Testing NINES

Dana Wheeles
University of Virginia

Abstract

As scholarly publishing transitions from a static medium (paper) to a digital one, a necessary tension has emerged between a static, or archival, approach to content, and the dynamic, ever-evolving requirements of online resources. Two years after the release of the NINES website (the Networked Infrastructure for nineteenth-century Electronic Scholarship), the first online peer-reviewing organization and hub for the aggregation of scholarly resources in nineteenth-century studies, the development team decided to evaluate the site's efficacy and update it accordingly. But because of NINES's unique position as a federation of scholar-driven (mostly primary-source oriented) sites as well as a forward-thinking software developer, it was essential that any changes to the site remain true to the institution's dual roles. This article explores the NINES redesign as an experiment in implementing usability studies and user-centered design to enhance its appeal within our community. This narrative is meant as a case study, one to be considered within the larger context of the questions and challenges faced by those creating, managing, and using digital projects.

Correspondence:
Dana Wheeles
E-mail:
dw6h@virginia.edu

1 Testing NINES

Like many online scholarly resources in the humanities, NINES was built by a group of talented and multi-skilled scholars engaging the help of knowledgeable software developers. As Jerome McGann and Bethany Nowviskie described in their 2005 whitepaper, NINES was designed as a federated model for integrating digital scholarship, allowing participating sites to maintain their structural autonomy while leveraging a standardized metadata format to bring all of them together in a central hub. Bethany Nowviskie designed the software that would power this site, called **Collex** <http://collex.org>, a Ruby-On-Rails application that allows users to search aggregated sites, collect, annotate, and tag the online objects they discover, and to repurpose those objects in illustrated, interlinked essays or exhibits. From its very conception, **Collex** was meticulously constructed through the use of storyboards, use-case scenarios and focus group testing. The NINES steering committee, a group comprised of scholars invested in the future of humanities work online, was one of the first groups to test **Collex** thoroughly and to provide feedback on how the site worked—not just for contributors, but users as well.

From its release in December of 2005 until 2008, **Collex** had a sleek interface styled in red, white and gray (Fig. 1). The central portion of the screen was dedicated to searching, with blanks for user-entered queries, and expandable lists of the various facets one could choose to browse the sites federated in NINES. Indeed, the design of the faceted browser itself, a tool that allows the user to continually refine the categories of their search in order to pinpoint resources from large pools of data, presented a singular challenge: the developers of Collex needed to adapt the power and flexibility of software generally utilized in the commercial sector (in online shopping, for example) to enable scholarly research. Rarely is the scholar looking for one single object, but rather a (potentially serendipitous) collection of relevant objects based on a targeted list of

