

Firefox Scholar 2.0:

**Building the Next Generation
Scholarly Research Tool**

A proposal to the Andrew W. Mellon Foundation

Center for History and New Media
MS 3G1
George Mason University
Fairfax, VA 22030
Phone: 730-993-4524
E-mail: dcohen@gmu.edu, roy@gmu.edu



<http://chnm.gmu.edu/>

1. Introduction

Over the past decade, tremendous progress has been made in putting scholarly and research materials online. Mellon Foundation-funded initiatives such as the invaluable JSTOR and ARTstor, major digital library projects from the Library of Congress, University of Michigan, and New York Public Library, commercial products from ProQuest and Thomson, as well as the efforts of thousands of smaller institutions, have placed vast digital collections at the fingertips of scholars, teachers, and students around the globe. Nearly every library, museum, and archive has at least put their catalog online, and many have digitized significant parts of their holdings.

While hundreds of millions of dollars have been spent to create these extraordinary digital resources, far less funding and effort has been allocated for the development of tools to facilitate the *use* of online resources. The web browser is by far the scholar's most commonly used digital tool—indeed, for the vast majority of scholars the browser is the only digital tool they use for their research, aside from word processing software and an email program. But this browser remains for the most part a passive window into materials on the web. (And commercial word processing and email programs offer few concessions to the specific needs of scholars.) In the same way that researchers demand extensive annotation and documentation in scholarly editions, they also need a more intelligent web browser with scholarly apparatus that enables and enhances high-quality research in a digital age and replaces awkwardly adapted commercial tools.

At the Center for History and New Media (CHNM) at George Mason University we recently received a grant from the Institute for Museum and Library Services (IMLS) to begin to build such a pilot version of a new, smarter web browser based on the popular open source Firefox browser. Rather than being a passive window, this new browser, which we are calling *Firefox Scholar*, will consist of a major extension to the browser package and will provide an active desktop for conducting digital research. *Firefox Scholar* will intelligently sense when its user is viewing a digital library or museum object, an online book or article, a library catalog entry, or a bibliographic citation; the browser will then capture—automatically, seamlessly, and accurately—metadata from the page, such as the creator and title of an object, and its copyright information. Moreover, *Firefox Scholar* will allow users to organize and sort this information, as well as store full copies of digital objects (not just their citation information) if so desired by the user and permitted by the rights management metadata.

Under IMLS funding we are working with our partner, VIVA (the Virtual Library of Virginia, which represents all libraries of the thirty-nine state-assisted colleges and universities as well as thirty of Virginia's private non-profit colleges and universities—a heterogeneous collection of institutions ranging from major research universities to small liberal arts colleges to rural community colleges) to create a “1.0” pilot version of *Firefox Scholar* that will focus on enabling users of college and university libraries in Virginia to create personal collections for their research. But *Firefox Scholar* 1.0 represents only a first step; scholarship in the next decade will be advanced not only by increased accessibility of research materials but also by the effects that occur when

materials are brought into networks of relationships and analyzed in complex ways by communities of scholars. This second generation of the web—sometimes dubbed “Web 2.0”—entails the aggregation and mining of distributed data (as seen currently in fairly rudimentary ways in the rise of “mashups” of information and documents from different servers, e.g., the placement of houses for sale on a Google Map) and a corresponding ability to extract and exchange the meaning and metadata stored in digital objects (seen most ambitiously in the efforts of those involved with implementing the “semantic web”).

Given these trends in the digital realm, we are seeking funding from the Mellon Foundation not only to expand the 1.0 version of *Firefox Scholar* into a universal bibliographic and note taking tool that not only recognizes the vast majority of catalogs and collections online, meets the needs of a wide range of researchers (including those with accessibility needs), and integrates with existing bibliographic and word processing tools, but also to develop our pilot project into a robust, “2.0” version of *Firefox Scholar*, which would take advantage of this social, semantic web. This version would enable users not just to create their own local bibliographic collections, but to share, aggregate and “mash up” those collections with other researchers and analyze their collections using next generation data mining and visualization tools.

The resulting software would foster scholarly collaboration through shared workspaces both server- and client-side, as well as leverage the network effects of a wide user base to facilitate the creation of new communities of practice around documents, collections, and disciplines. (*See media studies “use case” below.*) Moreover, by relying on the distributed intelligence of the network—a central characteristic of Web 2.0 and what Yochai Benkler calls “commons-based peer production”—we will no longer be reliant solely on the efforts of major libraries and commercial vendors to digitize the analog cultural heritage. Each scholar will be able to add his or her own “micro-collection” to the digital corpus. In short, the funding we request would allow us to develop the functionality of our current pilot project in both scale and kind, exponentially expanding its utility to the scholarly community.

2. *Firefox Scholar* 1.0

As we initially framed the problem more than a year ago to IMLS, libraries, museums, and commercial vendors have made enormous strides in putting the world’s cultural heritage at the fingertips of scholars, but they have done relatively little to enable researchers to seamlessly incorporate those materials into their scholarly work. In effect, the availability of online scholarly resources has vastly outpaced the existing scholarly digital tools. In the real world, scholars do not simply read, look, and listen; they actively take notes, make photocopies, shoot photographs, and draw connections among objects housed in different collections.

These scholarly practices, however, have not been well reflected by many (if not most) library and museum websites or even well-funded commercial initiatives. Most of these institutions have spent a significant amount of money on their sites, offering search

capabilities and digital resources that dramatically extend the reach of individual researchers. Unfortunately, once users find the objects for which they are looking, the utility of digital resources drops considerably. Researchers have cribbed together ad hoc systems to capture, organize, and annotate their findings, but with no conventions for storing or sharing them and not even easy means of locating notes taken a few months earlier.

We thought we should be able to do better. One doesn't expect a researcher to leave her pencil and notebook outside the library, forcing her to put down her book and run outside every time she wants to jot down a note; she is able to take notes in the library, at the site of her research materials. We came to believe that the same should be true of digital libraries and databases—a researcher should be able to create and annotate her unique personal collections from within the common denominator across all such sites: the browser itself. Moreover, we argued, a tool to do this should be available freely and openly, less for reasons of cost (although we suspected that such a tool would helpfully increase market competition for standalone applications such as EndNote), but because the library and museum community could further improve it and extend it in the best interests of themselves and their scholarly clients.

