Advanced uses of thesauri that are released with multiple services: Websites, APIs, and SPARQL endpoints

Marcia Zeng, Tao Hu
School of Information, Kent State University

NKOS Workshop, DC2017
From the exhibition of the Mogao Caves and the website of the exhibition, *Cave Temples of Dunhuang*, let’s start the exploration of information related to the Silk Road though the *Getty Thesaurus of Geographic Names (TGN)*.

http://www.getty.edu/research/exhibitions_events/exhibitions/cave_temples_dunhuang/index.html
Let's see how to use them through three examples:

1. HTML
2. XML
3. RDF

are available in multiple formats.

Getty Thesaurus of Geographic Names (TGN) and other Getty vocabularies
Example 1
Learn through TGN Website

“Mogao Caves” =>
“Related geographic places”
located on .... “Silk Road”

<table>
<thead>
<tr>
<th>Place Types:</th>
</tr>
</thead>
<tbody>
<tr>
<td>ancient site (preferred, C)</td>
</tr>
<tr>
<td>rock-cut architecture (C)</td>
</tr>
<tr>
<td>World Heritage Site (C)</td>
</tr>
<tr>
<td>historic site (C)</td>
</tr>
<tr>
<td>caves (C)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Related geographic places:</th>
</tr>
</thead>
<tbody>
<tr>
<td>located on .... Silk Road</td>
</tr>
<tr>
<td>(road)</td>
</tr>
<tr>
<td>Dunhuang (inhabited place)</td>
</tr>
</tbody>
</table>

Note: System of over 735 caves, with manmade alterations and excavation; 492 caves are decorated; located outside of Dunhuang city. Carved and painted beginning in 4th century CE.
TGN gives us ...
Learn from these entities through TGN, one by one
Learn from these entities through TGN, one by one, &,

If you want to download the dataset →
Palmyra

Source: http://vocab.getty.edu/tgn/7018835

<table>
<thead>
<tr>
<th>Subject</th>
<th>Predicate</th>
<th>Object</th>
</tr>
</thead>
<tbody>
<tr>
<td>tgn:7018835</td>
<td>rdf:type</td>
<td>gvp:Subject</td>
</tr>
<tr>
<td>tgn:7018835</td>
<td>rdf:type</td>
<td>skos:Concept</td>
</tr>
<tr>
<td>tgn:7018835</td>
<td>rdfs:label</td>
<td>Palmyra</td>
</tr>
<tr>
<td>tgn:7018835</td>
<td>rdfs:label</td>
<td>Tadmor</td>
</tr>
<tr>
<td>tgn:7018835</td>
<td>rdfs:label</td>
<td>Palmira</td>
</tr>
<tr>
<td>tgn:7018835</td>
<td>rdfs:label</td>
<td>Tadmur</td>
</tr>
<tr>
<td>tgn:7018835</td>
<td>rdfs:label</td>
<td>Palmyre</td>
</tr>
<tr>
<td>tgn:7018835</td>
<td>rdfs:label</td>
<td>Tamar</td>
</tr>
<tr>
<td>tgn:18835</td>
<td>rdfs:label</td>
<td>Kharāʾib Tadmur</td>
</tr>
<tr>
<td>tgn:18835</td>
<td>rdfs:label</td>
<td>Tedmor</td>
</tr>
<tr>
<td>tgn:18835</td>
<td>rdfs:label</td>
<td>Tabmur</td>
</tr>
<tr>
<td>tgn:18835</td>
<td>rdfs:label</td>
<td>Tudmor</td>
</tr>
<tr>
<td>tgn:18835</td>
<td>rdfs:label</td>
<td>خرائب تدمير</td>
</tr>
<tr>
<td>tgn:18835</td>
<td>gvp:broader</td>
<td>tgn:1002054</td>
</tr>
<tr>
<td>tgn:18835</td>
<td>gvp:broaderPartitiveExtended</td>
<td>tgn:6000039</td>
</tr>
<tr>
<td>tgn:18835</td>
<td>gvp:broaderPartitiveExtended</td>
<td>tgn:1000004</td>
</tr>
<tr>
<td>tgn:18835</td>
<td>gvp:broaderExtended</td>
<td>tgn:1002054</td>
</tr>
<tr>
<td>tgn:18835</td>
<td>gvp:broaderPartitiveExtended</td>
<td>tgn:7001526</td>
</tr>
<tr>
<td>tgn:18835</td>
<td>gvp:broaderPartitiveExtended</td>
<td>tgn:1000140</td>
</tr>
<tr>
<td>tgn:18835</td>
<td>gvp:broaderPartitiveExtended</td>
<td>tgn:7029392</td>
</tr>
<tr>
<td>tgn:18835</td>
<td>gvp:broaderExtended</td>
<td>tgn:6000039</td>
</tr>
<tr>
<td>tgn:18835</td>
<td>gvp:broaderExtended</td>
<td>tgn:1000004</td>
</tr>
</tbody>
</table>
Mogao Caves

Source: http://vocab.getty.edu/tgn/7029798

Download the dataset

http://vocab.getty.edu/tgn/7029798
Example 2.
Following these geographic places located on the Silk Road, using the geo-coordinators TGN provided, we get them on the map (through TGN’s API).
Example 2. Following these geographic places located on the Silk Road, using the geo-coordinates TGN provided, we get them on the map (through TGN’s API).
APIs

TGNWebServices

TGN Web Services

The following operations are supported. For a formal definition, please review the TGN Web Service WSDL:

- **TGNGetChildren**
  Return all immediate children of a given subject record

- **TGNGetMergedSubjectID**
  Return a list of records with old and new IDs that have been merged

- **TGNGetParents**
  Return parent hierarchy for a given subject record

- **TGNGetRevisionHistory**
  Return info on edits made within certain data based on input parameters