Literary and Linguistic Computing, Vol. 25, No. 4, 2010.
doi:10.1093/llc/fqq024

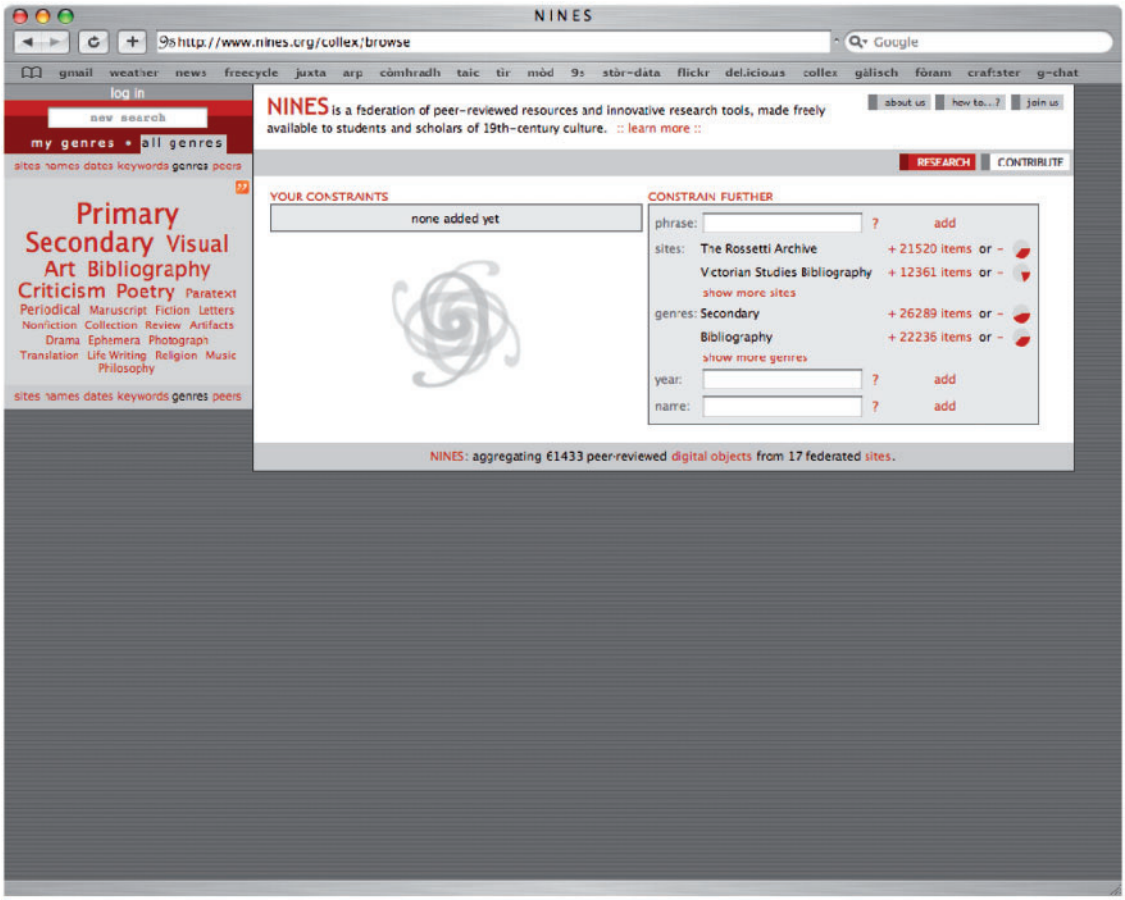

Fig. 1 The original **Collex** design

constraints. Collex needed to allow users to tunnel into our data quickly and efficiently, all the while exposing the numerous other materials available to them. Since searches in a faceted browser can become quite elaborate, a large and central portion of the **Collex** interface was set aside for keeping track of user queries.

The left-hand sidebar served as a reminder of the larger NINES community, displaying a tag cloud of popular keywords, dates, genres or names, depending on your preference. This was also the space where a user could manage his or her account: list collected objects, add or remove tags, and edit your personal information (Figs 2 and 3). In this way, a balance was struck between the user's ongoing research and his or her participation in the scholarly experiment that is NINES.

Over the course of three years, scholarly interest in NINES grew, as did the number of sites aggregated into **Collex**. When the number of digital objects reached 300,000, it became clear that it was time to revisit the site's design and evaluate its performance. Not only was the site's overall functionality in need of testing, but there was also very little feedback about the user experience. The community had grown beyond the confines of literary scholars with digital interests, to one made up of tenured faculty, junior scholars, graduate and undergraduate students, not to mention other developers interested in **Collex** as a generalizable application. Although

Fig. 2 A list of a user's collected objects with the genre of 'Visual Art'

Fig. 3 A tag cloud of all keywords added by users in NINES, sorted by popularity

the tag cloud continued to grow, and activity on the NINES site was apparent, we did not have any metrics in place by which to measure our success or to improve upon our deficiencies.

It became clear that NINES had reached an important milestone in the life cycle of digital resource management. Of the first wave of large Digital Humanities resources launched around the turn of the millennium, all but a few have since undergone considerable site redesign.[1] These changes (cosmetic as well as organizational) were not implemented because these sites had been poorly structured before—quite the contrary. It is precisely at that moment, when the resource has been active for some time and a community of users has begun to form around it, that it becomes crucial to revisit the analytics and prepare to adjust for user feedback. Ideally, resources for usability studies should be inscribed into grant applications and budgets in two phases: the first for the construction of an

intuitive and user-friendly site, and then toward the *end* of the first phase of development. 'Test early and test often', is the maxim in software development—and yet, in the realm of humanities computing, lack of funding as well as expertise can make it difficult to move beyond those first, early tests. The very existence of NINES is proof of the significance of these late evaluations: the idea for an aggregator of nineteenth-century resources emerged years before, during serious conversations about the structure of The Rossetti Archive <http://www.rossettiarchive.org>.

From 2006 to 2008, NINES did not have the resources to engage in regular, formal testing of **Collex**., but rather depended upon the generosity and goodwill of our steering committee and editorial board members. In addition to formal demonstrations of the software at our annual meetings, we sent out periodic requests to this group of roughly thirty scholars to log into their accounts and complete certain tasks, mostly oriented to searching, collecting and tagging objects in NINES. In June of 2008, however, we were able to extend our circle of testers beyond this group of 'super-users' thanks to several attendees of the NINES Summer Workshop, held that year at Miami University in Oxford, Ohio. Of these obliging and charitable colleagues, a number were ideal subjects for our experiment: ranging from graduate students to faculty and those trained in library science, the group was representative of NINES's overall audience. In addition, some of these participants had no prior experience with the **Collex** software, and were ideally suited to the task of showing us just how accessible the site truly was.