We therefore proposed to build a major open-source browser extension for libraries, museums, and digital databases—*Firefox Scholar*—that would enable the rich use of their Web collections with no cost—either in dollars, or probably more importantly, in secondary technical costs related to their Web servers. *Firefox Scholar* would relieve libraries and museums of the burden of setting up server-side applications for the storage of personal collections by their users). An open source descendant of Netscape, Firefox offers developers a truly cross-platform starting point from which to build new functionality into the browser window through extensions (using JavaScript) and new interfaces (using the XML-based XUL language). For our purposes, Firefox's extensibility and flexibility make it an ideal foundation; *Firefox Scholar* itself exists as a single extension, easily downloadable, installable, and upgradeable from within Firefox, which has rapidly gained popularity as a secure, feature-rich (but speedy) alternative to Internet Explorer. In the United States, Firefox has attained 16 percent of the market share for browsers in only twenty months since its 1.0 release; in educational settings this percentage is even higher and continues to grow at a fast clip.

Currently on its way to a beta release, *Firefox Scholar* 1.0 will significantly augment the capabilities of this popular Web browser. When a user loads a web page, a small icon appears alongside the URL if any digital library or museum objects are detected (much like the Firefox icon that indicates an available RSS feed). Clicking this icon highlights each object that can be “collected,” and subsequently clicking on a given object will save it into the user's personal collection along with the available metadata. If there are no recognizable library or museum objects on the page, there is an option to save the entire page as an object in its own right. (See *Figures 1 and 2.*)

Bibliographic Info

http://infosparc.gmu.edu/cgi-bin/Pwebrecon.cgi?v1=11&ti=1,11&S Google Search

GEORGE MASON UNIVERSITY   **University Libraries**
Library Catalog

Database Search Headings Titles Patron Login Request History Help

Database Name: George Mason University
Search Request: Left Anchored Title = american revolution
Search Results: Displaying 11 of 60 entries

previous next

Bibliographic Holdings MARC format

The American Revolution.

Database: George Mason University
Main Author: [Rankin, Hugh F.](#)
Title: The American Revolution.
Primary Material: Book
Subject(s): [United States--History--Revolution, 1775-1783--Sources.](#)
[United States--History--Revolution, 1775-1783--Personal narratives.](#)
Publisher: New York, Putnam [1964]
Description: [1st American ed.]
382 p. maps. 22 cm.
Notes: Bibliography: p. 360-374.

Database: George Mason University
Location: Fenwick item not available due to renovations, Summer 2006.
Call Number: [E203 .R3](#)
Number of Items: 1
Status: Not Checked Out

previous next

Record Options

Select Download Format Full Record Format for Print/Save

Done Scholar is loaded

The screenshot displays the Firefox Scholar application window. The top browser bar shows the URL `http://infosparc.gmu.edu/cgi-bin/Pwebrecon.cgi?v1=11&ti=1,11&S`. The Scholar interface is titled "Bibliographic Info" and "Scholar 1.0 Alpha". On the left, a sidebar lists project folders: "My Library", "Another Test Pr...", "Test Project", and "Yet Another Pro...". A red arrow points to the "Book" icon next to the selected record in the main list. The main list contains several entries, including "Biblical and classical myths", "Britain and the Americas", "British writers", "Blue book, Philadelphia furniture", and "The American Revolution" by Rankin. The right pane shows the detailed record for "The American Revolution" by Rankin, Hugh F., with fields for Title, Author, Date Added, Modified, Source, Rights, Series, Volume, Number, Edition, Place, Publisher, and Year. Below the list, the George Mason University logo and "University Libraries Library Catalog" are visible, along with navigation buttons like Database, Search, Headings, Titles, Patron, Login, Request, History, and Help. The search results section shows the database name, search request, and results count. Navigation buttons for "previous" and "next" are present. At the bottom, tabs for "Bibliographic", "Holdings", and "MARC format" are shown, followed by the title "The American Revolution." and detailed metadata including author, title, primary material, subject(s), and publisher.

Figures 1 & 2. Firefox Scholar automatically recognizes that the user is looking at a record for a book on the George Mason University catalog. Clicking on the book icon saves the citation information into the user's personal Scholar collection. The user can now begin to take notes on the book by selecting the notes tab, and can also tag the book or place it into specific project folders. Firefox Scholar will also capture the actual digital object—HTML, PDF, etc.—if it is available online and, then, allow the researcher to annotate that object.

To be truly useful to scholars, *Firefox Scholar* naturally needs to be able to recognize discrete digital objects, from MARC catalog records to digitized versions of physical objects in digital libraries and idiosyncratic museum exhibits. To make this work, we chose to implement object discovery using an open plug-in architecture, enabling users

to install the autodiscovery filters they need at a given moment while leaving open the ability to connect with future standards and technologies as they become available. In other words, without figuring out the complexities of how to connect to a z39.50 server or how to install “translators” for specific collections of interest to them, users merely go to a library, museum, or repository site and the browser lets them know that it can automatically save citation information or even whole documents or objects if a translator is available. If a translator isn’t available for a site, one can be made in about fifteen minutes (by either the institution or the individual user) and added to a central repository for the use of all researchers.

In cases where rich metadata is exposed, either through RDF or embedded microformats like OpenURL or OAI, *Firefox Scholar* will seamlessly ingest that information. Unfortunately, however, the vast majority of online resources are presented as embedded text within web pages, and pulling information out of the raw HTML (“screen scraping”) is still the only way to extract metadata from many pages (inelegant as it is). To that end, we have begun (with the encouragement of the Mellon Foundation) to take advantage of some of the work done at MIT’s SIMILE project, adopting their screen scraping architecture for semantic information extraction as well as their “Solvent” tool for creating scraping translators.

We believed when we wrote the IMLS proposal (and believe even more so now) that *Firefox Scholar* will have a profound impact on libraries and museums and on the work of scholars. It will relieve libraries and museums of the need to build personal collection tools for their users and greatly leverage the substantial investment they have already made in digitizing collection materials. But the greatest beneficiaries will be “ordinary” scholars; it will allow them to seamlessly collect, organize, and annotate digital materials from diverse, heterogeneous sources and enable better provenance and rights tracking for items collected in the course of scholarly research. Their students will also benefit—for example, the automatic metadata extraction will promote much more accurate citation practices among students. The overall philosophy is to make *Firefox Scholar* useful to the widest base of users, especially those with little computer expertise. As a result, we believe *Firefox Scholar* will bring scholarly research practice (and not just scholarly collections) into the digital age.

3. Building *Firefox Scholar* 2.0

Although we are excited about our progress on *Firefox Scholar* 1.0, we only intended it as a pilot—one that would not work with all the major research collections and one focused on libraries in the Commonwealth of Virginia. Moreover, as originally formulated, *Firefox Scholar* was essentially a personal silo for information. Recently, however, we have grown increasingly convinced that this development is only the first step toward the real promise of *Firefox Scholar*. The truly groundbreaking nature of this tool will only become evident once we are able to help individual users connect collections with those of other users, and with the world (and web) at large.

