Blue Mountain UI Spec

Process and Purpose

The purpose of this UI Spec are to better understand the goals of Blue Mountain's key audience and define interfaces that allow these users to reach these goals through a RESTful interface which offers renderings of data for both humans and machine consumption. Our process began with ten individual User Interviews conducted with researchers and librarians from a variety of academic backgrounds. The interviews helped us determine that the current Blue Mountain interface (Veridian) falls short in a number of areas, particularly in terms of finding "contextual patterns and relationships" or being able to answer certain research questions, many of which will be possible with a Linked Data approach. The current UI deliverables are this specification document, a set of RESTful URI templates, and a series of HTML-based, "shovel-ready" wireframes that provide a starting point for the next-generation Blue Mountain interface.

Scope

Our intent is to plant the seeds that will "let a thousand flowers bloom", so to speak. The interface is not only designed for end users, but also for developers and machines to extend the functionality and usability of the content. While there is great potential for the development of visualizations and community tools, the first step is to expose the data in standardized ways for others (and ourselves) to build upon and integrate with current and emerging research tools.

Use Cases

Interviews with ten Blue Mountain researchers revealed a number of insights about the role Blue Mountain can fill in art, periodical, and general humanities research and pedagogy. Many requested features [see appendix] can be delegated, for the time being, to existing applications. For example, the "creation of ad hoc image collections" can be handled by Pinterest, Mendelay, and to some extent, Zotero. Furthermore, "image comparison features" can be handled by Mirador. It came as a surprise that the features interviewees were most excited about had less to do with our initial design focus, a custom "METS-ALTO Viewer," and more to do with enhanced search functionality that is possible with a Linked Data infrastructure.

High-priority Features
- Full-text Search
- Entity Relationships discovery and cross-reference (perhaps using Stanford's Named Entity Recognizer)
- Geospatial search (perhaps using GeoNames)
- Cross-language search capabilities

- In-text highlighting of search terms with visual map of hit locations (see below)
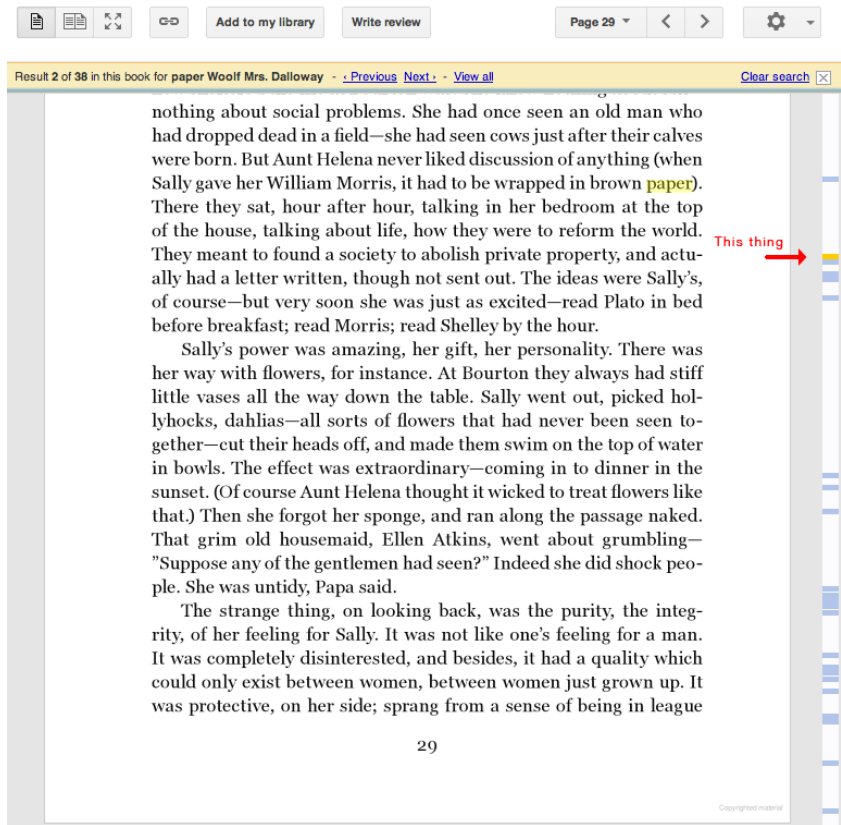
- Scoping of searches to types of content (such as "fiction", "articles", "adverts", etc.)
- Ability to download and print out issues in PDF format
- Aggregation of images for teaching and research purposes
- Datamining and Visualization APIs

