

From Records to Data with Viewshare: An Argument, An Interface, A Design

by Jefferson Bailey and Trevor Owens

Digital Humanities and Information Visualization

EDITOR'S SUMMARY

Viewshare is a platform designed to support the work of cultural heritage institutions in their efforts to share data, dynamically visualize their collections and access and discover items. It was developed as a free and open tool for partners of the National Digital Information Infrastructure and Preservation Program. Its central goal is to reveal patterns within collections to lead to new understanding. Viewshare features a simple and intuitive interface and limited requirements for technical skills, and it supports a wide variety of digital collections and metadata types. Its users establish custom metadata, import and enhance collection data and create their own interface to support dynamic and interactive views of their digital collections. The process is illustrated with the Fulton Street Trade Card Collection from the Brooklyn Public Library.

KEYWORDS

museum collection management systems interfaces
electronic visualization human computer interaction

Jefferson Bailey is currently a fellow focusing on digital preservation projects in the Office of Strategic Initiatives at the Library of Congress. Previously, he managed digital projects at Brooklyn Public Library and at Frick Art Reference Library and has done archival work at NARA and NASA. He can be reached at jbailey@loc.gov.

Trevor Owens is a digital archivist with the National Digital Information Infrastructure and Preservation Program (NDIIPP) in the Office of Strategic Initiatives at the Library of Congress. At the Library of Congress, he works on the open source Viewshare cultural heritage collection visualization tool, as a member of the communications team and as the co-chair for the National Digital Stewardship Alliance's Infrastructure working group. Before joining the Library of Congress he was the community lead for the Zotero project at the Center for History and New Media and before that managed outreach for the Games, Learning and Society Conference. He can be reached at trow@loc.gov.

Creating and manipulating an interface to a digital collection is fundamentally an interpretive act. Viewshare (<http://viewshare.org/>) is a free and open platform that intends to make it easier for archivists, librarians, curators and scholars working with discreet digital collections within cultural heritage institutions to engage in that interpretive act. Through work in digital preservation, the National Digital Information Infrastructure and Preservation Program (NDIIPP) team is continually reminded of the critical need our partners have to provide access to the collections they are working to preserve. Access remains an essential way to demonstrate the value of these collections. Viewshare grew out of a project designed to enable NDIIPP partners to provide access to their diverse digital collections, but has widened its scope to support the broader cultural heritage community's efforts to empower data sharing, the creation of dynamic collection visualizations and access and discovery of their digital collections.

In this essay, we describe some of the ideas and ideals that informed the design of Viewshare. This effort is, in effect, an argument for the design of this free, open-source tool. We will suggest how Viewshare's design helps empower cultural heritage stewards to create new ways to encounter, manipulate and explore digital collections. In the next section we briefly describe three goals that have emerged through our iterative software design process. They include setting a low barrier for entry and emphasizing usability, allowing users to work with heterogeneous data and helping to reveal emergent patterns in collection data. We then provide a walkthrough of how to use Viewshare and examine a sample collection in order to illustrate how these broader goals are enacted in the design and use of the application.

Values Embedded in Designing Interfaces

In developing Viewshare our goal was to create an application that is open, beneficial and widely adopted. To this end Viewshare leverages the existing open source Exhibit software as a lightweight structured data publishing mechanism [1]. Viewshare aspires to provide a dynamic tool useful to collection stewards and users who need to share, build and interpret digital collection interfaces. It accomplishes this task by designing features and workflow that can be utilized by non-technical users, by supporting ingest and augmentation of heterogeneous, extant data and by empowering the visualization, discovery and understanding of collection-wide trends and patterns.

Low barrier to entry. The software needs to provide collection managers without deep technical skills an intuitive way to articulate and visualize their deep knowledge of their digital collections. It needs to balance among the complex demands of being “sophisticated, robust, transparent, and easy to use” in order to attract a broad user base [2]. To this end Viewshare is designed entirely around a drag-and-drop interface.

Work with heterogeneous extant data. The software needs to support the heterogeneity of metadata across many different types of institutions and collections. It also needs to offer behind-the-scenes tools to transform this metadata into the kinds of data required for visual interfaces. To demonstrate the value of a range of digital and digitized collections, the tool needs to work with many different kinds of collections and with a heterogeneous mixture of extant content and metadata.