In addition to pulling metadata and material off the web and storing it locally, we envision new ways that this material could in turn be shared and combined with other material spread across the Internet. Such an approach accounts for the growing popularity of services such as del.icio.us (now owned by Yahoo), which allows people to share and tag web pages they bookmark, and the recent launch of Google Notebook, which provides (as the name suggests) a panel for taking notes on web pages and sharing those notes with others. Both of these products obviously only deal with online materials, however, and they do so in only rudimentary ways. They also are essentially unusable for critical scholarly resources such as JSTOR, since those gated materials often mask their web addresses in the proxies that universities use to ensure that users are properly logged in; for instance, at George Mason all of the articles in the *American Historical Review* have the same URL, <http://mutex.gmu.edu:2048/login>, so a user of Google Notebook taking notes on different articles would have them all unhelpfully lumped together by Google. By being able to combine the collection, tagging, and annotation of both print and online materials—gated as well as open—and in being able to store and share not merely a simple web address but the much more important (and defining) information such as the author and publication information, *Firefox Scholar* will satisfy the more complex and sophisticated demands of researchers, teachers, and students to go beyond the simple annotation of web pages offered by these commercial tools.

In doing so, *Firefox Scholar* joins the more expansive vision of a semantic-based platform (where the web doesn't exist merely to display digitized documents but rather to exchange in powerful ways the meaningful elements embedded in these documents), which is now coming into being with the substantial help of the Andrew W. Mellon Foundation. Various projects have explored pieces of this new online scholarly platform, using the Resource Document Framework (RDF) as a sort of Esperanto by which applications can pass data back and forth. Unlike Google Notebook or del.icio.us, which only handle (sometimes faulty) URLs, RDF provides a way to talk about, e.g., the important metadata of a journal article itself, rather than just its URL (if it even has one).

Yet Mellon's platform could be further strengthened by finding a way to let researchers—including and perhaps especially those who have never heard of the semantic web or some of the ideas circulating around "Web 2.0"—participate in, help to build, and see the benefits of this new platform. Since reference software packages like EndNote and RefWorks are familiar and extremely useful tools to these researchers, it makes sense that a reproduction of this tool in the browser could offer a natural entrée into this world. Moreover, such a tool would be critical for bringing offline materials onto the web in a standard, semantically rich format (as when a researcher scans and describes an archival document or annotates and tags a book) and to put those materials into relationships with online resources (which, for obvious reasons have been the focus of most semantic web and Web 2.0 efforts).

Once the researcher's references and notes are in *Firefox Scholar* and this browser-based software is linked to other pieces of the Mellon semantic web network, many forms of advanced digital research become possible, such as network effects,

autodiscovery, recommendations, data mining, and visualization. A user of *Firefox Scholar* who first selected it simply because it was free and integrated with his browser might discover new books or articles to read by having his collection scanned by a centralized server (see below). He could collaboratively build bibliographies and share them with others, find the most popular citations in a given field, or link into visualization or other data presentation systems easily by sending their personal *Firefox Scholar* collections into those digital environments. Just as a repository of semantic information becomes exponentially more useful as more users contribute and receive information, as the number of *Firefox Scholar* users grows, so will the power of the Mellon platform.

Our general principle here is a simple one: digital tools will enjoy the fastest and widest adoption if they meet immediate and concrete needs for users, particularly unsophisticated users. Obviously, the most important activities for scholars (after teaching) are gathering, organizing, annotating, analyzing, and citing research materials. And we believe that they will embrace a tool that facilitates that work. Our confidence in that is increased by the enthusiastic response that scholars, in particular, have given to numerous presentations that we have made about *Firefox Scholar* in the past twelve months, including large audiences at the meeting of the Association of Research Libraries and the University of Virginia's Summit on Digital Tools for the Humanities, and our own DC Area Humanities and Technology Forum. There has also been extensive (and positive) preliminary coverage of *Firefox Scholar* on numerous blogs and in publications such as the *Chronicle of Higher Education*.

In short, "this is exactly what I need" is what we have been told over and over. Although there has been some justifiable concern about the embrace of digital resources and tools in teaching or new modes of scholarly communication, scholars, especially in the tradition-bound realm of the humanities, have avidly taken up digital resources for their own research. As part of his work on the Mellon-funded ACLS Commission on Cyberinfrastructure in the Humanities and Social Sciences, CHNM Director Roy Rosenzweig did a limited survey of history graduate students and gathered a remarkable body of evidence about the enthusiasm that young scholars are showing for digital resources. But those new digital scholars (as well as their elders who are also embracing digital resources in growing numbers) need new digital tools like *Firefox Scholar*.¹

4. Deliverables for *Firefox Scholar* 2.0

Additional funding from the Mellon Foundation will allow us to fully realize the potential of *Firefox Scholar*, far beyond the work made possible by our pilot IMLS funding. Moreover, it will enhance our work on *Firefox Scholar* and broader Mellon Foundation initiatives in ways that will benefit both those who are building the next generation of web tools for research and education as well as the users who will be using those tools and *Firefox Scholar*. In brief:

- 1) *We will connect Firefox Scholar with other semantic web technologies and initiatives to increase the value of all of these projects, and to allow for more*

sophisticated data analysis and display. We will align our work more closely with Mellon-funded initiatives to the benefit of both *Firefox Scholar* and those initiatives, by including an architecture that allows information to be easily exchanged among semantic and data processing applications in RDF. We will also enable new data mining “modules” to be added to the software (in the same way that translators will be added), including a basic set of text mining and analysis tools in the 2.0 release. We will build alliances with a range of centers and developers working on digital tools in order to create a broad basis of support for *Firefox Scholar*.

2) *We will create a variety of mechanisms for both sharing and backing up user data and metadata. This sharing will foster the further development of semantic tools that meet the needs of users, facilitate collaborative research practices, and grow digital scholarly resources from the grassroots.* Researchers will be able to share folders, learn about new items that might interest them, and contribute to the semantic web. The range of online digital resources for scholars will, thus, grow through a grassroots, collaborative process. In addition, users will have the advantage of remote storage of their collections (either within their own institutions or at a centralized server maintained by CHNM) and the ability to synch with those server-based repositories.

3) *We will ensure that Firefox Scholar natively recognizes the fullest possible range of scholarly resources nation- and world-wide; our goal by the end of the project would be to cover 90 percent of the collections used most commonly by scholars.*

4) *We will create import/export and advanced synchronization tools.* Like EndNote, Firefox Scholar will be able to export in a variety of formats (Chicago Manual of Style, APA, MLA, etc.) to a variety of locations (word processors, web-based text editors, HTML, etc.). Using an open repository, we will allow others to add formatters for other bibliographic styles (reducing the cost of developing these myriad format styles by spreading the work across the user community). We will also create tools for importing Scribe, EndNote, and RefWorks collections into *Firefox Scholar*.