Much of the search functionality described above requires metadata that can sometimes be generated by modern natural language processing software, but will also require significant human effort. Therefore, it's worth noting that use cases covering the cataloging of the resources and administration tools have not been documented here. Furthermore, our use cases included researchers and librarians across multiple domains, but omit cases for undergraduate students or the general public.

Standard Data Models and API's

Thanks to recent standardization efforts, the scholarly research community is on the cusp of a new era when it comes to the access, discovery, and sharing of image-based resources. In order to take advantage of emerging standards and break free from the all-too-familiar siloed, bespoke local application model, Blue Mountain will publish data in ways that will allow it to be part of this broader community.

- IIIF Image and Presentation APIs

The Blue Mountain viewer will be built using the [IIIF Image API](#) and the [IIIF Presentation API](#). This decision allows Blue Mountain material to be used for purposes far beyond the scope and vision we have for this project.

- [Schema.org](#)

  Blue Mountain houses a rich set of data and images, making it simple to optimize discoverability by major search engines with tools like schema.org. Schema.org provides a collection of schemas that developers can use to markup HTML pages in ways recognized by major search providers, and that can also be used for structured data interoperability (e.g. in JSON).

- OpenAnnotation Data Model

- [Linked Data](#) / Custom Ontology?

  A Linked Data approach would be necessary for some of the most desired search functionality. For example, "What avant-garde periodicals does the collection have from Czechoslovakia?" There could be some fairly sophisticated search algorithm that happens here. For example, a periodical may not be cataloged as "avant-garde", but contributors may be and we can deduce that if a periodical has a number of "avant-garde" contributors, it should be returned. Furthermore, a country's name may change over the