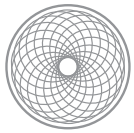

MIT Hyperstudio

US-Iran Project Software Overview

Christopher York
MIT Hyperstudio

MIT Hyperstudio 16-635 77 Mass Ave, Cambridge, MA 02139 **Christopher York -** yorkc@mit.edu http://hyperstudio.mit.edu

US-Iran Project Software Overview

Project space

The US-Iran Project software (v.1) was conceived as a general-purpose tool for scholars and laypersons with an interest in political or diplomatic history. It was designed to capture, in online form, some of the highly-collaborative approaches to evidence pioneered by Jim Blight and janet Lang in their "critical oral history" conferences—while still appealing to a wider audience of mainstream history enthusiasts. The initial installation of the software was loaded with documents relating to US-Iran relations between 1979-2005, but the intent was to support collaborative exploration of historical documents for a variety of topics. For example, with relevant source materials loaded it could be used to explore the 20th century local history of a town like Cambridge, MA; to trace the development of a political movement like the Taleban in Aghanistan and Pakistan; or to track happenings in a single family's genealogy.[1]

Because the software focuses on events and people as core points of interest for the past, it's particularly well-suited to studying 19th and 20th century diplomatic and political history. In these milieux, individual people and diplomatic events often played a strong role in shaping historical outcomes; the US-Iran software aids in identifying and charting them collaboratively. To take an example from Blight and Lang's work: by October 1962 US-Soviet tensions were so heightened that officials at each Cold War superpower recognized that launching a nuclear weapon near Cuba would almost certainly spark a rapidly-escalating international nuclear conflict. In this context, developing a clear and well-documented picture of the relevant actors (John F. Kennedy, Nikita Kruschev, Fidel Castro, Robert McNamara, among others) and the exact progression of events (e.g. when precisely did Soviet missiles arrive in Cuban waters? when during the conflict did Kruschev and Castro communicate, and with what result?) is vital to understanding how the crisis developed. In this case, the course of the conference revealed both new events and new configurations of people. The Americans were astounded to learn that

[1] For a succinct overview of critical oral history practice, see by Bruce J. Allyn, James G. Blight, and David A. Welch, *Cuba on the Brink: Castro, the Missile Crisis, and the Soviet Collapse.* New York: Pantheon (1993): pp. 3-14. Briefly, critical oral history conferences hinge on a series of round-table discussions involving policy-makers, scholar specialists, and period documentation. For example, participants in the 1992 Havana conference on the Cuban Missile Crisis included Cuban Head of State Fidel Castro, Robert McNamara, US Secretary of Defense under Kennedy, and Arthur Schlesinger Jr., former special assistant to Kennedy (among others); scholars with expertise in cold war international relations, such as conference director Jim Blight, Cuban historian Rafael Hernández Rodríguez, and Russian Latin American specialist Sergo Mikoyan; and a set of newly-declassified government documents relating to the Bay of Pigs conflict. The critical oral history process is highly collaborative, since it extends historians' conventional practice of source critique to all three of these elements. Policy-makers' gaps of memory and retrospective self-fashioning are corrected by documentary evidence and the interjections of specialist scholars. Scholars, motivated by their own research agendas and benefiting from historical hindsight, are brought back to the moment of diplomatic conflict by discussion among period policy makers. And gaps in the documentary record are filled by the hitherto-private revelations of policy makers and diplomats. The US-Iran software was designed to enable this collaborative process, in which community members together flesh out the details and timeline of a historical situation in an iterative fashion, progressively integrating more and different evidence.

Soviet missiles were already in place at the height of the conflict, and that Castro had urged their use in the event of an American attack.

But history is more than excavation of facts—it is also integrative and interpretive. From a scholar's point of view, events like the Cuban Missile Crisis represent pivotal moments in the course of Cold War diplomacy and lay bare the dynamics of an era. From a politician's point of view, they also represent "missed opportunities," or moments when broader understanding might have suggested a better, more humane course of action.[2] In both cases, the concrete products of historical research and discussion are (1) a clearer understanding of which were the key people and events, supported by citations to primary source documents or oral accounts; and (2) new personal appraisals of those personalities and events. As an example of the first case, conferences on the Cuban Missile Crisis coalesced scholarly and official consensus around new events or "facts"—that Soviet nuclear missiles were deployed in Cuban waters in October 1962, and that Castro had encouraged Kruschev to make use of them immediately in the event of a U.S. action. Prior wisdom had read the conflict as one of Cold War grand strategy, in which John F. Kennedy cooly countered Soviet diplomatic maneuvers. This overlooked a key event, the Soviet deployment of nuclear warheads near Cuba and the imminence of real nuclear conflict; and a key actor, Castro, who is now accepted to have played a central role in the progression of what was, in retrospect, a profoundly "Cuban" conflict. In the software, this new shared knowledge would be hashed out through iterative revision of shared timelines and biographies, by the entire community.

Now an example of the second case, interpretive narrative. Later critical oral history conferences on Vietnam led both to consensus around new events and facts—such as the realization that Hanoi did not play a coordinating role in massing resistance to American military activity—but also to new personal appraisals of the broader course events. In this case, former U.S. Defense Secretary Robert McNamara learned that he and his superiors had coordinated a military action leading to nearly 2 million Vietnamese deaths, in pursuit of an opponent they wrongly surmised was an organized arm of cold war Soviet Russia. In light of this and other revelations, McNamara realized that he and his American superiors were long culpable for crimes against humanity—particularly for the fire-bombing of Japan in World War II, but in Vietnam as well. In contrast to the excavation and discussion of clear timelines of diplomatic events and rosters of important actors, McNamara's appraisal represents a clearly personal re-interpretation of those events. As will become clear later, the US-Iran software was designed to support both: collaborative identification and shared critique of key events and actors (e.g. the arrival of warheads in Cuban waters), and personal re-appraisal and interpretation of those historical facts (e.g. McNamara's private revelations and counterfactual history).

