Narrative Section of a Successful Application

The attached document contains the grant narrative and selected portions of a previously funded grant application. It is not intended to serve as a model, but to give you a sense of how a successful application may be crafted. Every successful application is different, and each applicant is urged to prepare a proposal that reflects its unique project and aspirations. Prospective applicants should consult the Preservation and Access Programs application guidelines at http://www.neh.gov/grants/guidelines/HCRR.html for instructions. Applicants are also strongly encouraged to consult with the NEH Division of Preservation and Access Programs staff well before a grant deadline.

Note: The attachment only contains the grant narrative and selected portions, not the entire funded application. In addition, certain portions may have been redacted to protect the privacy interests of an individual and/or to protect confidential commercial and financial information and/or to protect copyrighted materials.

Project Title: Expanding the Medical Heritage Library: Preserving and Providing Online Access to Historical Medical Periodicals

Institution: Open Knowledge Commons

Project Director: Maura Marx

Grant Program: Preservation and Access Humanities Collections and Reference Resources

EXPANDING THE MEDICAL HERITAGE LIBRARY: PRESERVING AND PROVIDING ONLINE ACCESS TO HISTORICAL MEDICAL JOURNALS

3. NARRATIVE

Under the aegis of the Open Knowledge Commons (OKC), the Medical Heritage Library (MHL) proposes a two-year project to digitize and preserve historical American medical journals and to make them freely available publicly via the Internet Archive (IA). These materials are a rich source of information not only for research in the history of medicine but for interdisciplinary studies ranging from economics to civil rights to gender studies. Providing free access to them will enable scholars all over the country and internationally to pursue their studies more effectively and more thoroughly, using full-text search and other modern technological tools.

The participants in this proposal are well-qualified to carry out this new initiative, both because they hold significant collections from which to draw titles important for humanities research, and because they have developed expertise and shared procedures during an on-going monograph digitization project. Equally important, as members of HathiTrust they can assure that the digitized content will be preserved for continued public access into the future.

The Medical Heritage Library (http://www.medicalheritage.org) was created in 2009 through an OKC initiative to leverage the expertise of five leading medical/special libraries (the National Library of Medicine, the Francis A. Countway Library of Medicine at Harvard Medical School, the Harvey Cushing/John Hay Whitney Medical Library at Yale University, the Augustus C. Long Health Sciences Library at Columbia University, and the New York Public Library) to enable open access to medical content through digitization. New members as of 2010 include the College of Physicians of Philadelphia and the Alan Mason Chesney Medical Archives, John Hopkins University.

Four of these preeminent medical libraries are collaborating in the project proposed here: the Countway Library, the Cushing/Whitney Library, the Long Health Sciences Library (HSL), and the College of Physicians. The National Library of Medicine (NLM) will not receive funding through this proposal, but will assist the project participants by lending volumes that are missing from the participants' collections, and the remaining MHL members will supply volumes if NLM cannot. Brief descriptions of the MHL and the participants in this project are provided in Appendices A-G. Letters of commitment from the four participants are attached as Appendix P.

The participating institutions plan to digitize up to 200 journal titles dating from 1797 to 1923 for a total of approximately 7,350 volumes (2,152,000 pages). The project will include titles reflecting mainstream medicine of the time as well as those of the era's intriguing alternative and sectarian movements. The goal is to provide a cross-section of the main currents of American medical thinking and practice in the nineteenth and early twentieth centuries.

I. SIGNIFICANCE

A.   The Importance of Medicine and Historical Medical Journals for Humanities Scholarship

Perhaps no subject is so fundamental to understanding and interpreting the human condition as medicine. Such interpretations entail concerns regarding birth and death, health and disease, and the normal and the pathological. Around such concerns, systems of knowledge, patterns of care, and professions and para-professions have been developed in contexts conditioned by and conditioning related belief systems and political/economic configurations.

Over the past several decades, the history of medicine as a discipline has itself evolved from a history of doctors and scientific discovery to a social and cultural history of medicine, wrestling with such fundamental questions (in addition to traditional concerns regarding the production of medical knowledge) as: what have been the determinants of disease and its distribution over centuries and millennia; how have the professional and nonprofessional responses to disease been organized and distributed; what have been the meanings attributed to particular disease states, and how have such disease states themselves been constructed or framed around existing biological, technological, and sociological factors?

Subsumed within or existing alongside such concerns have been equally important ones regarding the history of medical education; race, ethnicity, and gender among both patients and practitioners; the contested domains between "orthodox" and "unorthodox" medicine; the equally contested domains between private practice and public health; the evolving meaning and implications of "global health" concerns; the organization and financing of the health care industry; and the ongoing cultural configuration of the "body" itself, with all its attendant religious, gender, and social implications.

Given the centrality of such concerns to the human condition, the history of medicine transcends disciplinary boundaries, extending to the interests of general and economic historians, classicists and archaeologists, historians of religion and art, linguists and literary scholars, philosophers and historians of ideas, and sociologists and policy-makers. As described above, major medical libraries in this country have come together to create the Medical History Library precisely to build and share knowledge in these areas through digital technologies. Monographs digitized by MHL and mounted on the IA site are already providing resources with which readers and scholars across a multitude of disciplines can examine the mutually embedded nature of medicine and society. The project proposed here will add access to journals.

Historical journals are among the greatest and most versatile resources in any medical library. While monographs were the initial means to preserve and transmit medical knowledge and information, the rise of journals – weekly, monthly, quarterly, and annual – allowed information to be exchanged and disseminated rapidly and provided physicians with critical current information and all manner of other content. One such rare example is the *Confederate States Medical & Surgical Journal*, published in Richmond in 1864 and 1865 "to lay the foundation of a Southern medical literature on a firm and enduring basis" and analogous to the *Boston Medical and Surgical Journal in the North*. In addition to scientific content, as the page image in Appendix J shows, issues of the *Journal* are filled with items of news and events relating to military surgery and wound treatment, the progress of the war, prosthetics, epidemics,

Confederate hospital reports, the supply of drugs from plants indigenous to the South, and news of medical research in Europe.