Thanks to Laura Mandell, Associate Director of NINES and Professor of English Literature at Miami University, NINES was granted access to a usability lab in the library, complete with a computer fitted with eye-tracking software. Using a camera mounted to the monitor, this application recorded the action of the participants' pupils in concert with their interactions with the site displayed before them. Eye-tracking software, though certainly not essential to successful usability studies, can offer a telling glimpse into the user's experience of your site: the attractions as well as the distractions. Web usability consultant Jakob Nielsen used eye-tracking software to pinpoint the now widely recognized phenomena of the F-Pattern of screen scanning (that is, most users will rapidly take in your site's content from left to right in the shape of a large F) and Banner Blindness (by which users learn to ignore certain portions of the screen to avoid advertisements)—concepts that have profoundly influenced design for the web in the past decade.[2] This technology was particularly useful in our case, not only because this was our first true usability test, but also because it allowed us to save the results for further in-depth review back in our offices at the University of Virginia.

With the technology in place, the NINES development team conferred about the aspects of the site we would test. A series of three scenarios were decided upon, guiding our five volunteers through the search interface, the collecting and tagging features, and finally, the newly unveiled authoring tool, called Exhibit Builder. Over the course of the test, we hoped that each action would build upon the last, encouraging the user to engage in more and more complex activities. The culmination of these tasks would be the basic template for a short essay based on federated materials, allowing us to gauge the viability of producing user-generated content in the current interface.

Very early in the test, we discovered that users were having trouble deciding where on the screen to focus their attention. Figure 4 shows a screenshot taken during one of the tests, with complete data from the eye-tracking software. The colored areas indicate places where the user's eye rested (with red marking the longest durations), and the red 'x' marks represent mouse clicks. From this screen, which documents a search for objects by John Keats, one can see that the user does not dwell on the search results listed below, but lets the gaze rest on the constraints above, and the overall number of results. The tag cloud at the left is almost completely saturated, clearly identifying it as an attractive feature, but one that distracted the user from the search in progress. This same pattern unfolded time and time again leading us to conclude that users were expecting the information in that left sidebar to interact with the main search function.

Fig. 4 Screenshot of participant engaged in Task 2: *Search for objects by John Keats*

They did not perceive it as the separate, community-oriented navigation system we had intended.

Although test participants were uniformly confounded by the presence of the tag cloud, their searches were relatively straightforward and successful. We were pleased to see one participant move beyond the search blank to experiment with the genre facets, demonstrating a measure of ease and adventurousness despite the complexity of options available to her. However, once the searches had been completed and objects collected, each and every one of our volunteers balked at the idea of taking the next step toward authorship. When instructed to perform the final task, 'Using the objects you have collected, create an illustrated essay', every single person in our study responded immediately with nervous laughter. Clearly, we had encountered some difficulties.

In retrospect, our mistakes were three-fold: first, writing an essay requires more than clicking on objects in an interface, and we had neglected to make it clear to the participants that they would be supplied with dummy text. Usability studies are most successful when the volunteer feels secure that it is

the site that is being tested, not **them**. Second, the creation of an essay in this online space was not necessarily the logical step we had assumed it to be. Academics and librarians have become more and more accustomed to digital search and retrieval mechanisms, making that aspect of **Collex** relatively familiar. However, the transition to using that same space for the authoring side of scholarship was far from intuitive. Realizing the importance of this divide led us to realize that we had engineered the site to feel too much like a search engine, without providing sufficient emphasis on authorship as an expected and encouraged end result. Lastly, questions of interface aside, we had asked too much of a single usability test. In this, our first study, we had packed too many concepts and features into the evaluation, over-loading the participants, and glossing over more subtle functions that could be ameliorated.

After reviewing these results, we determined that the application was fundamentally sound: its features were compelling and the searching capabilities attractive. In a way, **Collex** was *too* powerful, to the extent that it was intimidating. Everything about our interface, from the NINES home page to the **Collex** engine, had been conceived in scholarly and theoretical terms, and demanded a similar dedication from the user. And while NINES has always sought to appeal to the growing community dedicated to the mission of excellence in digital scholarship, we never intended to limit our users exclusively to its membership. We wanted the website to be a friendly environment: a place where both scholars and their students go to conduct research in the nineteenth-century studies online.

