User Stories - Best Practices for Authoritative Data Working Group

- Overview
- Use Case: User wants to find the URI of an external entity from within a metadata editor
- Language processing for Metadata Editor Use Case
- Cache of entire or significant portion of dataset with updates via retrieve a known concept using a consistent format across datasets
- Update cached data
- Common External Search Format/Ontology
- Auto-fill Batch processing
- Manual Batch processing

Overview

This refines the use cases with a focus on user stories for each use case.

See also: Initial Use Case Definitions - Search API Best Practices for Authoritative Data Working Group

Prioritization of user stories will be gathered through surveys.

Use Case: User wants to find the URI of an external entity from within a metadata editor

The Metadata Editor Use Case assumes that all users are English speakers and that searching they are searching for items written in English. See Language Processing Use Case below for language specific user stories.

ID	User Story
Catalogers	
NOTE: The following user stories are related to creating or editing resource descriptions typically within a metadata editor or similar application.	
c-1	As a cataloger, I want to edit an entity (e.g. work, instance, etc.) and add a link to an external URI from an external authoritative sources (e.g. LCNAF, OCLC FAST, etc.).
c-2	As a cataloger, I want to edit an entity (e.g. work, instance, etc.) and display a label from an external authoritative sources (e.g. LCNAF, OCLC FAST, etc.).
c-3	As a cataloger, I want to be able to enter the exact external authoritative label and get the URI from the external authority linked to the entity being edited. This applies when there is a unique authoritative term.
c-4	As a cataloger, I want to start typing a known external authoritative label and get the URI from the external authority linked to the entity being edited. This is left anchored type-ahead.
c-5	As a cataloger, I want to start typing a known variant external authoritative label and get the URI of the authoritative label linked to the entity being edited.
c-6	As a cataloger, I want additional information in the search that indicates that the term listed has a variant that matches the keyword typed.
c-7	As a cataloger, I want to type in the alternate identifier (e.g. Q label in wikidata, ISNI label, organization code etc.) and get the URI for those entities.