Nearly 250 medical journals appeared in America before 1850; these include the earliest printed proceedings and communications of medical societies and the first true American medical journal, the *Medical Repository* (1797-1824) – the beginnings of an outpouring of journals devoted to all aspects of medicine, surgery, dentistry, pharmacy, and public health, along with the rise of publications devoted to medical specialties such as neurology, psychiatry, obstetrics and gynecology, pediatrics, medical jurisprudence, veterinary medicine, and military medicine. Some current titles, such as the *American Journal of the Medical Sciences*, the *New England Journal of Medicine* and the *Journal of the American Medical Association*, continue or are directly linked to their nineteenth century forebears. There were also printed transactions of local, state, regional, and national professional medical organizations, such as the Boston Society for Medical Improvement, the Medical Society of New Jersey, and the American Medical Association. Important medical developments and discoveries during this period were first brought to public attention through journals: John Otto's first description of hemophilia and the female's role in transmission, "An account of an haemorrhagic disposition existing in certain families" (1803); William Wood Gerhard's, "On the typhus fever which occurred at Philadelphia" (1837), the first distinction between typhus and typhoid fever; Oliver Wendell Holmes' "The contagiousness of puerperal fever" (1843); and early discussions of ether anesthesia, such as Henry Jacob Bigelow's, "Insensibility during surgical operations produced by inhalation" (1846).

The nineteenth century also saw the rise of alternative and sectarian medical movements, such as botanic medicine, balneology, homeopathy, and eclecticism, along with fringe or pseudo-scientific areas, including phrenology, Grahamism, animal magnetism, and psychic research. All of these developments, movements, and disciplines are reflected in medical journals of the time, and many of these subject areas had their own professional organizations and societies that published their own transactions and proceedings.

As it entered the twentieth century, medical literature began to grow increasingly scientific and technical in nature. The earlier period covered by this project, however, due in part to the wide-ranging subject matter found in the journal literature, provides a rich resource of information in accessible language in a variety of subject areas comprehensible to researchers with no specialized medical background or knowledge of the terminology. Allied topics of interest to humanities researchers such as physical and beauty culture, exercise and diet, climate, marital and sex education, and the care and rearing of children are all well represented in medical literature of this period, along with local news not available elsewhere.

From an early period, massive bibliographic indexing projects were put in place to provide access to medical information in journals. The foremost of these tools – the *Index Catalogue of the Library of the Surgeon General's Office* and the annual compilations of *Index Medicus* beginning in 1879 – are the principal means of providing author, title, and subject access to the resources in medical journals. Now in some part digitized, these tools remain essential points of entry for finding and retrieving medical information. Full digitization of runs of seminal and also unusual medical journals will provide an even greater level of access to information

currently locked away in print form. Medical journals are full of information beyond or below the article level which never would have been indexed and made visible; this type of information includes editorials; obituaries and biographical accounts; book reviews; mortality and morbidity statistics; observations on weather and climate; and local news and information on medical organizations, schools, hospitals, dispensaries and clinics, and libraries; and advertisements, such as those shown in Appendix J.

Consider *The Water-Cure Journal*, active from 1845 to 1862, which was published by Fowler and Wells, one of the foremost phrenological firms in the United States. Its issues describe and depict life and treatment in hydropathic establishments but also contain articles and information on such far-ranging subjects as women's rights, the illness of historical figures, longevity, exercise and physical education, book reviews, female medical education, tobacco, temperance, and as shown in Appendix J, women's fashion and dress reform. Once digitized, all of this fascinating information becomes visible and accessible for researchers to mine in new and exciting ways.

Why This Project is Needed
The four project participants plan to draw on their individual collections and subject strengths to select important medical journals – based on the criteria given below – for two purposes: to digitize and make them available to the public, and to preserve the digitized versions. As is all too well known, most journal publications from the nineteenth and early twentieth century were printed on acidic paper that has gradually become brittle over the decades. Repeated use leads to cracked pages, while aging bindings begin to deteriorate and fail to protect the fragile pages. Eventually text is lost as pages break. All of the participants' collections suffer from these conditions, and as a result research suffers. Digitization provides a means to create a visual surrogate as well as offering greatly enhanced use.

But digitization alone does not guarantee survival of the content into the future, since digital files are themselves subject to corruption, loss, and obsolescence. The participants in this project will assure long-term preservation by submitting copies of the files to HathiTrust, a shared archive specifically established for the preservation of digitized texts. Last year HathiTrust was certified as a "Trustworthy Digital Repository" by the Center for Research Libraries through their rigorous Trustworthy Repositories Audit and Certification (TRAC) assessment program.[1] Only a small number of digital repositories have been granted this certification. The participants will also of course care for the original volumes, returning them to open stack areas where appropriate, and in other cases transferring them to off-site repositories where they will be stored in optimal environmental conditions while continuing to be available to researchers upon request.

The project participants plan to fully digitize complete runs of medical journals which have not yet appeared in digital form, and to supplement runs that have been digitized only selectively and imperfectly. Currently, digital versions of historical medical titles are scattered and coverage is far from exhaustive. Large-scale digitization efforts, such as the Google Books Project, have digitized some full volumes of individual titles but rarely include comprehensive runs, as well as omitting essential folded charts and plates. Although Columbia and Harvard carried out Google

---

[1] For details on the TRAC program, see http://www.crl.edu/archiving-preservation/digital-archives/certification-and-assessment-digital-repositories

digitization efforts, HSL and the Center for the History of Medicine of the Countway did not contribute volumes. Similarly, Cushing/Whitney did not contribute to Yale's large-scale project with Microsoft.

