

Collection Space Stuff

REST API

intro

How CSpace works: Use HTTP requests to connect with the server and query data from those pages

URLs should follow this format:

`https://<your-cspace-domain>/cspace-services/<authorityType>/<CSID>/items?pgSz=<pageSize>`

Base URL/CSpace Domain

The base URL connects to an individual CSpace instance. Our base URL is,

<https://ucbg.collectionspace.org/cspace-services>

Authority/Object Types

Then, in addition to the base URL, make a request from a specific source e.g.

orgauthorities

The following is a complete(?) list of authority types ([source](#)). Some are defunct:

citationauthorities, conceptauthorities, locationauthorities, orgauthorities, personauthorities, placeauthorities, taxonomyauthority, workauthorities

Append the authority type to the end of the base URL e.g.

<https://ucbg.collectionspace.org/cspace-services/orgauthorities/>

***In addition to accessing a list of objects through Authority Types, you can also access individual items by following the URI tag (highlighted for the below item). In this case append the URI to our base URL e.g.,

This XML file does not appear to have any style information associated with it. The document tree is shown below.

```
<?xml version="1.0" encoding="UTF-8" standalone="yes"?>
<ns2:abstract-common-list xmlns:ns2="http://collectionspace.org/services/jaxb">
  <pageNum>0</pageNum>
  <pageSize>40</pageSize>
  <itemsInPage>40</itemsInPage>
  <totalItems>35224</totalItems>
  <fieldsReturned>csid|uri|refName|updatedAt|workFlowState|rev|shortIdentifier|sas|proposed|deprecated|taxonRank|termStatus|termDisplayName</fieldsReturned>
  <list-item>
    <csid>42a9e8bb-13e5-426d-831c-6a6d638defa5</csid>
    <uri>/taxonomyauthority/c1662cc5-d458-4788-96ed/items/42a9e8bb-13e5-426d-831c-6a6d638defa5</uri>
    <refName>urn:cspace:botgarden.cspace.berkeley.edu:taxonomyauthority:name(taxon):item:name(201159)'Abelia'</refName>
    <updatedAt>2013-04-24T18:49:00.000Z</updatedAt>
    <workFlowState>project</workFlowState>
    <shortIdentifier>201159</shortIdentifier>
    <taxonRank>genus</taxonRank>
    <termStatus>accepted</termStatus>
    <termDisplayName>Abelia</termDisplayName>
  </list-item>
</abstract-common-list>
```

`https://ucbg.collectionspace.org/cspace-services/taxonomyauthority/c1662cc5-d458-4788-96ed/items/42a9e8bb-13e5-426d-831c-6a6d638defa5`

CSID/UUID

Each authority type has one or more CSID which corresponds to a different set of objects in the database. The CSID highlighted below corresponds to 'Common Taxonomy Authority' which can be found within the "display name" tags. Notice how there are two other CSIDs present on this page. This means taxonomyauthority has three subtypes which can all be accessed using taxonomy authority along with their respective CSID.