years, but with linked data we can also return items that are cataloged as "Czech Republic," "Slovakia," and even parts of "Prussia" and "Ukraine". Another example would be that other relevant suggestions could be made to supplement the search. For instance, if a user searches for a work of art, additional works by that artist or contemporaries, reviews, and critiques may be suggested as well. Linked data and the search algorithms it supports must be curated and managed in line with specific use cases.
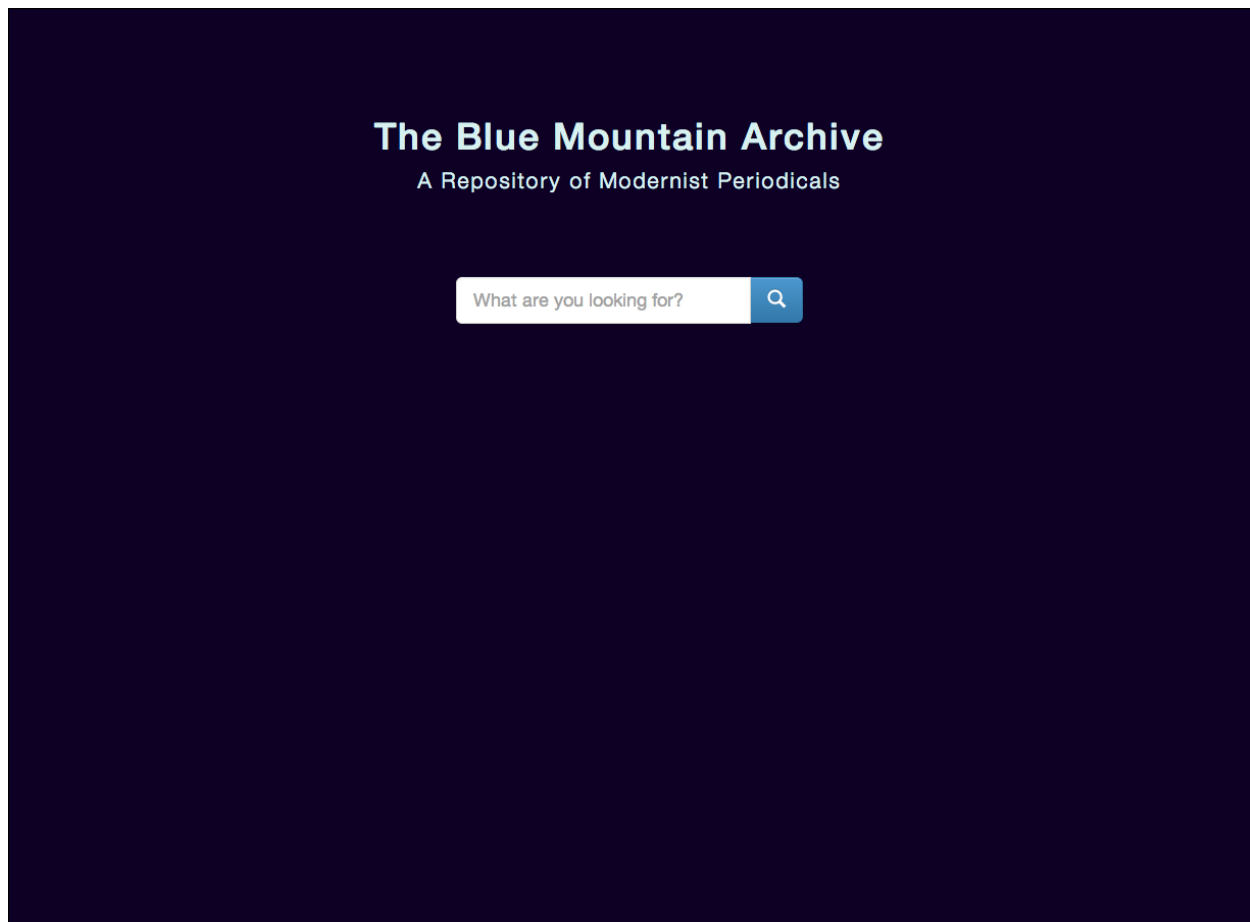
  Cliff and Natasha can talk more about their specific ontology... Contextual Relationships, GeoNaming,

Screens & RESTful URI Templates

Click on HTML-rendered screens to try interactive features. (Cliff - we should probably make a github.io page to host these, or at least point to the code at GitHub.)

http://bluemountain.princeton.edu
(Home page)

What are you looking for?

Periodicals

Lorem ipsum dolor sit amet, consectetur adipiscing

Names

Lorem ipsum dolor sit amet, consectetur adipiscing

Topics

Lorem ipsum dolor sit amet, consectetur adipiscing

Description

The Home Page provides a way to quickly search the archive for content, and offers a way to highlight and/or navigate to specific periodical, contributor, and topic listings. The dark area is intended to showcase any number of curated images that can be found in the collection.

/periodicals

Method	Content-Type/Accept	Extension	Produces/Consumes
GET	'application/xhtml+xml', 'text/html'	n/a	HTML representation of complete Blue Mountain periodical set, or, if params, results filtered by search params (paged)

GET	'application/xml'	{$uri}.xml	XML representation of complete Blue Mountain periodical set, filtered by optional params (paged)
GET	'application/json'	{$uri}.json	JSON representation of complete Blue Mountain periodical set, filtered by optional params (paged)

Description

Periodical Collection View provides a set of periodicals that are available in the Blue Mountain repository. Further filtering can be done based on "topic", "location", "language", "contributor", and "constituent" (text, music, image, chart).

