record-keeping, not only to historians, but to policy makers and administrators as well (Warnow et al, January 1982).

As "history in the making," the Human Genome Initiative is generating more materials than can possibly be preserved in toto. A documentation strategy will ensure that a usable record of the genome project in all its facets will be available to science policy analysts, government and corporate researchers, philosophers, sociologists, historians and other students of contemporary science.

Our goal will be to help preserve documentary materials that will facilitate understanding of the Human Genome Initiative in all its aspects, from its scientific innovations to its ethical dilemmas.

Specific Aims

The Arnold and Mabel Beckman Center for the History of Chemistry staff, guided by an expert committee and a key group of consultants and advisers, will:

* review existing records retention practices in institutions in which Human Genome Initiative research is underway.

* assess past, present and future uses of such records.

* prepare a series of reports outlining the roles of records creators, archival programs and potential users of these records in improving the documentary records.

* prepare a formal documentation strategy that will draw on this research, proposing policies and practices to improve archival documentation of the project.

* work with scientists, archivists, clerical staff and historians to implement these preservation strategies, sponsoring outreach programs (including seminars during site visits) to engage local support for the strategy.

* carry out oral histories with key players in the HGI, and prepare guides for oral histories to be conducted by other groups.

* sponsor a series of videotaped records to be deposited in relevant archives.

* develop an edited, published volume on the human genome in historical context.

* prepare a final analytical study that can be used by archivists to apply the lessons learned in the Beckman Center project to other initiatives in the biomolecular sciences.

The products will be consistent with those produced by the American Institute of Physics in its innovative and highly respected documentation programs: Handbooks for all parties, oral histories and guides for conducting oral histories, guides for records appraisal, and analytical summaries of the overall success of the strategy. In addition, the Beckman Center proposes to sponsor the production of videotaped records, through Ray Kondratas at the National Museum of American History, and the publication of a scholarly volume drawing on academic conferences scheduled at the Center.

The plan will affect scientists working on the genome project only in the sense that participating scientists will be asked to release materials identified as crucial to the archivist responsible for the records of their institution, whenever those materials are no longer of use to the laboratory. It will affect archivists and support staff at participating institutions somewhat more, since they will be asked to participate in this plan in their identification and selection of materials to be preserved.

The Beckman Center's work will not address directly the problem of storage and retention of scientific data, that is, the actual genetic maps to be produced by the genome project. Presumably such immediately useful materials will be preserved for scientific purposes without any attention from archivists. The Beckman Center will be concerned, however, with the documentation of how those maps were produced. This would include attention to the efforts by computer scientists to develop informatics programs that can accommodate the various levels of information to be contained in such maps (e.g. genetic, physical, STSs, etc.).

Genome project participants perceive themselves as "making history" and the project provides an excellent opportunity for a case study in preserving the records of an important and complex project in contemporary biomolecular sciences. Our purpose, then, is both to facilitate the preservation of significant records of the genome project, and to develop documentation guidelines that can be productively applied to other projects in the biological sciences. As students of contemporary science begin trying to make sense of the phenomenal developments in biomolecular sciences since 1945, the Beckman Center for the History of Chemistry will assist that effort by promoting the preservation of important materials generated by biomolecular science's most ambitious program, the Human Genome Initiative.

Background and Significance

Archivists have traditionally provided detailed descriptions of archival records for scholarly use. They have

generally launched collecting projects by conducting surveys of records. But surveys provide information only about what records exist. They do not suggest what documentation would be required for an adequate record.

Some archivists have begun calling on their profession to help solve a more analytical problem, that is, identifying what would be necessary to make the historical record more complete. The concept of a documentation strategy was proposed in the mid-1980s, and since then cooperative documentation strategies--plans that identify what materials should be preserved based on study of the phenomenon to be documented--have attracted considerable attention within the professional community of archivists.

Some of the most substantive and impressive documentation strategy efforts have been carried out in discipline history centers. Such organizations as the American Institute of Physics, the Charles Babbage Institute for the History of Information Processing, and the Beckman Center for the History of Chemistry have traditionally engaged in some archival activities, such as collecting and placing materials and promoting the preservation of historical records relevant to their fields. More recently, however, both the AIP and the Babbage Institute have sponsored archival documentation projects. The AIP has produced admired documentation strategies for astrophysics, solid state physics, nuclear physics, and multi-institutional collaboration in space physics. The Charles Babbage Institute is in the process of developing a documentation strategy for the history of computers. The Beckman Center for the History of Chemistry is therefore a suitable sponsor for a documentation strategy in the biomolecular sciences. The Beckman Center has undertaken extensive archival surveys for its earlier Polymer Project, and has a demonstrated commitment to the history of biomolecular sciences, including work on lab instrumentation in biochemistry, and an oral history project with the Pew Scholars.

