Franco American Portal Project

Building a discovery tool for Franco American collections

White Paper: Portal Design and Tool Expectations

UMaine Franco American Programs
Fall 2019

Susan Pinette, Ph.D
Director of Franco American Programs

Jacob Albert
Program Manager

Project Advisory Board:
Doris Bonneau - University of Southern Maine Lewiston-Auburn College
Keith Chevalier - St. Anselm College
Leslie Choquette, Ph.D - Assumption College
Libby Lipin - Assumption College
David Nutty - University of Southern Maine
Lise Pelletier - University of Maine at Fort Kent
Mary Rice-DeFosse, Ph.D - Bates College

1865 THE UNIVERSITY OF MAINE

Table of Contents

Project Background

The Franco American Portal Project (https://francoamericanportal.org/) is a five-university collaboration to build a discovery tool for Franco American collections. Our project addresses the need for online access to Franco American cultural history by: locating and identifying archival materials that concern French-Canadian and Acadian diaspora communities in the US Northeast, wherever these materials have been collected around the world; bringing together information about these collections and their contents in an organized, searchable, and culturally conscientious way; and making these collections and their contents accessible by directing the public toward their digital presence and the institutions that collect and preserve them. This project is funded by the University of Maine and is a collaboration between the Franco American Programs at the University of Maine, the Franco-American Collection at the University of Southern Maine Lewiston-Auburn College, the Acadian Archives/Archives acadiennes at the University of Maine at Fort Kent, the Msgr. Wilfrid H. Paradis Archives & Special Collections at St. Anselm College, and the French Institute/Institut francais at Assumption College.

Purpose: Design & Expectations

The Franco American Portal Project proposes to build an open-access, online tool for both scholarly and lay audiences. In order for this tool to be successful with these audiences, its design will require attention to details that meet the needs of a range of different users: academics, genealogists, educators, students, community researchers, and the general public. Our team will make design choices based upon these considerations; these choices will have impacts on portal aesthetics and usability, and may also have some bearing on the portal's functionality.

At the same time, the portal's design will need to respond to the needs of its users in ways that also attend specifically to the cultural materials that will make up the linked content of the portal. The broad scope of this content will be Franco American collections of diverse type and medium, with varying levels of description, located in repositories throughout North America.

This white paper will present design considerations our team has identified as crucial to the portal's success in bringing diverse content to a diverse user base.

Drawing upon our team's discussions, as well as upon what we have learned from similar online portals constructed for historic materials in other subject areas, we will use this portal white paper to:

- identify functional requirements for our portal and highlight portal features our identified audiences will expect for a functional tool catering to Franco American materials;
- reflect on aesthetic elements that may make sense for our project's needs; and
- speculate on design innovations that our project and its historical content could benefit from.

This white paper is meant to serve as a report on the state of our team's considerations for the design and functional requirements of the online tool we propose. As such, this document works toward a blueprint for our portal's construction. Indeed we are aware that most web design is iterative: many of our design decisions would not necessarily preclude their own undoing after their implementation. At the same time, we are also aware of the depth and breadth of our project and how, at virtually any point in the project, small design decisions can form the foundations for subsequent design decisions that may make future adjustments difficult, time consuming, and costly. Bearing these competing ideas in mind, we intend for this documentation of our design reflections not to be a strict determination of that design or any of its elements, but instead to be a guide.

This white paper will be software agnostic; our discussion of design will not be adherent to the particular limitations or features of any one type of content management or digital asset management software. Though it may comment on the benefits or pitfalls of certain technology brands, and though our team proceeds on the assumption that our tool will be produced with open source technology, the design elements discussed in this white paper will not commit the overall design of our portal to any software type(s).

Functional Requirements

Our portal project team has identified several baseline functional requirements for a portal tool that are in-step with the needs of Franco American archival materials, the needs of our projected audiences, and University of Maine policy for web accessibility. In order to satisfy these requirements, we will need to make certain decisions that impact the design of the portal. Some of these design implications will be illustrated in more detail later on in this paper.

Functional requirements of the proposed portal as of December 2019 are:

- ***Americans with Disabilities Act (ADA) compliance***
According to the University of Maine Division of Marketing and Communications, all University webpages are required to be accessible according to the US federal government's ADA Section 508 Priority 1 guidelines, to assure that all users can access and benefit from University of Maine web content. Because the Franco American Portal Project is a web project sponsored by the University of Maine, and because we are committed to making our portal accessible to diverse audiences, we will ensure, to the best of our ability, that our portal meets the US federal government's ADA section 508 priority 1 guidelines. These guidelines can be read here: https://webaim.org/standards/508/checklist (Center for Persons with Disabilities, 2019).

- ***Multilingual support***
The Franco American Portal Project is distinctive in that the records, collections, and materials that will comprise the searchable content of the portal occur in English, in French, or in both languages. The design of the portal therefore must accommodate and support the display of French, English, and bilingual content. Because the portal will not host these records and collections as complete digital objects themselves, but will simply link out to these objects and to the digital repositories where they are hosted, this means that the "searchable content" representing each item or collection in the portal will be catalog records or finding aids of these linguistically diverse materials.

At baseline, our discussions of accommodating French and English content have included the potential of cataloging item and collection metadata into the portal in the language in which that material occurs. For example, an item summary of a letter written in French would occur in French; the title of a book written in English would occur in English; etc. In other words, the content of metadata fields would be represented from the item itself, and not translated from its original language. Naturally, due to the linguistic variations in the materials that the portal will include--especially when items occur in both languages, when items occur in no language (like some photographs), or when items have corresponding translations--this approach may not be universally adequate. Nevertheless, we know at the very least that we are committed to accommodating the linguistic needs of French, English, and bilingual materials in order that their appearance in the portal 1)makes them more accessible to the wider public, and 2)does not impact the manner in which their original language gives them meaning and helps shape their material identity or cultural context.

At the same time, because of the types of materials it will include, this project would benefit from addressing the needs of audiences that read only English, only French, or both. Certain design elements of the portal would therefore improve user experience if they were able to be adjusted linguistically, depending on the linguistic needs or preferences of each individual user, without altering the shape of the portal's contents. This means that certain language included in the structure of the portal, depending on the language selection of the user, could appear in either English or French. These language areas would likely include metadata field headers, portal navigation tabs, search facets, and all narrative language on static pages describing one or another aspect of the portal itself. These language areas would also likely include Library of Congress Subject Headings (LCSH), whose French translations can be made available to information professionals (insert citation here); and topics, subjects, themes, or other culturally-specific terms developed within the portal's controlled vocabulary. In terms of design, we see this type of multilingual support mechanism often in the form of small flags at the top of a user's screen (USA or UK representing English; France representing French; etc.), or a two-letter alpha code (EN representing English; FR representing French), or a "Language" dropdown, where a user's selection determines the language of the site's interface.

To extend this approach, the language areas just described could also go so far as to include the metadata itself, the "searchable content," for items and collections. In this case, for example, a letter written in French could be described both in English and in French, or an English item's title could be translated into French, with the user's interface language selection determining the language in which any individual user's content display will appear. These language areas will likely not include, however: geographic place names, author/creator names, and other proper nouns; or dates. Keeping in mind the precautions described above in service of maintaining cultural context, if this type of language presentation takes place, it would be essential to indicate the original language of the item or collection, where applicable, so as not to disappear its language of origin.

Certain other issues will arise in working to provide multilingual support for this portal. For example, the portal's search tool will have to accommodate the occurrence of accent marks on letters in French. In this case, it would be important to ensure that one searching the word "ecole" ("school," in English) would yield results including the word, "école," whose first "e" includes an accent mark.

- *Faceted searching*

Faceted searching has become a standard feature of OPACs, digital libraries, and other online databases, not to mention popular e-commerce outlets like Etsy, Target, Amazon, and others. Facets (sometimes called "limiters") allow users to refine one's initial search query visibly. A common use of facets in libraries, for example, is when a user performs a simple keyword search (ie., "science"), and is directed to a large list of results that the user can then narrow by one or more categories that are applicable to the listed results (ie., "date," "theme," "subject," "media type," "creator/author"). Facets also often display a numerical indication of the frequency of a result within a certain category, and allow a user to drill down within their macro search results based on particular choices within these categories. In supplementing traditional boolean search strategies, facets make database research an activity that is much simple, more intuitive and, therefore, more accessible to audiences of diverse levels of familiarity with online

databases. Online commerce has popularized facets in ways that make manipulating them second-nature to online consumers and so, too, to many online library users.

The Franco American Portal Project team's discussions about facets have been less about whether to include them in the portal's design--which we unanimously agree upon--and more about how to include them: what categories they should contain, what they should look like, and how they should be structured.

One approach to the structure of facets would be to list all metadata categories as faceted categories, depending on the project's metadata profile. But some categories for facets that have been discussed specifically for inclusion are those that relate to the nature of Franco materials. Some of these are relatively standard: **item type**; **item language**; **date**. Because of the nature of our project as "portal," an essential category would also be **institution** or **repository,** in order for a user to narrow a search by the physical place where certain materials are held. Other categories that have been discussed and are also somewhat common, but may be adjusted to reflect the particular needs of our portal are: **location**, which for example might include not only LCSH city and state or province, but neighborhood or parish identifier, or historical Canadian region of origin; **occupation**, or a refined LCSH, which includes information about job types or popular forms of employment; **family name** or **surname**, which may be different from the name(s) reflected in other metadata fields for one object or another; **theme**; **historic event;** and others . Some of these latter categories would require the development of controlled vocabularies particular to Franco American materials and, arguably, particular to this project. It is the team's full intention to develop this vocabulary in order to improve and increase the number of access points to the materials within the portal, as well as to ameliorate one's search processes by standardized sets of common search vocabulary.

Our team has also discussed the structure of facets. Certain databases restrict facet hierarchy to a single level. This means, for example, that beneath the facet heading, "Subject," would be a list of subjects one could choose from. A multilevel facet hierarchy would allow for individual category items to have subheadings. This may be useful, for example, in thinking about items geographically: beneath a "Location" facet heading, a list of states might be expanded to include sublists of cities; cities might be expanded to include sublists of parishes or neighborhoods; and so on. We recognize the value in establishing a sort of hierarchical facet structure, however we have yet to determine if it would best suit our project and, if so, what form this might take.

It is also worth noting that certain softwares and software stacks include faceted searching natively; others do not. Omeka, for example, a popular open source software for archives and collections, and the software selection for our beta portal, does not include faceted searching within its basic install; Omeka requires applications or other backend work around facets in order for this feature to be included in public display. [See an example of an Omeka tool with facets here: http://discovery.civilwargovernors.org/ (Kentucky Historical Society, 2019).]