As we began to sketch out the redesign, our primary goal was to break up the **Collex** interface into its component features, giving each one its own important area of the site. In many ways, our task was much simpler than the one that faced the original designers of **Collex**—in the intervening years, the number of faceted browsers had multiplied considerably leading to the emergence of web conventions for their styling and organization. Drawing attention to the social aspects of NINES was our primary goal, however, and ensuring that even the most casual visitors to our site were aware of the user-generated content as well as the aggregated resources.

In a series of design meetings, our lead developer, Nick Laiacona of Performant Software Solutions, NINES Director Andrew Stauffer and I identified what we believed to be the most important features of the site and drafted a new primary navigation system.

The result was a set of wireframes, or mock-ups with no style information, that exemplified the new structure. On the new home page (Fig. 5) we made a concerted effort to elide the information about NINES as a community-driven, peer reviewing institution with the features of **Collex** itself. A prominent search blank served as an invitation to stay and explore the site, while illustrated and extensively hyperlinked text below could introduce new users to the various operations possible within the site. The tag cloud was separated from the search page altogether, and a tab was set aside for 'Exhibits' as well. A blog was added under the 'News' tab to indicate the dynamic nature of NINES, and under 'About' one could find information about the editorial boards, peer review principles and our various outreach efforts such as the summer workshop.

The implementation of these changes on a test site, however, led to concerns that we still had not done enough to connect research in NINES to other possible acts of scholarship. The main tabs proved useful in exposing the site's components, but their very nature implied a separation, or disconnect, between the modules when in fact scholarly work is a fluid endeavor.

After exploring social media sites such as Facebook and Flickr for successful examples of user-generated content management, the NINES team conceived the *My 9s* page, a private homepage within NINES (Fig. 6). On this page, all authoring, editing and collecting efforts are centralized, breaking down the borders between search, analysis, and authorship. From this dashboard, users can craft their own identity in NINES (including links to an institutional home page or curriculum vitae), always maintaining complete control over how much of that information is shared. Tags, saved searches, recently collected items and annotations are all easily accessed in this space making re-mixing and