The documentation strategy approach proposes that functional analysis of the subject to be documented should help shape the selection of materials for preservation. In the case of the Human Genome Initiative, this analysis would begin with the fundamental question: What activities can be encompassed under the title of "human genome project"? Some obvious activities include administration of funding, dealing with the needs of the political system (Congressional hearings), day-to-day research in the laboratory, theorizing, creating appropriate forums for debate, publishing, producing new technologies crucial to the project, writing grants, and exploring ethical and philosophical questions.

As MIT Archivist Helen Samuels has noted, some human activities leave records and some do not. Yet the human activities that do not leave a paper trail are often of considerable interest. A documentation strategy explores a scientific project in terms of activities--what do the scientists

and administrators do--rather than in terms of records generated. This approach leads naturally to a different way of conceiving of what needs to be preserved.

A documentation strategy can lead archivists to seek out information about those kinds of activities that might not appear "naturally" in the written record, such as explications of the unwritten rules of genome laboratory etiquette, or the personal (craft) skills necessary to maintaining a pure cell line, cloning DNA, or working with restriction fragment length polymorphisms or polymerase chain reactions. Videotaped records, oral histories and sociological observation can all play a role in documenting such activities.

A documentation strategy can also address the pressing problem of the changing nature of scientific communication. Electronic mail, for example, and the increasing importance of other less formal means of communication are having a significant impact on the kinds of records social scientists have traditionally exploited. Genome project investigators commonly depend on electronic communication. At the University of Pennsylvania, for example, where a plan for collaborative work on mapping chromosome 22 is underway, the PennNet system is a crucial part of keeping the far-flung participants in the project informed. How can this system of communication be documented?

Bruce Lewenstein, who has been directing a documentation program for the debate over cold fusion, has grappled with this problem by downloading E-mail records from the USENET bulletin board, and preserving the "Cold Fusion Newsletter" produced electronically by physicist Douglas R. O. Morrison. Yet the cold fusion group has still not resolved the question of how to handle these records, whether they should be stored on computer disk or printed out, and how they should be made available to users (Lewenstein, 1990). Any effort to document the HGI will have to take into account the importance of electronic mail to contemporary scientific communication, and one of our goals will be to suggest how such electronic data can most profitably be stored and used.

A documentation strategy can also explicitly target materials that will provide insight into the collaborative nature of the genome project. Specifically, the relationships between academic science and private industry promise to generate materials of considerable historical and immediate interest. For example, Leroy Hood's group at Cal-Tech developed genome technology that became the basis of a new biotechnology firm, Applied Biosystems, which took on the marketing and distribution of VLSI-based systems for the analysis of genetic sequences. And some genome centers being reviewed for funding include one or more industrial collaborator(s), often with computer information processing firms that can help the scientists solve problems of sequence analysis. This involvement of private industry is not unprecedented--private industry has played an increasingly

important role in all varieties of modern science throughout the century. Yet the mechanics of how these collaborations work--the questions of intellectual property and academic freedom--are of crucial importance to our understanding of contemporary science. Mark Pearson, director of Molecular Biology at DuPont, which is developing technologies for work on HGI, has observed that the traditional differences between academic and corporate science are eroding in molecular biology (personal communication, 5 November 1990). How will this affect the field? And what can we learn from the interaction of corporate and academic science in this particular project?

The center's plans will be coordinated with the work of other groups interested in related questions. For example, the Center plans to work with the National Reference Center on Bioethics Literature at Georgetown University, which is currently developing a proposal to serve as a repository for documentation of the ethical and legal questions raised by the HGI. The in-depth program at Georgetown, focusing specifically on bioethical issues, will complement and enrich the Beckman Center's broader analysis of the entire range of documentary problems posed by the project. Similarly, other studies dealing with particular issues, such as the plan at Cal-Tech to explore instrumentation, will complement the Beckman Center's work. And Everett Mendelsohn's program (at Harvard) to bring together a consortium of scholars and archivists interested in the HGI will be coordinated with the center's own academic conferences. The Center can provide the broad perspective that will encourage a coordinated effort across many institutions and disciplines.

The Center's plan will not involve proposing new archival centers, nor will it involve extensive preservation of archival records of the HGI at the Center itself. Rather, the plan will promote the efficient use of existing systems and the appropriate deposition of papers in the home institutions of individuals involved in the project.