- ***Controlled vocabulary***
As mentioned above, one or more controlled vocabularies developed and/or implemented for our portal will help in the organization and discovery of portal contents.

At the very least, our portal project team is committed to using Library of Congress Subject Headings (LCSH) to describe materials featured in the portal. LCSH will assure not only that materials are described with detail, but that their descriptions adhere to certain nationally recognized standards. We are also committed to activating the French translation of all English-language LCSH headings for use in our portal's French-language interface.

We also aim to include other standard vocabulary in other metadata fields that occur within our Metadata Application Profile as adapted from DPLA and Dublin Core. For example, the "Item Type" field would be filled with one of several standard item types; the date field with common and recognizable numerical forms; etc.

We also recognize the simultaneous difficulty of describing Franco American materials completely and accurately by using only LCSH and other standard means of description. As we have noted elsewhere in the pursuit of a Franco American portal, the Library of Congress did not designate the subject heading "Franco-Americans" to apply to Americans of Acadian and French Canadian descent until 2008, and remains deficient in other categories for signalling idiosyncrasies. Certainly other means of describing our community are more accurate and representative of that community than LCSH terms, which are often meant to encompass broad categories. Nevertheless, the use of both LCSH and a controlled vocabulary will enable us to both participate within national standards while in our own way, supplementing those descriptive practices with a vocabulary that is unique to our project. This is a crucial aspect of the portal: making materials available by terms that are particular to those materials and to Franco American communities.

What might these controlled vocabularies include? As mentioned above, we have already acknowledged a handful of categories whose diversity and complexity will command our attention to developing sets of standard language--for keyword searchability, on the one hand, and for standard facet manipulation, on the other. These categories are **occupation, surname**, **theme**, and **historic event.** In each of these categories, our team will generate, grow, develop, and control a list of terms and phrases that enrich our portal users' search results, and that clarify these results' relevance to the search intentions of that user.

In terms of design, the appearance of controlled vocabulary in the portal would occur in facets, and so would depend very much on design decisions relating to the structure and appearance of those facets. Controlled vocabulary could also be the basis for exhibits, or other curated groupings of materials, which could occur adjacent to the main function of the portal and highlight segments of the portal's contents.

- *Display of multiple item types & item-level metadata*

The proposed portal will include links to a range of different item types: text, sound, image, moving image, finding aid. Because the content of the portal will be largely metadata and not digital objects or facsimiles of digital objects, the portal will need to determine if it will choose to accommodate records for different types of items in different ways aesthetically and, if so, how this might occur. At baseline the portal will include thumbnail images of each item to which it will link, regardless of item type. Items

without images will be represented by a stand-in image that indicates its type. Item-level metadata for each item will adhere to the metadata application profile developed by the portal team for digital items.

- ***Display of collection finding aid content & collection-level metadata***

The nature of the proposed portal also prioritizes the need to feature and display collection-level metadata in a way that makes searching for and viewing a collection both meaningful and user friendly. Certain portal models and digital repositories often take to simply including collection finding aids as an individual item type, and leaving a finding aid's textual content out of plain view. [See for example the Digital Transgender Archive's inclusion of the genre, "Finding Aid" (2019).] A keyword search does not search through the text of the finding aid unless that text is otherwise represented in the item record.) In this case, a finding aid and, therefore, a collection is described with "item metadata" : title, creator, date, etc.

Our project acknowledges that a large portion of materials to which the portal will link are a) not yet digitized, or b) not yet even described at the item level at their home institutions. These facts, in our view, should not be impediments to the online discovery of these items. It is therefore the intention of this portal project to create meaningful, unique structural presentations of collections so that their content is readily searchable and viewable. This presentation could include using "collection metadata" to describe collections: title, creator, scope and contents, inventory/box list, provenance, custodial history, etc. To put it differently, a collection's "finding aid" and could be represented in the portal. Placing a collection finding aid's contents and collection-level records in unique conversation with other item records and item-level metadata could enhance these collections' discoverability with some immediacy, without forcing our project to pause until one or more items have been digitized and described on the micro level.

Despite all of this descriptive work, there remains a question of how to display item records within the context of their collections. In other words, collection finding aids may include content lists, or inventories, where items contained therein have their own item records and, therefore, their own descriptive metadata. In cases where items have been described on the item-level, it may also be important to relate those items to their specific collections that have been described on the collection level. One area for growth and innovation within our portal project and within archival description more generally may be to explore how to display collections and collection-level metadata concurrently with their items and item-level metadata.

It is possible to represent one collection with its own record and metadata just as it is possible to represent one item with its own record and metadata. But, how might we best represent an item and a collection when an item occurs within a represented collection? A white paper entitled, "Aggregating and Representing Collections in the Digital Public Library of America," drafted by the DPLA Archival Description Working Group, attempted to answer this question (2016). In their paper the working group also provided design recommendations for the digital display and appearance of object-collection relationships, keeping in mind the rich professional differences between digital library descriptive practice for item-level objects, and traditional archival description that links together items within a collection according to provenance. Though the apparent graphical recommendations in the paper were not implemented to the DPLA public user interface (see figure below), and the working group itself "deemed

non-digital material out-of-scope for the report since it is outside of the collection guidelines for DPLA" (10), they are helpful for our project's purposes in at least one major way: they affirm the concern of representing the relationships between items and collections both descriptively and graphically has commanded time and resources of digital collections aggregator projects on a national level (13).
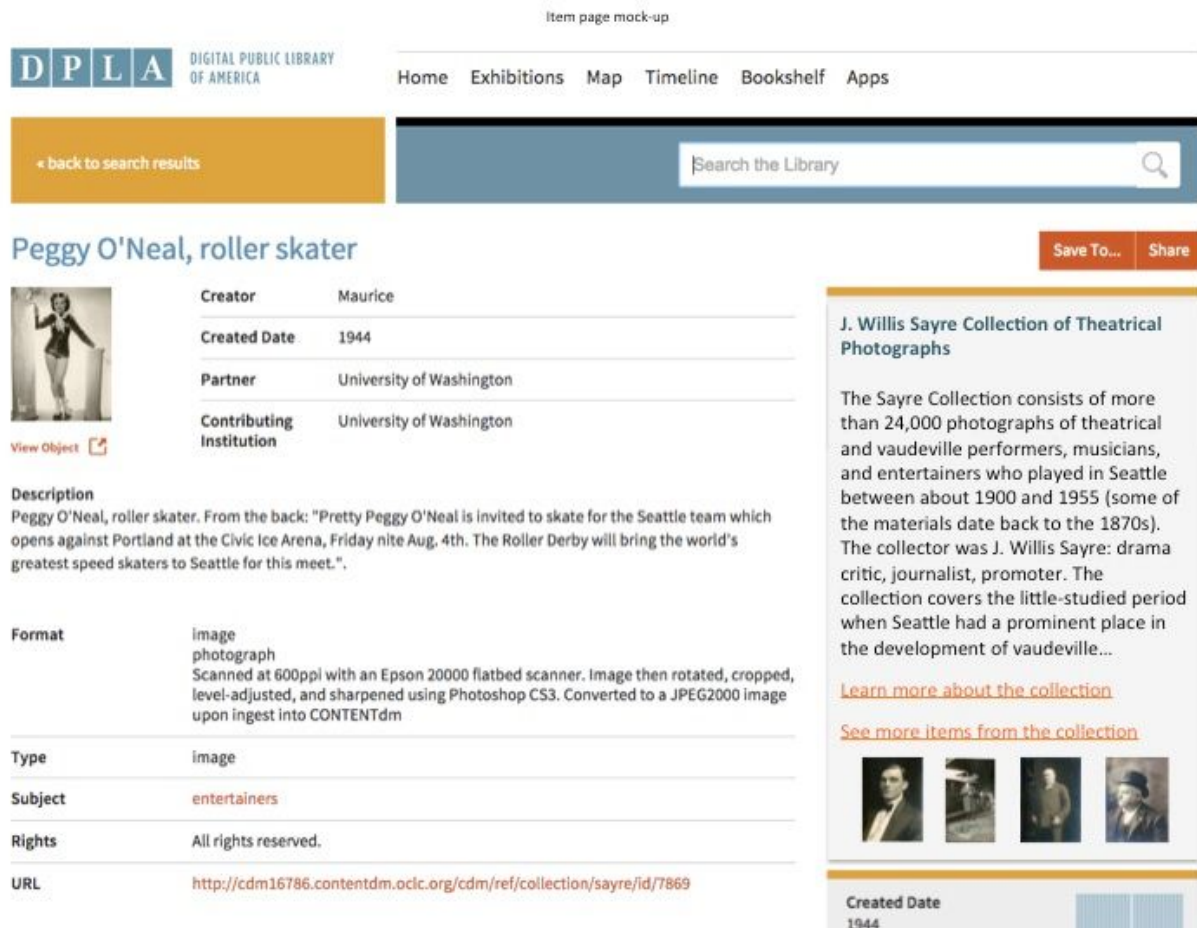
"Mock-up of a sample DPLA [item-level object] record showing Collection description," from DPLA Working Group, "Aggregating and Representing Collections in the Digital Public Library of America," p.29.

A noteworthy example of where such an object-collection relationship has been graphically displayed is the recently updated Digital Collections from the Northwestern University Libraries (https://dc.library.northwestern.edu/). Another example of this relationship clarified is apparent in Archives West (http://archiveswest.orbiscascade.org/), a production of the Orbis Cascade Alliance, where finding aids and digital objects are searchable through one interface together and displayed hierarchically on two levels, where applicable. (See an Archives West example of that particular relationship here, with digital object links couched within a collection: http://archiveswest.orbiscascade.org/ark:/80444/xv88922/op=fstyle.aspx?t=k&q=baseball&d=1 .) In their own way, certain iterations of ArchivesSpace show object-collection relationships as well. (For example, see this folder as an aggregated object within its collection at UMaine: https://archives.library.umaine.edu/repositories/3/archival_objects/238866 .)

- ***Capacity to harvest metadata from many institutions, discrete systems, & unique metadata profiles***

The primary function of the portal is to provide a single online space to link content that exists in other disparate spaces. This means that the bulk of metadata content that will live and be searchable within the portal will be harvested from other archives, repositories, and digital spaces. Currently, item- and collection-level metadata for our beta portal is harvested in one of two ways: either manually, or via OAI-PMH harvesting. In both cases, additional labor is required to check, format, and in some cases supplement another institution's metadata to fit within the profile of the portal.

