Samvera Vocab Manager: Functional Requirements

This document outlines the functional requirements for a Samvera Community vocabulary manager. Gathered use cases show the need for an application that can manage both predicates and controlled vocabularies for use in Samvera application development and local implementation. The functional requirements outline how a Samvera Community vocab manager needs to work for people and machines. Finally, development phases are suggested for initially producing a functional Samvera Vocab Manager and enhancing its capabilities to work for the Samvera community over time.

NOTE: This is a preliminary set of functional requirements from the URI Management Working Group. The Functional Requirements section currently focuses on creating and hosting predicates (defined by Development Phases 1-3 and Use Cases within those Phases). The use cases and development phases include the need to create and host controlled vocabulary terms as well. Defining functional requirements for creating and hosting controlled vocabulary terms will be future work for this working group.

Use Cases

- Community application development (Sufia, CurationConcerns, PCDM) that has need for predicate that does not exist. - Phase 1
  a. Predicates needed for Sufia/CurationConcerns: [link]
  b. Predicates added for embargo/leasing to Samvera: [link]
- Predicates for Local development (BPL example for Internet Archive identifier) - Phase 2
  a. [link]
    i. See Option 2, row 36.
  b. Might be that local predicates could actually be used by others so they could become community-used
- RDF that doesn’t work well for Samvera/Fedora environment and doesn’t map to other existing RDF ontologies (PREMIS) - Phase 4
- Accepted vocab (standard) not available through RDF (GeoRSS) - Phase 4 or 5 (depending if Classes are needed)
- Provide an API optimized for autocomplete or batch strings-to-things - Phase 5
  a. Returned results should be in relevance order
- Local vocabulary management - install, host, and manage vocabularies locally using Samvera Vocab Manager code (opposite side of Questioning Authority gem) - Phase 6
Functional Requirements

For people:

- See human-readable views of URIs, providing the following for predicates (based on required information from Predicate Decision Tree to request new predicate):
  - URI
  - Label
  - Comment
  - Usage example
  - Creator
  - Date created
  - Date modified
- See groupings of predicates by rdf:type identification (see Predicate Decision Tree), provide UI with lists of related predicates
- View relationships: skos:closeMatch, skos:exactMatch, skos:broadMatch, skos:narrowMatch, or skos:relatedMatch (this would get us to the 5th star in the 5-star Open Data categories)

For machines:

- Harvest predicate information/negotiate content in various machine-readable formats/serializations (JSON-LD, Turtle, N-Triples)
  - Harvesting should be available per predicate and the entire set

For people and machines:

- URI pattern: http://[URI base/domain name]/predicate/[predicate name]
  - Slash Recipe #5 - with modifications to allow serialization in formats other than rdf/xml and to allow for RDF serialization access to dereference a URI for a single property
    - TTL (application/x-turtle and .ttl extension)
    - JSON (application/json and .json extension)
    - N-Triples (text/plain and .nt extension)
    - HTML (text/html and .html extension) - human-readable only
  - Recipe decision explanation: http://id.loc.gov/ontologies/bibframe.html
    - Search for “shelfmark” to see there is a “shelfMark” (predicate) and a “ShelfMark” (object class).
    - Adding /predicate/ in URI path to help avoid this sort of collision
• Limited downtime, status monitored and timely staff response required (within 1 business day, Monday-Friday); make full vocabulary serialization available in single file via Github

Development Phases

1. Host predicates only - Ontology file placed in web accessible directory (for example, git repo)
2. Host predicates only - Request and review process in place and manual/external
3. Host predicates only - Web application that allows creation of predicate based on required info through form input; management module that allows for editing and deleting of properties, managing access to management module
4. Host predicates only - Request and review process part of web application
5. Host predicates and controlled vocab terms (possibly classes) - Request and review process part of web application
6. Provide way to download and locally host Samvera Vocab Manager instance

Mock-ups

The following are examples of what a possible implementation of the vocabulary manager could look like. The examples use two properties (arkivoChecksum and importURL) that may or may not be actual Samvera predicates. Each example screenshot is linked at the end of this document.

The route /index.html would list all properties in HTML with links to a complete RDF serialization of the same list in N-Triples, JSON-LD, and TTL (turtle). (See Appendix A).

Each property has its own HTML page showing metadata about the property term itself and includes links to the same data in the following RDF serializations: N-Triples, JSON-LD, and TTL.

One example property is the HTML page for arkivoChecksum and its serializations: N-Triples, JSON-LD, TTL. (See Appendix B).

Another example property is the HTML page for importURL and its serializations: N-Triples, JSON-LD, TTL. (See Appendix C).