Before planning this project, the participants assembled an informal list of out-of-copyright American medical journal titles and investigated whether the titles are available freely and completely online or whether they are candidates for digitization by this project. The study confirmed what is well known to librarians and researchers: existing collections of digitized medical journals are incomplete and often difficult to use. Some titles have been digitized but are available only for a fee; some subscription resources, such as the *American Journals Series Online*, provide access only to articles within selected publications and ignore editorials, book reviews, and other publication apparatus. Six digital collections were searched – two for-fee sites (ProQuest and JStor) and four free sites (PubMed, Google Books, Internet Archive, and HathiTrust) – and the results show that each has a limited number of journal titles in the area of medical history, and most have incomplete sets. Metadata also is often incomplete, not only as regards volume and issue information but also in tracking title changes, so that researchers are led to believe that a journal has ceased publication rather than having simply changed its name.

Ability to search varies among collections, but none is entirely user friendly. While searching for a specific article in many search engines can be very efficient, there is nothing quite so frustrating for a scholar as trying to track down a sequence of volumes from the same journal title. A scholar may, for instance, wish to investigate the coverage and development of a concept over time by checking through journal runs of a given period, volume by volume in chronological order. Google, for instance, does not easily support this kind of approach to research. Collections such as ProQuest and JStor have related material of interest to medical historians scattered over various topical sections, making searching cumbersome. HathiTrust is free and is relatively easy to search but has a limited number of titles available in this subject area. PubMed, while comprehensive in its coverage of modern journals, has fewer historical titles. Moreover, it is primarily a citation index, especially for older titles; it includes relatively few links to full-text articles. It is also complicated to search and is best used with the guidance of a librarian. (See Appendix H on the difficulties encountered in trying to compile a list of digitized journal titles and volumes.)

Internet Archive is easy to search in terms of locating specific titles but determining volume or issue number can require clicking on each individual volume if the metadata does not specify the enumeration. As the vendor selected both to digitize and to provide access to the journals online, the managers of IA are concerned with improving its search and display mechanisms. In their letter (Appendix L), they state that they will support development in these areas. At the same time, since part of the difficulty is caused by inadequate metadata provided by libraries for their digitized volumes, the MHL partners assure that they will provide IA with consistent metadata that includes volume numbers and dates, to prevent this unsatisfactory situation from arising with MHL project titles.

Selection Criteria
The contributing partners, with the assistance of MHL's Scholarly Advisory Committee (described in Appendix A) will choose titles from American medical, dental, and public health

journals, emphasizing titles that ceased publication before 1923, although the pre-1923 volumes of some titles that continued publication may also be selected. Work on compiling an initial list of possible titles began in 2011 as partners mapped their collection strengths and complementary subject areas and began to choose specific titles of value for future research. The preliminary list is provided in Appendix I; input from scholars will be sought in developing the final list of titles to be digitized.

Criteria for including titles in the population from which the partners will choose are as follows.

- Historical significance. The historical, social, or cultural importance of the topics to which the title is devoted is of primary concern – does it focus on subjects of likely current or future interest to scholars? What is the value of a particular title to the study of those subjects or issues – does it offer information or a point of view that is authoritative? What aspects or trends of the contemporary scene does it present? Might it document a change in paradigm? Publications representing polemic, sectarian, or unusual points of view have a clear value in providing a context for mainstream developments.
- Editors and authors. The relative importance, authority, or reputation of editors (personal, learned society, research institute, government agency, other) is a major criterion. In the same way, the scholarly standing and importance of the authors of individual articles are involved in assessing the significance of the title over all.
- Duration. Duration can be a sign that a title was valued as a source of important information. It also provides scope for the study of how topics developed over time, and for chronological comparisons of attitudes towards difference groups and issues.
- Rarity and Age. The relative rarity of a title affects its availability – was this title subscribed to widely and held in a large number of repositories? Or was it short-lived, so that it has become relatively rare over time as other copies have deteriorated or been lost? Older titles not only offer insights into earlier periods that are less well documented, but they also tend to be rarer due to smaller initial press runs and attrition over the years. Assuming a scarce and/or early title also has historical significance, its digitization may well take higher priority over materials more commonly held and more easily obtained, even if only scattered holdings have survived.

For titles that meet these criteria, priority for selection will be given to (1) titles that have no digital version available; (2) titles that began publication before 1850; and (3) titles that have been digitized incompletely.

B. Description of the Selected Collections

The participants in this project propose to digitize, make accessible, and preserve up to 200 historical medical periodical titles, equivalent to approximately 7,350 volumes and 2,152,000 pages.

The scope of the world's published medical heritage literature is extensive. The total is conservatively estimated by project participants to comprise at least 1.5 million volumes, about one half of which (750,000) are pamphlets (including dissertations), about one sixth are monographs (250,000), and the remaining third are journal volumes (500,000). Among the 130

American institutions and over 40 others worldwide that have significant collections of this published literature, the project participants collectively hold a substantial portion. They include three of the world's leading collections in the history of health care and medicine (Countway and Cushing/Whitney as active participants and NLM as a source of lending). Together, the participating libraries have over 400,000 volumes collected over 200 years to support the work of scientists, researchers, and students throughout the world.[2] Brief descriptions of the collections of the four participants are available in Appendices D-G.

The combined journal holdings of the four participating libraries represent a near-comprehensive collection of American medical journals. However, no single library holds the complete corpus of legacy literature, even within individual sub-domains. The proposed project will pull journal titles from this uniquely comprehensive assemblage, allowing the resulting digital resource to extend beyond the already high value of the local collections. Thus, while each of the four libraries will assume responsibility for specific titles it holds in full or nearly full runs, they will be able to draw upon the National Library of Medicine to provide missing volumes, complemented by other non-participating MHL partners such as the New York Academy of Medicine Library and the William H. Welch Medical Library at the Johns Hopkins Medical Institutions in instances where NLM may not be able to provide fill-in items.