Ideally, the back end of our portal would include a mechanism by which metadata could be harvested from digital repositories seamlessly. Though this characteristic and functionality would have little to no bearing on the public frontend of the portal website, it would indeed streamline the process by which metadata content is added to and becomes searchable within the portal.

- ***Support for automated metadata gathering & manual metadata supplement***

While it is possible to imagine that a mass harvest of metadata from participating institutions would populate the portal with content (via OAI-PMH or other mechanisms), we must also remember that this metadata is very much contingent upon the materials and labor provided by these participating institutions, including the specific curatorial and cataloging decisions that these institutions make in stewarding and describing their own collections.

Our portal project team must codify a process whereby metadata can be initially harvested from participating institutions in the shape and form which the institutions seek to represent their materials. There therefore must also exist a means whereby harvested metadata can be either translated into the portal's controlled vocabularies, or supplemented with vocabulary that is local to portal facets - or both - in order to maximize the discoverability of culturally specific materials within the portal.

At the same time, our portal must also address the need for subsequent harvests at these same institutions in order to account for new in-scope materials, metadata updates at local repositories, or other changes to digital content that an institution would want reflected in the portal space. In other words, our portal project team must establish a process in order to make sure that the contents of the portal are as up-to-date as the in-scope contents of the repositories to which the portal links. This process will likely involve establishing agreements (MOU) with contributing institutions that ensure open communication between the portal team and the contributing institution, in order that both parties can maximize the benefit that each provides to the public for discovery of Franco American materials.

- ***Other features***

Other functional features relevant to our project but which our team has not yet explored in full or determined are:
  - Full-text content search of linked textual items
  - Capacity to store digital objects alongside item records, for institutions that do not have digital repositories
  - Bookmarking/Backpacking, for users to save search results
  - Organization and description of clipfiles

Design Case Studies

Over the past five years, Franco American Programs at the University of Maine have worked with a variety of partners toward conceptualizing and building an organizing and discovery tool for Franco American collections. The shape of this concept has evolved over time with evolutions in our team and in our thinking, and with innovations in digital asset management technologies. The scope and breadth of this potential tool and its contents have similarly evolved as prospects for digital connectivity between disparate institutions have rapidly improved throughout North America. Particularly in the past decade, cultural heritage institutions have owed much to communities working to enhance open-source software tools - Islandora, Omeka, Samvera, and others - designed for the benefit of academic and public online research.

The concept of an online portal that links archival materials from disparate repositories is not a new concept. But over the past year, the Franco American Portal Project team has worked to focus its own notion of a Franco American primary source portal through the lens of recent other projects that have successfully developed tools for similar purposes. Expansive projects like Europeana (https://www.europeana.eu/portal/en), a major portal to European cultural heritage institutions and their materials, and the Digital Public Library of America (DPLA) (https://dp.la/), a similar project linking institutions and materials in the United States, have made familiar the possibilities of collaborative and collocated access to digitized primary source materials. Other, smaller projects resemble these two in meaningful ways. In fact, for four consecutive years, Chicago's Center for Research Libraries (CRL) has awarded its Primary Source Award for Access to four different, smaller portal projects that share in Europeana and DPLA's conceptual field: awards from 2016 to 2018 were given to new portals themselves, and the 2019 award was given to a digital newspaper project that relies upon a portal based in Texas (http://www.crl.edu/primary-source-awards).

Projects like these have shaped our team's general approach to thinking about digital asset management and archival discoverability in Franco American communities. We have evolved our concept of what a collaborative Franco American primary source discovery tool could and should be: from a localized bucket that keeps local sources in a single closed space, to an open, centralized web of disparate institutions whose various materials contribute to discussions in related subject areas wherever they are physically and digitally held. Indeed, our project's advisory partner institutions in some ways resemble these initial "localized buckets" of content; the portal project itself represents the open, centralized web that links our content together. In this way, our portal project is an effort in developing a network of heretofore unlinked primary sources, placing them in conversation with one another, and celebrating and drawing attention to the home institutions that do the essential work of stewarding these materials and making them publicly available.

We have seized the opportunity to study and reflect upon a few specific subject-based discovery tools, both for their functionality and for their appearance. Literature reviews and discussions with scholars or technologists involved in these and other projects have influenced the shape and trajectory of our own, and have given us a host of questions to weigh and test as we move forward in linking Franco American

collections online. Many of these questions appear answered earlier in this draft, in the form of decisive functional requirements; some other questions remain unanswered, or unasked. All of these questions have design implications.

The following subsections will briefly discuss two projects that have advanced our project team's thinking about our own project: ***Umbra Search African American History*** (University of Minnesota Libraries; funded by CLIR, IMLS, & Doris Duke Charitable Foundation) and ***Explore Chicago Collections*** (Chicago Collections Consortium; funded by Andrew W. Mellon Foundation and others). We will discuss unique features that each project brings to discussions about the design of our own project, and share resources for learning more about these projects and their development. We will also introduce two other similar projects with design illustrations only, and share resources for learning more about them: ***Digital Transgender Archive*** (College of the Holy Cross; funded by the American Council of Learned Societies and others) and ***College Women*** (Bryn Mawr College, Mt. Holyoke College; Wellesley College; Smith College; Barnard College; Vassar College; and Radcliffe Institute for Advanced Study at Harvard University; funded by the National Endowment for the Humanities).

Model #1: *Umbra Search African American History:* https://www.umbrasearch.org/

- Marcus, Cecily and Sarah Carlson. "Out of the Shadows: Bringing African American Digital Collections Together in *Umbra Search African American History*." *Open Library of Humanities*, *4*(2), 17. 2018. DOI: http://doi.org/10.16995/olh.279
- Marcus, Cecily, and Trevor Munoz. *Umbra Search and the Future of Black Digital Platforms*. Dec. 2018. *hcommons.org*, https://hcommons.org/deposits/item/hc:21919/.
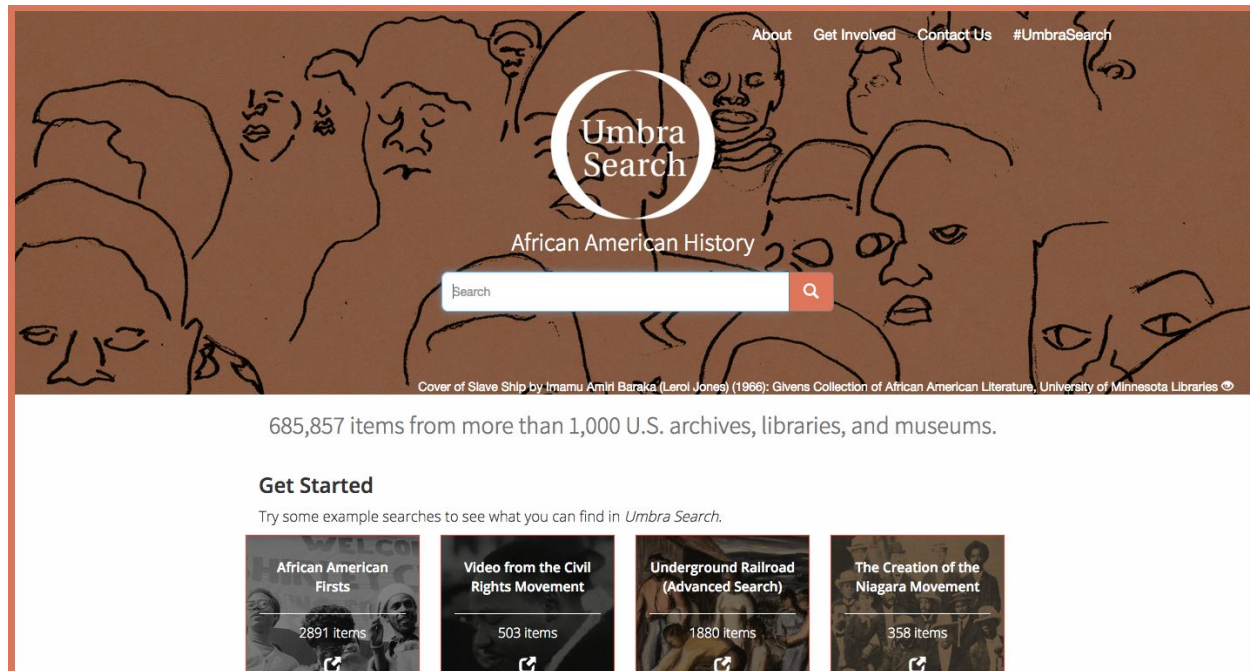
Front page, *Umbra Search African American History* (November 2019)

Umbra Search African American History was the first aggregator tool of its kind to come to the attention of our project team within the context of our own portal concept discussions. According to their website, "*Umbra Search African American History* makes African American history more broadly accessible through a freely available widget and search tool, umbrasearch.org; digitization of African American materials across University of Minnesota collections; and support of students, educators, artists, and the public through residencies, workshops, and events locally and around the country." Operating on dual principles of value and scarcity - as its creators remark, "American history is African American history," and much of African American history is absent from US schools - *Umbra Search* raises awareness of African American archives through the lens of their unique production, preservation, and exclusion narratives (Marcus and Carlson, 2018).

What excited us about the *Umbra Search* search tool was its ability to bring together a range of primary sources from a range of institutions into a single search. Though the project is rooted at the University of Minnesota and has incorporated digitization of African American materials across UM collections, *Umbra Search* keenly provides access points to hundreds of thousands of records across hundreds of institutions with digital African American archival content.