The HTML pages would show all the metadata for a given predicate. The Usage Example text comes from the property skos:example.
### Predicates

<table>
<thead>
<tr>
<th>Predicate</th>
<th>URI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Akkio Checksum</td>
<td><a href="http://projecthydra.org/ns/predicate/akkioChecksum">http://projecthydra.org/ns/predicate/akkioChecksum</a></td>
</tr>
<tr>
<td>Import URL</td>
<td><a href="http://projecthydra.org/ns/predicate/ImportURL">http://projecthydra.org/ns/predicate/ImportURL</a></td>
</tr>
</tbody>
</table>

**Formats:** N-Triples, JSON-LD, TTL
Arkivo Checksum

http://projecthydra.org/ns/predicate/arkivoChecksum

<table>
<thead>
<tr>
<th>rdflabel</th>
<th>Arkivo Checksum</th>
<th>English [en]</th>
</tr>
</thead>
<tbody>
<tr>
<td>rdfscomment</td>
<td>An MD5 checksum for a bitstream uploaded by Zotero’s subscription service, Arkivo.</td>
<td></td>
</tr>
<tr>
<td>rdfsseeAlso</td>
<td><a href="https://github.com/projecthydra/sufia/issues/1327">https://github.com/projecthydra/sufia/issues/1327</a></td>
<td></td>
</tr>
<tr>
<td>dcterms:creator</td>
<td><a href="mailto:spr7b@virginia.edu">spr7b@virginia.edu</a></td>
<td></td>
</tr>
<tr>
<td>dcterms:issued</td>
<td>2018-11-11</td>
<td></td>
</tr>
<tr>
<td>dcterms:modified</td>
<td></td>
<td></td>
</tr>
<tr>
<td>dcterms:description</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Other Formats: N-Triples, JSON-LD, TTL

A repository object has an MD5 checksum generated by Arkivo that should be stored with the bitstream.

importURL

Appendix A — Sample Output for /index.html

As N-Triples

# ss:arkivoChecksum example
# Automated information

<http://hydra.vocab.manager/predicate/arkivoChecksum> <http://www.w3.org/2004/02/skos/core#example> "A repository object has an MD5 checksum generated by Arkivo that should be stored with the bitstream.cbr"&lt;http://example.org/object&gt;&lt;http://example.org/object&gt; &quot;C7F2C181A6A7A332C72483C48D5E202D&quot; .
<http://hydra.vocab.manager/predicate/arkivoChecksum> <http://purl.org/dc/terms/creator> &lt;mailto:spr7b@virginia.edu&gt; .
# Required information

# ss:arkivoChecksum example
# Required information

<http://hydra.vocab.manager/predicate/arkivoChecksum> <http://www.w3.org/2004/02/skos/core#example> "A repository object has an MD5 checksum generated by Arkivo that should be stored with the bitstream.cbr"&lt;http://example.org/object&gt;&lt;http://example.org/object&gt; &quot;C7F2C181A6A7A332C72483C48D5E202D&quot; .
<http://hydra.vocab.manager/predicate/arkivoChecksum> <http://purl.org/dc/terms/creator> &lt;mailto:spr7b@virginia.edu&gt; .
# Optional information
As JSON-LD
{
  "@context": {
    "ns1": "https://www.npmjs.com/package/arkivo",
    "dcterms": "http://purl.org/dc/terms/",
    "rdf": "http://www.w3.org/1999/02/22-rdf-syntax-ns#",
    "rdfs": "http://www.w3.org/2000/01/rdf-schema#",
    "skos": "http://www.w3.org/2004/02/skos/core#",
    "xsd": "http://www.w3.org/2001/XMLSchema#"
  },
  "@graph": [
    {
      "@id": "http://hydra.vocab.manager/predicate/arkivoChecksum",
      "@type": [
        "rdfs:Resource",
        "dcterms:Dateline"
      ],
      "dcterms:creator": {
        "@id": "mailto:spr7b@virginia.edu"
      },
      "dcterms:issued": {
        "@type": "xsd:dateTime",
        "@value": "2016-11-11"
      },
      "ns1:isDefinedBy": {
        "@id": "https://www.npmjs.com/package/arkivo"
      },
      "rdfs:comment": {
        "@language": "en",
        "@value": "The URL from which content hosted in a cloud storage service is imported."}
    }
  ]
}
"@value": "An MD5 checksum for a bitstream uploaded by Zotero's subscription service, Arkivo."
},
"rdfs:label": {
  "@language": "en",
  "@value": "Arkivo Checksum"
},
"rdfs:range": {
  "@id": "xsd:string"
},
"rdfs:seeAlso": {
  "@id": "https://github.com/projecthydra/sufia/issues/1327"
},
"skos:example": "A repository object has an MD5 checksum generated by Arkivo that should be stored with the bitstream.<br/>
&lt;http://example.org/object&gt;\nhttp://hydra.vocab.manager/predicate/arkivoChecksum&gt; &quot;C7F2C101A6A74332C72483C48D5E202D&quot;
},
{ 
  "@id": "http://hydra.vocab.manager/predicate/importURL",
  "@type": [
    "rdf:predicate",
    "rdf:Property"
  ],
  "dcterms:creator": {
    "@id": "mailto:sp7b@virginia.edu"
  },
  "dcterms:issued": {
    "@type": "xsd:date",
    "@value": "2016-11-11"
  },
  "rdfs:comment": {
    "@language": "en",
    "@value": "The URL from which content hosted in a cloud storage service is imported."
  },
  "rdfs:label": {
    "@language": "en",
    "@value": "Import URL"
  },
  "rdfs:seeAlso": {
    "@id": "https://github.com/projecthydra/curation_concerns/issues/561"
  },
  "skos:example": "A repository object was imported from Google Drive cloud storage.<br/>
http://example.org/object&raquo; \nhttp://hydra.vocab.manager/predicate/importURL&raquo; 
https://drive.google.com/file/d/0B_3cn7zuYfw-UVVhTmdTc2VYd1k/view?usp=sharing&raquo;"
}
}