C.   Relation to Similar Collections

While there are other significant medical history collections in North America, such as the Louise M. Darling Biomedical Library at UCLA, the Lane Medical Library at the Stanford University School of Medicine, the Wangensteen Historical Library in the Bio-Medical Library of the University of Minnesota, the Becker Medical Library at the Washington University School of Medicine, the Dittrick Medical History Center at Case Western Reserve University, and Ebling Library of the University of Wisconsin-Madison, these collections are, in the main, smaller in scope and established more recently than most of the MHL partners, and so are less likely to hold journal titles of significance not already owned by at least one of the project participants, who have built journal collections of national and international scope over a longer period. Certainly these complementary collections hold titles of local or regional interest absent from the MHL collections that would be of value in future digitization efforts, and the other collections will benefit from access to the digitized historical titles provided by the proposed MHL project, since it will enable their researchers to consult and use the digitized journals without costly traveling or utilization of document delivery services.

D. Audience and Use

Medical publications are of great value to historians of medicine and to other disciplines interested in medical developments and changing medical knowledge, and assist in tracing continuity from the past to current medical ideas. The wide range of subjects represented in nineteenth and early twentieth century medical publications will also be of utility to historians and researchers in women's studies, education, climatology, geography and natural history,

---

[2] February 2008. Directory of History of Medicine Collections. National Institutes of Health. *United States National Library of Medicine.*

vegetation and botany, temperance, chemistry, ethics, leisure studies, diet and exercise, sociology, local history, and printing and publication history.

Scholarly interest in the history of medicine monograph titles digitized by the current MHL project is evident. Columbia and Harvard began gradually loading files into IA from the Sloan-funded project as of March 2010, where they can be searched either as part of the entire IA corpus or as a separate collection at http://www.archive.org/details/medicalheritagelibrary/. As of late June 2011, over 10,060 volumes had been made public and had garnered over 255,000 hits, even though a large portion of the total have been available for only a few months and the NLM, Cushing/Whitney, and New York Public Library titles have yet to be loaded. As the site's content continues to grow, and as MHL becomes better known among researchers, the rate of use will expand ever more quickly.

While online uses can easily be counted, use of printed journal volumes is difficult to quantify, since in most of the partner libraries journals are still housed on the open stacks for in-house consultation and do not circulate. However, as of 2008 a portion of the Columbia HSL journal runs have been transferred to the off-site facility, ReCAP, and requests for those items can be tracked. Since July 2008, almost 10,500 requests for medical journal volumes have been received from Columbia and outside users, indicating the high degree of interest in the journal literature.

The MHL partners all report that pre-1923 American medical journals are in constant demand by scholars. Recent examples include an undergraduate student researching nineteenth century cholera outbreaks in the Caribbean who found valuable citations to articles in the contemporary journals. More advanced researchers have tracked discussions of yellow fever outbreaks using the American medical press to determine what kinds of public health interventions were popular in different communities. Others have explored medical arguments as they unfolded within the pages of journals, as physicians wrote back and forth contesting each other's views on questions such as the most effective interventions for treating cases of diphtheria or rabies. Ethics disputes carried out within the pages of medical journals have been the basis for research projects, as physicians leveled charges and counter-charges against each other using the public forum of the journals to air their views. Recently one researcher made use of the full run of a short-lived local medical journal that was edited by a single person between 1879 and 1881 to write a book both about the editor himself and about the evolution of local medical practice as viewed through the lens of this one particular journal. Earlier this spring, another researcher spent close to six weeks examining late nineteenth and early twentieth century runs of journals produced by the German-American medical community in New York, looking at the ways in which newly arrived émigré physicians constructed new communities, explored medical questions, and developed organizations and institutions to meet their professional needs.

By creating digitized versions of their historical medical journals, the MHL libraries will strongly affect how researchers use them. Scholars working with print journals must rely on tools such as the *Index Catalogue of the Library of the Surgeon General's Office* to track down citations or backtrack from more recent references to locate relevant earlier information, or else they must browse through the cumulated years of volumes to find the articles that matter to their studies. Jacalyn Duffin, writing of her efforts to track historical articles, states "Had those

journals been available electronically, the work would have been completed much more quickly and with far less expense – and probably with greater accuracy" (letter of support, Appendix O). Research of this type into the primary information offered by medical history journals will be immensely facilitated by the digitization project proposed here, first, by making the digitized journals freely available to scholars world-wide, and then by providing full-text search of creating machine-readable versions whose content can be manipulated in ways not possible with print. The proposed project will provide free and permanent widespread access to a rich array of resources to be mined through computer tools now available to humanities scholars.

II. HISTORY, SCOPE, AND DURATION

The Medical Heritage Library as now constituted encompasses two earlier initiatives – the MHL itself and the Digital Spine Group (DSG) – that had overlapping membership and a common interest in elevating the visibility, accessibility, and utility of history of medicine collections. Formation of the DSG began in the fall of 2009 when the College of Physicians, with the support of the Council on Libraries and Information Resources (CLIR), disseminated a prospectus for collaboration to potential partners. Catalyzed by the Open Knowledge Commons (OKC) in April 2009, the National Library of Medicine, the New York Public Library, and Countway, Cushing/Whitney, and HSL submitted a proposal to the Alfred P. Sloan Foundation to support digitization of 30,000 public domain volumes from their rare book holdings. The Sloan Foundation granted funding of 1.5 million dollars in December 2009, enabling the establishment of the Medical Heritage Library; the digitization effort has now entered its second year. The two groups joined together in 2010 as the Medical Heritage Library to further their shared vision.

The MHL's mission, informed by scholars and special collections librarians who interact with researchers from many disciplines, is to promote free and open access to quality historical resources in medicine. The goal is to provide the means by which readers and scholars across a multitude of disciplines can examine the interrelated nature of medicine and society, both to inform contemporary medicine and strengthen understanding of the world in which we live.

Over several meetings in 2010-2011, CLIR has played a central role in facilitating discussions across the broader membership to further shape the MHL's vision and goals. The partners have been working on several fronts and are continuing to formulate and seek support for new initiatives in scholarly engagement, resource discovery and delivery, and digitization.