5) *We will use Firefox’s embedded system for skinning to ensure that Firefox Scholar meets the highest level of accessibility.*

6) *We will start to work with scholarly associations, institutions of higher learning, libraries and archives, and database providers to test and disseminate Firefox Scholar on a national, rather than local (Virginia) basis and begin to build a critical mass of user communities around Firefox Scholar to insure its wide use and long-term sustainability.*

1. *We will connect Firefox Scholar with other semantic web technologies and initiatives to increase the value of all of these projects, and to allow for more sophisticated data analysis and display.*

Additional funding will allow us to invest substantial time and resources up front to make sure that *Firefox Scholar* is fully compatible with (and in fact leads naturally toward) the RDF-enabled vision of the future semantic web. As a powerful “killer app” for semantic web technologies, *Firefox Scholar* will help shift SIMILE’s advanced research and development work from proof-of-concept and command-line installations (as many SIMILE projects are at the current time) to widespread use among scholars, students, and researchers, where the underlying technology is hidden by the friendly confines of the browser. In particular, we see *Firefox Scholar* demonstrating the importance and power of the semantic web through dissemination to diverse scholarly communities, especially in the humanities.

Currently we are using Firefox 2.0+'s SQLite backend for data storage. (Firefox 2.0 is scheduled for release in fall 2006, but we have worked from the start with their earliest alphas.) Although we still plan to use this relational database for reasons of speed and simplicity, we will engineer the software so that any records stored in a *Firefox Scholar* collection could be interoperable and combined with any other SIMILE-based (or other) RDF data. We will code *Firefox Scholar* to enable users to easily send items or whole collections to MIT's Semantic Bank, quickly filling that important resource with thousands of records. We would also build on the Solvent code to promote an interchangeable page-scraper format for materials not yet in semantic format.

More generally, we would work with other Mellon fundees to develop a common bibliographic RDF—a standard that still needs to be established, as we have learned from our discussions with various project leaders as well as leading digital librarians. We anticipate beginning this standardization using SIMILE’s RDFizers, especially the lowest common denominators of resulting code from MARC/MODS->RDF and BibTex->RDF. As the user notes for that latter translator make clear, however, a real RDF for citations has yet to emerge. We would widely publicize the final specification and leverage *Firefox Scholar* to press the adoption of the RDF in other applications.

Moving *Firefox Scholar* into this RDF environment will also enable new kinds of scholarly research, for instance, by enabling visual mapping of *Firefox Scholar* collections using Tuft's VUE technology or SIMILE’s Welkin visualization tool. A researcher who has put together a folder containing biographies of those who landed at Jamestown in 1607 could send this collection into VUE and instantly see the relationships among those people. A Shakespeare professor could show his students clusters of the Bard’s works (based on his bibliography of the primary texts) and their relationships to the works of Marlowe, Milton, and other authors.

Although the openness and interoperability of *Firefox Scholar* would encourage the use of such external programs, we suspect that most scholars will not immediately take advantage of these powerful analytical tools simply because they are unfamiliar with data mining techniques. To foster this new form of computational scholarship (which will undoubtedly be critical in an age where scholars are swamped by digitized materials), we will build into the 2.0 release of *Firefox Scholar* a basic set of text analysis tools (we have less expertise in visualization and will leave that to the work of Tufts, MIT, and others). Not only will users be able to do full-text searches of their

collections; with this data-mining menu they will be able to extract statistically significant terms and phrases from a document; show word frequencies; find dates in a document; and highlight a particular item and then find other items that are closely related. We have already undertaken some text-mining experiments of our own in connection with our H-Bot software agent (which automatically answers natural language questions about history by statistically analyzing large databases of text) and Syllabus Finder (which profiles course materials and finds them instantly for users).²

In addition to moving *Firefox Scholar*'s data architecture into RDF-compatible formats, we will also move other pieces of the project, such as annotations of stored documents, into RDF so they too could be shared among a community of scholars and accessed and analyzed by a variety of semantic tools. For example, we have begun (again with encouragement from the Mellon Foundation) active and fruitful collaborative discussions with David Wiley at Utah State University. Their Scrumdidilyumptious web service will be able to extract semantic relationships between items in *Firefox Scholar* (e.g., it will learn from a *Firefox Scholar* user that Proclus's *Commentaries* is a critique of Euclid's *Elements of Geometry*). In turn, users of *Firefox Scholar* will be able to query the Scrumdidilyumptious server to see if there are any web pages related to the item they are highlighting, and what those relationships are. Through a combined RDF or RSS feed from the Scrumdidilyumptious server and the *Firefox Scholar* server a researcher could keep abreast of a specific topic of interest to them, for instance, tracking items that users have tagged or related using specific keywords.

Another key collaboration we will pursue is with Sakai, which is developing an open source Collaboration and Learning Environment for higher education. We believe that *Firefox Scholar* support could readily be built into Sakai. As such, it will provide an easy-to-use research tool for undergraduates and graduate students, from introductory English composition to advanced research seminars. On the simplest level, it would ensure rigorous and accurate citation of sources, a perennial problem for instructors. But it would also make it possible for *Firefox Scholar* collections to be posted to a class website by a professor or student (as well as downloaded from that website).

We expect that this is just the tip of the iceberg. For instance, we have recently met with representatives from JSTOR about text mining their databases; such tools could be integrated into *Firefox Scholar* to enable new forms of scholarly research. We have already presented preliminary reports on our work in meetings and phone calls with developers from MIT, Yale Library of Medicine, the University of Sydney's Archaeological Computing Laboratory, the National Archives and Records Administration, the developers of the Scholar's Box at U.C. Berkeley, the Digital Library Federation, and Utah State's Center for Open and Sustainable Learning. We will continue to forge relationships with these and other projects and institutions as they see how *Firefox Scholar* fits in with their work and as we see where we might adopt certain widespread standards to make the software as interoperable and powerful as possible.

2. We will create a variety of mechanisms for both sharing and backing up user data and metadata. This sharing will foster the further development of semantic tools that

meet the needs of users, facilitate collaborative research practices, and grow digital scholarly resources from the grassroots.

The sharing of metadata, bibliographies, annotation, and digital objects would create a virtuous circle that would encourage other scholars, teachers, and students to use *Firefox Scholar* (and by extension, other advanced digital research tools that take advantage of the semantic web). We will develop multiple approaches to the sharing of metadata and digital objects and to providing external backup for local *Firefox Scholar* collections.

One strategy for fostering sharing and providing backup is a centralized server, which we are currently calling *Scholar Central* (but for which we are actively considering alternative names, particularly to reflect the fact that we will also offer non-centralized options; see below). This might be similar, for example, to Share Your OPML (Outline Processor Markup Language) (share.opml.org), which allows individuals to upload their personal collections of RSS feeds they read and then find new feeds they might be interested in or other users that share their interests. *Scholar Central* will serve similar functions for *Firefox Scholar* users. Users will be able to upload and tag specific personal *Firefox Scholar* collections, and in turn receive recommendations for books, articles, artworks, and other items that might interest them based upon their synchronized libraries. This server will have RSS and RDF feeds for topics and allow for collaborative bibliography making.