In sum, the software's design goal was to provide a tool for the collaborative excavation of key people and events (with primary source citations), while providing space for individual—and perhaps even polemical—interpretations of those facts. This approach is particularly useful for diplomatic, political, or institutional history using 19th or 20th century sources that offer a profusion of concrete evidence about specific people and events. It is less useful for exploring systemic demographic or cultural shifts where individual people and incidents are less important than cumulative changes, which express themselves gradually in the lives of large groups (e.g. the effects of demographic explosion in 17th century Europe on diet, as grains replaced meats; or the relationship between Islam and black slavery in 16th-18th century Africa, where Muslims often refused to enslave fellow "peoples of the Book").[3] One might address these changes by supplementing the software with modules that focus on aggregating demographic evidence or close-reading specific documents rather than focusing on

---

[2] See Jim Wright and janet Lang's introduction to *The Fog of War: Lessons from the Life of Robert McNamara*. Lanham, Md: Rowman & Littlefield (2005) for a more thorough discussion of the differences between scholars' and participants' engagement with critical oral history.

[3] See Jeff Ravel's Comedie Française Registers Project for an example of a Hyperstudio project that focuses on systemic demographic changes more than charting actions and events surrounding a select set of important historical actors. For the examples listed see, respectively, *Fernand Braudel, The Mediterranean and the Mediterranean World in the Age of Philip II, trans Sían Reynolds. New York: Harper and Row (1972)*; and *Paul Lovejoy, Transformations in Slavery: A History of Slavery in Africa. Cambridge: Cambridge University Press (1983).*

specific events and historical actors. However, the software's emphasis on supporting historical inferences by encouraging users to link directly back to primary source evidence will be useful in any documentary project focused on the past.

Software Design / Humanistic Process

This section will describe both how the US-Iran software design aimed to support the process described above, and how it sought to bring the collaborative process of critical oral history to the wider community of history enthusiasts, both scholarly and popular. It will also outline how the software aimed to support basic to advanced users with several different levels of engagement, from source retrieval to collaborative event editing and essay writing.

A short overview of general archival practice and its digital equivalents will help situate these design decisions. Researchers —scholarly and popular, professional or student—approach an archive of historical documents as a source of evidence about the past. In many cases, particularly when the historical topic is relatively obscure or the subject of government secrecy, little is actually known about even the basics: what happened, where, and to whom. The hope is that reading previously unexplored primary sources will illuminate a milieux we knew little about, or that it will cast new light on well-accepted interpretations. In the face of a relative paucity of easily-accessible primary evidence on most historical topics, traditional historians assume their practice is mostly empirical in nature: one locates evidence about events, people, and places in the past, and evaluates it. Based on individual judgment, the strongest evidence is taken as a positive indication of "what really happened," insofar as that is knowable. Without citation to primary source evidence or well-accepted publications based on primary source evidence, most historical statements will be regarded at best as ungrounded and at worst as likely falsities. Hence it's vital that any historical statement in a new domain be supported by citations to archival primary sources, so that later readers can evaluate the statement's reliability before assuming in turn that it is "true."[4] Likewise the logical way to confirm the veracity of a given statement is to return and re-evaluate the documentary evidence in support. In general, the assumption is that one works, in a roundabout process, from large masses of circumstantial archival evidence, towards a clear map of the progression of events and the various historical actors involved in them.

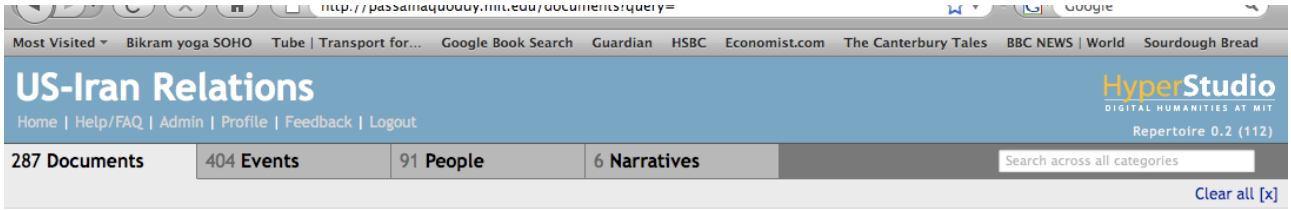

Figure 1. Navigation and Workflow

The US-Iran software's navigation (Documents/Events/People/Narratives) was intended to capture this general workflow, from primary source evidence, to the past set of events and actors we can make concrete statements about based on those

---

[4] These observations are largely self-evident, especially to a modern audience weaned on scientific approaches to data and argument. However, archival practice introduces a number of considerations. Some ancient history, for example, is known only via treatises written centuries after the event, by authors who could not have been witnesses, or may be based on primary sources that have since been lost. In these cases there is the question of evaluating evidence from secondary sources. And for any document, whether deemed "primary" or "secondary," there is the question of authorial bias and hidden rhetorical intent, particularly if the document served a political purpose. Beyond these considerations, there is for any individual piece of evidence a question of whether it is sufficient to justify a given historical inference, and whether it is representative of larger trends or just an atypical case that happens to have been preserved. For these and other questions of traditional historical source critique, a good reference is Martha Howell, *From Reliable Sources: An Introduction to Historical Methods*. Ithaca: Cornell University Press (2001). Most of these nuances can be left for the users of historical software to debate; the salient point here is that history as an academic discipline assumes a stance that is simultaneously sceptical of unsupported inferences and strives for empirical methodology, often more strenuously than other humanistic disciplines. Hence, software that strives to support collaborative excavation of historical facts from source documents will need to begin with support for primary source citations.

sources, and then to personal and idiosyncratic interpretations (see figure 1). Hence, as one moves further on in the progression, the historical picture becomes more and more interpretive, but hopefully also richer and nuanced. Users who are only interested in locating particular documents by keyword search or cataloging data (e.g. kind of document, author, date), but are not interested in historical collaboration can limit themselves to the 'Documents' category but still have a productive research experience. By providing exact image replicas wherever possible, with handwritten marginalia etc. we reassure users they are working from sources that are "almost the real paper," and allow them to make their own evaluation of paleography and other stamps and marks. However, by allowing users to search the full text of documents (where OCR permits), and refine based on document type and date, we still provide a basic approach to drill down and find relevant primary evidence related to a topic (figures 2, 3).