- **TGNGetSubject**
  Return all data element associated with a VCS subject

- **TGNGetSubjectBriefReports**
  Returns subset of data elements associated with a VCS subject

- **TGNGetSubjectTerms**
  Return termID and all terms based on subjectID

- **TGNGetSyncSubjectID**
  Return current subject ID for a given subject ID

- **TGNGetTermMatch**
  Return result count, preferred term and matching terms

TGNGetSubject:

Return all data element associated with a VCS subject

**Test**

To test the operation using the HTTP POST protocol, click the 'Invoke' button.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>subjectID</td>
<td>7031416</td>
</tr>
</tbody>
</table>

**SOAP 1.1**

The following is a sample SOAP 1.1 request and response. The placeholders shown need to be replaced with actual values:

```
POST /TGNService.asmx HTTP/1.1
Host: vocabservices.getty.edu
Content-Type: text/xml; charset=utf-8
SOAPAction: "http://vocabservices.getty.edu/TGNGetSubject"

<?xml version="1.0" encoding="utf-8"?>
    xmlns:soap="http://schemas.xmlsoap.org/soap/envelope/"
    xmlns:tns="http://vocabservices.getty.edu/"
>
    <soap:Body>
        <TGNGetSubject xmlns="http://vocabservices.getty.edu/">
            <subjectID>string</subjectID>
        </TGNGetSubject>
    </soap:Body>
</soap:Envelope>
```

```
HTTP/1.1 200 OK
Content-Type: text/xml; charset=utf-8
Content-Length: length

<?xml version="1.0" encoding="utf-8"?>
    xmlns:soap="http://schemas.xmlsoap.org/soap/envelope/"
    xmlns:tns="http://vocabservices.getty.edu/"
>
    <soap:Body>
        <TGNGetSubjectResponse xmlns="http://vocabservices.getty.edu/">
            <TGNGetSubjectResult>
            </TGNGetSubjectResult>
        </TGNGetSubjectResponse>
    </soap:Body>
</soap:Envelope>
```
This XML file does not appear to have any useful content. It appears to be a generic structure without specific data or information.
Example 2.
Following these geographic places located on the Silk Road, using the geo-coordinators TGN provided, we get them on the map (through TGN’s API).
... find places in a boundary
(through the “near/adjacent to” relationship
provided by TGN)
... launch the search for museums’ collections through a place on the map.
“Dunhuang”

Europeana:

MET:
... or connect with Wikipedia, Europeana, or Google Arts & Culture (which features content from over 1200 leading museums and archives)
Procedures

(1) Obtain the "Silk Road" ID from TGN;

(2) Using the ID to request through Getty Vocabularies Web Services APIs to get data of locations related to the "Silk Road"; obtain the datasets in XML format;

(3) Extract preferred name, non-prefer name, type of place, and Coordinates from XML file; store data in the PostgreSQL spatial database;

(4) Present data on and interact with the defined geographical area on the map using Google Map API.
Example 3

Find certain place types around the Silk Road through a **LOD Sparql Query platform**

http://vocab.getty.edu/

I would like to get something like "caves on or around the Silkroad",
or, "UNESCO Heritage sites on or around the Silkroad".

All these place types are defined in the Getty Art and Architecture Thesaurus (AAT)
Go to: http://vocab.getty.edu/

Choose “Queries”

4.18 Places by Type Within Bounding Box

Let's specialize the previous query and look for castles around The Netherlands, we get 170:

```sparql
PREFIX ontogeo: <http://www.ontotext.com/owlim/geo#>
SELECT DISTINCT *
WHERE {
  ?place skos:inScheme tgn: ;
  gvp:placeType/ (gvp:placeType/gvp:broadernGenericExtended) [rdfs:label "castles (fortifications)"@en];
  foaf:focus [ontogeo:within (50.78715 3.389722 53.542265 7.169019)];
  gvp:prefLabelGVP [xl:literalForm ?name];
  gvp:parentString ?parents
}
```
Query a specific place type (e.g., caves) in a geographic boundary.

Replace the place type you choose, put your geo coordinators.

```
select distinct * {
  ?place skos:inScheme tgn. ;
  gvp:placeType | (gvp:placeType/gvp:broadergenericExtended) [rdfs:label "caves"@en];
  foaf:focus [ontogeo:within(24.75083 28.95778 43.80722 108.92861)];
  gvp:prefLabelGVP [xl:literalForm ?name];
  gvp:parentString ?parents}
```
Query a specific place type (e.g., caves) in a geographic boundary

Got the results & downloadable datasets:

Caves within (24.75 083 28.95778 43.80722 108.92861)
Query a specific place type (e.g., World Heritage Sites) in a geographic boundary. Got the results & downloadable datasets:

World Heritage Sites within (24.75083 28.95778 43.80722 108.92861)
Summary

1. Learn through TGN Website
   [http://www.getty.edu/research/tools/vocabularies/tgn/](http://www.getty.edu/research/tools/vocabularies/tgn/)
   (Anyone can do!)

2. Following these geographic places located on the Silk Road, using the geo-coordinators TGN provided, get them on the map (through TGN’s API).
   (Need someone who can play with API and write a little bit Java.)
   - The APIs are available to any institution having a login, which may be obtained by writing to [vocab@getty.edu](mailto:vocab@getty.edu). See details in the [Web Services User’s Instructions](http://www.getty.edu/research/tools/vocabularies/tgn/) (PDF).

3. Find certain place types around the Silk Road through a LOD Sparql Query platform
   [http://vocab.getty.edu/](http://vocab.getty.edu/)
   (Anyone can use the template to query; follow our simple demos; knowledge of SPARQL queries will be ideal.)