



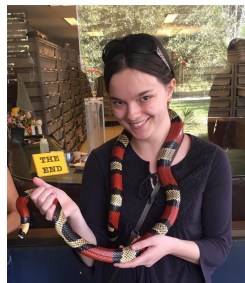
# Finding Field Station Data for Research Use



**Erica Krimmel**  
Digitization  
Resources  
ekrimmel@fsu.edu



**Gil Nelson**  
Project Director  
gnelson@  
floridamuseum.ufl.edu



**Cat Chapman**  
Biodiversity  
Informatics  
cchapman@  
floridamuseum.ufl.edu



**David Jennings**  
Project Manager  
djennings@  
floridamuseum.ufl.edu



**Molly Phillips**  
EODI  
mphillips@  
floridamuseum.ufl.edu



**Searching programmatically...  
...by using the iDigBio search API**



# What is an API?

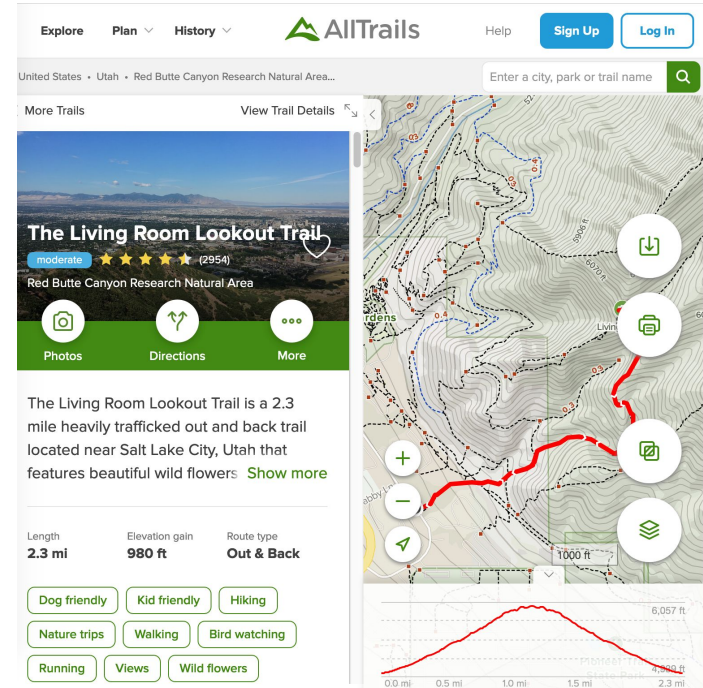
Application Programming Interface



An organization

has a system

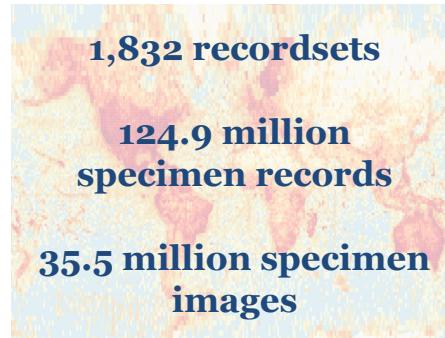
that external users can interact with





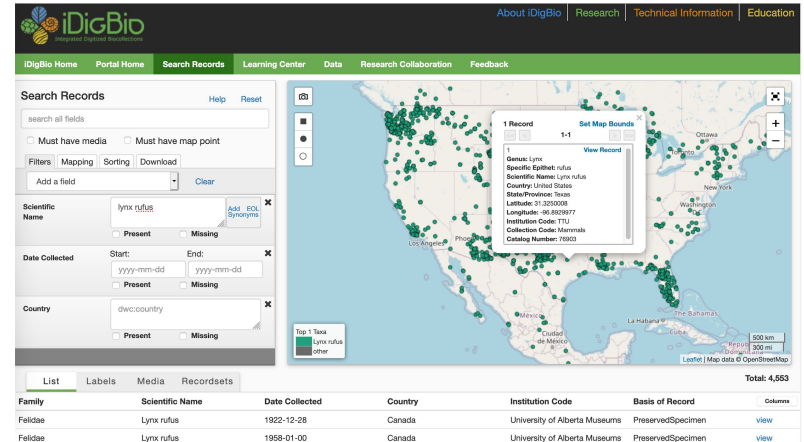
# What is an API?