Umbra Search's front page prefaces this aggregator/single search function with a bright and centralized search bar. The project's introductory page also provides selected, curated entry points to portal content with links to thematic example searches. Much like the DPLA, many of whose harvested African American materials are included within *Umbra Search* and its search parameters, *Umbra Search* incorporates an easily navigable search interface powered by Blacklight with a select number of major facets: keyword, creator, item type, contributing institution, collection, and geographic location (see "Search Results Page" below). Search results include materials from all of *Umbra Search*'s contributing institutions or content hubs with relevant content. After selecting an item, users are then shown an abbreviated item record that includes a list of related keywords, several suggestions for related content, and a large orange button ("View") linking the user back to the digital object and a more complete metadata record at the item's contributing institution or service hub (see "Item View Page" below).
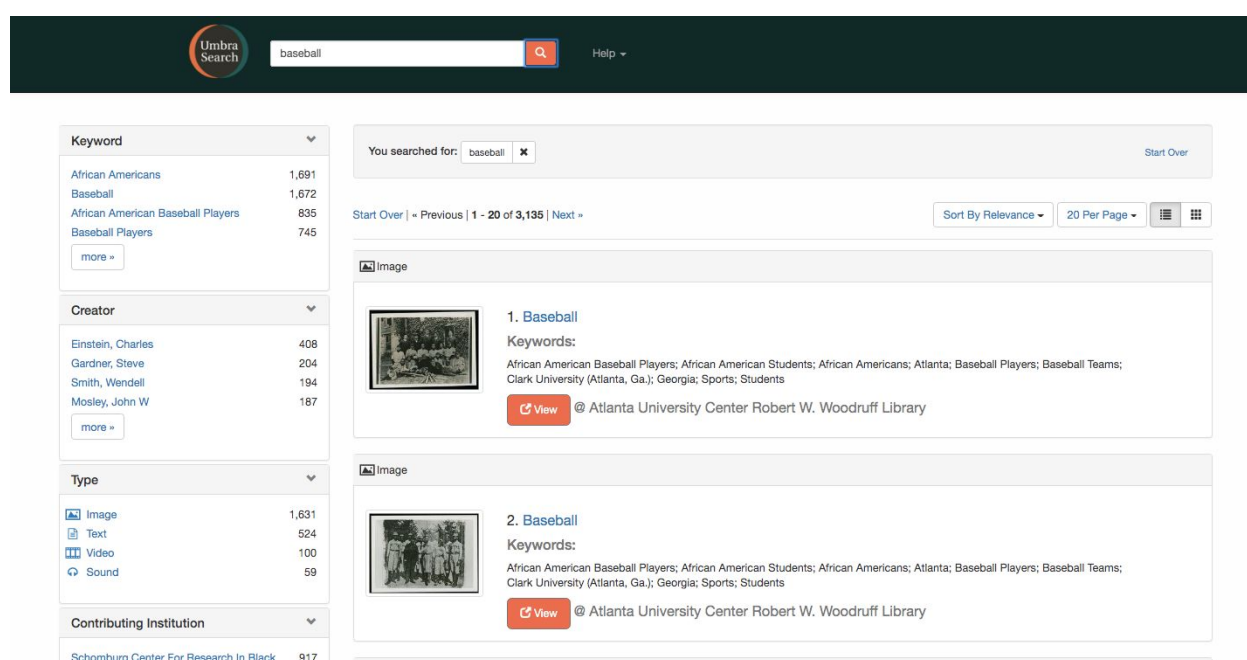
Search results page, "keyword: baseball," *Umbra Search African American History* **(accessed August 2019)**

It is also important to note that *Umbra Search* item records provide specific information about the originating record and its creating institution. These records appear on "item view" pages also with information about the collection that has provided broader access to this originating record, and the location from which *Umbra Search* harvested the item. (For example, a record of a photograph provided by a small institution but collected by DPLA and harvested from a DPLA hub may include links to that small institution, to the DPLA, and to the DPLA hub.) Social media icons also provide user opportunities for sharing *Umbra*-discovered content.
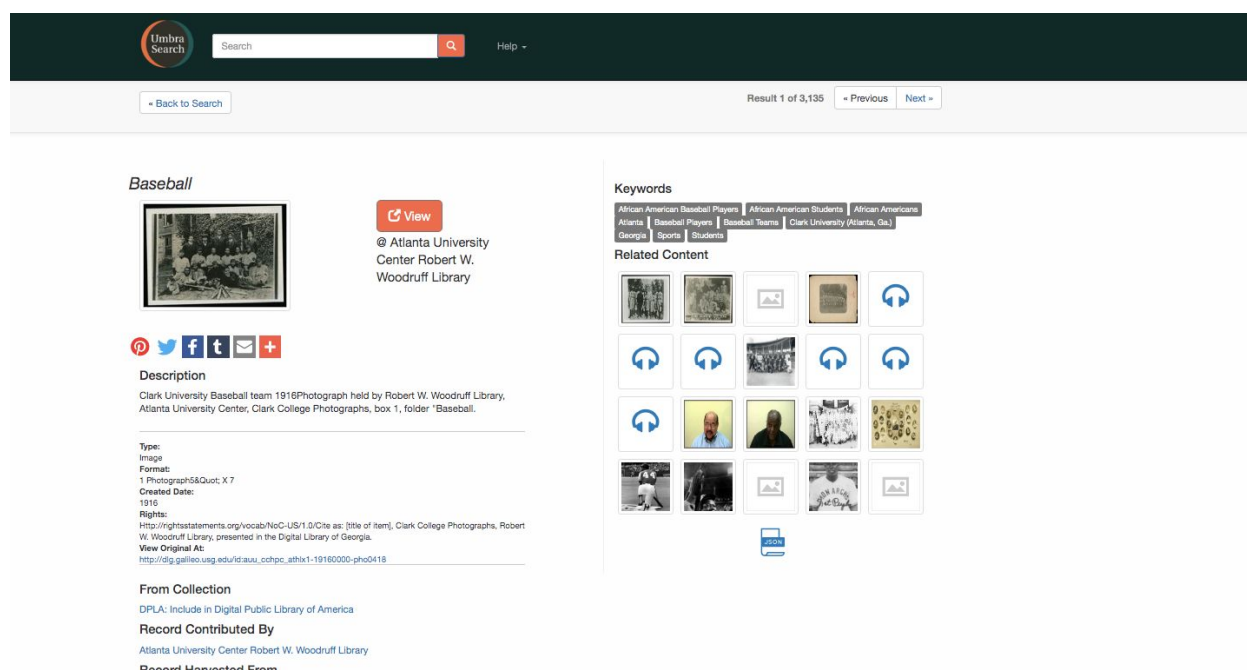
Item view page, "Baseball," *Umbra Search African American History* **(accessed August 2019)**

Umbra Search does not immediately provide a digital object's complete record or the digital item itself, but instead generates access to these things. In this way, *Umbra Search* functions as an intermediary: a user's search endpoint is not located within *Umbra Search*, but with the institution whose materials are contributed to *Umbra Search*. This fact has several meaningful implications:

- *Umbra Search* functions as an aggregator of metadata and a search portal for that metadata.
- Instead of "collecting" complete archival digital objects unto itself as a repository would, *Umbra Search* provides support for, and gives attention to, the contributing institutions and harvesting services that provide storage of and/or initial access to these digital objects. The digital object that *Umbra* curates and shares is a version facsimile of much smaller size and quality than the original.
- The burden of storing a complete archival digital object and its complete metadata record is not assumed by *Umbra Search* but by the digital item's contributing institution or service.
- *Umbra Search* harvests metadata from contributing institutions uniquely, depending upon the technology of those institutions.
- The item records within *Umbra Search* are also highly dependent upon the institutions from which it harvests, their descriptive choices, and the resultant metadata.
- *Umbra Search*'s harvests must occur periodically over time, in order to account for new in-scope records created by contributing institutions.
- *Umbra Search* must have harvest specifications that narrow its catchment area and retrieval pool according to a subject scope and scope of material types that *Umbra Search* determines.

The content pool for *Umbra Search* is currently limited to digital items only and, within those items, four item types: image (still image; photograph), text, video (moving image), and sound (audio). More specific

information about item formatting is included in the item's description field, or occurs in the item's full online record provided by the item's digital host.

Umbra Search's harvest strategies, as indicated on the project's main page, have also lent to a very large pool of materials: as of this document's writing, *Umbra Search* provides access to over 685,000 items. As with any mass harvest strategy, the possibility of yielding false positive search results becomes higher with the greater number of items discovered, but the total number of true positive materials also becomes much higher. Weeding out false positives may require manual labor in addition to the labor required to clean or supplement metadata appropriately for the shape and needs of the discovery tool itself.

Model #2: *Explore Chicago Collections*: https://explore.chicagocollections.org/

- Case, Mary M. "Meeting the Challenge of Collocating Metadata: Explore Chicago Collections." *Technicalities*. November/December 2017. http://hdl.handle.net/10027/22740
- Seneca, Tracy, et al. "Erasing Borders Through Digital Discovery: EXPLORE Chicago Collections as the Foundation for Collaboration Among Libraries, Archives, Museums (and Others)." *Bucknell University Digital Scholarship Conference*, Oct. 2016, http://budsc16.scholar.bucknell.edu/files/2016/11/Chicago-Collections-Erasing-Borders-through-Digital-Discovery.pdf
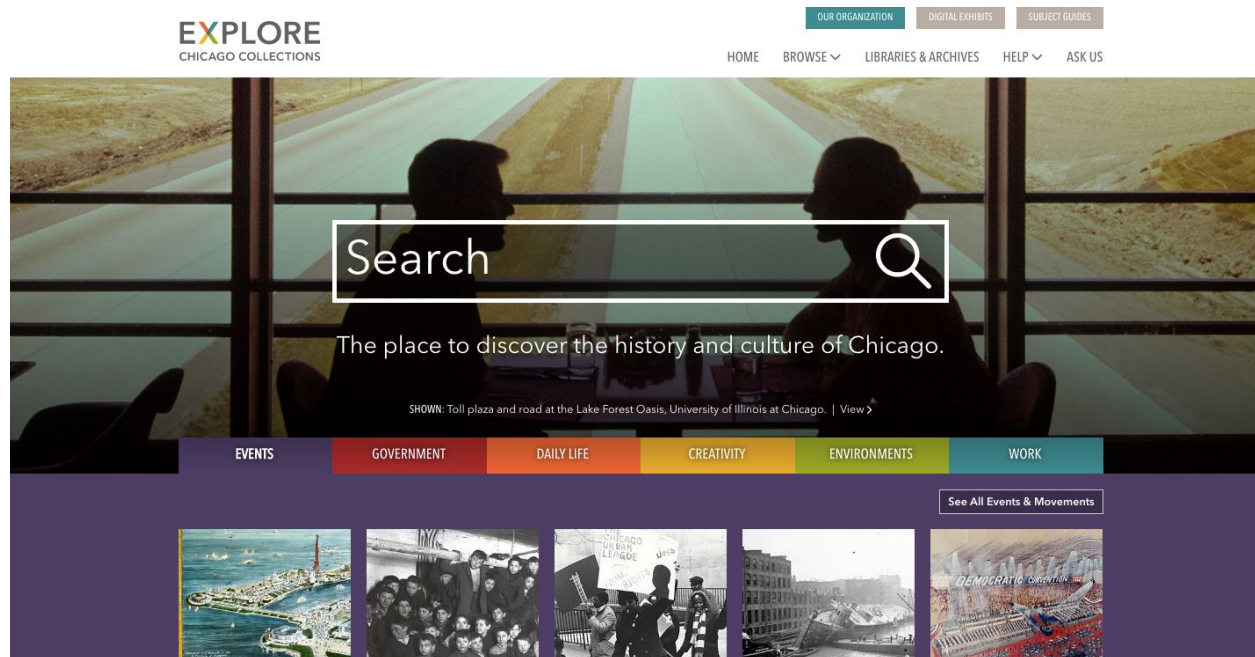
Front page, *Explore Chicago Collections* (accessed August 2019)

The Chicago Collections Consortium (CCC, https://chicagocollections.org/) is an organization made up of "non-profit organizations that maintain vital collections of books, letters, photos, or maps related to the Chicago region." CCC now comprises 15 governing members, 5 participating members, 26 members, and 1 partner (CCC, 2019). After four years of funded planning and technological development, in 2015 CCC launched its award-winning discovery tool, *Explore Chicago Collections,* to provide online access to member collections.