The goal of a documentation strategy is to preserve materials capable of helping to answer a broad range of possible research questions. It is impossible to predict the specific questions the records of the project will be used to explore. But a documentation strategy that attends to all the varied activities of a project such as the Human Genome Initiative is more likely to ensure the preservation of a record that will be of use to science policy analysts, historians, sociologists and scientists themselves.

The suitability of the Human Genome Initiative for such a documentation study is indisputable. The project has attracted scholars from many fields and interest in both the science itself and its social implications is extraordinarily high. Several scholars have already begun placing the project in a broader historical and sociological context. Stephen Hilgartner at Columbia University is working on a detailed sociological

analysis of the genome project laboratory; Daniel J. Kevles at Cal-Tech has been exploring social implications in light of his earlier work on the history of the American eugenics movement; M. Susan Lindee, at the University of Pennsylvania, has been preparing a study of the professional impact of the genome project on the field of human genetics; Lindee is also working with Dorothy Nelkin (New York University) on a study of genetic imagery and metaphor in popular culture that will suggest how findings from the HGI can be expected to be interpreted by "consumers" of genetic information; John Beatty at the University of Minnesota has explored the origins of Department of Energy funding for the genome project; and Michael Fortun at Harvard has begun analyzing published histories of the genome project as a form of literary text. Clearly scholarly interest in the project is high and can be expected to increase as the project progresses. In addition, the implications of the findings expected to emerge from the genome project have been the subject of dozens of popular articles (Hall, 1990; Joyce, 1987; McCauliffe, 1987; del Guercio, 1987) and several popular books (Duster, 1990; Nelkin and Tancredi, 1989; Yoxen, 1984; Angier, 1988). Several published analyses have explored the HGI as science policy and Robert Cook-Deegan is in the process of completing a comprehensive study of policy-making on the project (Cook-Deegan, 1990). Scientists themselves have recognized the importance of public understanding of the project, suggesting a high level of awareness within the scientific community of the possible long-term implications of this project (Alberts, 1985; Cantor, 1990).

The HGI promises to be a much studied and analyzed scientific project, and there are substantive reasons for this high interest. The project is a prototype for "big science" in biomolecular sciences. It is reshaping several scientific fields, affecting public expectations of medical genetics, and promising to revolutionize scientific understanding of human heredity.

And yet the documentation of this project, which can be expected to continue to attract historians, science policy analysts and other students of modern science, also poses unusually complex problems. The variety of institutional settings, including academia, national laboratories, laboratories of several sectors of private industry, and a growing array of conferences and new journals mean that records will be both scattered and subject to widely different systems of preservation. And the broad implications of the project suggest the need to document activities in other sectors, including the activities of critics of the genome project, such as biochemists Martin Rechsteiner of the University of Utah and Bernard Davis of Harvard University and microbiologist Michael Sylvanene at the University of California at Davis, and the bioethical debate.

Design and Methods

An expert six-member committee will oversee the development of the Beckman Center's documentation strategy. Members will be historian John Parascandola of the National Library of Medicine; Corporate Archivist Jeffrey L. Sturchio of Merck & Co.; genome researcher Beverly Emanuel of Children's Hospital of Philadelphia, whose project to map chromosome 22 is the focus of a proposed Genome Center at the University of Pennsylvania; archivist Joan Warnow-Blewett of the Center for the History of Physics, historian Daniel J. Kevles of Caltech, and molecular biologist Mark Pearson of DuPont.

This committee will provide professional input from the various potential users of a documentation strategy and will facilitate contacts within their communities.

The project's first goal will be to identify in as much detail as possible the activities that can be taken to constitute the genome project. This will define the field to be surveyed, and will, unavoidably, also identify those aspects of the project that will not be a focus of Center interest (e.g. scientific genome maps, Congressional records). This initial survey will also suggest some limitations on the preservation of materials directly relevant to the plan, such as second generation technologies that do not differ substantially from the first generation. While complete records might be preserved for the original version, the second version might require little or no documentation (See Hackman and Warnow-Blewett, 1989).

The initial survey will also consider the practical problems of working with various archives of the institutions involved. Private industry represents a particularly complex problem in records preservation, since this is an area where the automatic "systems" of preservation already in place are least likely to meet the needs of scholars. Corporations tend to preserve documentation with very different priorities from those that guide academic preservation. While most biotechnology firms can be expected to preserve (for some period of time) technical documents related to R&D, and materials relevant to patent applications, they are not likely to preserve the correspondence files of administrators and scientists, or the records of internal disputes over market strategies or scientific collaboration. Very few of even the largest American corporations maintain formal archives or hire full-time archivists. Yet the role of industry in scientific and technological development in the United States has been extremely important. While the Beckman Center is unlikely to overcome all the obstacles to preserving a clear and accessible record of the role of biotechnology firms in the genome project, the Center can at least be a vocal advocate for improved documentation.