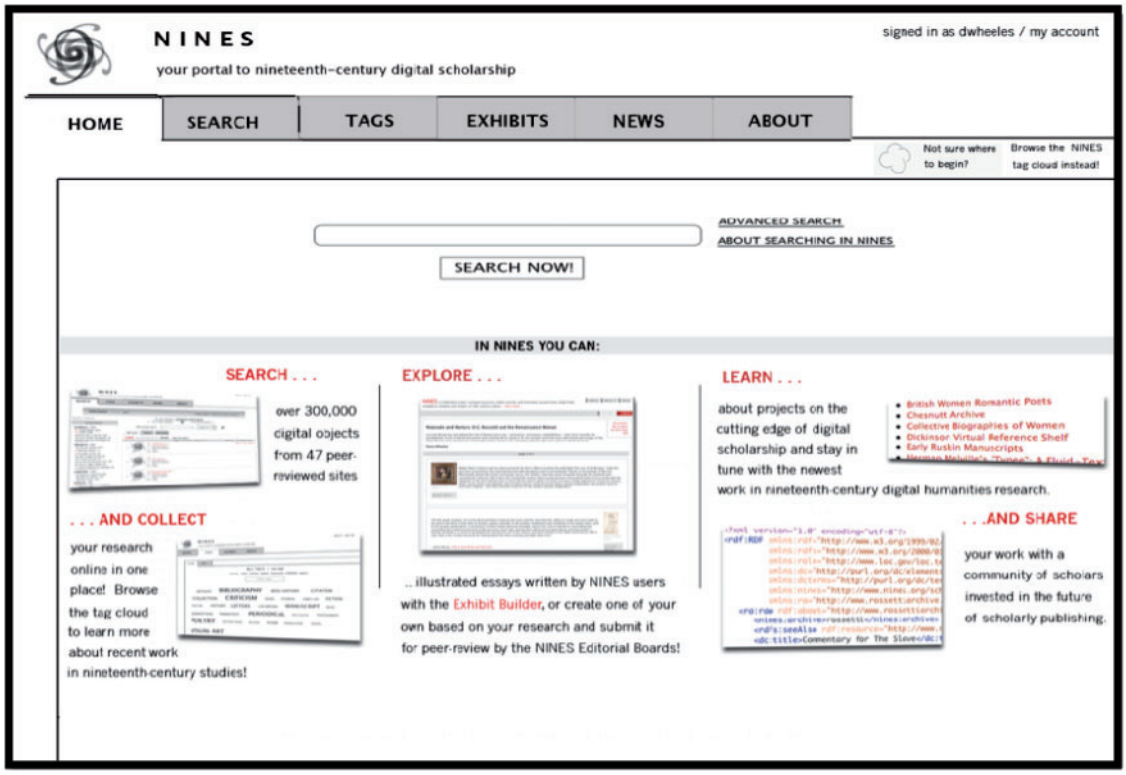

Fig. 5 Wireframe of redesigned NINES home page

re-purposing for future publication a much simpler operation. Exhibits, which had previously been a mysterious and disconcerting option, could now be offered as a reasonable outcome of work within NINES.

Before the new site design could be released, another usability test was necessary to make certain that the new enhancements did not disrupt previously functional features. Although the first study had provided us with invaluable information about user experiences in NINES, we did not want to duplicate our prior missteps. A formal study was arranged under the auspices of the University of Virginia Library, in which organizers and proctors of the exam would be library staff, and therefore less biased in their approach to and interpretation of the site. In all, nine volunteers were recruited from the general faculty and student body of the university, and a small sum of money was set aside to provide incentives for participation (in our case, a free coffee at a nearby café). The version of NINES to be tested looked much like the wireframes in Figs 5 and 6: we wanted to be sure that the users were paying attention to basic structure and functionality, not graphics or color.

In this test, participants were guided through a series of tasks and asked to think aloud during the session, allowing the facilitator to take note of their experience. Each session lasted sixty minutes, during which the facilitator took extensive notes on the aspects of the site that worked as expected, needed improvement, or provided considerable difficulty.[3] This report became the basis for numerous enhancements on the site especially with regards to the faceting sidebar (were users availing themselves of it?), account creation (did they take note of it as

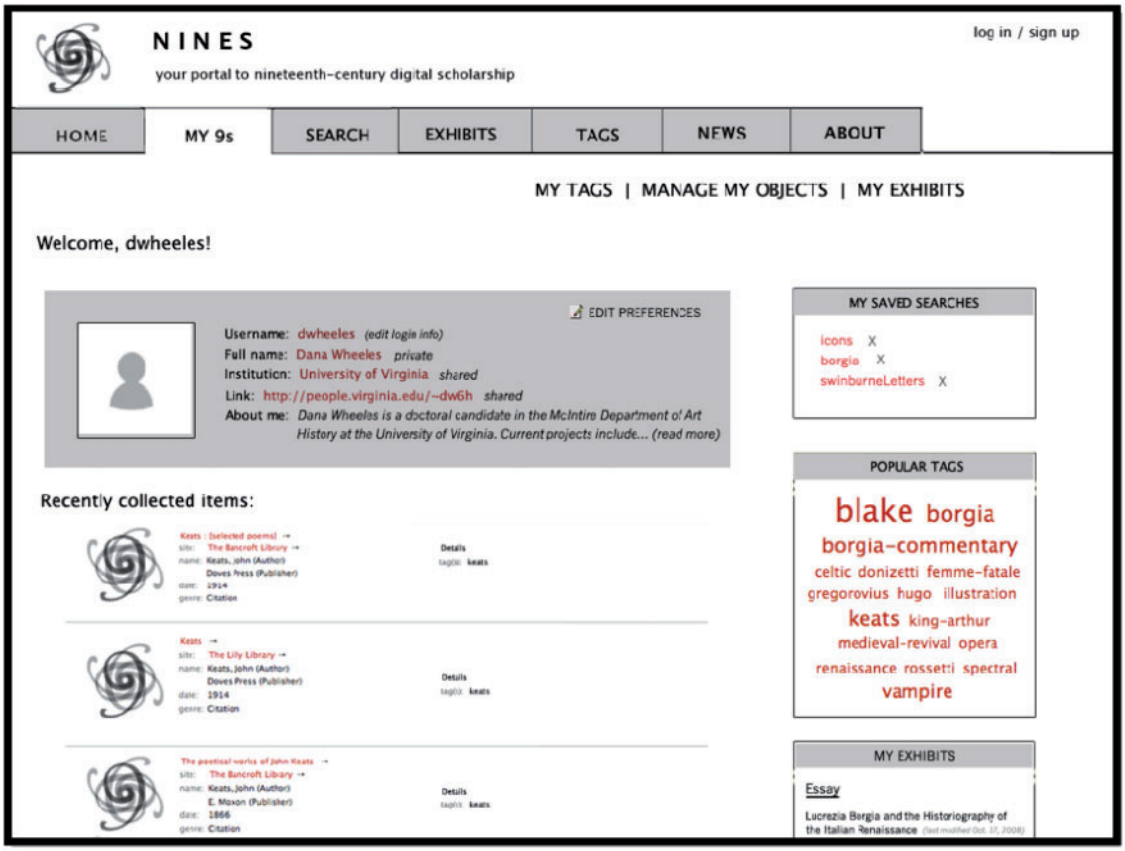

Fig. 6 Wireframe of new 'My 9s' page

an option?), and the site's credentials and mission (were the institutional aspects of NINES made clear?).