Much like *Umbra Search*, *Explore Chicago Collections* is a research portal that provides a gateway to archival materials at disparate institutions; in this way, the system's logic is a common one. *Explore*'s front page, like that of *Umbra* and other like projects, is basic and familiar: a large central search bar with navigational links in a header toolbar; below there are selected, curated entry points to content with links to thematic example searches. Bright square images not only invite users into the collections but suggest a handful of historical access points for linked materials: themes (Events, Government, Daily Life, Creativity, Environments, Work) with accessible subthemes (eg. 1933 World's Fair; Theater; City Planning); Neighborhoods; Cities; and Names.

Outside of its content scope--materials related to the Chicago region--and after its familiar-looking frontpage, what makes *Explore Chicago Collections* really unique are the ways in which it chooses to shape itself as a gateway to materials, to arrange them, and to represent them.

According to the project team, a primary objective for *Explore* from its inception was to provide a roadmap to where--precisely--certain items and collections physically existed in the greater Chicago area. According to the *Explore Chicago Collections'* website "About" page, "Until now, researchers had to know (or guess) in advance which libraries were likely to have relevant materials." In a 2017 article for *Technicalities,* Mary Case wrote of the origins of CCC and the *Explore* project with a similarly basic concern:

> ...the first major project of the CCC was to build a portal through which researchers, students, and the interested public could discover more easily what materials were in our members' Special Collections related to the history and culture of Chicago. We knew that not all finding aids were online and not all finding aids were consistent within an institution, let alone across institutions; many collections were unprocessed; researchers had to go from institution to institution just to discover which held relevant collections; and even special collections librarians in the city did not know what other institutions held. A quick and dirty spreadsheet developed early on of collections from several members indicated that collections had actually been split among several institutions by donors in the past without current staff realizing it.

To view the *Explore Chicago Collections* tool with this original concern in mind gives deeper meaning to some of the design and display choices within the tool itself.

Search results page, "keyword: baseball," *Explore Chicago Collections* **(accessed August 2019)**

The *Explore* "search results" page structure resembles *Umbra*'s in several ways: a "sort by" dropdown appears to organize search results; several facets on the left side of the screen help to narrow search results by one or more fields; search result items, many with selected metadata, are arranged in the main portion of the screen. This is a familiar design. But what makes *Explore* unique?

The contents of the search results page are very specific to the *Explore* project. From the outset we see that *Explore*'s search results are limited to two distinct digital object types: **digital images** (eg. photographs) and **archival collections** (eg. collection finding aids). These types also occupy their own uniquely divided spaces on the search results page: the first five returned digital images appear in a narrow section at the top of the search results, organized left-to-right in thumbnails, without text, and with an accompanying link to a page that features all returned images; the titles of resultant archival collections appear in their own section, below the images, in a top-to-bottom list form with a brief description, related topics, and an identified contributing institution.

Explore's actual results are therefore at once narrower in type than *Umbra*'s selected four types, yet somewhat broader than *Umbra*'s results in the way that they allow for an archival collection's full finding aid text to communicate the contents of an entire archival collection, and for the complete text of that finding aid to be searchable in the context of the portal. This fact has several meaningful implications:

- *Explore* has the capacity to support both digitized and non-digitized content.
- *Explore* does not require that content be digitized in order for it to be searchable.
- *Explore* informs users of the physical location of both digitized and non-digitized content, but does not give users immediate access to non-digitized content--only to their description in finding aid form.
- In these ways, *Explore* places emphasis on both the **value of digital content** in providing immediate access to it, and the **value of institutions** in directing users back to the physical repositories that hold archival collections described in *Explore*'s searchable finding aids.

Explore's facets are also unique. The "Topics" limiter is populated with "six top-level" topics and 88 sub-level themes--all terms determined by the project's Controlled Vocabulary Task Force (Case, 2017). The "Neighborhoods" limiter is just as unique and significant to material discovery: over 180 neighborhoods in and around Chicago are tagged to *Explore*'s archival collections.

From the search results page, *Explore*'s subsequent object pages appear uniquely according to item type: **digital image** and **archival collection.**

A digital image object page expands the selected thumbnail image into a clearer picture, roughly 500 square pixels in area, with an accompanying rich metadata record that includes the following fields: Collection ID; Creator Names; Title; Dates; Physical Description; Access and Use Restrictions; Collection Summary; Terms Governing Use and Reproduction; Names; Subjects; Geographic Coverage; Genre; Geographic Name; Finding Aid URL. The top of the page features the item's title, a list of major topics and themes assigned to the image by CCC, and a graphic linking the user to the member/contributor page on the CCC site where all of that member's contributed content is featured.

Item view page, "[Baseball player Honus Wagner...]," *Explore Chicago Collections* (accessed August 2019)

An archival collection object page is, in some cases, very similar to the digital image page; in other cases it is very different. For an abbreviated archival collection page, the page resembles the digital image page without the thumbnail; metadata fields may be fewer than those for digital images, but the terms are the same. Most importantly here, an archival collection record points a user to the contributing institution where the physical collection can be explored in its entirety.

Collection view page, abbrev., "Cook County...Records," *Explore Chicago Collections* (accessed August 2019)

For pages with a fuller collection description, the archival collection object page is also much fuller. A lefthand sidebar provides the collection page its own navigational table for a richly detailed, seemingly DACS-compliant finding aid with the following sections: Descriptive Summary; Restrictions; Terms governing use; Provenance statement; "Please cite this collection as"; Collection Summary; Biographical/Historical note; Related material; Control Access; Organization and Arrangement of Collection; and Inventory. As in the abbreviated collection page, this finding aid's link back to to host of the physical collection is a simple but significant feature in the effort to provide online access points to archival materials that are not yet digitized.

Collection view page, complete, "Chicago White Sox and 1919 World Series baseball scandal collection, 1917-1929, bulk 1920-1924," *Explore Chicago Collections* (accessed September 2019)

Recognizing that their hybrid digital/physical approach to communicate archives access might demand of its users a particular skillset, *Explore* has also created a "Help Guide" to introduce novice users to its own portal and the portal's contents. Not only this, but *Explore* also demystifies the physical archive for new researchers by sharing the basic steps of visiting and exploring a repository in-person. With a highly readable and basic guide, the *Explore* project encourages and fosters its use, mitigates the need for trained CCC staff to respond to user questions, and gets out in front of a number of potential user issues before they occur.

Help Guide, "Overview," *Explore Chicago Collections* **(accessed September 2019)**

Model #3: *Digital Transgender Archive*: https://www.digitaltransgenderarchive.net/

- Rawson, K. J. "Rhetorical History 2.0: Toward a Digital Transgender Archive." Transcript. 2013. http://www.enculturation.net/system/files/RawsonRhetHistScript.pdf
- "Overview: History & Purpose." *Digital Transgender Archive.* Web. Accessed September 2019. https://www.digitaltransgenderarchive.net/about/overview
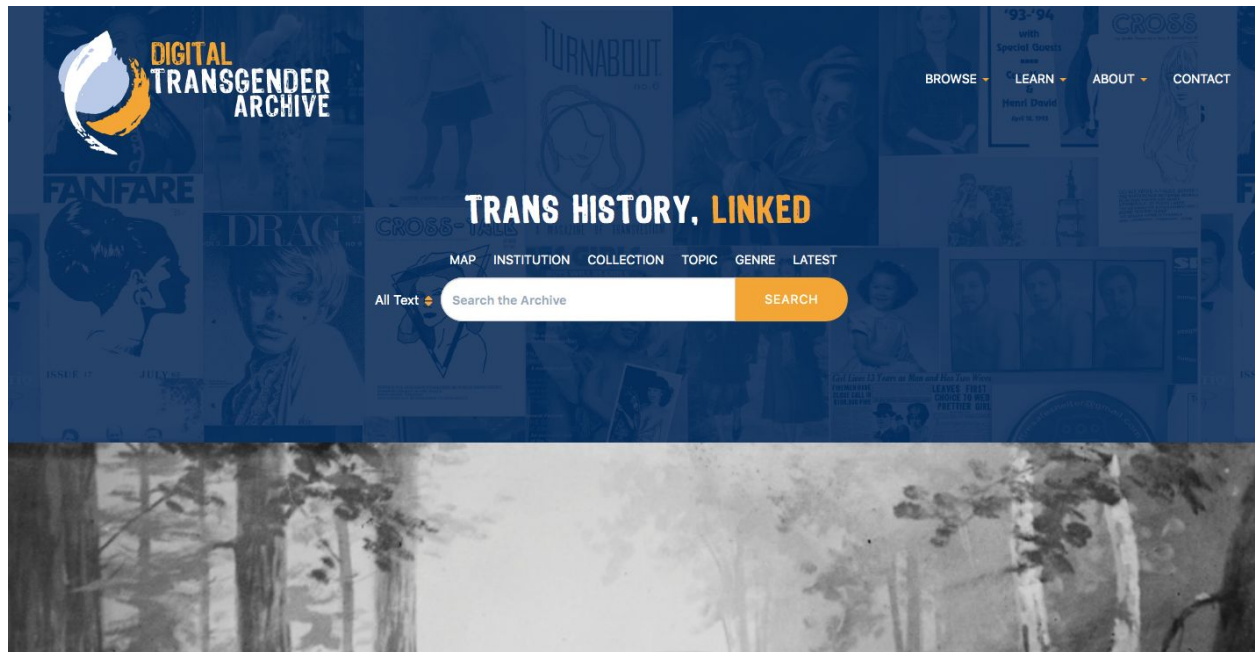
Front page, *Digital Transgender Archive* (accessed August 2019)
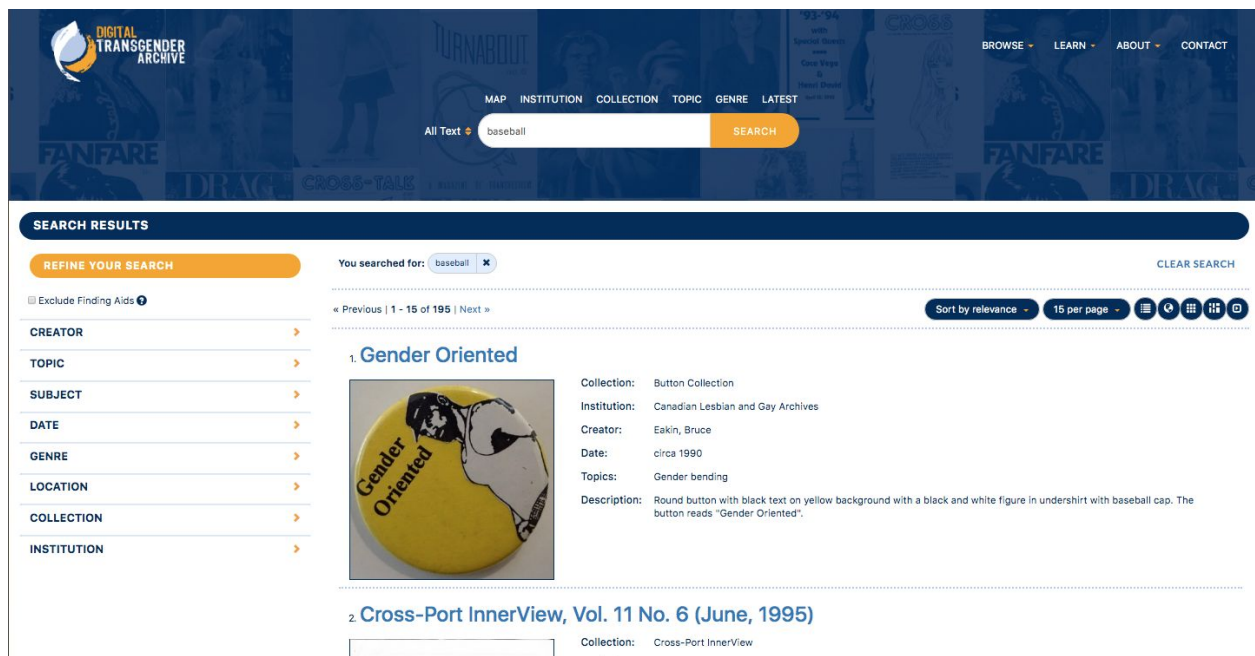
Search results page, "keyword: baseball," *Digital Transgender Archive* (accessed August 2019)
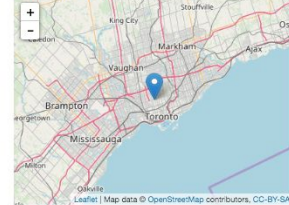
Item view page, "Gender Oriented," *Digital Transgender Archive* **(accessed August 2019)**