Scholar Central will also run a set of APIs so that other Mellon-funded tools can extract RDF-encoded information from it, and deposit information into it, as in the Scrumdidilyumptious case above. Users of SIMILE's Haystack will be able to browse *Firefox Scholar* collections online. In short, this server will enable new kinds of semantic-based autodiscovery, recommendation, and other automated systems for enhancing scholarship.

We will implement this server potentially using parts of SIMILE's Semantic Bank as the backend, while also linking with Cite-U-Like, Connotea, and other online services. By aggregating the research collections, notes, and annotations of thousands of scholars, this server-based software (in effect, a del.icio.us for scholars) could offer an exciting wealth of the kind of community-building tools often developed by larger corporate sites like Flickr, Amazon.com, or Yahoo, but not currently available to researchers in the humanities and social sciences. These include recommendations of other scholarly communications in their field; other forms of collaborative filtering and tagging of books, articles, and other scholarly resources; RSS and RDF feeds on disciplines and subfields in the humanities; backing up of personal collections; and sharing (group-restricted and fully open) of these collections and specific folders. In this effort (as in dissemination), we will work closely with scholarly associations in the humanities and related fields.

Although the centralized server has some obvious advantages, we also recognize that it will not be a solution that appeals to all users. We expect that some (perhaps many) institutions and organizations may wish to serve as the remote backup for the

bibliographies, notes, and files of those affiliated with them, rather than sending full copies of these research corpuses to CHNM. Moreover, some individuals may be uncomfortable with sending their data to a centralized server. We will therefore build into the *Firefox Scholar* client two “remote server” settings that will be similar to those in email clients (such as POP and IMAP). The first setting will specify the web address of the server to which a user’s personal collection should be synched (if any). The second setting will specify a location to which just the metadata of a user’s collection will be synched—for the purposes of data mining and other social features such as feeds and recommendations. (This will allow us to provide some of the advantages of a centralized recommendation system without the disadvantages of full centralization of all data and files; for finer-grained privacy control, we will also allow users to “expose” *Scholar* collections on a folder-by-folder basis.) Both of these settings could be empty if a user does not wish to partake in any of these server-based features.

Institutions and groups will, thus, be able to create distinct “Scholar Communities,” based upon their affiliation or interests, in one of two ways. First, *Scholar Central* will provide subdomains that can be administered by someone from the group or institution. For instance, a *Scholar Central* subdomain could be run by the American Historical Association, with recommendations and bibliographies shared more tightly among the members of the organization. In this case we will provide the authentication and administration tools (as well as the hosting for remote backups). In the future, we might provide this service for a modest yearly fee, thus providing at least partial funds for *Scholar Central* beyond the term of this grant.

Second, institutions or groups could set up their own remote servers, providing a new web address for the first setting mentioned above. Because of the complexity of *Scholar Central*—a computationally intensive social website run on an elaborate server architecture distributed across multiple machines—we do not believe that we will be able to build a full turnkey, downloadable package for an exact replica of *Scholar Central* under the terms of this grant. (One of reasons for the difficulty of creating such a package is that the integration of the software and the hardware will be critical; in this way, *Scholar Central* will differ from blog or forum software packages.) Nevertheless, we will provide a reference model and documentation for building a replica of *Scholar Central* for those institutions and organizations that have the capabilities for this more complex work. For others, we will provide a simplified system for creating an institutional silo using an off-the-rack WebDAV-based solution. This more basic installation will not include the more sophisticated features of *Scholar Central* (which users of alternate silos may still get by keeping the second preference setting on *Scholar Central*). It will still, however, enable the remote use of *Scholar* collections. For instance, a student at a university that has a WebDAV storage system in place could work on their *Scholar* collection from a library terminal and then later synch their updated collection onto their personal computer. Many educational institutions are already running WebDAV in some form, so we believe this solution could be very easily implemented.

And of course, as an open source, well-documented project, it will be easy for other developers to build new storage and data-mining mechanisms using our clearly defined

database structures and RDF. In short, we have no intention of restricting individual and institutional users of *Firefox Scholar* to using *Scholar Central*, and we will do everything possible to provide alternatives for sharing and storing *Firefox Scholar* data and files.

While we want to give users and institutions other options besides the centralized server, we also want to offer those users who are interested the advantages of centralization, especially the ability to share digital objects from personal collections. Other semantic web projects focus on what is already on the web—taking web pages and making them “smarter”—but we believe that *Firefox Scholar* has a unique potential to take offline materials and bring them (or in some cases at least their metadata) online, into semantic relationships with born-digital materials, and to encourage scholarly collaboration and sharing. Because it is built into the web browser but deals with print and physical objects in addition to documents and images on the web, *Firefox Scholar* enables scholars to build comprehensive personal collections, to forge associations between items online and off, and to share with colleagues or even strangers across the globe what they digitize (assuming it is in the public domain), annotate, and cite.

For the first time, scholars will be able to easily preserve and share research notes with other active scholars as well as with the next generation of scholars, who might want to re-examine a problem or topic mined by an earlier researcher. In addition, book and manuscript reviewers will be able to more closely scrutinize the evidentiary base for scholarly projects and claims. *Scholar Central* will be one advanced nexus for this critical exchange. The draft report of the ACLS Cyberinfrastructure calls for the creation of “extensive and reusable digital collections,” while also recognizing that cost constraints will prevent many important, but specialized, collections from ever being digitized. *Scholar Central* provides one way to leverage the dispersed energies of tens of thousands of scholars (and amateurs) to grow the digital corpus.

The ability to share and collaborate in this way is one of the advantages of a centralized server. It would be possible to build a fully decentralized system that provides recommendations and other advanced features that would match the power of *Scholar Central*, but we do not believe this will be possible under the scope of this two-year grant proposal. But, as noted above, we will provide alternatives to this centralization for institutions and individuals that wish to create different repositories and group-based *Scholar* collections.

3. *We will ensure that Firefox Scholar natively recognizes the fullest range of scholarly resources.*

Because the IMLS funding is only for VIVA schools (Virginia libraries), our 1.0 pilot does not focus on developing page scrapers for non-Virginia digital collections (with exceptions for major libraries and collections like the Library of Congress and widely used resources like Amazon). With Mellon funding we would do many more of these in-house, and work nationwide with developers and even average users to produce hundreds of scrapers for digital resources for scholars (using Solvent linked to our

Firefox Scholar repository). We would make sure the software worked with all major institutional online catalogs by building in auto-recognition of the main commercial catalog software, such as Sirsi, Voyager, and Aleph.