## Application Programming Interface



An organization

has a system



Search Records

search all fields

Must have media  Must have map point

Filters Mapping Sorting Download

Add a field  Clear

Scientific Name: lynx rufus  Add EOL Synonym

Date Collected: Start: yyyy-mm-dd End: yyyy-mm-dd

Country: dwc:country

1 Record

Genus: Lynx  
Specific Epithet: rufus  
Scientific Name: Lynx rufus  
Country: United States  
State/Province: Texas  
Latitude: 31.3230089  
Longitude: -95.8829977  
Institution Code: TTU  
Collection Code: Mammals  
Catalog Number: 79303

Family	Scientific Name	Date Collected	Country	Institution Code	Basis of Record	Columns
Felidae	Lynx rufus	1922-12-28	Canada	University of Alberta Museums	PreservedSpecimen	<a href="#">View</a>
Felidae	Lynx rufus	1958-01-00	Canada	University of Alberta Museums	PreservedSpecimen	<a href="#">View</a>

that external users can interact with



# What is an API useful for?

An API facilitates programmatic data access, which enables...

- Automating data retrieval
- Documenting procedures within your code
- Running analyses in a way that is reproducible
- Bringing data from different sources together, e.g. via:
  - ◆ multiple APIs
  - ◆ multiple downloaded datasets
  - ◆ your own data
  - ◆ any combination of the above

And, you can bring your own programming language!



# For example...

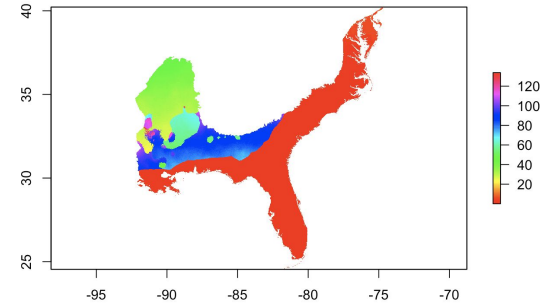
Retrieve the data

Clean the data

Bring in more  
data

Do the analysis

Visualize the  
analysis





# What is an endpoint?

An endpoint is an address, often a URL, where you can find a particular API

```
https://maps.googleapis.com/maps/api/directions/
```

↳ Let me ask Google Maps for directions

```
https://maps.googleapis.com/maps/api/geocode/
```

↳ Let me ask Google Maps to find a place on a map

```
https://search.idigbio.org/v2/search/
```