Item view page, "A Guide to the Linda and Cynthia Phillips Papers, 1964-2000" [Finding Aid],
Digital Transgender Archive **(accessed August 2019)**

Model #4: *College Women***:** https://www.collegewomen.org
- Pumroy, Eric, et al. "History of Women's Education Open Access Portal Project." *Library Staff Research and Scholarship*, Jan. 2015, https://repository.brynmawr.edu/lib_pubs/17.
- McGonagill, Evan, et al. "College Women: Documenting the Student Experience at the Seven Sisters Colleges." *Library Staff Research and Scholarship*, Jan. 2019, https://repository.brynmawr.edu/lib_pubs/21.
- Redmond, Jennifer et al. "History of Women's Education Open Access Portal Project: Narrative," Grant #11452367, *National Endowment for the Humanities Division of Preservation and Access: Humanities Collections and Reference Resources Grant Program*, July 2013, https://www.neh.gov/sites/default/files/inline-files/bryn_mawr_college_history_of_womens_education_open_access_portal_project.pdf .

Front page, *College Women* **(accessed August 2019)**

Search results page, "keyword: boston," *College Women* (accessed August 2019)

**Item view page, "Agora," *College Women* (accessed August 2019)**

Models #1-#4 Graphic Overview and Comparison

Many of the basic features of the four portal models are similar. Looking at them side-by-side makes this plain in terms of search display, search results display, and item display. These similarities suggest design trends in digital archival display that might benefit our project in a few distinct ways:

- **Recognition** - audiences who confront an interface they have seen before are likely to be less confused and feel more familiar with it
- **Ease** - these designs are striking in their simplicity and their ease of entry: a keyword on the frontpage, or one of many links to several suggested entry points
- **Longevity** - simple, widely used display arrangements tap into a general flow of contemporary information resources that exist across a range of fields. These arrangements seem to be poised for stability and longevity in their general acceptance and use, or at least in step with trends that capitalize on community web developments based on existing technologies.

See the following similarities in the four shared models:

Front pages:
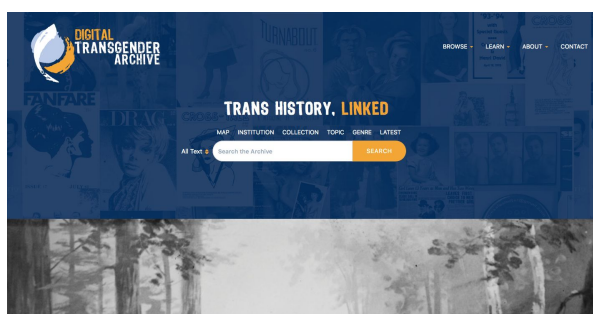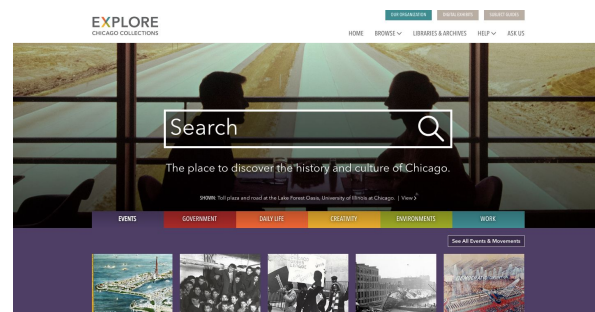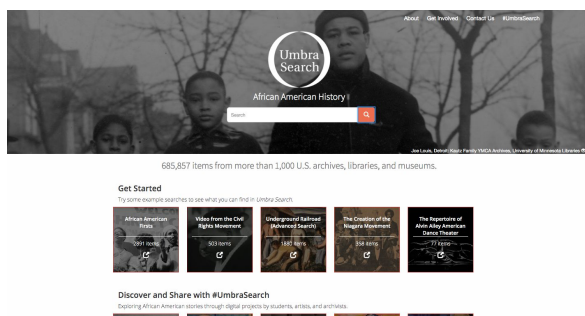
All four models present their title logo and a menu at the top of the screen. All four include the menu in the top right corner, and three of four present their project name logo at the top left. Each project contains specific, relevant imagery from the project in its background, while locating a large keyword search bar at the center of the page, adjacent to a project tagline. Most colors are subdued and not terribly bright, with a few exceptions in highlighted areas. Scrolling down each project's front page, users will find guided entries into each project's content: some with graphical representation, some with keyword, some with both.

Search results pages:
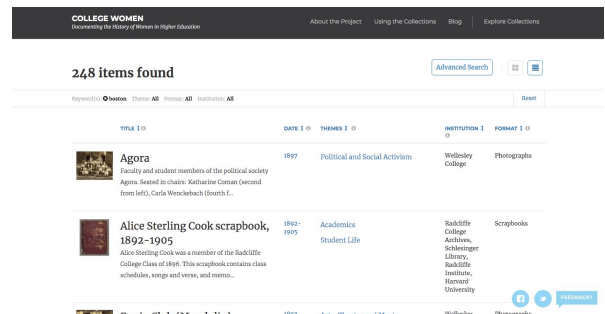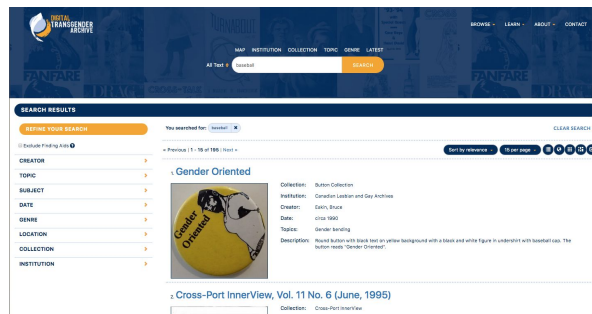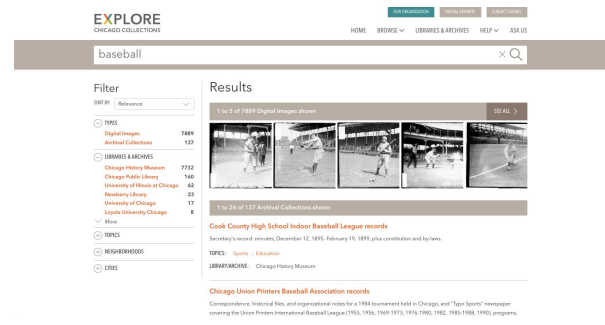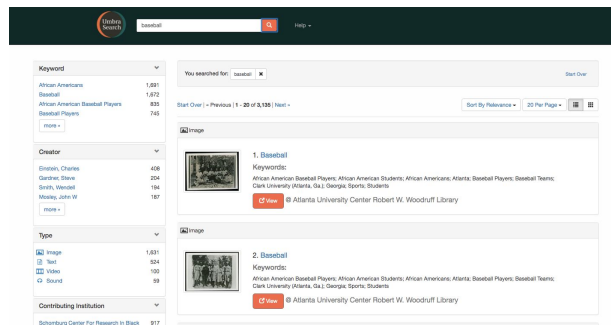
Search results pages from each of the four models are also strikingly similar. Three of the four offer faceted searching with facets organized at the left side of the screen, approximately 25% of the screen width, and search results presented to the right, occupying the remaining 75% of the screen. The exception here is *College Women*, whose search results are presented either in a sortable table (seen above, bottom right), or in an icon grid, depending on the user's preference. In all cases where the result is an item, search results are accompanied by a thumbnail image. In the case of *Explore Chicago Collections'* collection results, there are no thumbnail images, but there are titles and brief descriptions of collection contents. All four projects maintain their logo and menu at the top of the screen; three of four also maintain a functional search bar.