c-8	Refined to: As a cataloger, I want to be able to search for a term and then make another search using a broader or narrower term for the original term as a new search term.
	Original wording: As a cataloger, I want to be able to search for a broader term in a hierarchy and get a list of narrower terms from which to select. NOTE: Some systems have seen performance issues in actual implementations. Catalogers generally know what they are looking for.
	Reason for change: It was felt by the group that the intent of the two statements is the same, but the wording of the refined version provided more clarity.
c-9	As a cataloger, I want to be able to see broader and narrower terms when the authority is hierarchical.
c-10	As a cataloger, I want to be able to step into broader or narrower terms when the authority is hierarchical in order to see and potentially select from other terms on the hierarchy.
c-11	As a cataloger, when I am unable to find what I'm looking for in an authority lookup, I want to be able to search an authority source in an external site by clicking on a link to its native search UI.
c-12	As a cataloger, I know some keywords, other attributes related to the entity that are not in the primary or variant label (e.g. occupation, resource type, etc.), that will help me locate and select an authoritative entity.
c-13	As a cataloger, I want search results to contain highly relevant terms for my keyword search based on standard indexing approaches. Actual relevancy is subjective.
c-14	As a cataloger, I want transparency in the approach for indexing to be clear. (e.g. exact match on primary label, stemming, which fields are searched, etc.) May vary between authorities.
c-15	As a cataloger, I want search results listed in rank order as determined by standard indexing approaches.
c-16	As a cataloger, I want to choose how search results are returned (e.g. left anchored, keyword indexing rank, or as yet unknown approach)
c-17	As a cataloger, I want to see, for each entity that appears in results of my keyword search, which of the fields that were searched triggered its inclusion in those results. (e.g. keyword was in the variant label, occupation, or descriptions instead of the primary label)
c-18	As a cataloger, I want to see contextual information (e.g. variant labels, occupation, birth date, etc.) about the search results that distinguishes it from other, similar-looking results, to help me to select the correct authoritative entity and to recognize false positives. The context may be drawn from authoritative entities and real world object entities based on what is available in the authoritative data.
c-19	As a cataloger, I want to be able to filter search results to a specific date range for a field on the authoritative entity (e.g. birth date, death date, etc.).
c-20	As a cataloger, I want to be able to filter search results to a specific class type (e.g. a corporate name, person name, meeting name, etc.; manifestation, item, expression, etc.).
c-21	As a cataloger, I want to be able to filter on specific fields in the search results (e.g. occupation, resource format, agent, etc.) This is a filter of results after they are returned similar to a facet.
c-22	As a cataloger, I want to be able to specify in the search limiting results to a keyword in a particular field (e.g. an advanced search that passes in 'occupation includes humorist').
c-23	As a cataloger, I want search results to be returned quickly, so that I can catalog efficiently, generally seen as sub-second results or some indicator that a longer search is being processed.
c-24	As a cataloger, I want to be able to request additional search results if what I am looking for isn't visible in the current set of results displayed, e.g. I didn't get it in the first 10, so give me 10 more results (aka pagination).
c-25	As a cataloger, if a request fails, it should fail in a reasonable amount of time and reply gracefully providing a reason for the failure (e.g. Time Out, No Result Found, No Exact Match, etc.)
c-26	As a cataloger, I want information displayed in the editor UI to match the information in the authoritative source. This primarily impacts editing and displaying of entities where the data in the authoritative data has changed (e.g. split, merged, renamed, deleted) to be sure that cataloged data remains accurate over time.
c-27	As a cataloger, I do not want authoritative data to be automatically updated when the data is changed in the authoritative source. I want to control if and when that information is updated.
c-28	As a cataloger, I want to determine whether the entity I'm searching for doesn't exist in the authoritative source that I'm searching.
Addendum - remaining stories were added during the creation and refinement of the final summary document	
c-29	As a cataloger, I want to type keywords related to the entity I want to find and receive back a list of results ordered by relevance
c-30	As a cataloger, I want to be able to return to previous pages as I page through results.
c-31	As a cataloger, I want to be able to filter search results based on a range of values for a field supporting range search (e.g. call numbers within a range, numeric values within a range)
c-32	As a cataloger, I want to select a broader or narrower term from the extended context to be the result. This user story requires the URI of the broader or narrower term to be included in extended context.
Application Developers (UI and Backend)	

d-1	As a UI developer, I want to provide a selection widget that enables a cataloger to type in keywords and see a list of entities sorted in rank order.
d-2	As a UI developer, I want to provide a selection widget that enables a cataloger to type in a left-anchored string and see a list of entities that start with the typed string sorted alphabetically. This could be implemented as an autocomplete.
d-3	As a UI developer, I want to provide a widget that shows search results that the end user can understand by displaying a representative label. This is typically the preferred label of the resources in the search results, but not necessarily, notably in the case of identify management systems (e.g. ISNI, VIAF, etc.)

DISCUSSION: If, for example in ISNI, the representative label is the ISNI number, you would need additional context to be able to distinguish between entities. As a UI developer, I want to show the user something and the user wants this to be understandable. I don't want to have to create this differently for each authority. In general, the response from the service should have not just the representative label, but also additional information. Ex. if the user does a search and the search hits a variant, from the UI perspective, you want to help the user why they are seeing the results and how it came from the variant. |
| d-4 | As a UI developer, I do not want to decide what label to display as the representative label whether the results come from a vocabulary with a clear preferred label or an identity management system. |
| d-5 | As a UI developer, I want to show users additional information (aka extended context) to help them make an accurate selection.

DISCUSSION: Not sure that authority provider can give rank order results, that it may be more appropriate for the application to allow the user to decide based on the extended context. |
| d-6 | As a developer, I don't want to have to understand the extended context, I just want to display it to the end user. I don't want to have to treat each service uniquely.