We will create particularly tight integration with Mellon-funded collections such as JSTOR and ARTstor, enabling user authentication and the ability to take notes in the *Firefox Scholar* panel while viewing a work of art or a scholarly article. We will also create translators for other major resources such as Making of America, California Digital Library, WorldCat, PubMed, History Cooperative, Project MUSE, Lexis/Nexis, InfoTrac, ProQuest Research Library and Historical Newspapers, Dissertation Abstracts, National Gallery of Art, and other leading museums, libraries, and digital repositories (e.g., New York Public Library, Cornell, University of Virginia, Harvard University, the National Library of Australia, the British National Library, and dozens more). Ultimately we will aim to cover 90 percent of the resources used by scholars in the humanities and social sciences.

4. *We will create import/export and advanced synchronization tools.*

We will create export tools so that, like EndNote, *Firefox Scholar* can export in a variety of formats (Chicago Manual of Style, APA, MLA, etc.) to a variety of locations (word processors, HTML, etc.). Using an open repository, we will allow others to add formatters (using Firefox's simple XML and CSS-based configuration) for other bibliographic styles.

These export deliverables would include the ability to connect *Firefox Scholar* with Word, OpenOffice, EndNote, ProCite, CHNM's own note taking program (Scribe), and the Open Document Format, and possibly new "Web 2.0" web applications such as Writely (now owned by Google). We will also make it possible for scholarly bibliographies to be dynamically updated as references are added and deleted. Because we expect that users of existing note taking and reference tools will want to shift to *Firefox Scholar*, we will make it easy to import existing references and notes from the most popular packages. We will also add the ability to import and export collections via RDF, XML, tab-delimited text, and SQL dumps, and to import collections from EndNote and RefWorks.

Import/export tools are important not just for ensuring compatibility with other existing citation management software. For example, allowing users' *Scholar* collections to reside in multiple places (such as an institutional remote storage site or *Scholar Central*) will require us to provide a sophisticated, highly granular, conditional import/export system between locations. This interface will appear, for instance, when a user synchs their collection to *Scholar Central* or its decentralized alternatives. Users will be able to merge items manually, choosing either the server or client version in the case of conflicts or duplicates, or even choosing to merge some metadata from the server (e.g., if the server has publication data missing from the client version of items), and will be able to append new items either before or after existing items. We will implement simpler merges as well, where duplicates are automatically purged and either the server

or client data is given preference as the governing version. Such settings will be on the preferences panel, so that users with specific, constant requirements (such as always pushing information from the client to the server) can maintain their collections with one click. Like iTunes, we will also build a “find duplicates” menu option into the client to clean up collections and to deal with the inevitable messiness of name variations and other common metadata problems. Given the database back end to *Firefox Scholar*, we believe that it will also be possible to do global find-and-replace by fields for more substantial corrections or updates to a personal collection.

5. *We will build Firefox to ensure the highest level of accessibility.*

Any modern scholarly tool needs to be for *all* scholars, including those with disabilities and different ways of approaching and learning from materials. Meeting accessibility expectations means avoiding both a lowest-common denominator user interface that is good for no one as well as a parallel design where alternate views are treated as second-class (and often neglected as the software gets updated). Ideally, software should adapt on the fly in precisely the ways that make sense for each user, maintaining the richness and configuration that each user desires and learns from most readily. As Jutta Treviranus and Colin Clark of the Adaptive Technology Resource Centre at the University of Toronto have pointed out, true accessibility entails “just in time” adaptability of software such as Sakai to different user needs and ways of learning and “viewing.”

Although the software environment of *Firefox Scholar*—a mix of JavaScript, XUL (XML), CSS, and SQLite (all part of the new Mozilla “trunk”)—will prevent us from using the Java-based work of ATRC’s TransformAble modules (at least for the browser client), the remarkably flexible Firefox platform will allow us to implement ATRC’s commendable accessibility philosophy. Firefox’s separation of presentation, layout, and language from the underlying code—maintained in the “chrome” directory used extensively by *Firefox Scholar*—makes on-the-fly user interface (UI) changes possible. We will be able, for instance, to relatively easily create different language versions of the column titles or to alter the look and feel of the browser and its extensions (as hundreds of developers have already done with Firefox “skins”).

With Mellon funding, the “main” (default) view of *Firefox Scholar* would merely be one alternative (English, certain column layout and size) among many. The *Firefox Scholar* website would promote the variety of UIs and show how to use them. And as open source software based on XML and CSS changes to affect look and function, others will be able to create their own skins and functionality, such as the ability for audio browsers to read aloud items in a collection. We will maintain a central repository (much like Mozilla already does for Firefox skins) for these other UIs.

Moreover, as Treviranus and Clark point out, smart accessibility also entails providing on-the-fly reconfigurations of information when students fail to grasp standard presentations. Our “smart folders” that automatically group items by other algorithms show how this could be done within *Firefox Scholar*. We will work with ATRC and other

accessibility minded institutions to seed *Firefox Scholar* with some of these algorithms (e.g., listing items by frequency of use), and to include in the documentation ways for creating other filters.

6. *We will start to work with scholarly associations, institutions of higher learning, libraries and archives, and database providers to test and disseminate Firefox Scholar on a national, rather than local (Virginia) basis and begin to build a critical mass of user communities around Firefox Scholar to insure its wide use and long-term sustainability.*

As we have noted, we have been funded by IMLS to develop *Firefox Scholar* largely as a pilot for VIVA. With funding from Mellon we would make it national, work with scholarly associations, especially in the humanities and social sciences, as well as libraries, archives, and database providers to ensure that this becomes the one tool scholars need and use. One simple strategy will be to build alliances with these organizations, institutions, and companies to have them suggest that users “try Firefox Scholar.” In our recent meeting with JSTOR, they were receptive to this approach. And we have had similar discussions with the National Archives, with which CHNM has just signed a formal cooperative agreement. Under this grant, we would move ahead on forging dozens of these alliances.

“Marketing” has always been a weakness of academic digital projects. This is hardly surprising, since this is a topic that most scholars, especially in the humanities, view with great suspicion. But unless a tool like *Firefox Scholar* is widely adopted, it will fail to achieve its most ambitious goals and to sustain itself. Thus, marketing and dissemination are at least as important as programming and interface design in the success of *Firefox Scholar*.