This XML file does not appear to have any style information associated with it. The document tree is shown below.

```
<?xml version="1.0" encoding="UTF-8" standalone="yes" ?>
<ns2:abstract-common-list xmlns:ns2="http://collectionspace.org/services/jaxb">
  <pageNum>0</pageNum>
  <pageSize>40</pageSize>
  <itemsInPage>3</itemsInPage>
  <totalItems>3</totalItems>
  <fieldsReturned>csid|uri|refName|updatedAt|workflowState|rev|shortIdentifier|sas|displayName|vocabType</fieldsReturned>
  <list-item>
    <csid>47c2a276-868b-482d-888d</csid>
    <uri>/taxonomyauthority/47c2a276-868b-482d-888d</uri>
    <refName>urn:cspace:botgarden.cspace.berkeley.edu:taxonomyauthority:name(common)'Common Taxonomy Authority'</refName>
    <updatedAt>2013-04-24T01:52:06.000Z</updatedAt>
    <workflowState>project</workflowState>
    <rev>277</rev>
    <shortIdentifier>common</shortIdentifier>
    <displayName>Common Taxonomy Authority</displayName>
  </list-item>
  <list-item>
    <csid>e773d53a-d65b-4b6a-bd2c</csid>
    <uri>/taxonomyauthority/e773d53a-d65b-4b6a-bd2c</uri>
    <refName>urn:cspace:botgarden.cspace.berkeley.edu:taxonomyauthority:name(plantsales)'Plant Sale Taxonomy Authority'</refName>
    <updatedAt>2013-04-24T01:52:07.000Z</updatedAt>
    <workflowState>project</workflowState>
    <rev>9</rev>
    <shortIdentifier>plantsales</shortIdentifier>
    <displayName>Plant Sale Taxonomy Authority</displayName>
  </list-item>
  <list-item>
    <csid>c1662cc5-d458-4788-96ed</csid>
    <uri>/taxonomyauthority/c1662cc5-d458-4788-96ed</uri>
    <refName>urn:cspace:botgarden.cspace.berkeley.edu:taxonomyauthority:name(taxon)'Default Scientific Taxonomy Authority'</refName>
    <updatedAt>2013-04-24T01:52:06.000Z</updatedAt>
    <workflowState>project</workflowState>
    <rev>10066</rev>
    <shortIdentifier>taxon</shortIdentifier>
    <displayName>Default Scientific Taxonomy Authority</displayName>
  </list-item>
</ns2:abstract-common-list>
```

If you have access to a python interpreter, you can use this code to find the first CSID. Simply replace the url with the url of interest and fill in your username and password.

```
import requests

from requests.auth import HTTPBasicAuth

import re

url = "https://ucbg.collectionspace.org/cspace-services/taxonomyauthority"
username = "<Your prod username>"
password = "<Your prod password>"

response = requests.get(url, auth=HTTPBasicAuth(username, password))

csid = re.search('<csid>(.*?)</csid>', response.text)

print(csid.group(1))
```

```
test.py > ...
1 import requests
2 from requests.auth import HTTPBasicAuth
3 import re
4
5 url = "https://ucbg.collectionspace.org/cspace-services/taxonomyauthority"
6 username = "<Username>"
7 password = "<Password>"
8
9
10 response = requests.get(url, auth=HTTPBasicAuth(username, password))
11
12 csid = re.search('<csid>(.*?)</csid>', response.text)
13 print(csid.group(1))
14
```

Items/Page Size/Page Number

Finally, append the items and page size variable to your url

items?pgSz =< number of entries to display > &pgNum =< page Number >

Your final URL should look like this

https://ucbg.collectionspace.org/cspace-services/taxonomyauthority/47c2a276-868b-482d-888d/items?pgSz = 100&pgNum = 0

Doing Stuff

Example

The following is an example of a python program which utilizes URL construction to return all the determinations in the database

```
HTTPRequests.py > ...
1  import requests
2  from requests.auth import HTTPBasicAuth
3  import pandas as pd
4  import numpy as np
5
6
7
8  url = "https://ucbg.collectionspace.org/cspace-services/taxonomyauthority/c1662cc5-d458-4788-96ed/items?pgSz=3000"
9  username = "<username>"
10 password = "<password>"
11
12 output = []
13
14 for i in range(30):
15     url = f"https://ucbg.collectionspace.org/cspace-services/taxonomyauthority/c1662cc5-d458-4788-96ed/items?pgSz=3000&pgNum={i}"
16     response = requests.get(url, auth=HTTPBasicAuth(username, password))
17     if response.status_code == 200:
18         output.append(response.text)
19         i += 1
20     else:
21         print(f"Failed to fetch page {i}: {response.status_code}")
22         break
23
24
25
26 Df = pd.DataFrame(output, columns=["Response"])
27 X = Df['Response'].str.extractall(r"<termDisplayName>(.*?)</termDisplayName>")
28 X[0].to_csv("Determinations.csv", index=False)
```