One way of handling differences between authority ontologies is to provide configurations identifying ldpath to each piece of extended context in the result graph. |
| d-7 | As a UI developer, I want the selection widget to allow the user to ask for more results (i.e., page through results). |
| d-8 | As a UI developer, I want to provide the end user with a search interface based on the authority's ontology and allow the user to select which field they want to search. (aka advanced search)

DISCUSSION: This implies some understanding of the unique aspects of each authority. Perhaps a better approach would be to define what fields you want to search and connecting the fields of the ontology to the search fields. This doesn't take advantage of the linked data aspect where the data could be more variable. Are we asking for consensus at the data model level? Maybe we can abstract out some by mapping at the service level some common fields to the ontology of the authority. ISNI approached this by using general ontologies (e.g. schema, skos) that allow for an easier way to come to common understanding of fields across multiple authorities. |
| d-9 | As a backend developer, I want a consistent way to specify parameters that is the same for all authorities (i.e. standardized API request). Want to avoid one off code requiring if-then-else processing. |
| d-10 | As a backend developer, I want data returned in a consistent way across all authorities. Minimally the data structure of extended context (e.g. lists, nested lists). Are there non-textual data like images or

DISCUSSION: Having fields be structured the same is what is important. We want to be able to display the results without needing to interpret them. One exception would be to handle by datatype (e.g. image). |
| d-11 | As a backend developer, I want data to be returned in the same format from all authorities (e.g. normalized json, json-ld, or something else) to allow for one processing approach for all authorities

There is a request for normalized json preferred over json-ld.

DISCUSSION: The last thing the developer wants is json-ld. It is very difficult to work with. Not always consistent json. Querying RDF data is complicated. A normal json structure is easily processed by most developers without special knowledge. Inconsistent or bad libraries for dealing with json-ld. Json-ld is not always constructed consistently making it difficult to process as simple json. |
| d-12 | As a backend developer, I want to identify the search depth to follow in the broader graph in the authority (e.g. follow links up to 3 levels deep). This has implications for performance if the levels are too deep. The goal of this user story is to have authorities provide enough data to help the end-user make an accurate selection.

Ex. ldpath: (madsrdf:hasAffiliation/madsrdf:organization/skos:prefLabel) | (madsrdf:hasAffiliation/madsrdf:organization/rdfs:label) | (madsrdf:hasAffiliation/madsrdf:organization/madsrdf:authoritativeLabel) :: xsd:string |
d-13	As a developer, I want to receive a permanent URI for each entity to uniquely identify the entity.
d-14	As a backend developer, I want to have documentation of authority search API describing parameters and the structure of the data that will be returned. (e.g. See example documentation at Agrovoc and OCLC.)
d-15	As a backend developer, I want to be able to support left anchored search and keyword search.
d-16	As a UI developer, I want to know whether an authority supports left anchored search and/or keyword search so I can provide a UI widget to support switching between these approaches.
d-17	As a backend developer, I want to receive results marked with a ranked order. There are various approaches to do this in RDF (e.g. RDF sequence, identified by predicate, etc.) This is typically driven by keyword search and produced by the provider's indexing approach.