`/periodicals/*` is *overloaded*. If `$id` then a single Periodical. If a topic ID or slug, then

Periodicals with that topic:

`/periodicals/${topic-id}?sort=(${field}-(asc|desc))&rpp=${xs:positive`

`Integer}&start=${xs:nonNegativeInteger}`

/periodical/[id]

Method	Content-Type/Accept	Extension	Produces/Consumes
GET	'application/xhtml+xml', 'text/html'	n/a	HTML representation of a Blue Mountain periodical

| GET | 'application/xml' | {$uri}.xml | XML representation of Blue Mountain periodical |
| GET | 'application/json' | {$uri}.json | JSON representation of Blue Mountain periodical |

Description

Single Periodical View intends to give an overview of the periodical including icon, description, basic metadata, publication information (dates, locations, and frequency), and contributors.

/periodical/[id]/issues
(see also /issues for unscoped list of issues)

Method	Content-Type/Accept	Extension	Produces/Consumes
GET	'application/xhtml+xml', 'text/html'	n/a	HTML representation of complete Blue Mountain issue set, or, if scoped or filtered, results

			filtered by search params (paged)
GET	'application/xml'	{$uri}.xml	XML representation of complete Blue Mountain issue set, or, if scoped or filtered, results filtered by search params (paged)
GET	'application/json'	{$uri}.json	JSON representation of complete Blue Mountain issue set, or, if scoped or filtered, results filtered by search params (paged)

Description

Issue Collection View provides a set of issues that are available in the Blue Mountain repository, with scoping on periodical. Issues can be grouped by year, place of publication, volume, and editors. Further filtering (faceting) can be done based on contributor, place of publication, and format.

/issues/[id]

The Blue Mountain Archive | Periodicals | Names | Topics ▾ | Search 🔍

Home / Periodicals / Broom / Issues / Vol. 1, Issue 1

Broom Volume 1, Issue 1

[👁 View Online] [⊕ Download ▾] [↪ Share ▾] [❝ Cite This ▾]

Publication Dates
Nov. 1921

Editors:
Harold A. Loeb
Alfred Kreymborg

Published in:
Rome, Italy

Circulation:
500

Price:
Single issue: $0.50
Subscription: $5.00

Format:
33cm x 52cm

Typography:
Helvetica, Futura

Paper Type:
Hand made from recycled newspaper.

Contributors

64x64 **Man Ray**
Cras sit amet nibh libero, in gravida nulla. Nulla vel metus scelerisque ante sollicitudin commodo.

64x64 **Marcel Duchamp**
Cras sit amet nibh libero, in gravida nulla. Nulla vel metus scelerisque ante sollicitudin commodo.

[More people ...]

Read full text? Amazon-like "Look Inside"?

Table of Contents

Untitled Woodcut (p. 1)
E. Prampolini
Pre-raphaelites dada intervention art hard-edge painting biedermeier rococo post-structuralism.

Table of Contents (p. 2)
Pre-raphaelites dada intervention art hard-edge painting biedermeier rococo post-structuralism.

'Lake' (pp. 3)
Bayard Boyesen
Pre-raphaelites dada intervention art hard-edge painting biedermeier rococo post-structuralism.

Drawing of Igor Strawinsky (pp. 4)
Pablo Picasso
Pre-raphaelites dada intervention art hard-edge painting biedermeier rococo post-structuralism.

Marnia (pp. 5)
Wallace Gould
Pre-raphaelites dada intervention art hard-edge painting biedermeier rococo post-structuralism.

Hunger: a Dublin Story (pp. 6-17)
James Stephens
Pre-raphaelites dada intervention art hard-edge painting biedermeier rococo post-structuralism.

Blue Mountain Project © 2014.