↳ Let me ask iDigBio to look for certain specimen records

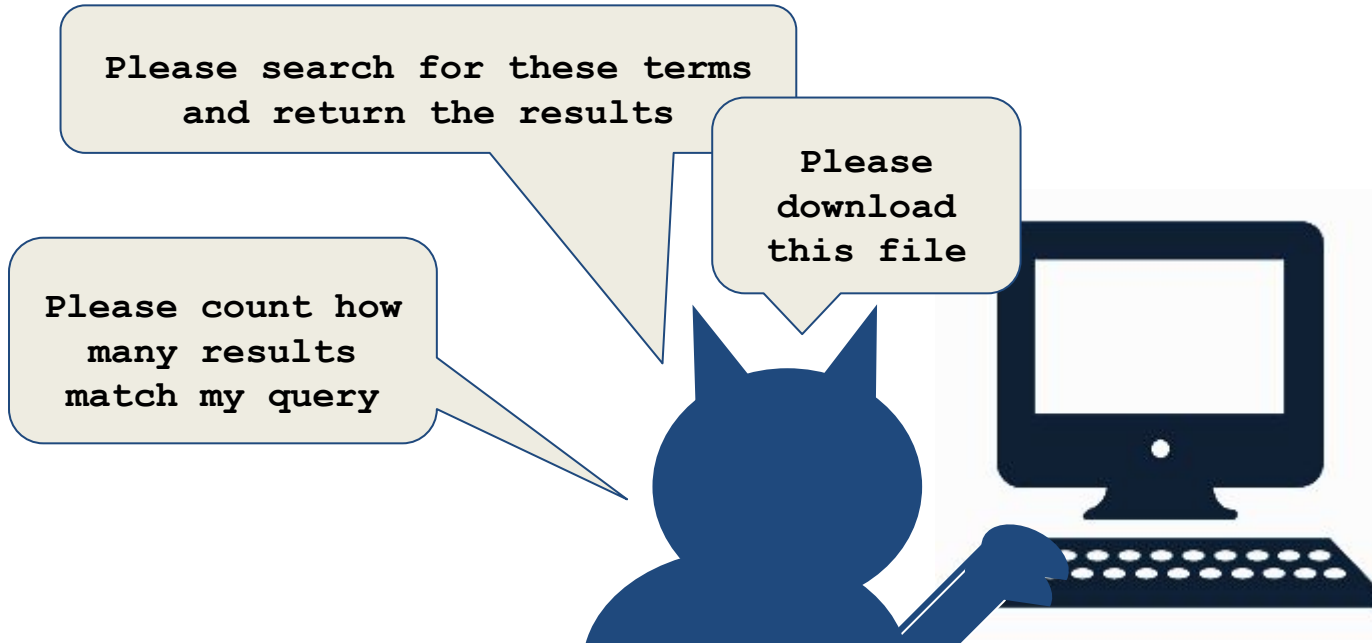
```
https://search.idigbio.org/v2/download/
```

↳ Let me ask iDigBio to download certain specimen records



# How do I talk to an API?

You talk to an API by making a request, either through a user interface or programmatically, e.g. via R or Python.

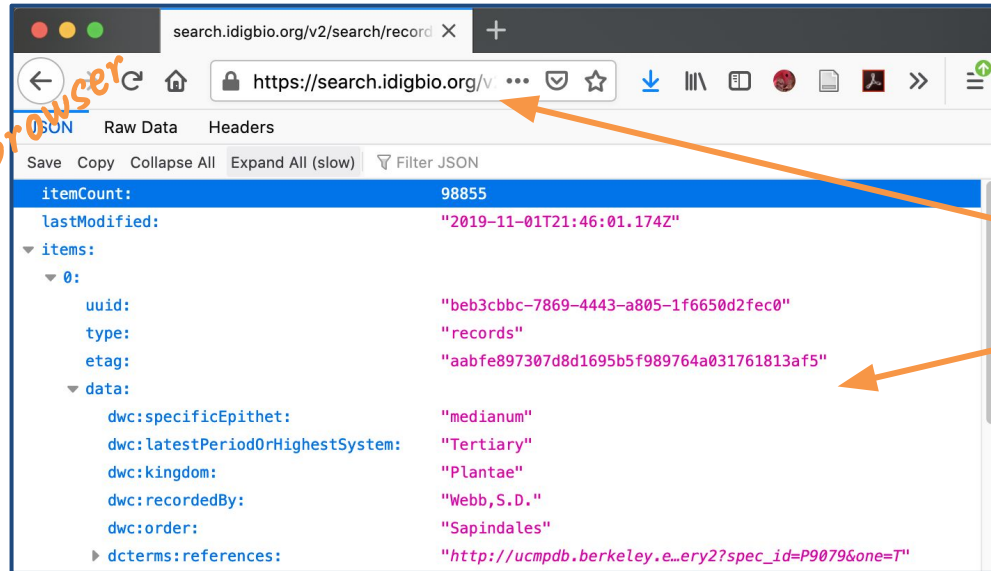






# What does a request look like?

You can talk to an API by making a request, often including parameters, to an endpoint