d-18	As a normalization backend developer, I want data to be returned as linked data allowing for configurations that map ontologies to a normalized json format.
d-19	As an application backend developer, I want to receive data in a consistent format that allows for a single approach to interpret results. (e.g. needs of an Application like Sinopia)
d-20	As a normalization backend developer, I want to receive search results as resources in a linked data graph.
d-21	As a backend developer, I want to receive search results as a simple list of URIs. From those I would be able to fetch individual resources for more information. (May have performance issues when fetching many individual resources.)
d-22	As a backend developer, I want to be able to request the level of detail (e.g. extended context vs. simple response) and format (e.g. linked data graph, simple JSON structure). This could be fulfilled as a single API with parameters or as separate APIS.
d-23	As a UI developer, I want to provide additional information in the selection widget about entities to improve the accuracy of selection.
d-24	As a backend developer, I want to receive additional information about an entity in the same predictable format for all authorities (e.g. json)
d-25	As a backend developer, I want to receive what facets are available to support browse by facet and advanced search.
d-26a	As a backend developer, I want to pass the type (e.g. class, entity, or other type of classification field) to the provider API to filter search results.
d-26b	As a backend developer, I want to pass an exact value or range of values for fields supporting range search (e.g. dates, call numbers, numeric values) to the provider API for filtering of results by the provider.
d-26c	
d-27	As a backend developer, I want to receive facet values that can be filtered by (e.g. language is the facet, values are English, German, Spanish, etc., and counts for each value) . The facet information should include values for the entire result set and not just the current page's result.
d-28	As a backend developer, I want to receive pagination information with search results (e.g. first result position, last result position, current retrieved results count, total count in full result set), such that, I can request the next page of results.
d-29	As a backend developer, I want to receive pagination information following a standard (e.g. JSON-API Pagination)
d-30	As a backend developer, I want to be able to fulfill all search requests. (i.e. source authorities respond to all requests with excellent uptime)
d-31	As a backend developer, I want to quickly show search results to users (e.g. sub-second, specific threshold TBD)
d-32	As a backend developer, I want to update cached labels and URIs as changes are made to the authoritative source data. Want to receive change management feed of changed resources (e.g. added, renamed, moved, split, deleted, etc.)
Addendum - remaining stories were added during the creation and refinement of the final summary document	
d-33	As a UI developer, I want to provide a selection widget that enables a cataloger to type in exact text of the primary label of the entity they are looking for in an external authority and receive the URI if found or a notice that it doesn't exist if not found.
d-34	As a developer, I want to store the URI of the selected entity as the value.
d-35	As a developer, I (optionally) want to cache the display label of the selected entity to facilitate performant display of data later without having to retrieve it from the authoritative data source based on the URI.
d-36	As a UI developer, I want to provide a widget that allows lookup based on a label in the results (e.g. clicking a label in extended context will repeat the lookup search with that label as the query text). This can be used to search based on broader and narrower, but may also allow a lookup based on data in other fields as well.
d-36	As a UI developer, I want to include a link to the authoritative data provider's native search UI at the external authority's website.
d-37	As a developer, I want to receive keyword search results in rank order or with a rank predicate on which the data can be sorted.
d-38	As a developer, I want to provide a link to authoritative data provider's documentation on their data and indexing process. This link may be in our application documentation.
d-39	As a developer, if available, I want to pass through available information about how data was selected.
d-40	As a UI developer, I want to display to users why a search failed, if possible, and a generic Time Out message if not

d-41	As a UI developer, I want to return control of the application to the user in a reasonable amount of time when a search fails to return a result.
d-42	As an application architect, I want the option to host provider data in a local cache to improve uptime and other characteristics of data access
d-43	As a cache developer, I want to be able to keep the cache up to date by taking partial updates from the data provider.
d-44	As a cache developer, I want a regular predictable update schedule for downloadable data.
Providers	
p-1	As a provider, I want my data to be used.
p-2	As a provider, I want to understand how my data is being used in order to better support access (e.g. analytics on API usage, identify why an API is not used, etc.)
p-3	As a provider, I want to understand the license under which our data is being used (e.g. commercial, open source, etc.)
p-4	As a provider, I want to provide enough data to facilitate accurate selection by the end user without providing an overwhelming quantity of data. (e.g. make an accurate selection from information in the search results)
p-5	As a provider, I want to allow the user to control the quantity of data returned (e.g. simple list, extended context, all available data)
p-6	As a provider, I want to provide enough data for an application to create widgets that provide access to our data. (e.g. facets, filtering, quantity of data, etc.)
p-7	As a provider, I want users to be able to identify and use the right API for the task they are trying to perform.
p-8	As a provider, I want users to connect to my server in a way that efficient and effective, without using excessive band width.
p-9	As a provider, I want to be able to put in place restrictions (e.g. throttling, timeouts, etc.) that prevent automated connections from overloading our services.
p-10	As a provider, I want to be able to communicate to users why restrictions were put in place when an overload is reached. (e.g. receive explicit information in the response to their query)
p-11	As a provider, I want to document the sources of the data in our system (e.g. documentation on a website not necessarily provided through the API)
p-12	As a provider, I want our API(s) to be performant (e.g. sub-second queries, specific threshold TBD) This may lead to multiple APIs that are specialized or return different levels (quantity) of data.
p-13	As a provider, I want access to our data to be available 24-7 with occasional outages for maintenance that will be announced in advance. (e.g. website with status info, email subscription, slack channel, etc.)
p-14	As a provider, I want to provide data for download facilitating local caching either full download or a selection of data. Some considerations impacting whether to download and what can be downloaded include privacy of data, quantity of data, frequency of access, etc.
p-15	As a provider, I want to keep download files up to date providing downloads on a regular, predictable schedule (e.g. quarterly, monthly, weekly, or daily). Defining the schedule will depend on aspects of usage and frequency of data changes.
p-16	As a provider, I want to be able to provide partial downloads with only modifications since last download update.
p-17	As a provider, I want to provide notifications of changes to data (e.g. new, deleted, deprecated, moved, split, and merged entities). This can be done as an ongoing feed, a registry with a queryable API, or as a single file.