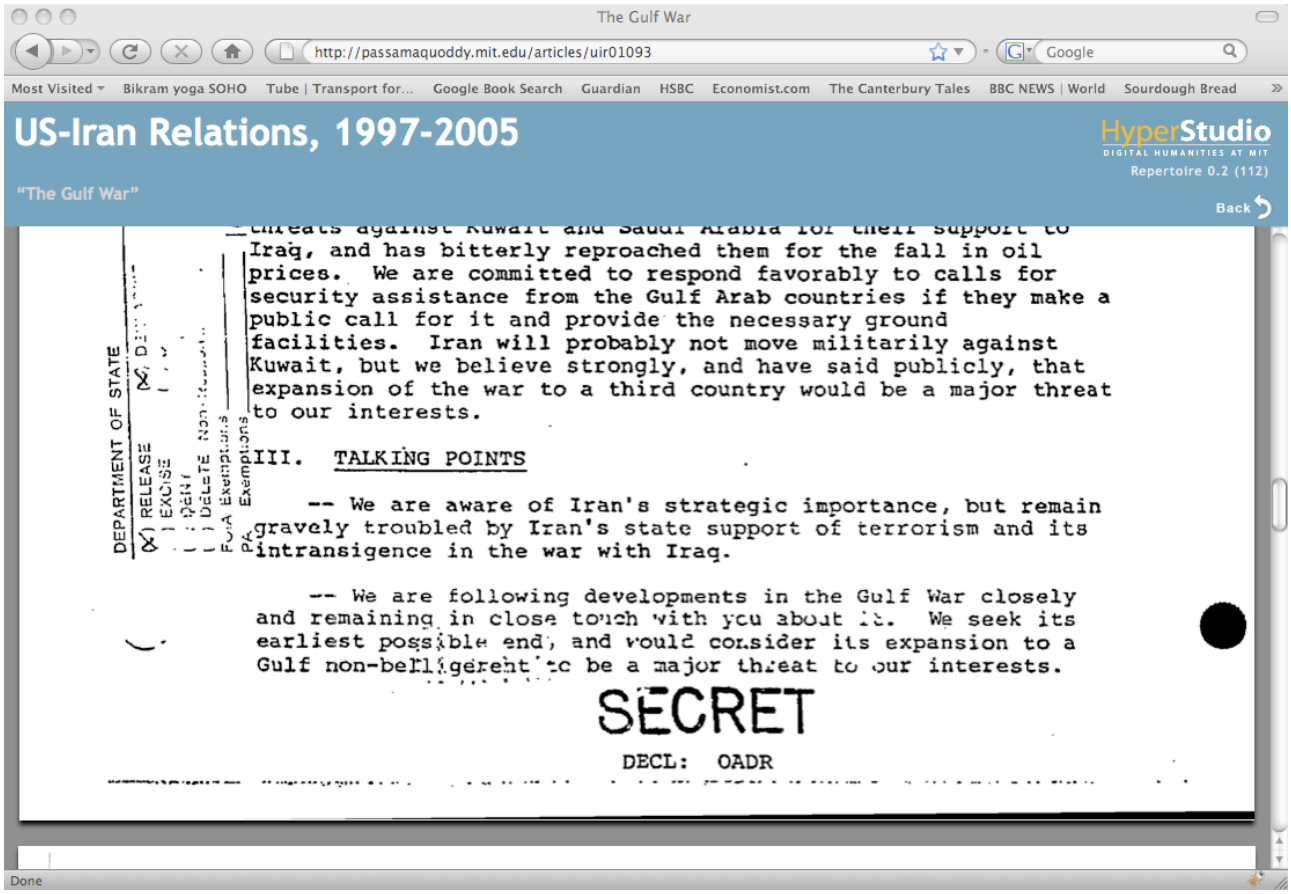

Figure 2. Document detail view, with marginalia, etc.