directly in your browser

request

results

```

{
  "itemCount": 98855,
  "lastModified": "2019-11-01T21:46:01.174Z",
  "items": [
    {
      "uuid": "beb3cbbc-7869-4443-a805-1f6650d2fec0",
      "type": "records",
      "etag": "aabfe897307d8d1695b5f989764a031761813af5",
      "data": {
        "dwc:specificEpithet": "medianum",
        "dwc:latestPeriodOrHighestSystem": "Tertiary",
        "dwc:kingdom": "Plantae",
        "dwc:recordedBy": "Webb, S.D.",
        "dwc:order": "Sapindales",
        "dcterms:references": "http://ucmpdb.berkeley.e...ery?spec_id=P9079&one=T"
      }
    }
  ]
}

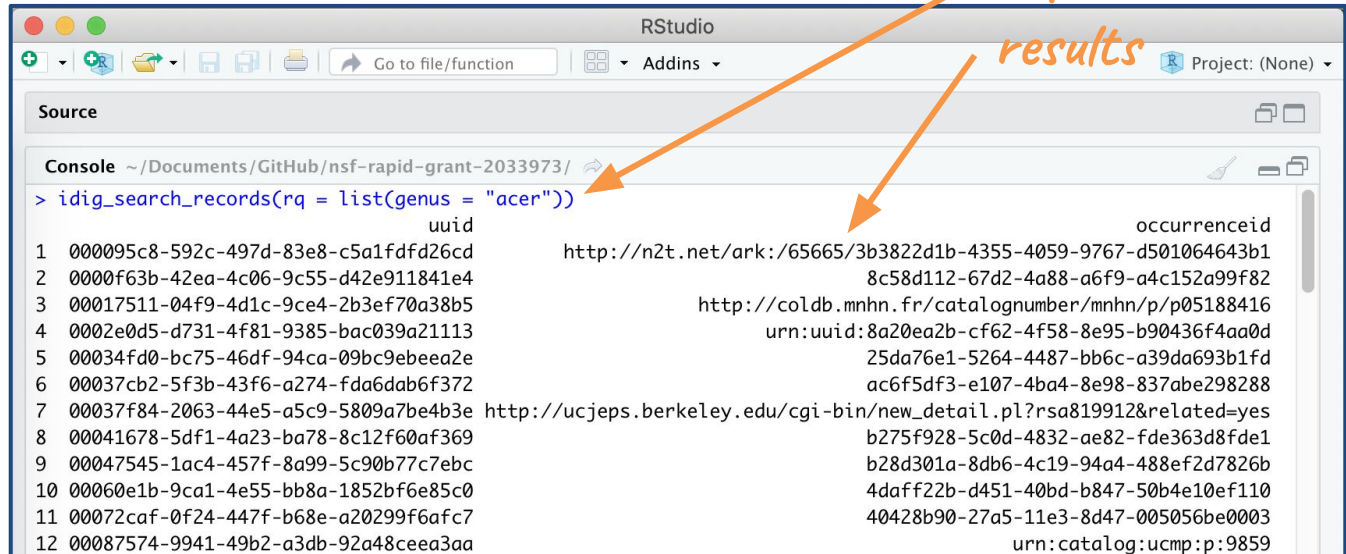
```

[https://search.idigbio.org/v2/search/records?rq={"genus":"acer"}](https://search.idigbio.org/v2/search/records?rq={)