p-18	As a provider, I want to provide a representative label for each entity returned in search results.
	DISCUSSION: I want one field in the authoritative entity to be a human readable, meaningful representation of the entity that can be displayed to users to identify the entity for selection. The challenge here may be the difference between an authoritative source and an identity management system. For example, ISNI does not provide a display label. The question is who is responsible for creating a meaningful label. Do we need "a" label or "the" label? Example, VIAF has multiple labels in varying scripts. Do you request label in a particular script? Example, label for a work that identifies that work may need to include contributors in addition to the title. One way to address the need for additional data to be able to make a decision is to include additional context along with the identified human readable label that are displayed in separate fields as opposed to a single joined data field. The label likely won't be unique. Are we making a distinction between authority management system where there is an authorized access point that has a preferred unique label and other approach where data may not be standardized?
	DISCUSSION:
	- ISNI - Would not like to have a single representative label. Would prefer provider to provide all information and the application would highlight relevant information returned from the API. Example, if searching for author, know you want the author name from the ISNI response. - Language can complicate this. And context can also complicate results. Some names may be counter intuitive. - Searching at ISNI has a preference list that shows a label. Challenge is very many optional elements in the data structure. - Potential solutions that impact provider API - 1) Have different views into ISNI with different configurations on the QA level (e.g. name focused) - Two types: person, organization - Field is easy to determine, but it may be repeated many times primarily in different languages. The 'primary' is generally selected as English version. Prefer LOC first, then VIAF. - Multiple contributors for a single entity. - Looking at API receiving a parameter that identifies a preferred language. - Wikidata - Preferred label is selected in terms of language. When you create something, you can specify a preferred label. Otherwise all labels are equal. - By language labels. - In sparql queries can define language fallback chains. - There is a primary label for most entities, but there are a few that do not have labels in any language. In that case, it defaults to the id. - Search - based on keyword search of label and aliases (alt label) - Display - label + fallback (language specific label or ID if not label) - ShareVDE - preferred label is identified. - MeSH - NLM available in English only. Others support in other languages, but not supported by NLM. - Has variant labels. - Have one preferred label and can have several variant labels. - LOC - ID has both situations where in most situations there is a clearly defined preferred label in English. Also has names which at not language tagged. Some encode script, but not language. - In BF works and instances, there isn't a preferred label. LOC does a constructed label for their own convenience so they don't have to compute the label downstream. Created for their own use. - RWO may not have an unambiguous preferred label. - Suggest service provides label + URI. - Editor lookup shows multiple results from suggest server. Hovering over a result shows a context modal with more information gathered with a second request.