Method	Content-Type/Accept	Extension	Produces/Consumes
GET	'application/xhtml+xml', 'text/html'	n/a	HTML representation of a Blue Mountain issue
GET	'application/xml'	{$uri}.xml	XML representation of Blue Mountain issue
GET	'application/json'	{$uri}.json	JSON representation of Blue Mountain issue

Description

Single Issue View intends to give an overview of the issue including cover icon, description, basic metadata, publication information (date, location), table of contents, and contributors. Users have the option to view or download the complete issue, or range of pages from this screen.

/issues/[id]/manifest
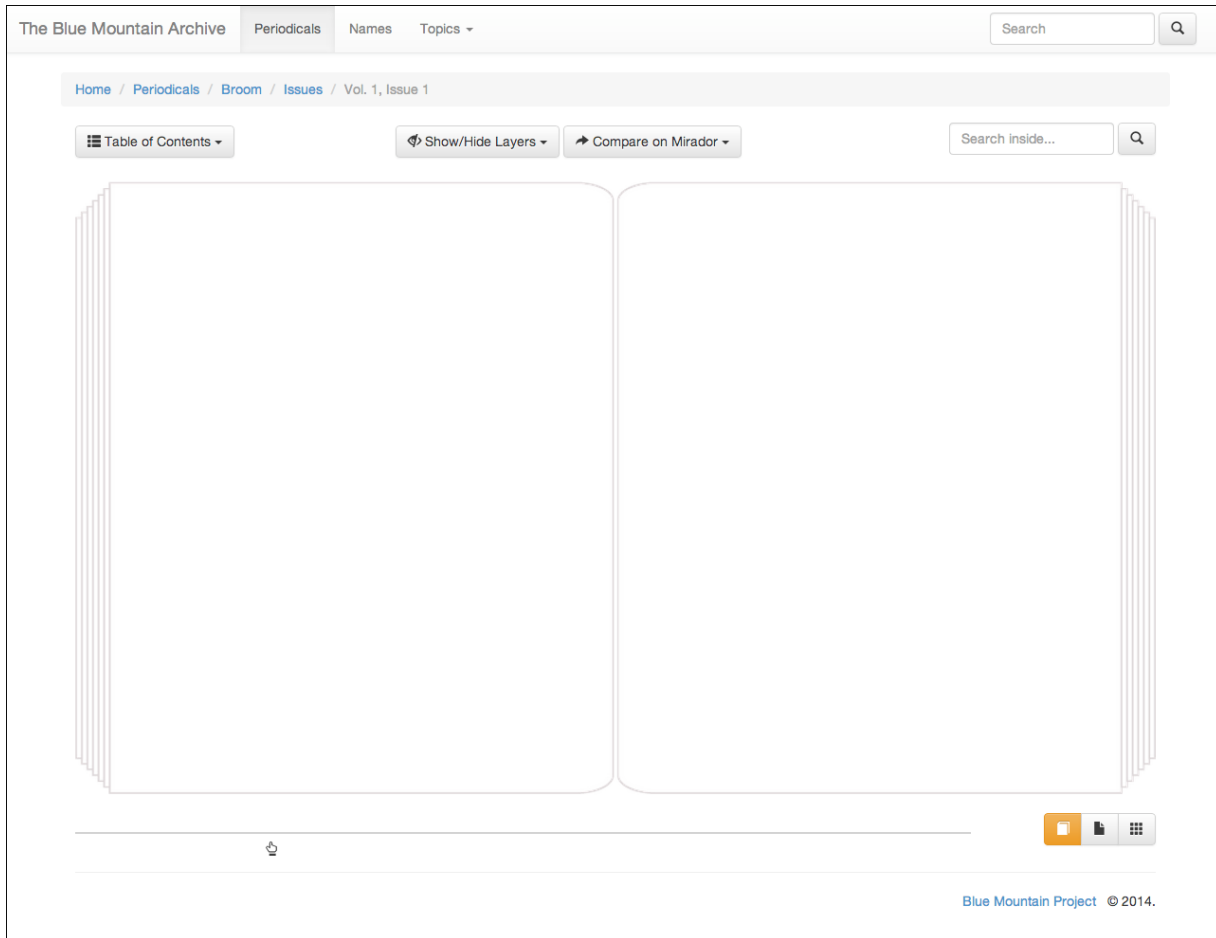
Method	Content-Type/Accept	Extension	Produces/Consumes
GET	'application/xhtml+xml', 'text/html'	n/a	HTML representation of a Blue Mountain issue manifest reader
GET	'application/xml'	{$uri}.xml	XML representation of Blue Mountain issue