Additional funding
In April 2011, the MHL received a Level-One Digital Humanities Start-Up Grant from the National Endowment for the Humanities. This grant will support planning activities among ten institutions and a scholarly advisory committee to continue developing the MHL. This funding will allow exploration into three areas critical to the long term usefulness and sustainability of the MHL: development of an innovative, expanded (and expandable) partnership that has both sufficient structure and flexibility; incorporation of significant collaboration and scholarly engagement as business-as-usual methodologies; and planning for the development of digital toolkits and approaches that have utility for scholars using digital medical heritage resources. The latter will include interviews with scholars about their research and teaching methods and

preferences, information that is key to improving discovery and increasing the impact of digital resources. The project furthers the MHL's mission to enhance the accessibility and utility of unique research resources in the history of medicine.

The group has also begun investigating possible funding sources for additional planning and implementation activities, such as digitization and preservation of other collections and creation of a digital platform and user interface, include the Institute of Museum and Library Services, the Andrew W. Mellon Foundation, and the Alfred P. Sloan Foundation.

Progress to date
Over the past twelve months, the MHL has made progress on a number of fronts. As of this writing, more than 10,060 monographs have been uploaded to the Internet Archive (IA); nearly 5,000 more have been digitized and are awaiting processing and deposit. Subject areas include general public health topics, psychiatry, popular medicine, medical directories, forensic medicine, and therapeutics, as well as surgery, anatomy, and physiology. The 'browse list' of topics on the MHL's IA homepage (http://www.archive.org/details/medicalheritagelibrary) demonstrates the breadth of the history of medicine, ranging from 'Abattoirs' to 'Zulu War, 1879.'

MHL content has generated approximately 255,000 downloads since the first deposit in early 2010. The single most downloaded book (with more than 875 downloads) is Volume 2 of *Per il XXV Anno Dell'Insegnamento Chirurgico di Francesco Durante nell'Università di Roma, 28 Febbraio 1898*, edited by Roberto Alessandri. While he is little known outside neurosurgery circles in this country, Durante (1844-1934) was a pioneering surgeon, esteemed teacher, and leading political figure in his native Italy. The volume's rarity outside Italy – OCLC locates only four copies of the set in North America and another in Paris – is likely a factor in its popularity. There is a demonstrated need for electronic access to even the most seemingly esoteric publications.

The MHL has established two-way communications with peer and user communities in a number of ways. The webpage at http://www.medicalheritage.org and Facebook presence (Medical Heritage Library) are frequently updated; MHL also communicates with a broad base of peers and users via a group of nine humanities listservs. Content on the webpage is highly relevant to a growing body of users, currently averaging 125 per week, 65% of whom are new. Visitors stay for 4.5 minutes on average and 54% visit multiple pages. Now that the mix of information provided via the website has demonstrated its effectiveness, MHL will begin a advocating for its inclusion on the link lists of academic programs and repositories.

As a first step in strengthening relationships with scholars and promoting the use, MHL partner representatives presented a session at the American Association for the History of Medicine's annual meeting at the end of April 2011. Speakers placed the MHL resources in the context of digital resources in the history of medicine to an audience of scholars. The session, which was oversubscribed and closed to aspiring registrants, will be featured in the AAHM's summer newsletter. The MHL's Scholarly Advisory Committee, composed of academics in the digital humanities, history, and history of medicine, will continue to be helpful in guiding the MHL in developing delivery strategies and tools, content selection, and communications with users.

A user survey is now available on the MHL's home page in IA and on the MHL website (http://www.medicalheritage.org/?page_id=281). It is being distributed to users at MHL partner repositories and will shortly be sent to peers and users via listservs. The numbers of books digitized and downloaded are important output measures; still more critical are outcomes – the impact digital resources have on the work of users. The MHL partners continually seek ways to assess these benefits.

Future plans
In the near term, the MHL plans to increase the availability of quality digital resources in the history of medicine. In November 2010, to better understand how peer libraries' digitization activities should influence its future projects, the MHL distributed a survey via listservs to repositories' holding medical heritage materials.[3]

The objectives were, first, to locate already-digitized materials that could be include in or point to from the MHL site, both to reduce duplication and to support users' research; and, second, to identify possible collaborators for future digitization efforts. Of the 62 unique complete responses from different types of repositories, mostly medical school or college and university libraries, nearly 80% had digitized materials, though those materials were not necessarily publicly accessible. Prints and photographs were the most common type of materials digitized, followed by archives. Nearly all the respondents wanted to digitize more, both because of user demand and because they now had the internal capacity to do so. Those that could not continue digitizing cited lack of staff and funding. Other questions concerned locations of digitized files (mostly local or institutional), possession of regional or subject area collections that were good candidates for digitization, and funding sources. The MHL will pursue additional information about public availability of these and other digitized materials, regional or subject area collections, and degree of involvement in interface development, as well as seeking more respondents throughout the U.S. and Canada. Based on the survey results, MHL will shortly approach two potential contributors of digitized content to learn more about how collaboration could be arranged.

Like the digitized books now available through the MHL, content digitized in this project will be freely available to all users, searchable, and downloadable. The MHL seeks to enhance the usability of content through additional interface and tool development. Important components of this effort are scholarly engagement, understanding how best to support for teaching and research, and awareness and re-use of advances made by our peers and those working in related fields. Drawing on the scholarly engagement activities completed during the Level-One Digital Humanities Start-Up Grant, MHL intends to seek funding for application development. While the journals to be digitized through the present proposal will of course be made fully available to research through conventional means, the success of the other MHL efforts to broaden new forms of access and discoverability has the potential to provide exciting additional enhancements.

---

[3] Summary analysis of survey results can be found at http://www.medicalheritage.org/?p=413.