p-19	Questions about triple store, SPARQL queries, indexing, etc. for providers

Addendum - remaining stories were added during the creation and refinement of the final summary document	

p-20	As a provider, I want to implement an API that searches for an exact match on a primary label and possibly variant label.
p-21	As a provider, I want to implement an API that performs a left anchor search for a primary label (and possibly variant label) that starts with the query text and returns search results in alphabetical order. NOTE: This could be implemented using wildcards with left anchor text followed by wild card. May need to provide additional context for variant label matches to provide the primary label with the results.
p-22	As a provider, I want to implement an API that accepts keywords to search for matching entities on select fields (e.g. primary label, variant label, etc.) as appropriate for the authority being searched.
p-23	As a provider, I want to provide relevancy ranking of results that are returned, ranked according to an indexing approach that is appropriate for the authoritative data.
p-24	As a provider, I want to implement an API that performs a boolean search across multiple fields with values targeted for specific fields in the API query (e.g. occupation=humorist&primaryLabel=Twain). Explore passing values related to fields or using GraphQL endpoint.
p-25	As a provider, I want to assign a unique URI to each entity in my authoritative data.
p-26	As a provider, I may not be able to provide a human understandable label and will provide the URI instead in search results. (See also p-18.)
p-27	As a provider, I want to include extended context for hierarchical terms including the URI for the term to allow users to select a term that isn't the primary result.

p-28	As a provider, I want to support specifying page size as part of the search API request (e.g. page[size]=4 will return the first 4 matches
p-29	As a provider, I want to supply pagination information with results including links for moving between pages (e.g. links to current, first, last, previous, and next pages following json-api specification on pagination; See also json-api pagination example)
p-30	As a provider, I want to supply the number of matching results (or possibly pages of results) (e.g. numFound=74 specified under meta when using json-api standard).
p-31	As a data provider, I want to support search of my data at my website.
p-32	As a provider, I want to support filtering of results by type by providing a parameter in which the type can be passed through the API (e.g. class, entity, or other type of classification field). We recognize that the complexity of the ontology representing the data may have implications for how this is supported.
p-33	As a provider, I want to include rank order with keyword search results. (e.g. rank predicate with rank value)
p-34	As a provider, I want to convey information about how search results are generated. This may be in the form of documentation of the indexing process on the provider's website.
p-35	As a provider, I want to include information on why a specific result was included in a set of search results. (e.g. matched primary label, variant label, descriptor, etc.) This could be challenging when data is returned as linked data and requires further investigation on how to pass this back to users.
p-36	As a provider, if possible, I want to return results in a reasonable amount of time and provide a message if this cannot be done (e.g. Time Out, No Result Found, No Exact Match, etc.)
p-37	As a provider, I want to support filtering of results based on a range of values for a field supporting range search (e.g. range of dates, call numbers within a range, numeric values within a range) (e.g. date_filter=04-23-1564, date_filter=04-01-1564:04-30-1564, date_filter=1972-1980, etc.). (Exact syntax TBD)
p-38	As a provider, I want to return a meaningful result when the data in the range value field is missing, partial, or inconsistent. This is complex and requires further investigation to determine strategies for dealing with data integrity issues.

Language processing for Metadata Editor Use Case

Processing of language is important and complex. It has been extracted out into a separate use case to better understand language issues. These user stories still relate to usage within a Metadata Editor.