			manifest in METS ALTO
GET	'application/json'	{$uri}.json	JSON representation of Blue Mountain issue manifest in IIIF Presentation format

Description

Manifest View offers the complete manifest of the issue as specified by the IIIF Presentation 2.0 API. The HTML view will be rendered as a "Viewer" with options for 1-up, 2-up, and thumbnail views, using the IIIF Image 2.0 API for image requests via the Loris IIIF Image Server, with per-page deep zoom functionality via OpenSeadragon. The Viewer will be METS ALTO aware, so as to allow for keyword and region highlighting, as well as article continuation. The Viewer will also allow for annotations to be viewed and made using the Open Annotation data model. Comparisons among images will exist via Mirador.

Next Design Steps

Search Results

Due to time limitations and the complexity and unique nature of search-based use cases, Search Results screens were not prototyped in this phase. These screens will ideally show sophisticated search options (i.e. "include annotations in results"), as well as relationships among the items. It should be noted that these should be geared toward specific use cases, such as exposing or suggesting relationships among creators, critics, and works to better understand public response or influences at the time of creation.

Relationships can be made both within Blue Mountain and externally through linked data sources. For example, through VIAF we know that E.E. Cummings and William Carlos Williams were coauthors. We can draw relationships between the two, even if their interactions did not take place within the Blue Mountain corpus. Further usability testing will be needed to determine the utility of that approach.

Topics

Many researchers want to know what topics and domains a resource covers up front, and topics can be a key starting point for many scholars. It's important to make some decisions about local topics for browsing purposes. Some thought should go into the categories, and these categories should also be mapped back to authority records to make data sharing easier.

METS-ALTO Viewer

While a developer could potentially extend the OpenLibrary BookReader to be METS-ALTO compliant, it may be more advantageous to take the time to develop a single page app from the ground up that has a native bias towards the technologies mentioned in the manifest section of this document.

In addition to front end design tasks, work will also have to be put into transforming METS-ALTO into the IIIF Presentation API Manifest.

Guide to Data APIs and Visualization How-to (with examples)

One of the key goals of the Blue Mountain project is to encourage and facilitate a community of developers and digital humanists to use the data in novel ways. The data api's need to be exposed (i.e. http://bluemountain.princeton.edu/developer). Furthermore, at least one example "how-to" should be offered based on an actual use case. Users should be encouraged to share back visualizations, apps, and mashups they develop so their work can be further promoted via the Blue Mountain blog.

Admin/Cataloging Workflows

(note to Cliff/Natasha: is DocWorks the ultimate admin solution? How does this jive with the preservation workflow that is being developed via Hydra?)

Admin users would ideally have the same view as everyday users, with extended functionality to edit the data, upload and reorder images "in place". This allows for continuity of the interfaces and less design work. Data entry forms should link into Linked Data sources to make it easy to identify named entities as early as possible through auto-suggest interfaces.