# What does a request look like?

You can talk to an API by making a request, often including parameters, to an endpoint



```

RStudio
Go to file/function
Addins
Project: (None)

Source

Console ~/Documents/GitHub/nsf-rapid-grant-2033973/
> idig_search_records(rq = list(genus = "acer"))
      uuid                                     occurrenceid
1  000095c8-592c-497d-83e8-c5a1fdfd26cd      http://n2t.net/ark:/65665/3b3822d1b-4355-4059-9767-d501064643b1
2  0000f63b-42ea-4c06-9c55-d42e911841e4      8c58d112-67d2-4a88-a6f9-a4c152a99f82
3  00017511-04f9-4d1c-9ce4-2b3ef70a38b5      http://coldb.mnhn.fr/catalognumber/mnhn/p/p05188416
4  0002e0d5-d731-4f81-9385-bac039a21113      urn:uuid:8a20ea2b-cf62-4f58-8e95-b90436f4aa0d
5  00034fd0-bc75-46df-94ca-09bc9ebeeae2e      25da76e1-5264-4487-bb6c-a39da693b1fd
6  00037cb2-5f3b-43f6-a274-fda6dab6f372      ac6f5df3-e107-4ba4-8e98-837abe298288
7  00037f84-2063-44e5-a5c9-5809a7be4b3e      http://ucjeps.berkeley.edu/cgi-bin/new_detail.pl?rsa819912&related=yes
8  00041678-5df1-4a23-ba78-8c12f60af369      b275f928-5c0d-4832-ae82-fde363d8fde1
9  00047545-1ac4-457f-8a99-5c90b77c7ebc      b28d301a-8db6-4c19-94a4-488ef2d7826b
10 00060e1b-9ca1-4e55-bb8a-1852bf6e85c0      4daff22b-d451-40bd-b847-50b4e10ef110
11 00072caf-0f24-447f-b68e-a20299f6afc7      40428b90-27a5-11e3-8d47-005056be0003
12 00087574-9941-49b2-a3db-92a48ceea3aa      urn:catalo:ucmp:p:9859
  
```

*in a programming environment, like R*



# What is an API useful for?

An API facilitates programmatic data access, e.g. using the iDigBio API in R

*Document procedures  
within your code*

*Automate data retrieval*

*Run analyses in a way  
that is reproducible*

```
# Load package to access iDigBio API
library(ridigbio)

# Retrieve records for specimens identified as being in the genus Acer and
# collected in Utah
data <- idig_search_records(rq = list(genus = "acer",
                                     stateprovince = "utah"),
                           fields = c("uuid",
                                     "institutioncode",
                                     "collectioncode",
                                     "catalognumber",
                                     "stateprovince",
                                     "county",
                                     "locality",
                                     "geopoint"),
                           limit = 1000)

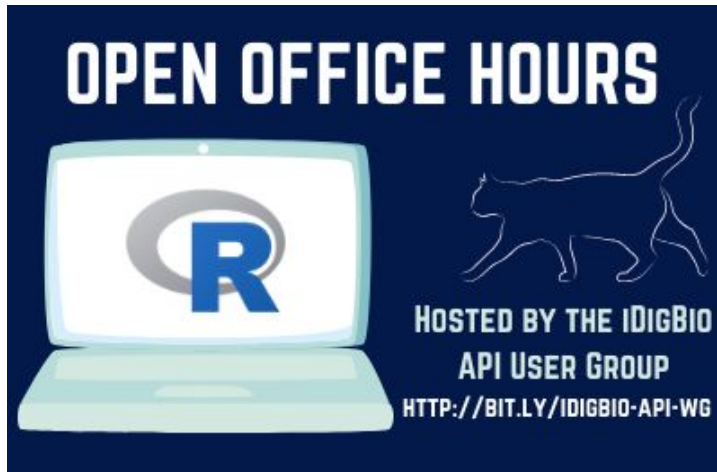
# Begin data cleaning and analysis...
```



# Resources to learn more

iDigBio office hours: an informal drop-in session where anyone is welcome to bring their questions or ideas about using tools such as the iDigBio API to work with biodiversity occurrence data.

*Every 2nd and 4th Wednesday of the month at 3:30pm Eastern, **next is October 14th***



- Focus on using the R language
- Often will do code demos but experience is not required or expected
- More details at <https://bit.ly/2Z5iYul>
- Example API code at <https://bit.ly/bio-spm-data>



# Thank you!



[www.idigbio.org](http://www.idigbio.org)



[facebook.com/iDigBio](https://facebook.com/iDigBio)



[twitter.com/iDigBio](https://twitter.com/iDigBio)



[vimeo.com/idigbio](https://vimeo.com/idigbio)



[idigbio.org/rss-feed.xml](http://idigbio.org/rss-feed.xml)



[idigbio.org/events-calendar/export.ics](http://idigbio.org/events-calendar/export.ics)