As Turtle
@prefix dcterms: <http://purl.org/dc/terms/> .
@prefix ns1: <https://www.w3.org/2000/01/rdf-schema#> .
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .
@prefix skos: <http://www.w3.org/2004/02/skos/core#> .
@prefix xml: <http://www.w3.org/XML/1998/namespace> .
@prefix xsd: <http://www.w3.org/2001/XMLSchema#> .

<http://hydra.vocab.manager/predicate/arkivoChecksum> a rdf:Property,
rdf:predicate ;
rdfs:label "Arkivo Checksum"@en ;
Appendix B — Sample Output for ArkivoChecksum

As N-Triples

```ntriples
#ss:arkivoChecksum example
# Automated information

# Required information
<http://hydra.vocab.manager/predicate/arkivoChecksum> <http://www.w3.org/2004/02/skos/core#example> "A repository object has an MD5 checksum generated by Arkivo that should be stored with the bitstream.&lt;br/&gt;&lt;http://example.org/object&gt;&amp;quot;C7F2C101A6A7A332C72483C48D5E20D&amp;quot;" .

# Optional information
```

As JSON-LD
"@context": {
  "dcterms": "http://purl.org/dc/terms/",
  "rdf": "http://www.w3.org/1999/02/22-rdf-syntax-ns#",
  "rdfs": "http://www.w3.org/2000/01/rdf-schema#",
  "skos": "http://www.w3.org/2004/02/skos/core#",
  "xsd": "http://www.w3.org/2001/XMLSchema#"
},
"@id": "http://hydra.vocab.manager/predicate/arkivoChecksum",
"@type": [
  "rdf:predicate",
  "rdf:Property"
],
"dcterms:creator": {
  "@id": "mailto:spr7b@virginia.edu"
},
"dcterms:issued": {
  "@type": "xsd:date",
  "@value": "2016-11-11"
},
"rdfs:comment": {
  "@language": "en",
  "@value": "An MD5 checksum for a bitstream uploaded by Zotero's subscription service, Arkivo."
},
"rdfs:isDefinedBy": {
  "@id": "https://www.npmjs.com/package/arkivo"
},
"rdfs:label": {
  "@language": "en",
  "@value": "Arkivo Checksum"
},
"rdfs:range": {
  "@id": "xsd:string"
},
"rdfs:seeAlso": {
  "@id": "https://github.com/projecthydra/sufia/issues/1327"
},
"skos:example": "A repository object has an MD5 checksum generated by Arkivo that should be stored with the bitstream.\n
\n&lt;http://example.org/object&gt;\n&lt;http://hydra.vocab.manager/predicate/arkivoChecksum\&gt; &quot;C7F2C101A6A7A332C724B3C48D5E202D&quot;\n"
}

As Turtle
@prefix dcterms: <http://purl.org/dc/terms/> .
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .
@prefix skos: <http://www.w3.org/2004/02/skos/core#> .
@prefix xml: <http://www.w3.org/XML/1998/namespace> .
@prefix xsd: <http://www.w3.org/2001/XMLSchema#> .