Reveal emergent patterns in collections. The software needs to enable visualizing cultural digital collections as unified sets of data and not just as discreet, individual objects. Visualization is thought of as a process for revealing and illustrating knowledge. Recognizing work from humanities scholars on visualization, we see the value of working with visual interfaces as part of a process in which we create new knowledge and understanding. According to Drucker, the process of visualization can be “generative and iterative, capable of producing new knowledge through the aesthetic provocation” [3, p. 41]. In embracing this subjective and iterative notion of

visualization, the Viewshare platform attempts to make it easy for librarians, archivists, curators and others working with cultural heritage collections to exercise and surface patterns in their collections.

The Viewshare Workflow

Viewshare’s workflow is designed to support these particular goals. Users import and augment existing collection data, iteratively build interfaces to their collection data and ultimately are able to share the interfaces and views that they have created. We will briefly explain the steps in this process and how they are connected to the design goals and then further illustrate those connections in an example.

Viewshare interfaces are built entirely on user-uploaded metadata. Recognizing the heterogeneity of collection data, Viewshare allows multiple methods of importing data. Users can build or work from existing simple spreadsheets or MODS (Metadata Object Description Schema) records or import Dublin Core metadata via OAI-PMH (Open Archives Initiative Protocol for Metadata Harvesting). To make this data usable, Viewshare includes a set of data augmentation tools to work from this extant data. For example, Viewshare enables users to derive latitude-longitude coordinates from plain-text geographic place names and then use these coordinates to plot their items on a map. Similarly, plain-text expressions of date information can be used to derive ISO 8601 formatted dates for plotting items on a timeline. With its ease-of-ingest and data augmentation features Viewshare understands and facilitates the use of the unique and sometimes idiosyncratic nature of cultural heritage collection metadata. At the same time it also allows users to enhance this metadata in order to enable the creation of dynamic interfaces.

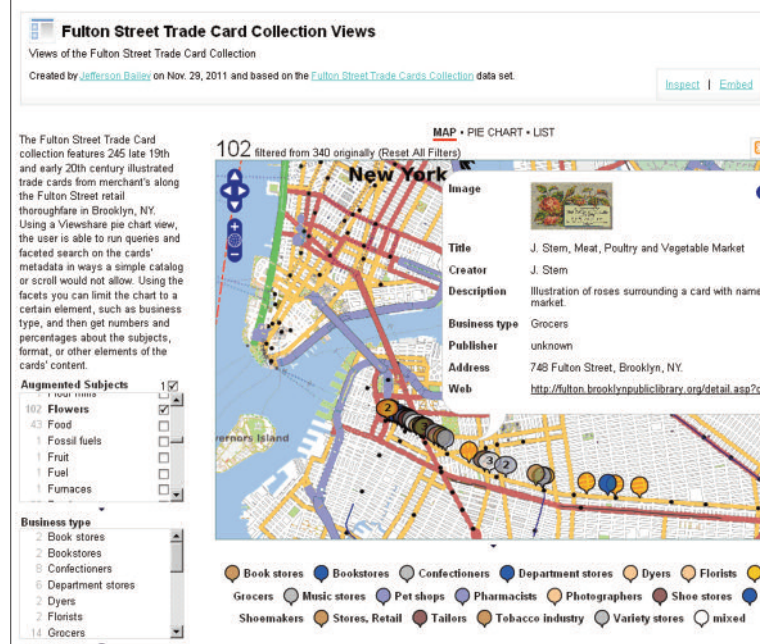
After importing and augmenting collection data, users can begin creating interfaces. The tool’s primary purpose is building dynamic, interactive views of digital collections. Through a drag-and-drop workspace, users can create multiple views including maps, timelines, charts and other dynamic visualizations. Users can then choose which facets they want to include, and these facets will be available for each view, creating unique ways of manipulating the data presented in each of them.

For instance, in a collection of postcards, a tag cloud facet set to display subject information will show the relative frequency of the subjects throughout the collection. If a user clicks on one of those subjects, Viewshare will limit the display of whatever view they are using to show only the objects associated with that term. As a user selects the data values they want to use in a given facet, and the particular views they want to display, they can use the “show preview” function to continually toggle back and forth between building their interface and a fully functional preview of what their resulting interface will look like. In this way, the tool supports an iterative and exploratory approach to creating these interfaces.