A field archivist, to be employed full-time throughout the course of the project, will be responsible for much of this

initial work. He or she will report to the principle investigator, and take direction from the principle director and the advisory board. The field archivist's reports on each trip will provide data for the production of the formal documentation strategy.

In the second phase, the Center staff will begin outreach activities and oral history interviews, and the program for videotaping the uses of various technologies will begin.

Ray Kondratas, historian at the National Museum of American History, will be working with the center on the production of four or five videotaped records dealing with instrumentation in the HGI. Kondratas will begin work on the project in the second year of funding. He has played a key role in the Smithsonian Videohistory Program which has been sponsored by the Sloan Foundation since 1986, and he has completed a videotaped study of LeRoy Hood and the DNA sequencer. He has access to relevant technical teams, and experience in production and planning of such records. The resulting videotapes will be deposited at the home institution of the persons featured, at the Smithsonian, and at the Beckman Center. Kondratas could also play a role in our effort to document in detail a few key labs as "probes."

Our strategy will presumably suggest that extensive records should be preserved for a select few "typical" laboratories or working groups in various institutions. Ideally these groups would reflect the broad range of genome research underway. For example, included should be an informatics research group; a major academic Genome Center, such as Washington University; and a relevant working group at one of the National Laboratories where Department of Energy genome research is underway, possibly Los Alamos. At these sites, in these labs, and with the concurrence of local archivists, a wide variety of documentary materials could be identified as worth preserving. These records would include those traditionally deemed worthy of preservation, such as research notes, publications, personnel and policy records. But the records preserved would also include those traditionally seen as transient or disposable, such as records of colloquia and lab visitors, records of equipment purchases and repairs, and even information on the size and layout of the laboratory. Commonplace internal memos and truly run-of-the-mill documents could be preserved as permanent records in these special cases, as in-depth "samples" of the day-to-day business of genome research.

Participants at these "typical" labs would also be encouraged to suggest other materials for preservation in appropriate archives. And the Beckman Center will encourage the production of film or videotaped records, and encourage scholars interested in the genome project to work with "typical" designated lab groups. In addition, oral histories could be conducted with both scientists and with the lab staff engaged in

the day-to-day work of sequencing. While it would be impossible to preserve all these materials for all the various laboratories and research groups involved in the project in one way or another, the selective preservation of a few "typical" labs will provide a rich fund of data that may prove to be of extreme historical interest. The records of these designated laboratories could provide information valuable to scholars of social history, laboratory methods, and to sociologists. It might also be of some interest as an historiographical experiment, a test of the efficacy of sampling in preserving the records of contemporary science.

We do not propose that the Beckman Center staff itself will conduct all these oral histories, produce all these videotapes or be responsible for the preservation of all other materials. Rather, the Center's goal will be to identify a suitable "typical" laboratory location, then work with the archivist(s) and staff responsible for the records produced by that laboratory to implement the strategy at the local level. This might mean, for instance, that some oral history interviews will be conducted by Center staff members, but it will more likely mean that the Center's guide to oral history interviews will be used by the group responsible for document preservation at that particular laboratory.

In conformity with its plans for the production of a scholarly publication relevant to the project, the Center is also interested in using the project to explore some specific historiographical questions. While documentation strategies do not generally include such interpretive questions, the Beckman Center's plan is part of a larger program to develop greater understanding of biomolecular sciences. Our interest in the Human Genome Initiative has been shaped by these interpretive questions, and some components of the documentation strategy, such as oral histories and videotaped records, will draw on these questions. The Beckman Center's overall research program, "Instrumentalities, applications and conceptual frameworks in the transformation of the biomolecular sciences," is summarized in Appendix A of this proposal.

The vivid public debate about the genome project raises several important questions that might not be readily accessible through the standard sources preserved by laboratories, universities and private industry. For example, how has this debate affected scientific practice? And how exactly do scientists perceive the relationship between their work and the ethical and social implications of their results? James Watson, in several published interviews, has expressed concern about the possible misuse of genetic information by employers, the legal system, or the school system. Yet these concerns do not seem to have dampened scientific enthusiasm for the project. Nelkin and Tancredi, in an analysis of how institutions currently use predictive biological information, have documented the powerful social and economic forces that can be expected to encourage the

creation of a class of "genetically unemployable" persons, or to legitimate exclusion practices based on anticipated costs (Nelkin and Tancredi, 1989). If scientists perceive these ethical problems as beyond the boundaries of "science," can such concerns nonetheless shape scientific practice?