Fortunately, CHNM has considerable strengths in marketing, advertising, and dissemination, a reflection, in part, of our long-term focus on reaching a dispersed community of K-12 and college students and teachers as well as the general public. That we have one of the most heavily trafficked history websites—with 300 million hits and 10 million visitors annually—puts us in a good position to promote *Firefox Scholar* through links, banners, and cross-promotion. We are also very experienced in techniques of search engine optimization and directory placement that are crucial to bringing traffic to a website. As a result, many of our sites are among the top two or three sites in their category in their Google ranking. The co-PIs on this project (Dan Cohen and Roy Rosenzweig) have written the very first guide directed specifically at building audiences for scholarly websites in their book, *Digital History*, and we will draw extensively on that work in promoting *Firefox Scholar*. CHNM’s family of websites also includes the *History News Network*, a site that averages a quarter-million unique visitors each month. We will also leverage the use of existing CHNM tools to build the initial user base for *Firefox Scholar*. In particular, *Firefox Scholar* will be a compelling upgrade for the nearly 10,000 people who have downloaded a copy of our free Scribe reference manager and note taking application.

Mellon funding will also allow us to get started on two additional key dissemination strategies. First, we will build alliances with national scholarly associations to reach their members through newsletter articles, email announcements, direct mailings, and conference presentations. CHNM has a good start on this because its director, Roy Rosenzweig, has served on the national board of three major scholarly associations. CHNM is also an affiliate of the American Historical Association (AHA), and we currently partner with the AHA and the National History Center in our “Doing Digital History” workshops.

Second, we will hire a full-time “evangelist” for *Firefox Scholar*. This outreach coordinator will straddle the library/museum and academic worlds, and will run workshops and demonstrations for both content providers and academic users. He or she will be particularly active in online venues and the blogosphere, building scholarly awareness about *Firefox Scholar* and forging ties with related projects and with libraries and database providers. He or she will also moderate our beta tests and evangelize the final software; in short, he or she will be the public face of *Firefox Scholar*.

Mellon support would also allow us to devote significant attention to user support and testing by producing more substantial documentation for “average” users. CHNM will organize national focus groups early in the process, beyond the local focus groups currently funded. This is critical for gathering essential feedback and for promoting mass adoption. We will ensure easy installation/downloads of the software and will verify this through pilot testing and focus groups. To be sure, a two-year grant will not allow the time or funds for the full-fledged dissemination and alliance-building effort that we believe is necessary, and we will later seek additional funds to support that effort. But we still believe that we can make a strong start under this grant.

5. Use Cases

a. Basic Use: A graduate student starting work on a dissertation prospectus

A history graduate student is beginning work on a dissertation prospectus on the ratification of the Federal Constitution. Although ratification is a well-studied topic, her focus is on the *process* of ratification as a window into the nature of the development of participatory politics in the 1780s and 1790s. Her first step is to develop a basic bibliography and to read the relevant secondary sources on both the ratification itself and the larger historiographic and theoretical context of popular politics, political culture, and the development of the public sphere. To build her initial bibliography, she consults such basic resources as *America: History & Life*, *WorldCat*, *Dissertation Abstracts*, and the Organization of American History's *Recently Published Scholarship*. Although these are subscription-based databases (available through her library or her membership in the OAH), she is pleased to discover *Firefox Scholar* icon next to the URL, which means that her *Firefox Scholar* browser extension can automatically capture the citation information. She quickly assembles a core bibliography of 200 secondary works with the full metadata.

She then plunges into her secondary source reading. One key review essay ("Constitutional Dialogues" by Gary McDowell) is from the October 2004 *William and Mary Quarterly* available on the History Cooperative website. She captures the full HTML version of the essay in her *Firefox Scholar* database and begins highlighting key sections and writing notes the annotation field. Other promising articles (e.g., "We, Some of the People: Apportionment in the Thirteen State Conventions Ratifying the Constitution," from the *Journal of American History*; "Representation and Political Deliberation in the Massachusetts Constitutional Ratification Debate" in the *Journal of Rhetoric & Public Affairs*; "The Concept of Political Culture" in the *Journal of Interdisciplinary History* and "From 'Conflict' to 'Constitutional Question: Transformations in Early American Public Discourse" in the *Quarterly Journal of Speech*) are similarly available at JSTOR, Project Muse, EBSCO—all databases recognizable by *Firefox Scholar*. She imports the articles (whether in plain text, pdf, or HTML) into her *Firefox Scholar* folder and then highlights and notes specific passages. Naturally, not everything relevant is on line. *America: History and Life* refers her to large project at the University of Wisconsin called the *Documentary History of the Ratification of the Constitution by State*. She borrows the volumes from her library and takes notes in her *Firefox Scholar* database.

Further online excursions take her to many more online articles and to archival guides (such as ArchiveGrid) as well as lead her to more sources available only in libraries, although the growing number of older works in Google Print makes some of those available online as well. After three months of reading and reflection, she begins to write her prospectus. *Firefox Scholar* allows her to easily organize her materials into the major sections—theory, historiography, archival resources, and chapter structure as well as to quickly locate quotations that she remembers reading but for which she has forgotten the source. Because she is working in the discipline of history where the *Chicago Manual of Style* is the conventional format, she uses *Firefox Scholar* to seamlessly export her footnotes and bibliography into that format.

b. Collaborative interdisciplinary media research

A media historian doing research on popular responses to early radio at the American Library of Broadcasting (University of Maryland, College Park) discovers in a radio magazine from the late 1920s more than fifty letters to the editor defending broadcasting by local farmers against the more upscale programs developed by national radio networks. He copies them with his digital camera and adds them to his *Firefox Scholar* collection and tags them as "rural life," "1920s," "radio" and "opposition to networks." The historian is surprised—all secondary sources he has read and previously added to his *Firefox Scholar* database suggest universal public support for network broadcasting and criticism of desultory local programming. Maybe the letter writers were not real people? But he checks U.S. census records at a genealogy site, <<http://ancestry.com>>, and discovers that most of the letter writers did exist. The media historian scrapes the biographical metadata on letter writers into *Firefox Scholar*—they were mostly white male working- and middle-class listeners—and links

the data to the letters. Knowing that the radio magazine's copyright was not renewed, he then uploads his collection to the *Scholar Central* shared space for researchers.

Meanwhile, a legal scholar studying the Radio Act of 1927 has collected in his *Firefox Scholar* the earliest annual reports to Congress by the Federal Radio Commission, available from the Federal Communications Commission (FCC) website <http://www.fcc.gov/mb/audio/decdoc/annual_reports.html> and in the public domain as public records, along with her notes on Congressional debates on the Act and its purpose: support for centralized commercial network broadcasting. These reports suggest that in the late 1920s legislators and their constituencies supported network expansion and wanted to get rid of noisy local stations. The legal scholar uploads her data and annotations to shared *Scholar Central* server. It automatically does a search on "radio networks" and "1920s" and finds the "oppositional" letters to the editor previously uploaded. The legal scholar is intrigued but not convinced. On the one hand, she would never have thought to look for traces of public opinion in letters to magazines; on the other, Congressional reports seem a more reliable source. The legal scholar enters into an online discussion with the media historian, informing him that there is no trace of the "opposition" to networks in the Congressional record.