A Working Example: The Fulton Street Trade Card Collection

The Brooklyn Public Library’s Fulton Street Trade Card Collection, a collection of late 19th and early 20th century merchant trade cards from Brooklyn, New York, is an excellent example of a “hidden” special collection that generates significant patron interest and use; it is also a good demonstration of Viewshare’s ability to empower new ways of navigating, interpreting and understanding collections. This colorful, idiomatic group of hundreds of advertising cards offers many insights into the commercial and social world of turn-of-the-century Brooklyn [4]. Special collections such as this one are not just a grouping of individual items, but should also be thought of collectively as datasets documenting the trends, influences and styles of this time period and place. Using Viewshare, a collection manager or curator, with his or her deep knowledge of the contextual details of the collections, can create views that allow users to interact with this collection in ways not

FIGURE 1. Screenshot of Fulton Street Trade Card Collection views using “flowers” from the “Augmented Subject” facet as the search term (<http://viewshare.org/views/jefferson/fulton-street-trade-cards-collection/>).



possible via a static web interface; and those views can also empower users to uncover collection-wide relations not evident or interpretable through traditional online gallery display or through item-by-item browsing. A working example of the view we are about to describe is available on the Viewshare website (<http://viewshare.org/views/jefferson/fulton-street-trade-cards-collection/>) and an image (Figure 1) at left.

By uploading a spreadsheet of collection data that includes links to the web-accessible image files, a collection manager can begin building new interactive views. After deriving points of latitude and longitude for the cards, the user can create a map view. The map view shows the locations of each card’s business. A clickable pin on the map allows users to see a thumbnail image of the item and select metadata elements. When users add

a facet to the view, they can click on any facet element, such as the Augmented Subject element “flowers,” as demonstrated in Figure 1. The map will then update to show only the location of the trade cards with flower imagery. Adding other facets such as date or subject will allow a user to further manipulate the display. A pie chart view can also be constructed from this same data set. The pie chart view defaults to a creator-defined metadata element. Each pie slice, when clicked, shows the calculation of the total number of items with that particular element in the collection and its percentage of the collection as a whole. For example, in a chart built on the element describing business type, the pie slice for “confectioners” when clicked will show there are 15 trade cards from “confectioners” representing 4.4% of the collection. Faceting can be used on charts, too, so if viewing a pie chart of subject headings, a user can then apply the business type facet

and see that the 15 confectioners used four different image subjects in their trade card advertising images. Switching to a gallery view, a user could use the same faceting and see thumbnails of those specific cards. As users create and iteratively interact with their collections they are simultaneously building an interface and deepening their own understanding of the collection and the collection's metadata.

Interfaces for Visualization and Understanding

There are many ways to build interactive interfaces and visualizations of cultural heritage collections. However, the time, cost and expertise necessary for creating these kinds of interfaces leave them outside the reach of many cultural heritage collections and users. To this end Viewshare is intended to empower librarians, archivists, curators and other cultural heritage professionals to create interfaces that help illustrate the value of the content they are working to preserve and make accessible. By making it

easy to rapidly and intuitively create these interfaces, we hope that Viewshare can enable cultural heritage professionals to experiment with and explore their collections. In this respect, Viewshare aspires to John Bradley's characterization of another digital humanities tool: "It is meant to be a tool that blends so well into the task of the development of an interpretation... as to be almost invisible" [5, p. 263].

In sum, working with idiosyncratic extant collection data, Viewshare enables new ways to visualize and analyze cultural heritage collections. By providing these new levels of interactivity, it encourages users to see digital collections less as discreet items and more as broad, unified datasets that can be manipulated and interpreted to form new modes of understanding. In this sense, Viewshare supports the ongoing efforts to involve existing cultural heritage assets in the growing use of information visualization in digital humanities scholarship as it also provides cultural stewards a free, intuitive tool for the display and use of digital collections. ■

Resources Mentioned in the Article

- [1] Huynh, D. F., Karger, D. R., & Miller, R. C. (2007). Exhibit: Lightweight structured data publishing. In *Proceedings of the 16th International Conference on World Wide Web* (pp. 737–746). Retrieved November 22, 2011, from www2007.org/papers/paper161.pdf.
- [2] Borgman, C. L. (2009). The digital future is now: A call to action for the humanities. *Digital Humanities Quarterly*, 3(4). Retrieved November 22, 2011, from www.digitalhumanities.org/dhq/vol/3/4/000077/000077.html.
- [3] Drucker, J. (2010). Graphesis: Visual knowledge production and representation. *Poetess Archive Journal*, 2(1), 1-50. Retrieved November 22, 2011, from <http://paj.muohio.edu/paj/index.php/paj/article/view/4/50>.
- [4] Brooklyn Public Library. (2011). *Fulton Street Trade Card Collection*. Retrieved November 22, 2011, from <http://fulton.brooklynpubliclibrary.org/>.
- [5] Bradley, J. (2008). Thinking about interpretation: Pliny and scholarship in the humanities. *Literary and Linguistic Computing*, 23(3), 263-279.