The attention to ethical and social implications of the project is one institutional response to the public debate. Clearly many scientists are thinking about the social implications of the information expected to be produced by the genome project. And few scientific projects have been undertaken with so much public acknowledgement from the scientific community that their work will have moral and social meaning. But does this awareness shape what scientists do?

Other questions involve the relationship of funding to scientific legitimacy. Much of the debate about the genome project within the scientific community has included some discussion of the limitations of "big science" in biomolecular research. Some critics suggest that "big science" is less valuable (in terms of providing access to truths about the natural world) than science produced by a lone investigator independently working through a fundamental problem (Alberts, 1985).

If "good science" is science removed from the political system, what does that imply about most post-war science? Is "mission-oriented" science also "bad science," and what kinds of arguments are used to suggest this? What can these arguments tell us about the changing norms of the scientific community? These are questions about the values guiding the scientific community engaged in the genome project, and specific attention to these questions in the oral history interviews would be particularly appropriate.

Another set of questions focuses on the relationships between various communities engaged in the project--human geneticists, biochemists, computer technicians, corporate scientists, mathematicians, government administrators, health professionals, and so on. The project has inspired four new journals, and has created a new community of collaborators from different fields. Human Genome News, a publication of the National Center for Human Genome Research of the NIH, lists dozens of conferences, meetings and workshops focusing on the genome project in each quarterly issue. Professional societies involved in these conferences include the American Association of Artificial Intelligence, American Chemical Society, American Society of Human Genetics, Association of Minority Health Professional Schools, and American Electrophoresis Society. Government agencies include the U.S. Department of Agriculture, the National Science Foundation, the Office of Health and Environmental Research, Department of Health and Human Services, National Institutes of Health, Department of Energy, National Research Council, and Office of Technology Assessment. National

labs include Argonne, Los Alamos, Oak Ridge, Pacific Northwest, Lawrence Berkeley and Lawrence Livermore.

Few scientific projects, particularly in biology, have ever involved such broad collaboration. The impact of this collaboration will not be easily documented. But attention to this phenomenon, and recognition of its importance, might help us more clearly define what activities count as the Human Genome Initiative. Do turf battles count, and how can they be documented? What about the strategies scientists have used to deal with rapidly changing professional mores? Have generational conflicts within fields emerged? Maynard Olson's sense that the genome project represents a return of human genetics to its roots is of particular interest, suggesting how an older generation of scientists (now in their 70s and 80s) have influenced the younger group (under 50) that predominates in human genome research (Cook-Deegan, 1989). Does that mean our strategy should include some attention to the older group, even if these scientists are not directly involved in human genome research? What about the competitive rise of new journals and even new specialties? What tensions have emerged in the scientific community over collaboration? And where has it been most productive?

By attending to these questions, the Beckman Center staff will be sensitized to the importance of disputed collaborations (as they arise), issues that provoke problems, and specific projects or locations where collaborations have been effective (particularly collaborations between biotech firms and academic labs).

We do not mean to suggest that such interpretive questions will entirely guide our strategy. But they have helped to shape our interest in the project and will continue to play a role in our work.

Plan of Action

The Beckman Center for the History of Chemistry's interest in the Human Genome Initiative grew out of the Center's Pew Scholars oral history project. At the suggestion of the chairman of the advisory committee, Nobel laureate Joshua Lederberg, the Pew Charitable Trusts developed an oral history project focused on the promising young scientists it supports for four years, the Pew Scholars. The Beckman Center was selected to oversee the project based on its extensive experience in oral histories. Center staff members and post-doctoral fellows have been conducting interviews with these young scientists on a regular basis: to date, oral histories have been completed with more than sixty Pew Scholars (typically assistant professors of the biomolecular sciences at institutions such as Salk, Scripps, Rockefeller, Caltech and Harvard).

The productivity of the interviews suggested the value of a more comprehensive effort to encourage documentation of

contemporary biomolecular sciences. The Beckman Center therefore began exploring the possibility of preparing a documentation strategy for post-war biomolecular sciences, eventually focusing on the Human Genome Initiative as a suitable special focus of historical and contemporary interest.