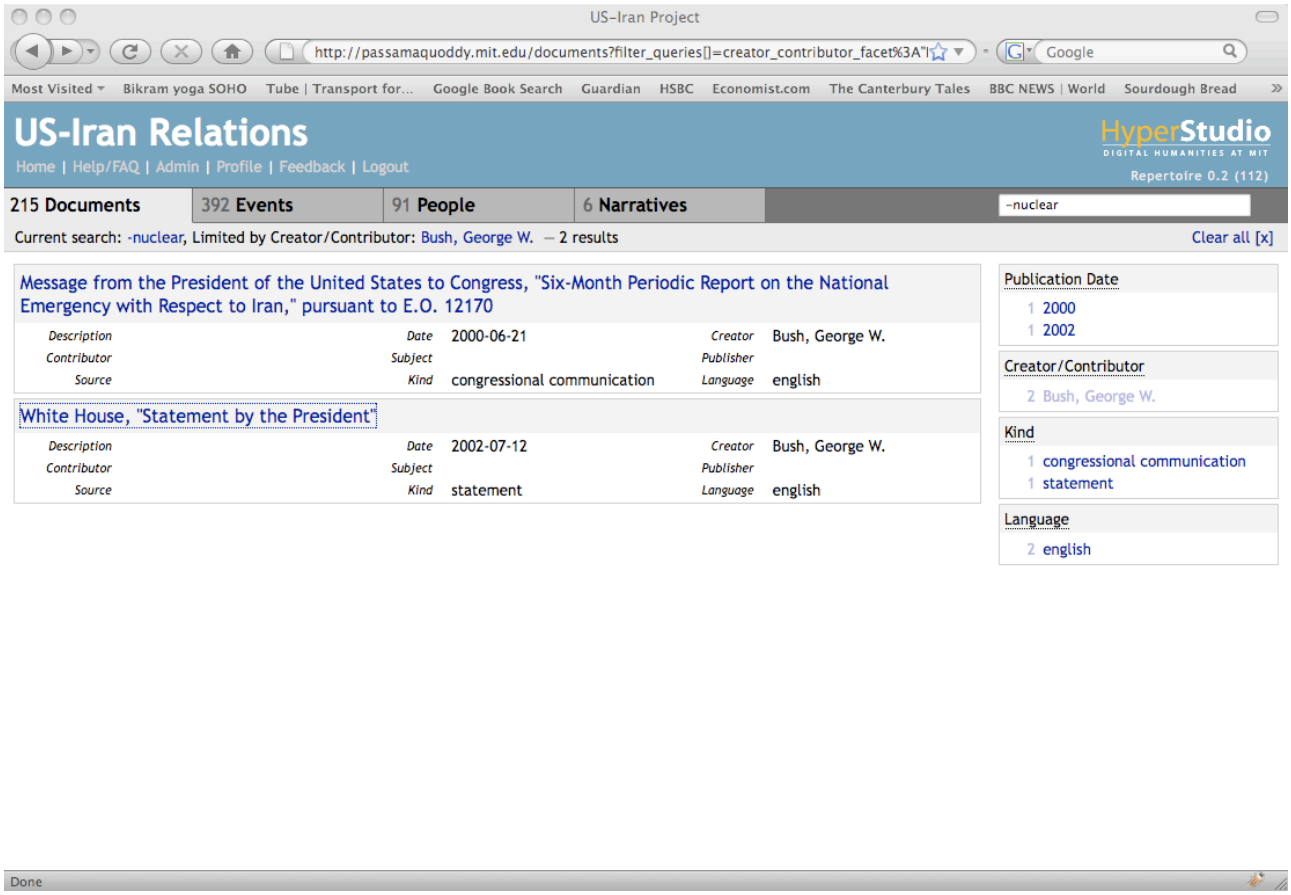

Figure 3. Document search and drill-down. This search shows items authored by Geroge W. Bush not including the word 'nuclear'—the user might be looking for evidence on Bush's views on foreign or domestic policy in a non-military context. The second document, for example, details his official stance on democracy and authoritarianism in Iran.

The US-Iran software's wiki model was devised to provide a digital metaphor for citations to primary sources (see figure 4) and to coordinate user collaboration. As in other semi-scholarly wikis (e.g. Wikipedia), citations appear as links in the published article, which allows readers to click directly back to the source document and evaluate how well-founded any particular historical inference is. And as in other wikis, users have edit access to all of the public pages (here, events and people entries), which encourages the community as a whole to revise incorrect information, add additional citations to supporting evidence, or question and revise out dubious material. The primary flow of collaboration among users occurs in iterative revisions of the same entry by community members.

Hence the basic collaboration premise for the US-Iran software: that over the course of time, individual pages describing a particular event or person would come to represent a self-regulated community consensus around the "basic facts," with support by primary source citation links.[5] As in other wikis, users can log in to revise or remove unfounded claims, to add evidence and citations, and to correct minor errors. Moreover, because the event and person records are all publicly visible and appear in the search interface, there is a strong incentive to re-use an existing entry whenever possible rather than starting a new one. For example, say there is already an entry for March 8, 1984 describing a United Nations decision to investigate claims that Iraq had used chemical weapons, with citations a newspaper article. A user who finds additional evidence in a government cable can add a new citation to the existing event entry and flesh out additional details the newspaper may have gotten wrong (perhaps because its reporters relied on circumstantial evidence). The two users have then collaborated together in establishing a basic timeline of events based on primary source evidence—and the community as a whole is in a stronger position, since a declassified government cable is generally a stronger source than a newspaper article. As later users work through the community's developing network of events and actors, those wishing to strongly

[5] This report suggests revisions to the wiki collaboration principle; see the next section.

substantiate their own points will link to events and person pages that are strongly-documented with primary source citations. Likewise, in evaluating another users' arguments, a first course of action would be to see how well-substantiated their timeline of events is (more on this interface in a moment).