<http://hydra.vocab.manager/predicate/arkivoChecksum> a rdf:Property,
  rdf:predicate ;
  rdfs:label "Arkivo Checksum"@en ;
  dcterms:issued "2016-11-11"^^xsd:date ;
  rdfs:comment "An MD5 checksum for a bitstream uploaded by Zotero's subscription service, Arkivo."@en ;
  rdfs:isDefinedBy <https://www.npmjs.com/package/arkivo> ;
  rdfs:range xsd:string ;
  rdfs:seeAlso <https://github.com/projecthydra/sufia/issues/1327> ;
  dcterms:creator <mailto:spr7b@virginia.edu> ;
A repository object has an MD5 checksum generated by Arkivo that should be stored with the bitstream.

http://example.org/object &quot;C7F2C101A6A74332C72483C48D5E202D&quot;.

Appendix C — Sample Output for ImportURL

As N-Triples

```xml
#ss:importeURL example
# Automated information
&lt;http://hydra.vocab.manager/predicate/importURL&gt; &lt;http://www.w3.org/1999/02/22-rdf-syntax-ns#type&gt;
&lt;http://www.w3.org/1999/02/22-rdf-syntax-ns#Property&gt;.
&lt;http://hydra.vocab.manager/predicate/importURL&gt; &lt;http://www.w3.org/1999/02/22-rdf-syntax-ns#type&gt;
&lt;http://www.w3.org/1999/02/22-rdf-syntax-ns#predicate&gt;.
&lt;http://hydra.vocab.manager/predicate/importURL&gt; &lt;http://www.w3.org/2001/XMLSchema#date&gt; "2016-11-11"^^&lt;http://www.w3.org/2001/XMLSchema#date&gt;.
# Required information
&lt;http://hydra.vocab.manager/predicate/importURL&gt; &lt;http://www.w3.org/2000/01/rdf-schema#comment&gt; "The URL from which content hosted in a cloud storage service is imported."@en.
&lt;http://hydra.vocab.manager/predicate/importURL&gt; &lt;http://www.w3.org/2004/02/skos/core#example&gt; "A repository object was imported from Google Drive cloud storage.&lt;br/&gt;&lt;http://example.org/object&gt; &lt;https://drive.google.com/file/d/0B_3cn7zuYfw-UVVhTmdTc2VYd1k/view?usp=sharing&gt;".
&lt;http://hydra.vocab.manager/predicate/arkivoChecksum&gt; &lt;http://purl.org/dc/terms/creator&gt; &lt;mailto:spr7b@virginia.edu&gt;.
# Optional information
&lt;http://hydra.vocab.manager/predicate/importURL&gt; &lt;http://www.w3.org/2000/01/rdf-schema#seeAlso&gt; &lt;https://github.com/projecthydra/curation_concerns/issues/561&gt;.
```

As JSON-LD

```json
{
    "@context": {
        "dcterms": "http://purl.org/dc/terms/",
        "rdf": "http://www.w3.org/1999/02/22-rdf-syntax-ns#",
        "rdfs": "http://www.w3.org/2000/01/rdf-schema#",
        "skos": "http://www.w3.org/2004/02/skos/core#",
        "xsd": "http://www.w3.org/2001/XMLSchema#"
    },
    "@id": "http://hydra.vocab.manager/predicate/importURL",
    "@type": [
        "rdf:Property",
        "rdf:predicate"
    ],
    "dcterms:creator": {
        "@id": "mailto:spr7b@virginia.edu"
    },
    "dcterms:issued": {
        "@type": "xsd:date",
        "@value": "2016-11-11"
    },
    "rdfs:comment": {
        "@language": "en",
        "@value": "The URL from which content hosted in a cloud storage service is imported."
    }
}
```
As Turtle
@prefix dcterms: <http://purl.org/dc/terms/> .
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .
@prefix skos: <http://www.w3.org/2004/02/skos/core#> .
@prefix xml: <http://www.w3.org/XML/1998/namespace> .
@prefix xsd: <http://www.w3.org/2001/XMLSchema#> .

<http://hydra.vocab.manager/predicate/importURL> a rdf:Property,
  rdf:predicate ;
  rdfs:label "Import URL"@en ;
  dcterms:issued "2016-11-11"^^xsd:date ;
  rdfs:comment "The URL from which content hosted in a cloud storage service is imported."@en ;
  rdfs:seeAlso <https://github.com/projecthydra/curation_concerns/issues/561> ;
  dcterms:creator <mailto:spr7b@virginia.edu> ;
  skos:example "A repository object was imported from Google Drive cloud storage."&lt;http://example.org/object&gt; &lt;http://hydra.vocab.manager/predicate/importURL&gt; &lt;https://drive.google.com/file/d/0B_3cn7zuYfw-UVVhTmdTc2VYd1k/view?usp=sharing&gt;" .