Major choices that distinguish each of the above search results page from the other are present in the metadata that each project chooses to foreground, to display alongside each item or collection as a preface to the item or collection's full record. Some include simply title and description; others include date, creator; others include the item's home institution. One--*Explore Chicago Collections*--simply presents an item thumbnail image for each item at the outset. (In order to see more item results and item titles, the user must select the "See All >" button.) These choices may seem basic, but they are also indicative of the major concerns and intentions of each project, and reflect a series of needs that each project identifies in its major audiences. The presentation of more or less information, the location of that information, and the manner in which that information is displayed are all reflections of project priorities, or project limitations, and reveal something about each project team's general aesthetic. How well that aesthetic matches the content of the project is up to the user to determine.

Item view pages:
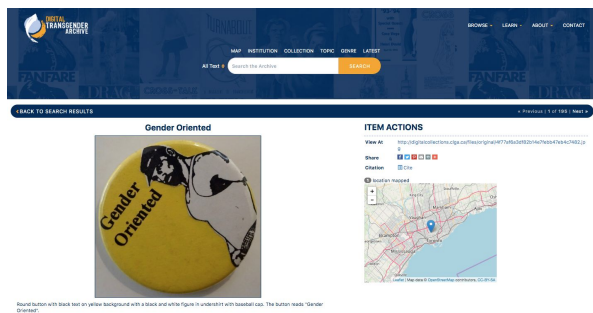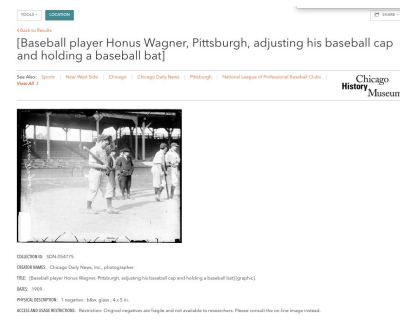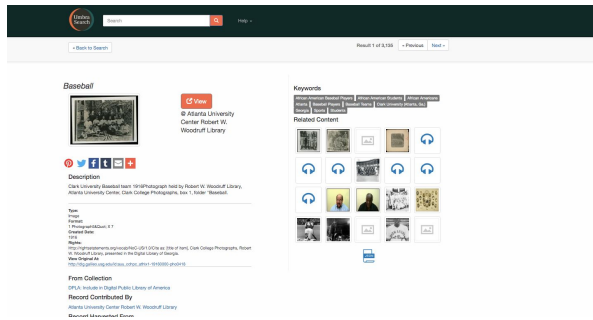
Item view pages resemble one another, too. The similarities are obvious: persistent project header, item thumbnail, item metadata profile. Most notably, each project highlights an item's host institution.

The differences, once again, are in what each project chooses to foreground. Half the models foreground social media item sharing. One foregrounds geography; one, related items; one, an item's metadata profile; and another, the item's related themes. In each item view case, none of the project's choices register unfamiliar to the user. This page in each project has been cut from a similar cloth. It is not too much to say that the Franco American Portal Project follows this trend.

Franco American Portal Project: Beta

In the proposal for our own project, a benchmark charge for project year two (calendar year 2020) was to develop a prototype portal for public consumption. This prototype would form not only a proof of our concept at the year-two stage of our collaboration, but would also be a public opportunity for us to continue to develop this concept at low cost with the input of interested users.

Over the course of project year one, however, it became apparent that our project's progress would benefit greatly from the early construction of a prototype tool, a "beta" portal, and from the experimentation that such a model would afford us. This beta site would allow for us to test out our initial concepts before moving forward with an investment in them. The site would also allow for us to publicly advertise the state of our work, for team members and the public to explore the technological position of our work relative to our projects needs and goals, and to increase web traffic through our own project URL: https://francoamericanportal.org.
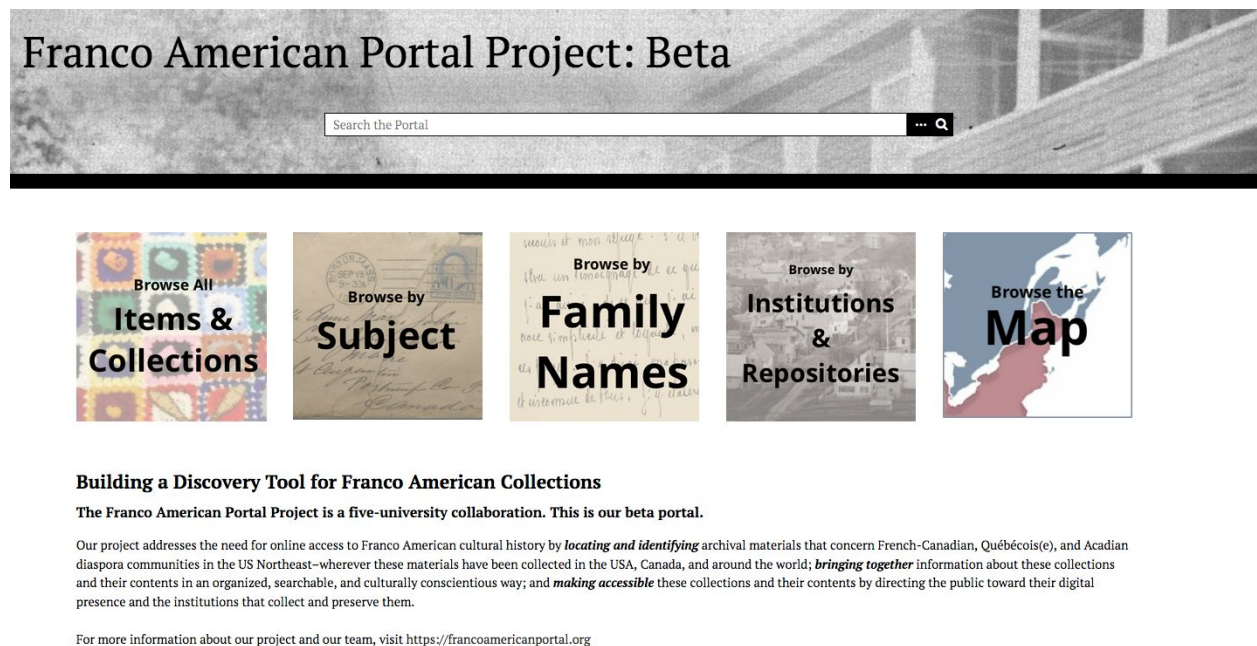
Front page, *Franco American Portal Project: Beta* **(November 2019)**

A subcommittee of our portal team conducted a brief software analysis that looked closely at five different software types for constructing a prototype portal. The subcommittee compared these softwares on a range of criteria: cost, ease, functionality, proprietary/open source, a handful of specific technical measures, and the appearance and success of certain comparable use cases. After review and discussion, for a variety of reasons, the subcommittee chose Omeka for building our beta portal. The team then secured access to third-party hosting, storage, and support services for building this initial tool, and for maintaining it over the course of its customization and development. Because our project is multi-institutional, the gesture of selecting a third-party host and providing each institution access to that host

was a nod to the truly collaborative, decentralized nature of the project, and was an effort in making sure that one institution's web services were not prioritized over any other. Third-party hosting also allowed our project the freedom to move and grow in its early stages with unmediated access to front- and back-end design management.

The initial Omeka install for Franco American Portal Project: Beta (https://beta.francoamericanportal.org) went public in May 2019.
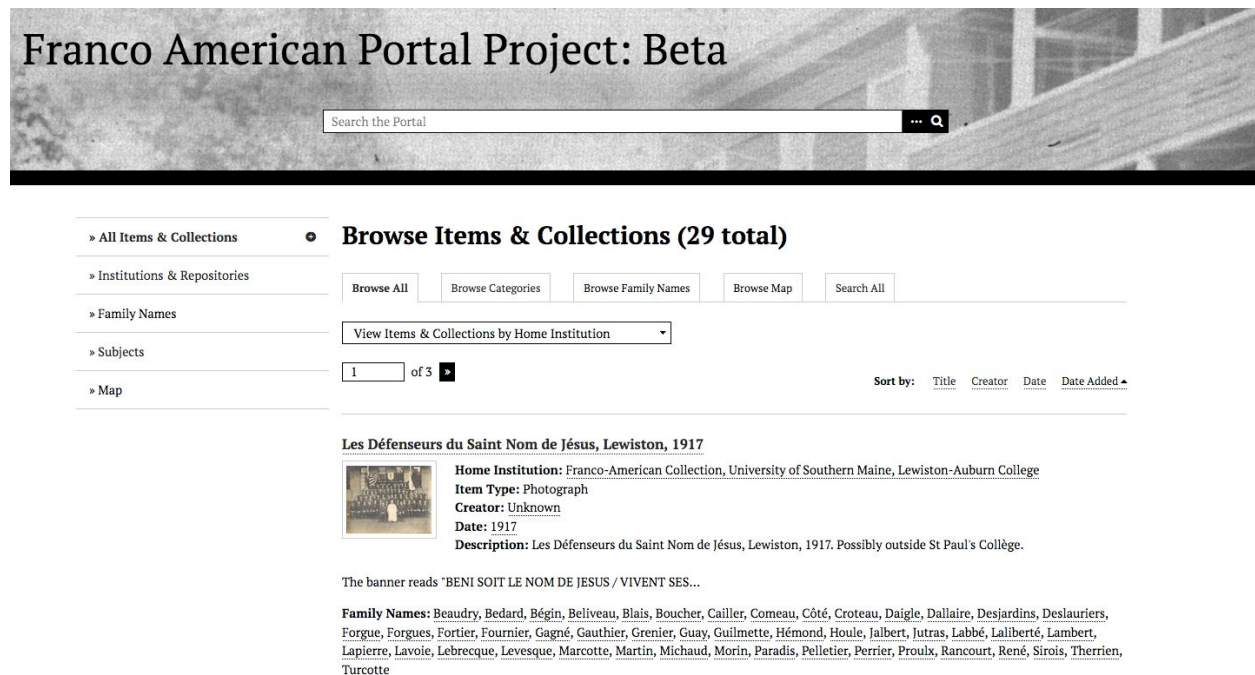
Items & Collections Browse page, *Franco American Portal Project: Beta* **(November 2019)**

From May 2019 to the present, our team has added archival content, customized web aesthetics, tailored native Omeka functionality to come closest to the project needs described above in our functional requirements, and adopted Omeka plugins to bring us nearer to those stated functional goals. Indeed, as illustrated in the beta portal screenshots featured in this section, this beta portal falls short of many of our requirements:

- At present, there is no multilingual interface. Multilanguage plugin options for Omeka will require extra resources to adapt to the current descriptive needs of the project
- At present, there is no faceted searching, hierarchical or otherwise. The beta portal's left-hand toolbar does indeed resemble the placement of facets, however these search options function quite differently than facets. They do not limit search results; instead, they perform an entirely new search of the portal's entire contents. Sort options at the top right enable users to sort search results to alphabetize or prioritize results, however users are unable to create search limits based on one category or another.