In March of 1989 a group of scholars including Christian Anfinsen and Paul Berg gathered at the Center for a symposium on "The Merging of Chemistry and Biology--Looking Backward/Looking Forward." In October of 1989 the Andrew W. Mellon Foundation funded an exploratory study of the revolution in the life sciences, and in March of 1990 the first meeting of archivists and scholars focusing on the biomolecular sciences initiative, now known as BIMOSI, was held at the Beckman Center. The group of visitors attending this meeting included historians Pnina Abir-Am of Harvard and Frederic L. Holmes of Yale; Helen Samuels, Institute Archivist at the Massachusetts Institute of Technology; Darwin Stapleton, Director of the Rockefeller Archive Center, and Jeffrey L. Sturchio, Corporate Archivist at Merck & Co.

This crucial meeting brought the project clearly to focus on the genome project as a suitable subject for a study of documentation and history in the making. The group agreed that the genome project was a particularly appropriate subject of a new Beckman Center initiative, one that could bring together many related strands of Center activities in a productive way.

In July 1990, the Center called together another group to explore the question of how to document the genome project. Participants included archivists Peter Hirtle of the National Library of Medicine, and Clark Elliott of Harvard University; historians John Parascandola of the National Library of Medicine, Spencer Weart, of the American Institute of Physics, and Seymour Mauskopf of Duke University, and NIH consultant Robert Cook-Deegan. This group explored the role of industry and academia, and discussed how to preserve those materials that pose known archival problems, such as unsuccessful applications for funding and the records of private industry.

In 1991, the Center plans to sponsor an academic conference on writing contemporary history, at which some of these issues can be further explored.

The Beckman Center for the History of Chemistry's qualifications as sponsor of a documentation project of the Human Genome Initiative include its close ties to the scientific and industrial communities involved in the genome project and its experience with oral histories in the biomolecular sciences. A research center with offices on the campus of the University of Pennsylvania, the Beckman Center is jointly supported by the American Chemical Society, the American Institute of Chemical Engineers and the University of Pennsylvania. Affiliated societies include the American Society for Biochemistry and Molecular Biology.

The Center's goals are to develop programs of oral history interviews that can document major developments in modern science; to locate historical manuscripts and archival records in the hands of individuals, societies, trade associations and corporations; to encourage the preservation of these records in appropriate repositories; to publish resource guides and historical materials; and to create traveling exhibits that can enhance public understanding of the history of chemistry, biomolecular sciences, chemical engineering and the chemical process industry.

The Beckman Center staff includes director Arnold Thackray, a professor in the Department of the History and Sociology of Science at the University of Pennsylvania; archivist-historian Stephanie Morris, a professional familiar with records in the history of science; research historian M. Susan Lindee, an assistant professor in the Penn H&SS department and a scholar in the history of biology; and James Bohning, a chemist who is responsible for an ongoing oral history project in the chemical and pharmaceutical industries. This staff works with private industry, professional scientific groups, academic scientists and historians and the professional archival community to promote public and scholarly understanding of the history and nature of modern science.

The Center staff has extensive experience in conducting oral histories and in exploring archival resources in the history of the biochemical sciences. Subjects of oral history interviews conducted by the center have included biochemist Mildred Cohn, on her studies of mechanisms of enzymatic reactions; Emil Smith on the chemistry of proteins; Harland Wood, on his work with polyphosphate metabolism and autotrophic growth; and Arnold O. Beckman on the development of modern instrumentation. In addition the Center staff has had the opportunity to work with outstanding young researchers who are establishing themselves as leaders in their fields, for example, Glenn Evans of the Salk Institute and Joseph A. Sorge, the CEO of Stratagene.

Drawing on its expert staff, and on the input from those involved in the planning conferences, the Center now proposes activities that fall into three basic categories. These are the production of formal and informal printed materials, outreach to the various communities involved, and the encouragement of scholarly work connected to the documentation of the genome initiative.

The center will engage a full-time field archivist to begin preliminary research. This will include a survey of relevant institutions, to collect information about current systems of records retention in private industry, academia and in the national laboratories. This archivist will be able to draw on some existing studies of archival systems, such as Joan Warnow's 1985 study of DOE records retention schedules (Warnow,

1985). The field archivist will prepare for, and conduct in the fall and winter of 1991, a series of site visits to the laboratories chiefly engaged in HGI research. During each visit, the field archivist will meet with the local records custodians (for example, records managers, archivists, librarians, and secretaries) to explore records retention schedules and practices. The archivist will talk with scientists working there, to better understand what they do and how they write it up. Discussions would be conducted with area repositories to determine which organizations would want to participate in the documentation strategy, and in what way.