III. METHODOLOGY AND STANDARDS

The proposal participants are not only leaders in their field, but also bring years of institutional experience in digitization to this project, including the ongoing Sloan-funded project, during which they have developed shared procedures that improve the efficiency and effectiveness of their efforts. The selected service provider, Internet Archive, is part of the Sloan-funded project and has worked successfully with several of the participants on a variety of earlier projects over the past five years.

A. Completeness of Periodical Runs

It is sadly often the case that libraries lack complete holdings of journal runs. Journals are extremely difficult for a single library to preserve, due to the exigencies of locating missing volumes, especially since many libraries refuse to lend whole journal volumes, even for the purpose of creating preservation copies. Organized cooperation is essential for the digitization of journals, and the MHL group is well situated to facilitate the complex processes involved. It is the intent of this project to produce a complete digital run of the public domain volumes of each title in the project if at all possible. The subject specialists of the member libraries have worked together to identify titles of research value to be digitized and an on-going effort is in place to compare the holdings of the group. The libraries with relevant holdings will contribute volumes as needed, enabling the participants to assemble runs by pooling their collections.

Each partner will take primary responsibility for a number of titles, digitizing its own holdings first. If there are gaps, the partners will make every effort to borrow fill-in volumes from the National Library of Medicine, which has agreed to participate in the project specifically in order to supply volumes to fill gaps, although it will not receive any funding from the project. A letter of commitment from NLM is provided in Appendix K.

If NLM cannot supply missing volumes, participants will seek them from each other and from the other MHL partners. If none of the partners can supply the missing volumes, an effort will be made to find a copy elsewhere if possible, and either borrow the volume to digitize or obtain a digital version from the owner.

B. Digitization by Internet Archive

Digitization will be performed by the Internet Archive at their facilities in Boston (for the Countway materials) and Princeton (for the College of Physicians, HSL, and Cushing/ Whitney materials), and all of the project titles will be made freely accessible on the IA website. Two of the partners in the current Sloan-funded project are already using IA as their digitization service provider. They have found the process efficient and the error rate very low, and the other two partners, who are carrying on the Sloan project digitization in house, have determined that using IA will be a more cost-effective and efficient approach than continuing with their in-house operations.

A 501(c)(3) non-profit that was founded in 1996 to build an Internet library, the Internet Archive's goal is to offer permanent access for researchers, historians, and scholars to historical collections that exist in digital format. IA makes volumes available at http://www.archive.org

for online reading and for download. They provide page views, plain text, both color and black-and-white PDF versions, a Kindle version, and several other formats. The IA public site enables simple title keyword searching as well as an advanced search capability. They provide a special entry page for the MHL titles at http://www.archive.org/details/medicalheritagelibrary. Researchers thus have the choice of limiting their searches to the MHL corpus or searching across all of IA's content.

Because IA has already created a site for MHL content that is continually being populated with material digitized by the partners, because it has committed to keeping that content freely available to the public in the future, and because it is already serving as the digitization service for two of the libraries participating in the current monographs project, it was selected as the digitization service provider for this proposed project without a bidding process. Their price of $.11 per page ($236,725 for 2,152,000 pages) is very competitive for bound journal volumes. IA's base digitization price is $.10 per page, but they have negotiated the $.11 rate because many of the volumes to be digitized are brittle and fragile. The price quote is provided in Appendix L.

C. Intellectual Property Issues

Only titles in the public domain will be digitized. If a title's run continues into the copyright period, only the public domain portion of the run will be included in this project. Metadata provided by IA and HathiTrust indicates which library contributed the digitized version, thereby informing scholars where physical copies of any subsequent in-copyright volumes may be found.

D. Post-Project Sustainability

All of the digitized titles will be added to the Internet Archive's public site. Online images and unlimited downloads are available free of charge to the public from the http://www.archive.org, and IA is committed to preserving its copy of the master files. Each of the libraries receives a copy of the files as well, which they are preserving separately, as discussed below.

E. Digitization Methodology and Standards

IA creates images in JPEG2000 format. They digitize standard-size materials in color at 300 dpi, using a Scribe scanning workstation composed of a frame that holds two digital single-lens-reflex (DSLR) cameras to enable capture of both the verso and recto pages of the book to be digitized, a floating V-shaped book-cradle whose angled design places minimal stress on books, a glass platen that is raised and lowered by means of a foot pedal, two banks of museum-grade lights that illuminate the book and a computer that captures the color images from the camera and performs some of the pre-processing. Scribe stations are calibrated and aligned before being used. A color target and a white card are shot with each book for reference for ICC-based color management. Before any digitized images are uploaded, the scribe operator checks for missing shots, crop/de-skew issues, accurate page labeling (i.e. title page, covers, TOC), and adds any notes regarding defects in the book (i.e. missing pages, tight binding). Page images are then uploaded to a central scanner management and image processing computer. A second quality assurance process is performed at the Scanning Center after the books are made available online.

Descriptive metadata
IA derives its descriptive metadata from MARC (Machine Readable Cataloging) records provided by the libraries whose volumes they digitize. Before digitization the project partner with primary responsibility will update its MARC record for the title as needed to assure that it correctly reflects title and publication information. All updated catalog records will subsequently be added to OCLC's World Cat for national and international access. Partners will also compile accurate holdings statements.

IA establishes a Z39.50 connection with the library's catalog in order to download the MARC records. The library uses IA's online Wonderfetch worksheet to identify the relevant MARC record ID numbers and to generate a packing list. When the volumes are received in the scanning center, they are reviewed and items that do not fit the digitization profile and similar anomalies are noted. The Wonderfetch list is used to locate each volume's corresponding MARC record and a quality check is performed to ensure that the item in hand and the MARC record match exactly. IA's Biblio Tool automatically creates a unique, persistent identifier for each item and the MARC record is attached to that identifier. The IA metadata clearly names the library whose physical volume has been digitized.

Library quality control
During the first year of the current Sloan-funded Digital Medical Heritage Library Project, the two partners using IA as their digitization agent (HSL and Countway) developed post-digitization quality control procedures. These procedures will be followed by all members of the project proposed here.