- At present, there only exists one item- or collection-level display. In other words, items and collections are displayed in the same way. Ideally, items and collections would be displayed each in their own way, with their own metadata profile and public elements.
- Similarly, there are no articulated part/whole relationships. In other words, there is currently no structure in use to represent the relationship between an item in a collection and the collection itself.
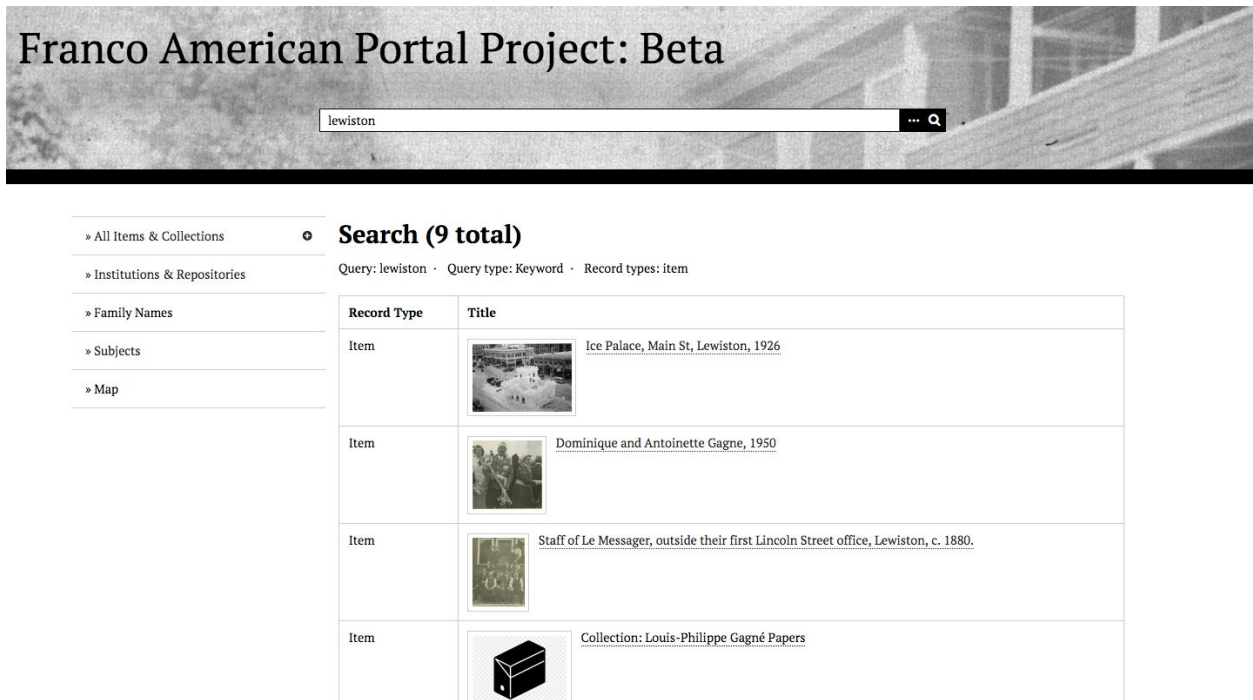
Search results page, "keyword: lewiston," *Franco American Portal Project: Beta* **(November 2019)**

Nevertheless, the beta portal represents a basic proof of concept, and begins to place the materials of each of our partner institutions in conversation with one another, directing researchers back out to wherever materials are held. Several functional and design elements begin to address the needs of Franco materials and the needs of our potential audiences, as determined by our project team:

- Each item and collection record is accompanied by hyperlinks to the host repository and to where the item or collection is featured at the host repository.
- As our team is especially interested in culturally conscientious points of entry into the archival materials featured in the portal, our beta portal enables users to access materials by family name, by geographic location, by the institution where the materials are held, and by Library of Congress subject heading.
- Beta portal content occurs in both English and French.
- Beta portal content can be limited for browsing by media type.

Franco American Portal Project: Beta

Search the Portal

» All Items & Collections
» Institutions & Repositories
» Family Names
» Subjects
» Map

Ice Palace, Main St, Lewiston, 1926

Home Institution
Franco-American Collection, University of Southern Maine, Lewiston-Auburn College

View Item at Home Institution
https://franco.pastperfectonline.com/photo/0636BD69-4B09-444B-A83B-824886791639

Title
Ice Palace, Main St, Lewiston, 1926

Subject
Franco-Americans

Franco-Americans--Lewiston (Me.)

Ice palaces

Item view page, "Ice Palace, Main St, Lewiston, 1926" *Franco American Portal Project: Beta* **(November 2019)**

In November 2019 our team began distributing an online survey to trusted members of our community in order to begin generating feedback on our concept and its first, basic iteration in the beta portal. As of the writing of this white paper, our team continues to gather feedback. We will reconvene with collected feedback data at the beginning of 2020.

Summary & Conclusions

The Franco American Portal Project began with the articulation of a shared need: an online space to help community and academic researchers find Franco American primary sources. Over the course of one year, with open communication, deep study, conversations with scholars and technologists, and reflection by our project team, our need transformed: a post-custodial, collocated portal space that places many institutions' collections in conversation, but also directs users to catalogs, repositories, and other websites that either house or describe Franco collections. Our project's mission transformed from simply highlighting Franco American primary sources also now to generating traffic and interest in the institutions that steward and preserve them.

The models that serve as conceptual and practical inspiration for our project have also, as described above, influenced the way that our project team thinks about our portal's design. In other words, we have created a beta portal with a feedback loop because we understand that web design is powerful when it is iterative, and particularly when it is responsive to the needs of its diverse audiences.

With an end product in mind, centered around our stated functional requirements, our intention over project years two and three will be to build a robust, complex, functional tool, and always with an eye toward how, over time, we may be able to improve its use and functionality.

One aspect of design that was assumed but not explicitly addressed in this white paper is aesthetics: How should our portal look and feel? What color scheme should it use? What fonts? Should we develop a logo? Should the portal itself have a new, catchy name? How will these choices about the portal's appearance impact its use?

Admittedly, up to now, we have opted to prioritize function over fashion. We have accepted aesthetic elements based on the ease of implementing the characteristics of our project that satisfy certain functional needs. In other words, how our project looks is currently, and entirely, an effect of how it functions. Many design choices over the next two years therefore will revolve around precisely how we want our portal to look and to feel, within its capacity to perform in ways that our team specifies. We look forward to this design work and, at once, recognize that it will be ongoing through the life of the project.

References

Case, Mary M. "Meeting the Challenge of Collocating Metadata: Explore Chicago Collections."
    Technicalities. November/December 2017. http://hdl.handle.net/10027/22740 . Accessed
    November 2019. Web.

Center for Persons with Disabilities, Utah State University. *WebAIM: Section 508 Checklist*.
    https://webaim.org/standards/508/checklist. Accessed November 2019. Web.

Center for Research Libraries. "Past Awardees | CRL." *Center for Research Libraries*.
    http://www.crl.edu/primary-source-awards/past-awardees. Accessed November 2019. Web.

Chicago Collections Consortium (CCC). [Various pages].*Chicago Collections.*
    https://chicagocollections.org . Accessed September 2019. Web.

College Women. [Various pages]. *College Women*, https://www.collegewomen.org/. Accessed November
    2019. Web.

Digital Public Library of America. [Various pages]. *Digital Public Library of*
    America, https://dp.la/. Accessed November 2019. Web.

Digital Transgender Archive. [Various pages]. *Digital Transgender Archive.*
    https://www.digitaltransgenderarchive.net . Accessed November 2019. Web.

Europeana. [Various pages]. *Europeana Collections*. https://www.europeana.eu/portal/?locale=en.
    Accessed November 2019. Web.

Franco American Portal Project. [Various pages]. *Franco American Portal Project: Beta.*
    https://beta.francoamericanportal.org . Accessed November 2019. Web.

Kentucky Historical Society et al. *Civil War Governors of Kentucky*.
    http://discovery.civilwargovernors.org/. Accessed November 2019. Web.

Marcus, Cecily and Sarah Carlson. "Out of the Shadows: Bringing African American
    Digital Collections Together in *Umbra Search African American History*." *Open Library of*
    Humanities, *4*(2), 17. 2018. DOI: http://doi.org/10.16995/olh.279 . Accessed November 2019.
    Web.

Marcus, Cecily, and Trevor Munoz. *Umbra Search and the Future of Black Digital*
    Platforms. Dec. 2018. *hcommons.org*, https://hcommons.org/deposits/item/hc:21919/ . Accessed
    November 2019. Web.

McGonagill, Evan, et al. "College Women: Documenting the Student Experience at the Seven Sisters
    Colleges." *Library Staff Research and Scholarship*, Jan. 2019,
    https://repository.brynmawr.edu/lib_pubs/21. Accessed November 2019. Web.

Northwestern University. *Digital Collections - Libraries - Northwestern University*. https://dc.library.northwestern.edu/. Accessed November 2019. Web.

Orbis Cascade Alliance. [Various pages]. *Archives West*. http://archiveswest.orbiscascade.org/. Accessed November 2019. Web.

Pumroy, Eric, et al. "History of Women's Education Open Access Portal Project." *Library Staff Research and Scholarship*, Jan. 2015, https://repository.brynmawr.edu/lib_pubs/17. Accessed November 2019. Web.

Redmond, Jennifer et al. "History of Women's Education Open Access Portal Project: Narrative," Grant #11452367, *National Endowment for the Humanities Division of Preservation and Access: Humanities Collections and Reference Resources Grant Program*, July 2013, https://www.neh.gov/sites/default/files/inline-files/bryn_mawr_college_history_of_womens_education_open_access_portal_project.pdf . Accessed November 2019. Web.

Seneca, Tracy, et al. "Erasing Borders Through Digital Discovery: EXPLORE Chicago Collections as the Foundation for Collaboration Among Libraries, Archives, Museums (and Others)." *Bucknell University Digital Scholarship Conference*, Oct. 2016, http://budsc16.scholar.bucknell.edu/files/2016/11/Chicago-Collections-Erasing-Borders-through-Digital-Discovery.pdf Accessed November 2019. Web.

Umbra Search African American History. [Various pages]. *Umbra Search African American History*, https://www.umbrasearch.org/. Accessed November 2019. Web.

University of Maine. "Franco American Reading Room | Fogler Library Special Collections and Archives." *ArchivesSpace*. https://archives.library.umaine.edu/repositories/3/archival_objects/238866. Accessed November 2019. Web.