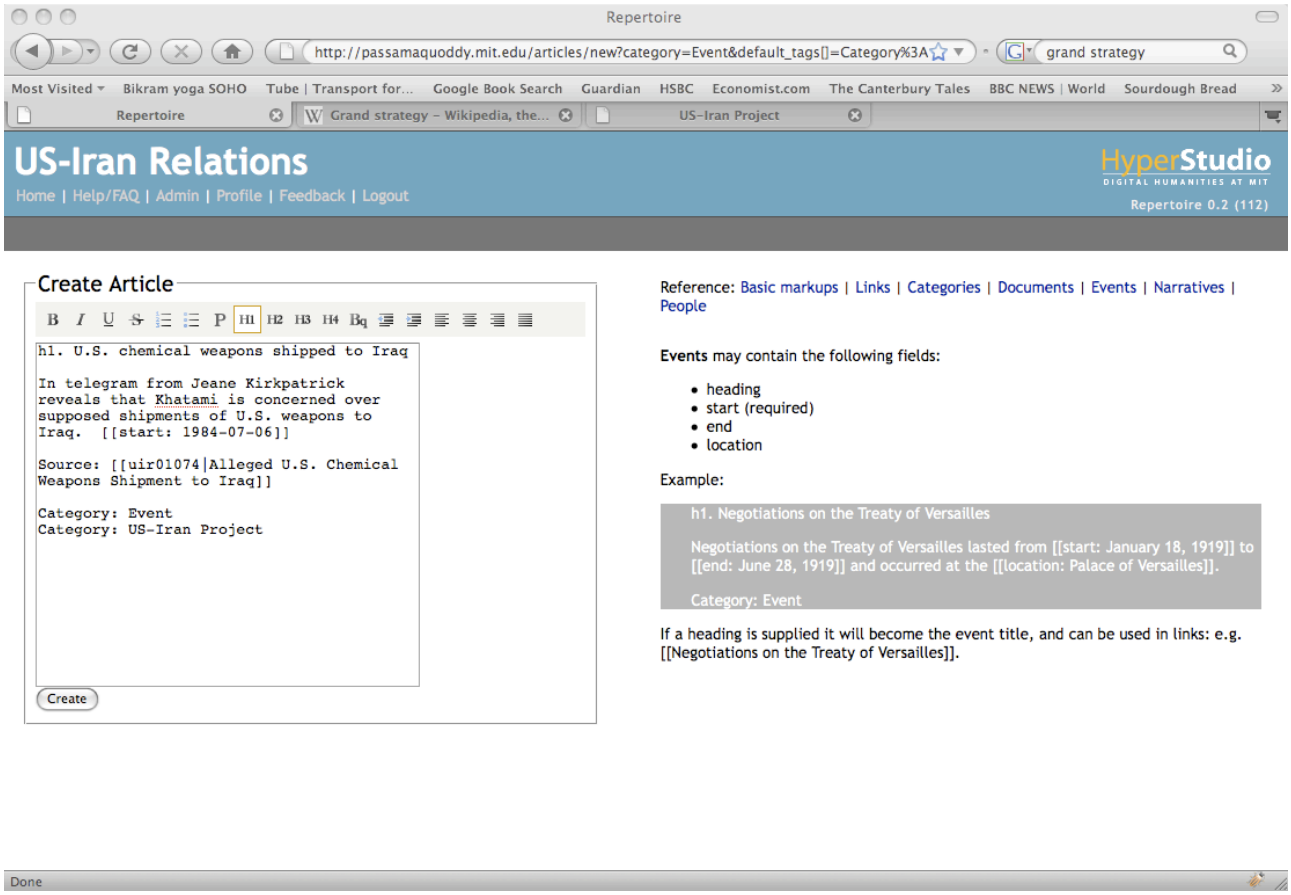

Figure 4. Editing an event. Note the primary source citation link.

The events and people tabs are intended to allow one to navigate the community's developing notion of "what happened": not the collection of primary source documents, but rather a constellation of putative happenings in the past, and the cast of people involved in them. Hence, exploring the content created by other users can be as illuminating than working in the repository of primary documents, since it crystallizes many individual archival insights. For example, figure 5 shows a timeline of events relevant to the topic of chemical warfare. Based on the rest of the community's timeline work, we can quickly identify when chemical warfare became a major concern (there is a cluster of events in 1984), and work through the progression to get a sense of how the diplomatic give-and-take surrounding chemical weapons developed. At any point, we could also follow citation links back to the primary sources that support a given event in the timeline, and reappraise others' work. In this way, the events and people categories become a quick way to explore the current state of the community's knowledge about a given topic, and chart it on a timeline. One would rarely view the collection of all events on a timeline, but would instead drill down on a particular topic to show a customized timeline. Note, also, how the cross-category search functions as a research tool: users can enter keywords for topics of interest without selecting a tab and result counts for each appear next to the category name, as primary source documents, events, people, or narrative essays. In this case, we also find that there are 53 documents containing the phrase "chemical warfare," and that no-one has yet written a person biography or narrative relating to this particular topic (the current user may be doing precisely that).