IA inserts a paper slip in each volume, specifying the bibliographic identifier for the digital version. As each returned item is received by the library, the paper slip is checked to verify that the bibliographic identifier matches the physical item. The returned IA excel spreadsheet is checked to verify that the identifiers match the bibliographic data provided by the library. A random 5% sample of each shipment is reviewed on the IA site. With original in hand, staff view every scanned page to assure that the item is complete, fully readable, and attached to the correct bibliographic record. Should significant problems be identified, the entire shipment is checked and if necessary re-digitized.

Enhancing discoverability
Improving discoverability and access to history of medicine materials is a primary goal of the MHL. As with the current Sloan-funded effort, titles digitized through this proposed project will be discoverable in several ways. Each library will make the digital versions of its own titles accessible through links to IA in its local OPAC, and the titles will of course be searchable on the IA site, where full text searching is provided. The digital files will also be submitted to HathiTrust, which provides full text searching across all of its content and full viewing of all public domain content; again, each library can provide links from its own catalog to the HathiTrust copies. Searches in Google will direct readers to the MHL copies in IA and HathiTrust, as will searches in WorldCat.

In a separate effort, MHL partners have applied to CLIR for a Hidden Collections Grant to develop a semantic application that works with search engines to transform EAD finding aids on

the fly. Users would be able to manipulate entity relationships described by the semantic format using visualization and analysis software. That effort, if funded, will advance the MHL's long-term goal of developing format-unified discovery of digital resources in the history of medicine (i.e. federated searching of books, archives, images, and other formats). The semantic finding aid would provide yet another means of discovering the journals to be digitized under this proposal, by providing the foundation needed to connect archival materials with both the journals and the monographs being digitized under the Sloan-funded project.

Preservation of the digital assets

In addition to making the digitized titles available through Internet Archive, where they will also be preserved for the future, each MHL partner will take responsibility for assuring preservation of its own files. Columbia, Harvard, and Yale are all members of the HathiTrust, and will individually deposit copies of their project titles there. The College of Physicians anticipates joining HathiTrust in the near future.

HathiTrust is a partnership of major research institutions and libraries working to ensure that the cultural record is preserved and accessible long into the future. In addition to making the MHL content freely available to the public along with the almost nine million volumes it has already preserved, HathiTrust is committed to bit-level preservation and format migration of materials created according to these specifications as technology, standards, and best practices in the library community change. They use standard and open content formats that meet community-accepted digital preservation standards. Details of their preservation program are available at http://www.hathitrust.org/preservation.

Preservation metadata

HathiTrust employs METS (Metadata Encoding and Transmission Standard) for its digital objects (Archival Information Packages) from all digitization sources. A "source" METS file is assembled from metadata provided to HathiTrust in the Submission Information Package, to which is added information about the content information from the time of its creation to the time it enters the repository. A "HathiTrust" METS file is created on ingest and includes a subset of the source METS file data, but is primarily a record of the object from the time it enters the repository forward. The source METS is kept for preservation purposes, while the HathiTrust METS is used for both preservation and access purposes (i.e., in both the Archival and Dissemination Information Packages). Preservation information included in the METS file is recorded using PREMIS (Preservation Metadata Implementation Strategies). The PREMIS implementation used for most volumes in the repository is PREMIS 1.0; however, the implementation used for more recently added Internet Archive-digitized volumes is PREMIS 2.0.

F. Collaboration and Coordination

As MHL members and partners in the Sloan-funded project, the four participant libraries already have a strong collaborative working relationship. The group holds weekly conference calls to coordinate activities and develop procedures, and makes use of a wiki hosted by Harvard to share documents, meeting agendas and minutes, and other information. The group also plans to meet in person once during each project year, either during the annual conference of the American Association for the History of Medicine or during a meeting of the full MHL membership,

whichever presents the most convenient scheduling.

IV. WORK PLAN

This work plan assumes that Columbia and Cushing/Whitney will each assume responsibility for one-quarter of the digitization, while the College will be responsible for 30% and the Countway for 20%. The total is estimated at approximately 2,152,000 pages amounting to 7,350 volumes. Each partner will take primary responsibility for specific titles. The library responsible for a given title will:

- o Repeat the search of the title to verify that it has not been digitized since the initial search effort conducted for this proposal. Titles that have been digitized in the interim will be removed from the project provided the digital versions are complete and fully legible.
- o Assemble its own holdings of public domain volumes and determine whether they can be digitized, based on physical condition.
- o Identify any gaps.
- o Request fill-in volumes from NLM and other MHL partners.
- o If necessary, attempt to locate fill-in volumes outside of the core group.
- o Update the library's MARC record as needed.
- o Add the title and volumes to the IA Wonderfetch database.
- o Ship the volumes to IA.
- o Receive the volumes back from IA.
- o Carry out quality control on every volume to assure bibliographic/metadata accuracy.
- o Carry out visual inspection of 5% of the digital files to assure completeness and legibility.
- o Notify IA and the project manager of any problems, and work with IA to resolve them.
- o Return borrowed volumes.
- o Reshelve its own volumes, including transfer to off-site facilities as appropriate.
- o Enter the digital version into its local catalog and OCLC.
- o Begin the process of submitting the digital files to HathiTrust to ingest for long-term preservation and to add to their publicly accessible collections.

The project will last two years. Repeating the searches for digitized versions of titles in order to avoid duplication of effort is expected to take seven months. Work on assembling runs that are as complete as possible will begin in the third month of the project, with catalog and holdings clean-up following and the other tasks necessary to prepare materials for shipment to IA. The actual digitization phase of the project will last approximately twelve months, including any corrections. The Countway plans to complete all of its digitization during Year 1, while the other libraries expect to accomplish about half of their digitization in each year of the project. The final months will be devoted to access issues and to ingest of the digitized versions into HathiTrust.