In planning a series of site visits, the preparations would include contacting the institution to be visited and determining who there has records retention responsibilities. This could be the university archivist, if the laboratory is so connected, the reference librarian, if in a privately conducted or corporate research center, or a records manager, in a corporate setting. The on-site person could then recommend other officials or individuals the field archivist should meet. This cooperation is essential to the success of the visit and to the effectiveness of the project. These contacts will help in planning the second round of site visits, the workshop-type presentations tentatively scheduled for the spring of 1992.

At the conclusion of the site visits, the field archivist would prepare a report on the record-keeping practices of the laboratory. This report would identify the level of documentation practiced (vital records only, or retention scheduling, for example) and if the practices in place would lend themselves to the artificially-detailed level of documentation desired of the representative or typical sites chosen as "probes." The availability of other local institutions willing to assist in preserving this level of detailed documentation would be considered as well.

By spring of 1992, the field archivist will have prepared a draft summary report of his or her activities. The Beckman Center will then formulate a preliminary documentation strategy report with the aid of other archivists serving on the advisory committee. The field archivist and the documentation strategist will be responsible for the production of a primary report, continually amended as the project progresses, that will guide the production of all other materials.

In year 2, the archivist and Beckman Center staff will also begin preparing guidelines for oral histories, and identifying suitable candidates to be targeted for oral histories. And it will begin a series of visits to relevant laboratories. Center staff members, including M. Susan Lindee, Stephanie Morris, Arnold Thackray, and the field archivist hired specifically for this project, will visit various government and academic laboratories, corporations, archives and scholarly centers. These visitors will both promote the relevance and

importance of a clear documentation strategy, and take advantage of feedback from those involved in the actual work to improve the documentation strategy plans. On-site visits will consist of a brief presentation by the Center staff member to the entire laboratory group (if possible) or archival staff, followed by a question and answer session.

In a few key locations, particularly those identified as "sample" locations in which a large volume of material is to be preserved, the center will organize more substantive training sessions with the archival and clerical staff. Center archivist-historian Stephanie Morris and the field archivist will meet with the relevant archivists and support staff to explore the practical problems of implementation of the Beckman Center plan. These meetings will provide an opportunity both for Center staff to explain the program to local archivists, and for local archivists to explain their own priorities and needs.

The center also plans to sponsor generalized seminars for lab personnel and scientists to enhance their understanding of the documentation strategy and to enlist their support at such laboratories. These seminars would be relatively short and infrequent, but could significantly improve documentation simply by calling scientists' and administrators' attention to the issue.

While the project staff and its advisory group will identify suitable subjects for oral history interviews and suggest relevant content for interviews, center staff (historians, post-doctoral fellows) will not conduct all interviews. Rather, the Center plans to conduct 16 to 20 "model" interviews, of fairly substantial length, with several levels of personnel involved in the genome project. These interviews, transcribed and distributed to appropriate archives, can then be used by archivists, historians, or researchers at other institutions to help plan and carry out productive oral history interviews with genome project researchers at their institutions.

At the end of year 2, in the summer of 1993, the Center will sponsor the first of two academic conferences focusing on the HGI in historical perspective. These conferences will bring together experts in documentation strategies in other fields (e.g. lasers, government labs) and historians, sociologists, anthropologists and ethicists interested in the genome project. The general theme of both conferences will be "The Human Genome Project as Self-Conscious History" and discussions will include attention to both archival and historiographical questions raised by the project. Participants will be asked to explore how documentation shapes their work, and to address questions of method and evidence in their papers. Anthropologists and sociologists who work in scientific laboratories, historians of science, ethicists and scientists with a particular interest in history will be invited to explore questions of documentation from their own perspectives.

This meeting will provide an opportunity for the Center staff to gain greater insight into scholars' perspectives on the project, and this feedback from an important group of "consumers" of documentation will enrich the formal documentation plans.

These meetings will result in an edited volume, "The Human Genome Initiative: Documentary Sources and Research Agendas," to be published at the end of the five-year project.

The Center will also sponsor a post-doctoral fellow and a graduate student each year, in years two through five. The Center will therefore provide important training opportunities for young researchers interested in the Human Genome Initiative, and the input of these young scholars will also facilitate the Center's work. The postdoctoral fellow will have an opportunity to conduct oral history interviews and will assist in the production of published materials and reports. The graduate fellow, in the Department of the History and Sociology of Science at the University of Pennsylvania, will assist in the production of resource guides and other materials, and could also play a role in on-site visits and oral history interviews. But primary importance will be placed on their scholarly work, which can enhance the overall goal of increasing understanding of the Human Genome Initiative.