We were able to respond immediately to a number of the recommendations offered by the usability team in the first phase of the release, launched in time for the MLA Conference in December of last year. As can be seen in the screenshot in Fig. 7, this version closely resembles the wireframes—mostly because the bulk of our efforts had gone into refining the functionality of the site.

In response to the student who wondered, 'How are the results displayed? Alphabetically or what?? Maybe they are organized by libraries??? I don't know...', we implemented a search sort function on title, name and date in addition to SOLR's relevancy factory.[4] Tags were exposed at the search result level, not just hidden away on the tag cloud, and a name browser implemented so that the results of any given search could also be explored based on the author, editor or publisher information provided in the metadata.

Having fully released the new site, complete with the *My 9s* page, Exhibit Builder and the blog, we turned our attention to the look and feel of the site. After having completed the first phase of the redesign in-house, we turned to professional web designers to craft a new motif: one that looked less slick and software-like, and more evocative of nineteenth-century studies. After all, a fully functioning website is useless if it feels cold and alien to its intended audience: it was time to respond to our users on a deeper level.

Working with Gibson Design Associates, we implemented a friendlier version of NINES; one that was intended, to quote Director Andrew Stauffer,

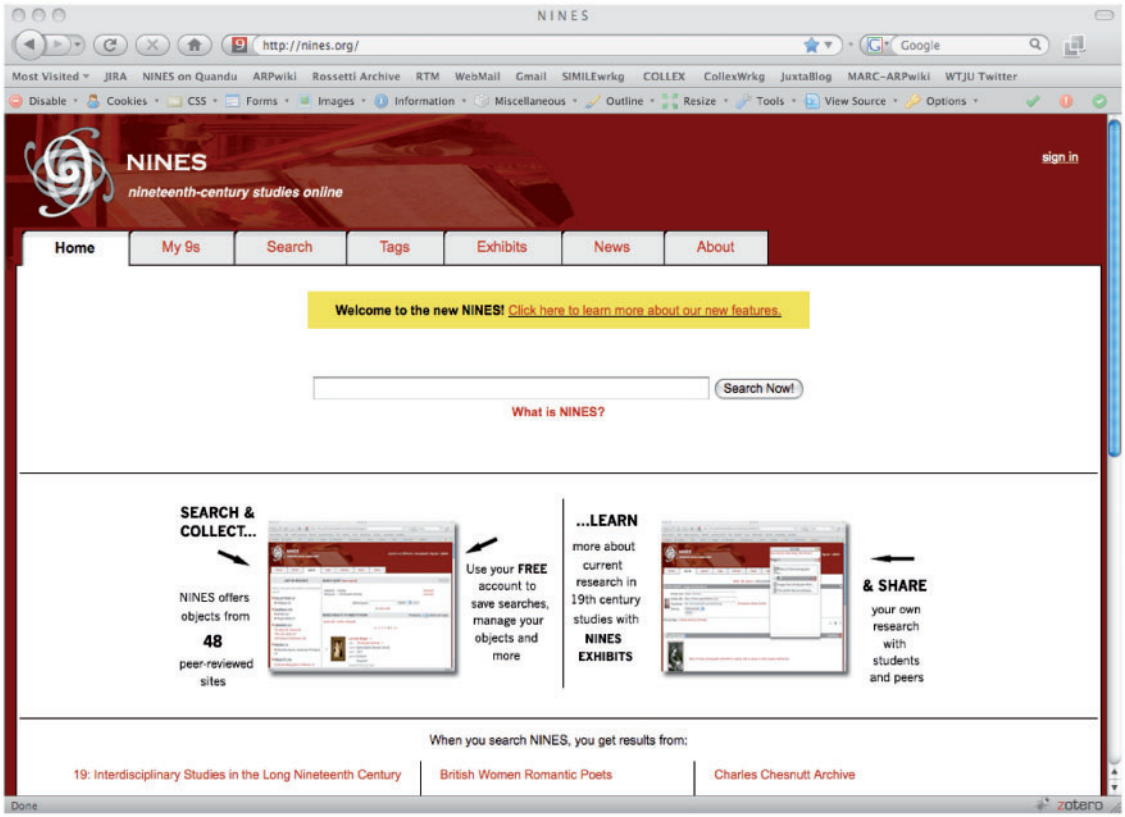

Fig. 7 The NINES Home page, December 2008

'to feel like settling into a big leather armchair.' The lively, yet cold color palette of the old interface was replaced by one of warmer shades of red and brown, and where possible, we added descriptive icons and tool tips to make the site m ore clear (Figs 8 and 9).