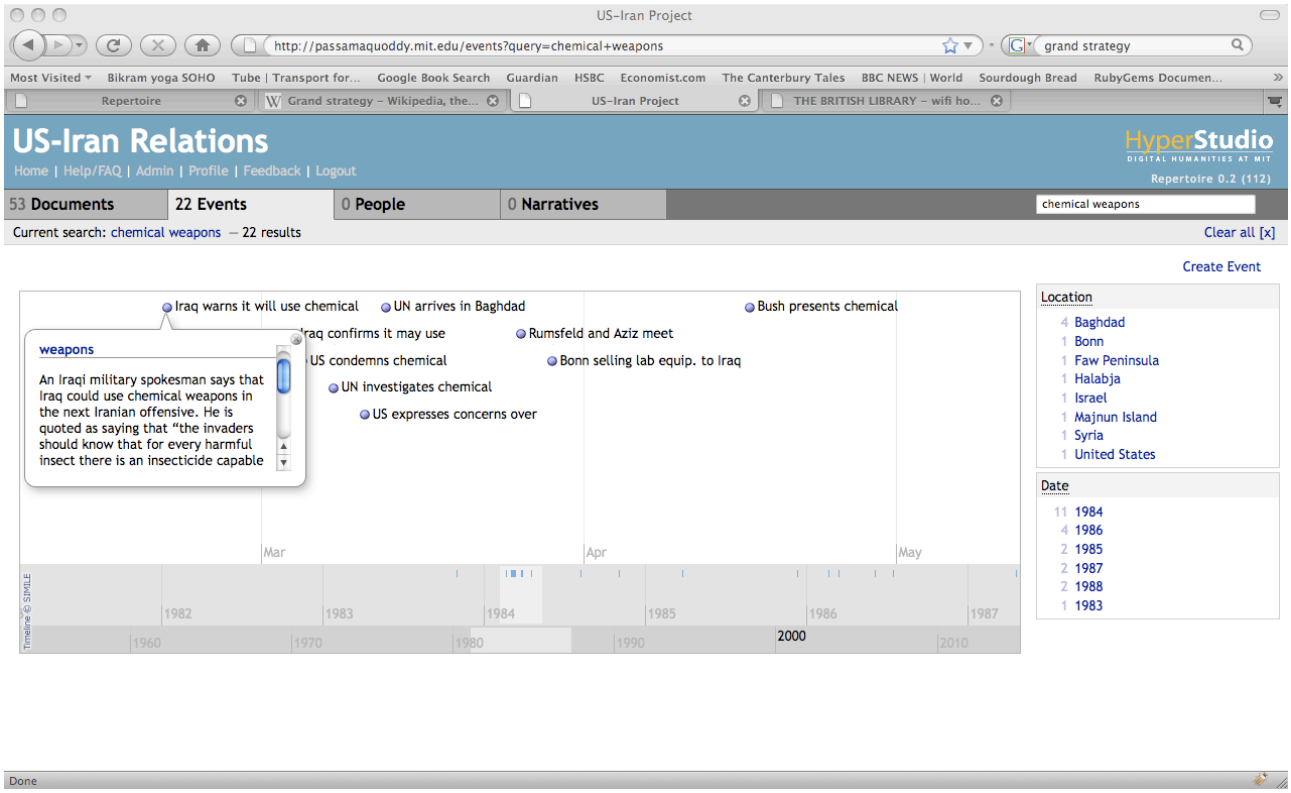

Figure 5. Drill-down search on the global community timeline

Being able to search through the community's list of events is a useful research tool, but what if one wants to create a customized timeline relevant to a particular person, or make a pointed argument based on a hand-crafted timeline? These functionalities are available in the "People" and "Narratives" categories, respectively. When looking at a person biography page, all of the events that reference this person's entry or are directly referenced in it appear in the timeline below (see figure 6). In this way, one can easily browse through a personalized timeline of events relevant to that particular historical figure—a biographical timeline of sorts. Likewise, when viewing a user-produced narrative, a customized timeline of events mentioned in that narrative appears at the bottom of the page (see figure 7). Later readers considering this essay can click on individual timeline events, document citation links, and person links to explore how this particular essay marshals evidence for its argument, or which pieces of the community's evolving body of knowledge are relevant to a given topic.

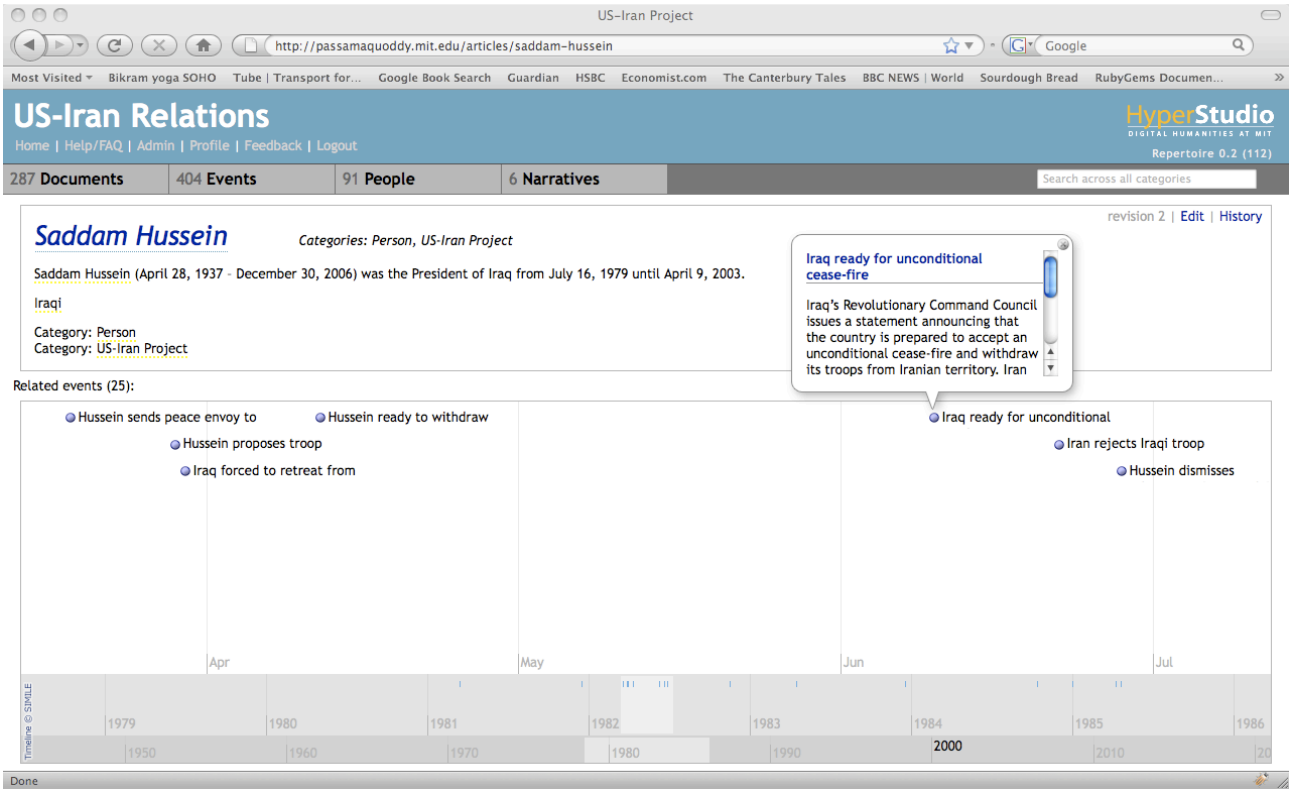

Figure 6. Example person entry, with supporting biographical timeline.