Required?	User Story
Catalogers	
	As a cataloger, I want to submit a query in my preferred language and see search results in my preferred language. Preferred language is generally set through browser's preferred language settings, application settings, or passed as a parameter. (e.g. My native language is German and I want to type keywords in German and see results in German.) This is dependent on the authority having data in that language.
	As a cataloger, I want to submit a query in the language of the entity for which I am searching. (e.g. The book I am looking for was written in Korean, so I want to type keywords and left-anchored string search in Korean.)
	As a cataloger, I want to be able to filter search results by language once the results are returned and displayed in the selection widget. DISCUSSION: (e.g. The book is written in Greek and I want to see the information displayed in Greek.) The user types in Greek and expects to find the Greek version of the title. May type non-Greek, but still want to have access to the Greek terms. This is complex and may need UI mockups that capture potential cases. Part of the results for this may be described by identifying which field triggered the results. May be good to have a UI widget that allows for filtering after the results are returned. May need User Studies to determine how best to do this. Want to lookup in a language I know but need to include a label in a language I do not know well. While cataloging works, the work may be in another language and need to be sure I've cataloged the one in the correct language. It would be good to collaborate with non-roman script working group to see what their use case is for this. For LOC, if the user searches in Greek, they will get Greek results back. Nearly all subject headings, names, lookups in general, they are all in English. The suggest search does not search variant labels. This can limit results that are returned.
Developers	
	As a UI developer, I want to provide a search widget allowing the user can specify the language to pass to the authority provider as part of the search (e.g. q=lait&lang=en).
	As a UI developer, I want to honor the browser setting for language. (e.g. If the user set their language in the browser to Spanish, then by default, searching is assumed to be in Spanish).

Required?	User Story
	As a developer, I want to be able to display the human readable label in the end user's native language (e.g. the user speaks Korean and want to display label in Korean). DISCUSSION: Does this differ from the cataloger's user case? As a developer, we need to be able to distinguish between how to display this to the user. How to allow user to select different languages? These seem more user interface oriented than API access by developers. Perhaps, this is more of a cataloger user story. The user story for the developer is that I want to be able to pass additional data in a parameter to the API that allows the authority provider to act on requests that limit the language of results returned. There may be other parameters passed through to the API that control results.
	As a developer, I want to be able to facilitate a user searching for things that are stored in a different language as the primary language. (e.g. book was written in Greek)
Providers	

Cache of entire or significant portion of dataset with updates via retrieve a known concept using a consistent format across datasets

Update cached data

Required?	User Story
Catalogers	
	As a cataloger, I want to be able to search across multiple authorities.
	As a cataloger, I want the display of data in the editing app to follow a consistent and predictable format.
Developers	
	As an application developer, I want a reliable source of data, such that, I am able to fulfill all search requests.
	As a cache developer, I want to store all or significant portion of authoritative data locally.
	As a developer, I want to control how data is searched.
	As a developer, I want to control what data is returned to and displayed in the end user application (e.g. extended context).
	As a developer, I want to control the format of data the editing application receives for the selection widget.
	As a cache developer, I want to update cached data as changes are made to the authoritative source data.
Producers	
	As a producer, I want to make my data available for download.
	As a producer, I want consumers of my data to use the most up to date version of the data.
	As a producer, I want to provide a change management API that provides information on changes to terms in the dataset (e.g. new, deleted, changed, deprecated, moved, split, merged, etc.).

Common External Search Format/Ontology

Required?	User Story
Catalogers	
	As a cataloger, I want to know that data drawn from multiple authorities have the same meaning.
Developers	
	As a developer, I want to use the same code to process data coming from multiple authorities.
	As a developer, I want to support searching across multiple authorities.

Producers	
	As a producer, I want data alignment between authorities.
	As a producer, I want to have an internal representation of my data that may be different from other authority providers.
	As a producer, I want to provide a mapping of my data to a community supported search ontology.

Auto-fill Batch processing

Required?	User Story
Catalogers	
	As a cataloger, I want auto-reconciliation processes to run that identify 'same-as' type relationships from other authorities.
Developers	
	As a developer, I want to run auto-batch processes once data has been entered.
Producers	
	As a producer, I want to support auto-reconciliation.

Manual Batch processing

Required?	User Story
Catalogers	
	As a cataloger, I want to load data in a reconciliation program (e.g. open refine) and apply changes across multiple entities (e.g. works, instances, etc.)
	As a cataloger, I want to be able to filter potential choices by date range.
Developers	
Producers	
	As a producer, I want to support an API that allows my data to work with popular reconciliation programs (e.g. open refine)