Priorities: Discovery, Access, and Open Data

It's all too easy to get bogged down in features and details. For example, one should not get distracted by developing data visualizations unless they are integral to the instantaneous understanding and relevancy of a resource. Instead, our role is to focus on the **discovery** of resources and relationships (exposure to search engines, enhanced relationship-oriented search interfaces, topics, open annotation, hydra integration), **access** to the content (viewer, implementation of IIIF APIs, compatibility with third party tools like Mirador), and publishing **open data** (publishing LinkedData and API documentation)

Appendix

Use Cases

- **As an Art History Researcher or Instructor**
  I would like to be able to create ad hoc groupings of images (including my own)
  So that I can refer to them later in one place or use them for teaching.
  (See Pinterest)
- **As an Art History Researcher**
  I want to examine two images next to each other
  So that I can compare similarities and differences.
  (See Mirador)
- **As an Art History Researcher**
  I want to be able to search across languages
  So that I don't have to manually search for (or be aware of) every variation a word may have.
  Example 1: A search on Istanbul should also bring up records that relate to Constantinople. (See Getty Thesaurus of Geographic Names.)
  Example 2: A search for "love" should bring back records in other languages that refer to the term: "amore", "aime", etc.
- **As an Art History Librarian**
  I want my searches to be able to locate publications/articles about a given genre (avant-garde) in a given region (Austria) in a given time range (early 20th century)
  So that I can tell patrons exactly what we have for any of the search criteria
- **As a Musicology Researcher**
  I want to be able to find contemporaries of a given person (artist, composer, author)
  So that I can understand the context in which the work was created
  (audience reaction to works was a related research question)
- **As a General Researcher**
  I want to be able to download a page, page-range, or arbitrary set of pages from a variety of resources
  So that I can share and manage my own research outside of Blue Mountain
- **As a Periodical Studies Researcher**
  I want search terms to be visualized with respect to where they fall in the periodical
  So that I can pinpoint terms and see patterns emerge
  (See Google Books vertical bar for search term location)
- **As a Periodical Studies Researcher**
  I want the terms I searched for to be highlighted, in context, on the page of the periodical
  So that I can pinpoint terms and see patterns emerge
- **As a Periodical Studies Researcher**
  I want the ability to scope searches to types of content (such as "fiction", "articles",

"adverts", etc.)
So that I can target my searches to the types of content I'm interested in
- As a Periodical Studies Researcher
I want a single search to return results from metadata, keywords/tags, full-text/OCR, and TEI
So that I don't have to perform multiple searches on the same terms
- As a General Researcher
I want to Discover Blue Mountain periodical titles via Google (and other major Internet Search Engines)
So that I don't have to know about the Blue Mountain Project to discover content
- As a Periodical Studies or Art History Researcher
I want to view a facsimile of the periodical as close to the original as possible
So that I can understand how the object (not just the words) functions and communicates "meaning and ideas". i.e.:
  - The physical qualities of the pages must be preserved:

    - The dimensions of the original must be inferrable from the surrogate;
    - The recto/verso and page spread relationships must be expressed;
    - Blank pages must be photographed and included
    - etc.

URI Templates

- /periodicals
- /periodicals/[periodical_id]
  - Note: A periodical contains a list of issues, contributors, and associated topics
- /periodicals/[periodical_id]/issues
  - Note: Issues scoped to a given periodical.
- /periodicals/[topic_id]/
  - Note: All periodicals with given topic
- /periodicals/[contributor_id]/
  - Note: All periodicals with given contributor
- /periodicals/[periodical_id]/issues/[topic_id]
  - Note: Issues with given topic scoped to a periodical.
- /periodicals/[periodical_id]/issues/[contributor_id]
  - Note: Issues with given contributor scoped to a periodical.
- /issues
- /issues/[topic_id]/
  - Note: All issues with given topic
- /issues/[contributor_id]/
  - Note: All issues with given contributor
- /issues/[issue_id]
  - Note: An issue contains a list of constituents, contributors, and associated topics

- /constituents
- /constituents/[constituent_id]
  - Note: Constituents are essentially "article-like-things". A constituent contains a list of associated tags [text, image, music, chart]
- /topics
- /topics/[topic_id]
- /contributors
- /contributor/[contributor_id]
- Place and Language will likely be a search filter/parameter