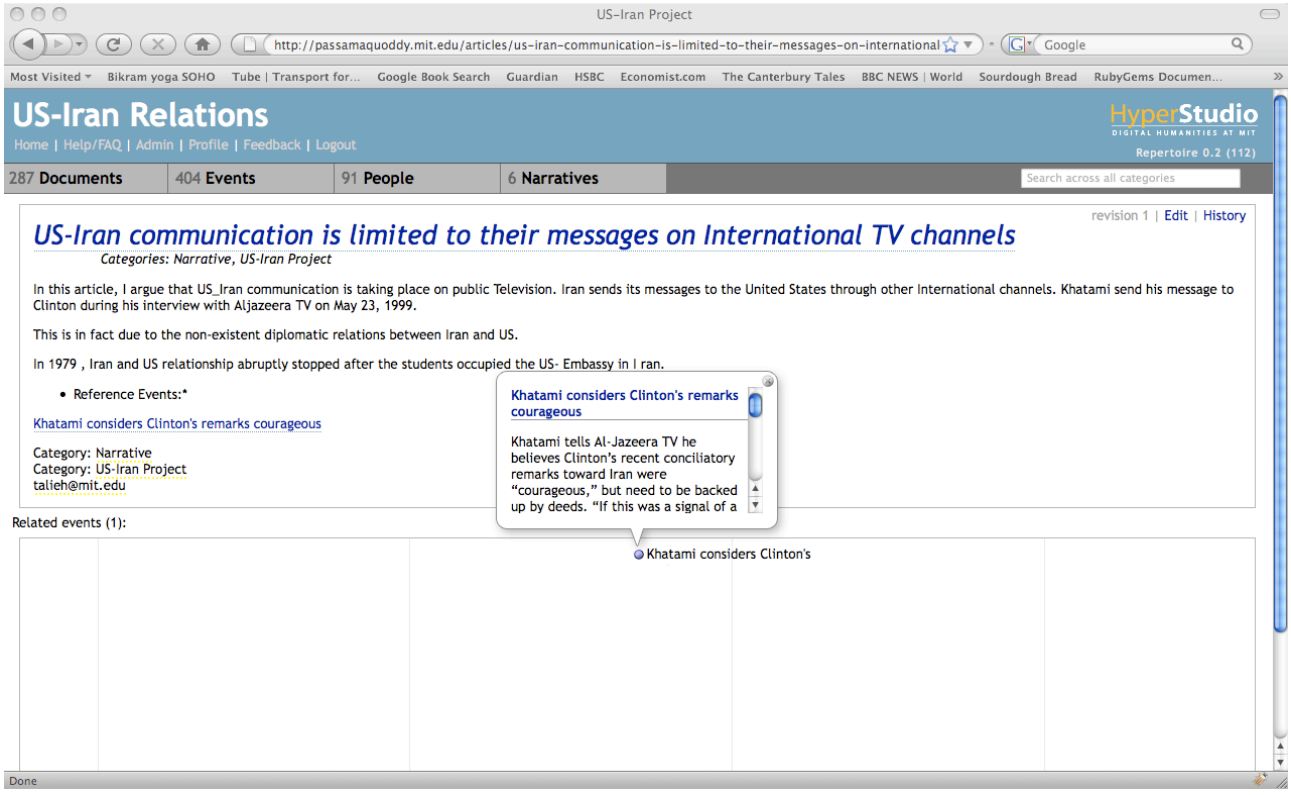

Figure 7. Example personalized narrative, with supporting timeline (this article cites a single event -- had it cited others they would be collected below as well).

This brings us to the final piece in the collaborative workflow supported by the software, the balance between communal work and individual work (see figure 1 again). As in an archival environment, the set of source "Documents" is shared by all users, and constitutes a fund of evidence about the past that can be mined by all users. The next two tabs, "Events" and "People" contain entries that can be iteratively improved by any user in the community, and so come to represent a shared body of knowledge about US-Iran relations. This is the nexus of collaboration for the software, as users evaluate and improve each others' entries in a shared space. Finally, in the last tab ("Narratives") the articles can only be updated by their authors. Here, one can present highly polemical or individually-argued posts to the entire community—it's a place to share new interpretations based on the shared fund of factual events and person entries. Thus, the software as a whole is designed to balance collaboration with individual expression, while still supporting the primary source citations that are the foundation of solid historical research. Here, it aims to bring the collaborative approach to the past pioneered by critical oral history conferences to a much larger community of history enthusiasts, whose work in the archives is typically conducted in isolation. By bringing both primary sources and a shared body of middle-of-the-road facts about people and events online, the software makes the process of archival research and knowledge production collaborative. Pursued on a large scale, with a good fund of primary source documents and an active community, this could have two transformative effects on historical research.

First, by focusing the attention of many users on specific event and person entries and the primary source citations that support them, it could raise the quality of our shared knowledge about the past—anyone can add supporting evidence for a specific event when he or she encounters it in the document repository, or append notes with citations to contrary evidence. The basic tasks of archival research (evaluating sources and determining whether evidence for a given event or interpretation is representative and sufficient) are democratized, and subject to the same network effects that make open-source software successful. A single engineer might not have time to write an entire software application, but can still contribute productively to the common goal by fixing a bug or adding a small feature. Likewise, a single user may not be up to writing an entire book on the past: reading archival materials, weighing evidence, and formulating a broad argument. But she might easily note that while a newspaper article previously regarded as a reliable source dated Mohammad Khatami's withdrawal from the Iranian presidential race to March 18, 2009, international cables document that he actually retracted on March 16—and update the event's date and add a new citation. In other words, ordinary people can contribute to the shared process of making prudent inferences about the past based on partial documentation, without being a research scholar or book author.

Second, this democratization of the basic process of archival research would place a correspondingly greater weight on historical interpreters to focus on innovative but well-supported arguments or narratives. With access to archival sources, individual users and enthusiasts have the tools at hand to evaluate whether an argument makes well-founded inferences. Likewise, because making citations to documentary evidence and well-accepted communal events is easy, essays of higher quality will tend to make liberal use of such citations in order to support their viewpoint. The online collaboration environment thus lowers the bar for participation in solid inquiry into the past, while making individual essayists more accountable for work. In the long term, to add value to community fund of non-specialist knowledge, scholars will need to be both more daring and more prudent in advancing new historical interpretations—more like literary critics and anthropologists and less like journalists and fact-checkers.

Critique

The first version of software was successful in identifying a novel and well-delimited problem domain in digital humanities, with an approach adaptable both for the core US-Iran critical oral history participants, but also to a wider community of researchers and history enthusiasts. It successfully allowed for several levels of engagement, from merely searching through primary source documents to contributing new event and person articles and revising old ones, to writing well-argued and -supported interpretive essays. It also achieved an interesting balance between collaborative or community-oriented activities and individual research and publication. However, in retrospect several design aspects could be dramatically improved or represent failed experiments.

General improments.