The prominent search blank was retained, as were the navigations tabs at the top. These were augmented by information about what one can do in NINES: a small tag cloud showing the most frequently used terms, a syndicated feed of the most recent NINES blog entries, and links to the most popular discussions ongoing in the newest Collex feature: the forum. A featured exhibit is prominent, showcasing our Exhibit tool, and a carousel was added, highlighting participating projects. With all these pieces in place, our site was finally ready to be a place for scholarship and community building, not just an appliance used to find resources.

Responses to the new NINES site have been generally positive, although a few have voiced concerns over the 'masculine' nature of the new look. Since the release of the redesign in 2008, the traffic to our site has continued to increase, from roughly two hundred unique visitors per month to almost four thousand.[5] Thanks to our newly implemented statistics records, we noted that many of our first-time visitors were directed to the NINES user tag lists from Google searches, a development that has led us to refine those pages to reduce their high bounce rate.[6] Most of the scholarly requests and feedback that we have received in recent months has been related to NINES as a publication model as well as a center for the pedagogical implementation of online scholarship. In response, we have decided to re-structure the main tabs once more, making a clear delineation between user-generated content

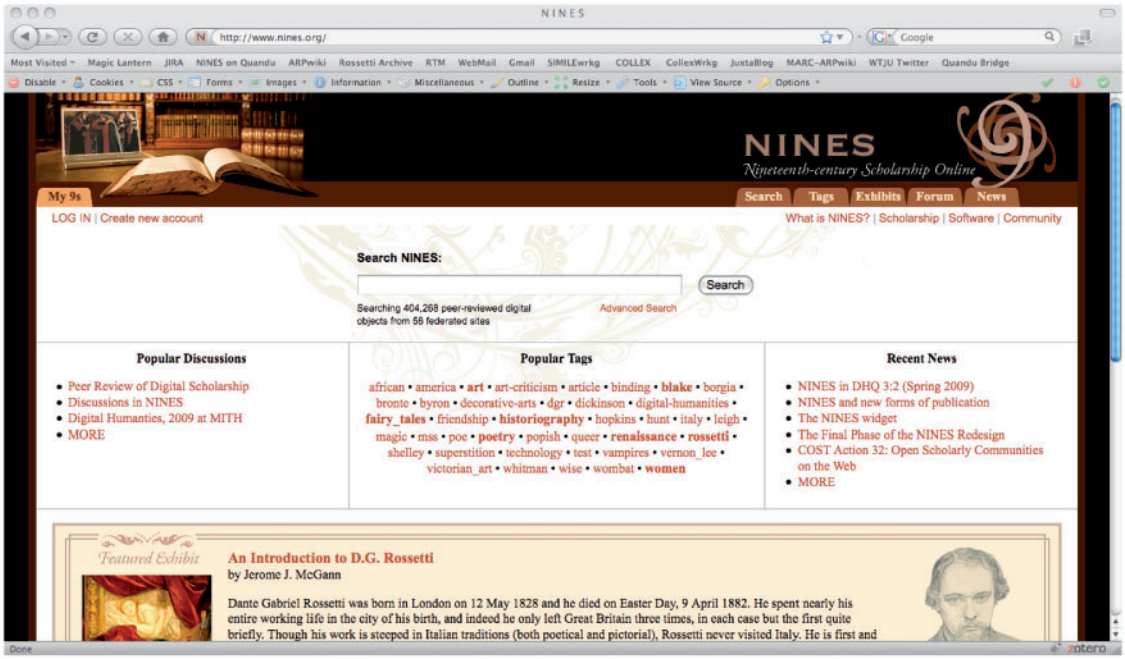

Fig. 8 The NINES Home Page, June 2009

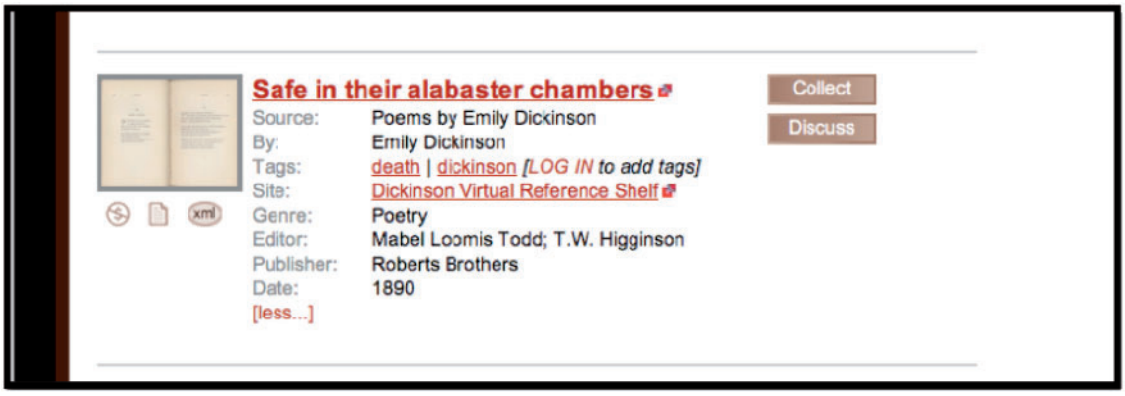

Fig. 9 Sample search result in the new interface, showing user tags in-line, the new 'Collect' and 'Discuss' buttons to encourage user activity, and descriptive icons

in NINES (flagging it as peer-reviewed or simply experimental), classroom use, and our original federated search interface (Fig. 10).

The main navigation of NINES has shifted to highlight 'Publications' (that is, journals and other contributors of peer-reviewed content created within our *Exhibit Builder* framework), 'Community' (experimental exhibits shared with the web and other groups oriented to the study of particular subjects in the long nineteenth century), and 'Classroom'

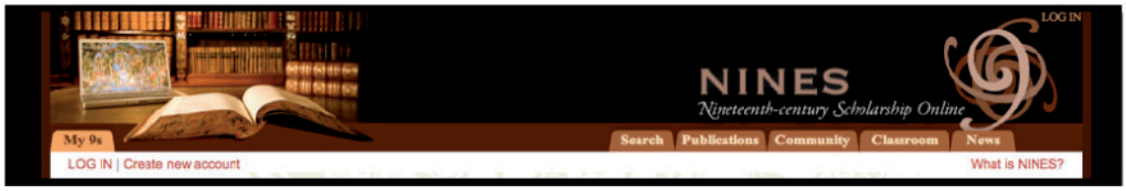