PROJECT TIMELINE	2012						2013												2014					
	Jul	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun
Publicize project, organize partner communications	■																							
Confirm & sign contract with IA for digitization	■	■																						
Hire and train staff as needed	■	■																						
Re-review selected titles & partner assignments	■	■																						
Retrieve volumes, collate runs, locate fill-in volumes				■	■	■																		
Clean up cataloging & holdings information					■	■	■																	
Prepare & ship batches to IA						■	■	■	■	■	■	■	■	■										
IA: digitize & return batches							■	■	■	■	■	■	■	■	■									
Perform quality control on returned batches								■	■	■	■	■	■	■	■	■								
IA: make corrections as needed									■	■	■	■	■	■	■	■	■							
Reshelve volumes, return borrowed volumes									■	■	■	■	■	■	■	■	■							
Add digital versions to local catalogs										■	■	■	■	■	■	■	■	■						
Submit digital files for ingest to repository											■	■	■	■	■	■	■	■	■					
Publicize project completion, write final report																				■	■	■	■	■

V. STAFF AND BUDGET

A. Staff

Maura Marx, Executive Director of the OKC, will have project oversight and will be responsible for overall project direction, including managing the contract with IA, overseeing the budget, paying invoices for digitization, and assuring that the project is successfully completed as described above. Before establishing OKC in 2008 to catalyze work supporting digitization of print collections, open access to digital materials, and the long-term preservation of such works, Maura Marx launched the Boston Public Library's digital library program and was responsible for system-wide delivery of digital services. She is currently the project director for the Sloan-funded digitization project. OKC will contribute 5% of her time to this effort.

Each library will contribute the costs for administrative staff members' participation in the project. Depending on local circumstances, libraries are contributing the time of project managers, subject specialists who will concentrate on the selection and publicity issues, and other staff such as day-to-day project supervisors or catalogers who will improve the quality of intellectual control of the titles to be digitized. Resumes of key personnel appear in Appendix M.

College of Physicians
- o Project Manager: Annemarie Brogan, Librarian, 5%
- o Subject Specialist: Robert Hicks, Measey Chair for the History of Medicine, Director of the Mütter Museum & Historical Medical Library, 2%
- o Cataloger: Sofie Sereda, 7%
- o Project Supervisor: Lori Jahnke, CLIR Fellow, 1%

Columbia University
- o Project Manager: Janet Gertz, Director of the Preservation and Digital Conversion Division, Columbia University Libraries, 5%
- o Project Supervisor: William Austin, Special Projects Librarian, Columbia University Libraries, 25%
- o Subject Specialist: Stephen Novak, Head, Archives and Special Collections, Augustus C. Long Health Sciences Library, 5%

Countway Library of Medicine
- o Project Manager: Kathryn Baker, Deputy Director of the Center for the History of Medicine, 10%
- o Rare Books Cataloger/Subject Specialist: Joan Thomas, 10%

Cushing/Whitney Medical Library, Yale University
- o Project Manager: John Gallagher, Deputy Director of Public Services, 3%
- o Subject Specialist: Melissa Grafe, John R. Bumstead Librarian for the History of Medicine, 5%

B. Budget

Funding from the Endowment is requested in the amount of $348,293. Approximately 68% of the requested funding will pay for digitization services provided by IA. As indicated above, Internet Archive will provide digitization services for this project at a price of $.11 per page. The remaining costs cover shipping to/from the scanning centers, and the cost of library support staff to retrieve and prepare volumes for digitization (barcoding, completion of inadequate cataloging records, identification and fill in of missing volume, loading and unloading of book carts – all the activities that are necessary to bring the books to the scanning facility), to return them to the shelf afterwards, and to perform quality control review of the digitized content.

The partners will contribute a further 44% toward the total project costs as cost share, including salary and fringe for administrative and supervisory staff from OKC and the four participants, shipping of volumes borrowed from NLM and others, travel costs for the annual meetings (assumed to be held in Philadelphia), long-distance charges for weekly phone calls among the partners, maintenance by the Countway of the project wiki and website, and ingest fees for HathiTrust that cover the costs of adding digital files into the collection. HathiTrust per-gigabyte fees have been computed based on the average size of IA files.[4] Details for each library are provided in Appendix N.

VI. DISSEMINATION

The products of this collaborative digitization project will be freely available at Internet Archive as part of the Medical Heritage Library. IA makes volumes available in a variety of formats for

---

[4] Please note that HathiTrust ingest fees are computed using their current pricing structure. This structure is expected to change in the coming year, but no decisions on changes will be made until October.

online viewing and downloading (e.g. page views, plain text, PDF, Kindle, and several other formats). While digitized content will be available immediately through IA and through HathiTrust once digitization is complete and the files have been ingested, the MHL is working toward development of a format-unified discovery environment that will support multi-disciplinary study of the history of medicine sources. The recent MHL submission to the CLIR Cataloging Hidden Collections program proposed applying semantic web technologies to archival description as a means to enhance metadata and to promote inter-operability among material formats and access environments. The journals to be digitized through the present proposal will be integrated with this access program to allow scholars greater flexibility in discovering sources and the analyzing content (e.g. text analysis, data mining, entity extraction).

In addition to promotion of the digital content and project news through the website at http://www.medicalheritage.org/, the members of the MHL are active in attending scholarly conferences and workshops. This activity is invaluable for soliciting user feedback and raising awareness of both the opportunities and issues surrounding digital scholarship. Recent examples of MHL activity include presentations at the American Association for the History of Medicine Annual Meeting (2010 and 2011), the CLIR Sponsors' Symposium (2011), and the Linked Open Data in Libraries Archives and Museums Summit (San Francisco 2011). Project members will continue to raise awareness through this program of scholarly and professional outreach for the term of this digitization project and beyond. The MHL also works with its Scholarly Advisory Committee to gather feedback on research trends and emerging methodologies in digital scholarship. Through our advisory committee and other outreach activities, we are building a community of scholars who can inform project development, as well as engage additional users who will benefit from the valuable digitized resources this project will create.