The Center is particularly interested in encouraging scholarly work on the role of instrumentation in the development of contemporary biomolecular science. The genome project is the first major biological initiative that has as its goal the development of new technology--indeed, its success depends on the development of new technologies that will lower costs. The development of common "languages" or landmarks on the genome, such as Sequence Tagged Site markers, the role of database systems in the production of multilevel integrated physical maps, the importance of mass spectrometry, yeast artificial chromosomes and bacterial artificial chromosomes can all be considered as problems of instrumentation. Other technological or instrumentation problems include software exchange--software within particular communities is not easily exportable--and the problem of hardware and operating system compatibility. The project depends on increasing the speed and lowering the cost of mapping, and therefore new technologies are crucial to its success. The human genome comprises approximately 3 billion base pairs, and if such a large section of DNA is to be sequenced, the cost per base pair should be below 50 cents to be cost effective. The current cost of DNA sequencing is estimated at $2 to $5 per base pair, which is too high to make a concentrated mapping of the human genome practical (NIH-DOE, 1990). Yet sequencing technology has improved dramatically in only the last two years, and machines that automatically identify the order of base pairs in prepared DNA are now available. The achievements even since 1987 are impressive. Some parts of the X-chromosome spanning more than two million base pairs are now contained in ordered clone

libraries; more than 10 million base pairs of the nematode genome have been continually mapped by overlapping clones; and while the longest continuous sequence to date is only 200,000 base pairs, large expanses of chromosome 4 and the T cell receptor region are now being sequenced. (Cook-Deegan, 1990).

New technologies providing greater access to the gene have historically had conceptual repercussions. Since 1940 the abstraction of a "gene" has been localized and made concrete, and each step of this localization--from the identification of DNA as the hereditary material in 1944, to the increasing use of starch gel electrophoresis in the 1950s, to the development of restriction fragment length polymorphisms in the 1970s, and to the use of the polymerase chain reaction in the 1980s--has helped to shape the modern conception of a "gene." Even the laboratory animals commonly used as genetic probes (Drosophila, E. coli) can be interpreted as constructed technology, as Rob Kohler's work on the use of Drosophila suggests (Kohler, forthcoming 1990). The Center hopes that either the graduate student or the post-doctoral fellow, or both, will have an interest in instrumentation and technology in contemporary biomolecular science.

Activities in years 3 and 4 will be similar to those in year 2--site visits, report production and amendments to the strategy, oral histories, and outreach activities--with the exception that the Center staff will begin a special program to reach corporate scientists and administrators and engage their support for a documentation strategy. This will include attendance by Center staff members at relevant industrial conventions, and the preparation of informational guides geared toward corporate needs.

A second academic conference will be held during year 4, the specific subject to be decided upon as the project develops.

The printed materials, to be produced by the field archivist, by the Beckman Center support staff, and by graduate students or post-doctoral fellows involved in the Beckman Center's project, will include:

* published resource guides providing information and assistance to those using and producing archival materials;

* formal recommendations for the inclusion of documentation guidelines in NIH and DOE application packets;

* guidelines for oral histories, including suggestions about who should be interviewed and about what kinds of questions are appropriate;

* a final report, to be published by the Beckman Center for the History of Chemistry, exploring the genome project documentation strategy as a case study in documenting

contemporary biomolecular science. This report will include a serious assessment of the Center's project, and will be useful to other organizations interested in planning and pursuing a documentation strategy in contemporary science.

All these materials will be readily available to interested scholars and professionals.

Conclusion

The Beckman Center for the History of Chemistry will act as a guiding force for documenting and preserving the record of the genome project. The Center staff and committee members will reach out to all communities involved to encourage cooperation and coordination of related activities.

We will apply innovative ideas about archival preservation to a major scientific project in ways that may be relevant to documentation problems in other projects. We plan this project as a case study in modern science and an opportunity for the scholarly community to think through the practical and intellectual problems posed by the volume and complexity of records produced by contemporary biomolecular sciences. While excellent work has been done in physics and other fields, the genome project promises to be the first documentation strategy project in molecular biology.

The genome initiative brings together scientists and technicians from diverse backgrounds. It promises to play an important role in both popular culture and the political system. And it is unique in its dependence on expected technological innovations that will decrease costs. The project therefore can provide insight into many important problems in contemporary biomolecular sciences, including industry-academic collaboration, political relevance and controversy, ethical and philosophical issues, social impact and technological complexity. It raises questions with a broad applicability to our understanding of modern science. Our program can help meet the immediate needs of science policy analysis and humanistic research and, of equal importance, contribute to understanding of the scientific process in the long term.