Fig. 10 The most recent change to the NINES site, in response to user feedback since the redesign launch in 2008

(content generated by professors utilizing NINES in their graduate and undergraduate classes).

Because of the way NINES is structured, we are only as strong as the community that we are able to build. Part of this process is built into our work as an institution: it involves the peer-review and aggregation of sites, and summer workshops to help new projects get off the ground. But our web interface and software tools are central to the ways that NINES will gather scholars together. Through the activities of this past year: usability studies, development meetings, discussions with scholars in the field, and pure trial-and-error, we have come to recognize the importance of re-tooling, and responding to our users—a process that is ongoing and constantly engaged with the social and technical ecology of the web.

An earlier version of this paper was published in the Digital Forum of the Journal of Victorian Culture, Volume 15, Issue 1, 01 April 2010, Pages 145–149.

Notes

1 The Whitman Archive <http://www.whitmanarchive.org/> is an excellent example of site redesign in which navigation became more prominent and accessible, exposing valuable pre-existing features to a wider audience.

2 Books on usability studies have proliferated in the computer science community since the mid-90s. See Jakob Nielsen, *Usability Engineering (Interactive Technologies)*, Academic Press, 1993.

3 Erin Mayhood, 'NINES Usability Report: Testing Performed 20–24th October 2008'.

4 NINES search results are, by default, sorted by 'relevancy' determined algorithmically by the number of times a search terms appears within a document. Learn more about SOLR, an open-source search server based on the Lucene Java search library, at <http://lucene.apache.org/solr/>.

5 The bulk of traffic to our site is gauged by Google Analytics, augmented by a secondary, Collex-specific set of statistics that tracks membership activity.

6 The tag cloud was originally conceived as a starting point for browsing NINES, leading users to list of objects that had been keyed with similar keywords. When we discovered that the Google search engine was directing a significant amount of traffic to these lists, we added more navigation information to them to ensure that visitors would be compelled to stay and explore the site.