- **Usability**. Poor usability in the wiki/markup editor for events, persons, and narratives is by far the software's biggest flaw. Here, incremental improvements would make a huge difference. Like first-generation wikis, the software relies on a proprietary markup format that has a learning curve too high for our target audience. Adding citations to other posts should be a simple matter of dragging and dropping the cited primary source paragraph or event onto the citing text that's being edited. Link citations should appear automatically inline or as footnotes, rather than requiring the user to enter Textile markup of any kind. Also, the system for embedding data fields in articles (example event: "On [[start: September 11, 2001]], terrorists destroyed the World Trade Center") should be discarded in favor of a set of user-friendly validating forms that accomplish the same end. Also, users absolutely must be able to cite specific pages/paragraphs in source documents (not currently possible), and be able to reference events by dragging from the timeline. In sum, many of the user-interface improvements that have been made to wiki and blog software in the past five years need to be integrated into the editor. Without this improvement, ordinary humanistic users would not spend enough time with the software to recognize its usefulness for research and teaching.

- **Collaboration model.** The wiki revision model, which allows all users to revise a set of shared posts, is in retrospect the wrong collaboration paradigm for this particular subject domain. Conventional wikis implement what might be called a "last man standing" approach to collaboration, as though it were a wrestling match: the last user to update an article has the final say, and can effectively quash all earlier community work, irrespective of quality. Moreover, the motivations and intellectual evaluation that lies behind each revision is sidelined in this model of collaboration—whereas really it constitutes a vital online forum for discussion between users about which inferences are best supported by the evidence at hand. In retrospect, the group revision model for collaboration around events should be discarded in favor of a weighted-votes model, which allows all users to create new entries and the prioritizes those that are most frequently cited as more "useful" and hence either more authoritative or controversial, and of greater general interest.[6] Also, user interaction around these interpretations should be moved out of wiki page revisions and into threaded discussions that link a series of contending event entries.

- **Greater evidence of community in UI.** Along similar lines, although the software is designed to be highly collaborative and community-centric, like most wikis it gives little sense of this when a user first logs in. To impart a feeling of collaborative enterprise, it would be helpful to foreground other users' recent activities, their discussions and comments on documents or timelines, and new materials in the front web pages. Users should get the sense they are participating in a community of like-minded enthusiasts.

- **Ability to upload new documents and form new communities.** Users should be able to upload their own primary source materials. Also, they should be able to form new community networks around them. This way we could support installations centered on different historical topics using the same software installation.

- **Technical refinements.** Some minor technical issues related to software packages we rely on need to be addressed.[7]

---

[6] This problem is common to wikis as repositories for vetted knowledge. For example, Wikipedia now implements a system of specialist editors in order to make sure important articles represent authoritative points of view. A weighted-votes system seems more useful than a Wikipedia-style appointed-editors system for this domain, because events are typically much smaller in scope than an encyclopedia article (rather, they are the *building blocks* of larger encyclopedia articles about the past). However, it might be worth experimenting with combinations of the weighted-votes and appointed-editors systems: one might weight citations from articles by respected authorities on a topic higher than citations from anonymously authored articles, for example.

[7] Simile timeline does not scale beyond ~500 events in a single timeline; it might be prudent to only show timelines for search results and essays instead of on the top events page. Similarly, the off-the-shelf solr faceting software proves inadequate to our needs since it can't index faceted values in real time or transactionally. Recent modifications to the Repertoire faceting module address these limitations by replacing solr with a database-based facet index.

- **Auto linking.** Person pages should be automatically linked to events and narratives whenever the person's name occurs —in other words, the user should only have to create citation links by hand where absolutely necessary. This refinement would also remove some difficulties wikis have with charting incoming versus outgoing links, and with identifying different pages to link to. It would also be nice to automatically register the production of most primary sources like telegrams, letters, etc. as events.

Improvements for critical oral history:

- **User discussions should be potential primary source material.** In traditional historical practice, critiquing archival sources is a central activity (see footnote 5). For example, only by investigating the context of their production does one learn that transcripts from the Spanish and Portuguese Inquisition were obtained with the aid of torture: as a result they must be read not only as straight confessions of witchcraft by the witness, but also as incremental accommodation to what the inquisitor *needed* his subject to confess. As another example of source critique, consider voting lists for the early United States: these in no way constitute accurate indicators of public opinion since they explicitly exclude women and slaves. Critical oral history expands on this process by adding oral accounts and discussions between the historical actors themselves to the conventional critique of documents by scholars. In the oral history paradigm, first-person oral accounts critique each other as sources by presenting different points of view on events they both participated in; and scholars and documents critique oral history accounts, correcting both prejudiced accounts from after the fact and the inevitable slippage of memory. The software should accommodate this increased collaborative element by allowing users to cite other users' comments and discussions as primary sources.

- **Track institutional organization.** At least in 20th century contexts, diplomacy is a balance of individual action and institutional organization and decision-making. The software should provide some way of clearly tracking people's affiliations with institutions over time. Because general information is available about most important historical actors elsewhere on the web, it would make sense to replace the individual entries in the "People" tab with specific timelines and institutional summaries backed by a table relating person, institution, position held, start date, and end date.

- **Contrast / compare timelines.** One activity of particular importance to the critical oral history community is comparing and discussing the difference between period actors' notion of which events were most important. In other words, in comparing their narratives of past events in order to see where politicians had radically different notions of which events were important or even what happened (as, for example, in the McNamara's realization that nuclear warheads had actually been deployed in Cuban waters). To support this, it would be very useful to add a functionality where users could drag two or more narratives onto a timeline to compare and contrast.

To prioritize: to achieve a software platform that's of general interest and wide usability, the usability improvements to the wiki editor would need to be implemented. Beyond this, while the existing collaboration model works, it would be better to switch to the weighted-voting model for events since it's both easier to understand than shared revisions and more sustainable. The remaining suggested improvements are listed in order of descending importance to a wide digital humanities audience. In general, however, these changes represent incremental improvements to an approach that's successful for the subject area and for an audience of scholars, students, and enthusiasts about the past—but needs refinement in the area of user interface usability and collaboration methodology.