This debate attracts the attention of a music scholar researching 1920s Chicago music at the National Archives. The scholar is in the middle of studying Chicago stations' appeals for better bandwidth to FCC for evidence of radio music programming for immigrant and black communities in Chicago. FCC records are copyright-free, so she has been able to digitize the key hearings and import them into *Firefox Scholar*. As soon as she uploaded them to the *Scholar Central* public space, it alerted her to the discussion about the 1920s stations. She re-examines her records and realizes that music station owners also blasted the networks with their audiences' support—hundreds of listeners' letters, in several languages, were submitted to the FCC. The music researcher informs the legal scholar and the media historian that she found additional evidence of popular opposition to pro-network legislation. She also realizes that the local music culture created by Chicago stations presented an alternative model to the larger legal and institutional history of American national broadcasting. By drawing on other scholars' sources and analysis, each researcher published a more nuanced, contextualized, and convincing analysis. Given these scholars' diverse disciplinary environments, such connections could be made only after their work was published without *Firefox Scholar*. Meanwhile, *Scholar Central* has begun to develop a valuable archive on early radio that will be used by future researchers.

c. A researcher exploring a specific topic using online archives

A historian of science is doing research for a book on early modern astronomy and consults the collections database of the Museum of the History of Science (MHS) at Oxford University <<http://www.mhs.ox.ac.uk/database/>>. Realizing that he is going to be using dozens of objects from the museum for his book, which is being assembled using *Firefox Scholar*, the professor quickly creates a scraper for the metadata and thumbnail images found on the search results page using the Solvent-based *Firefox*

Scholar wizard rather than manually cutting and pasting each field for each object. He then continues his search, grabs the items he would like to investigate further into his *Firefox Scholar* library, and sorts them into chapters within the organizing pane. He can take notes on image of astrolabes he finds from the sixteenth and seventeenth centuries, fine tune the metadata as his research gives him a better sense of the year in which each astrolabe was built (if the Oxford catalog only has a rough estimate of the period the instrument was built), and highlights details of the astrolabe to investigate further. To get a better sense of how the metalworkers who created these instruments were related to each other and influenced each other, the historian exports his personal collection into VUE.

Furthermore, the professor realizes, as the caveat on the MHS web page says, that the “on-line collections database is in an early stage of development.” To assist both the MHS and other researchers working in the history of science, he donates (with a single click) his scraper into the *Firefox Scholar* repository, where any other *Firefox Scholar* user can add it to their browser for their own research. In addition, as he has assembled the definitive list of early modern astrolabes, he posts this specialized collection bibliography to the *Scholar Central* exchange server, where anyone can view it, download it, and add their own notes.

A month after he uploads his MHS-based bibliography, he receives a notice from the *Scholar Central* server that images of similar astrolabes from Italy have been put online by the Istituto e Museo di Storia della Scienza in Florence. Another scholar working on Galileo who had a similar bibliographic profile on the *Scholar Central* server found these items. Together these scholars tag all of the astrolabes they have found with helpful descriptors. A new student of this subject may now quickly understand the range of scientific instruments available and their character without browsing multiple sites as “a novice in the archive” (in literary scholar Randy Bass’s phrase).

6. Project Timeline

Late September/early October 2006 – Initial 1.0 beta release to coincide with release of Firefox 2.0; implementation of Solvent for *Firefox Scholar*; repositories for scrapers goes live.

October - December 2006 – finalize bibliographic RDF standard and implement in the code, including import/export in RDF; meet with other Mellon project leaders to finalize other exchange issues; begin construction of *Scholar Central* (server cluster construction, design, scripting code); begin work on accessibility.

January 2007 – Official 1.0 release, with import from EndNote and other formats; begin work on data mining tools that will go into 2.0 (text analysis tools) and specific exchange requirements for using VUE, SIMILE visualization and browsing tools; test accessibility UIs and other skins.

Spring 2007 – 1.1 release with accessibility modules, seamless exports to other Mellon tools (just from the client, not from *Scholar Central*); limited beta for basic implementation (no recommendation systems but shared folders, RDF, and RSS) of *Scholar Central* (by invitation) and beta of 2.0 of *Firefox Scholar*.

Fall 2007 – *Scholar Central* with basic sharing opened to the public along with 2.0 release of *Firefox Scholar* (with full RDF support and advanced import/export features), with new data mining tools and code to work with *Scholar Central*; outreach coordinator begins attending meetings of scholarly associations, making other presentations, and evangelizing online. Release of alternatives to *Scholar Central* to developer community.

Early 2008 – Beta of auto discovery, recommendation systems, and other advanced features that require a critical mass of users on *Scholar Central*.

Summer 2008 – Full, final release of *Firefox Scholar* 2.0 and *Scholar Central* with advanced features.

7. Sustainability

CHNM is committed to nurturing *Firefox Scholar* and *Scholar Central* for the long haul. As noted below, we have already devoted considerable attention to developing free digital tools for historians and history teachers and we have sustained what we have developed for more than five years now. With the help of an NEH Challenge Grant, CHNM has raised a \$2 million endowment. That endowment (as well as our success raising other funds) allows us to guarantee that *Firefox Scholar* will be a permanent resource for scholars. Significantly, other major digital history projects—such as the award-winning *DoHistory*—have been given to CHNM because their creators believed they would find a more permanent home among other leading digital history projects and because CHNM has a reputation for sustaining its work for the duration.

The open architecture and standards of *Firefox Scholar* extensions will enable future development, and we envision the software as a living and sustainable project after its release, both enabling current users and facilitating a shift toward user-side personal collection systems that are interoperable across individual libraries and museums. To foster the future work of other developers, we will produce a document outlining the technical details of how *Firefox Scholar* works, and how developers of other organization and presentation tools might hook their applications into a *Firefox Scholar* collection. In addition, we will host a discussion forum for users and developers to share their knowledge and explore new directions for this software.

Like other open source projects, *Firefox Scholar* needs to find sustenance by building a community of developers committed to the project. We envision that CHNM will remain at the heart of that community, but we are also working hard to build allies and supporters throughout the scholarly and information technology communities. We have already made good progress in our preliminary discussions with academic developers such as Raymond Yee at Berkeley and Dan Chudnov at Yale. We have also begun serious

discussions with the developers of Firefox itself (the Mozilla Foundation and Corporation), and we see that tie as crucial to solidifying long-term support for *Firefox Scholar*. Moreover, one of the key activities of the “evangelist” hired to work on outreach for *Firefox Scholar* will be to build and further solidify ties with external communities of scholars and programmers that can help to sustain *Firefox Scholar* for the future. As noted above, we will begin these efforts at building alliances and sustainability under this two-year grant. But we believe that we will need additional funds and time to complete our work on this phase of the project.