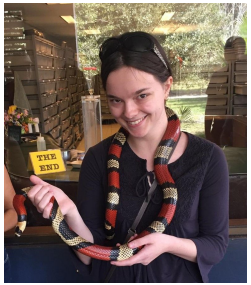
Finding Field Station Data for Research Use

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

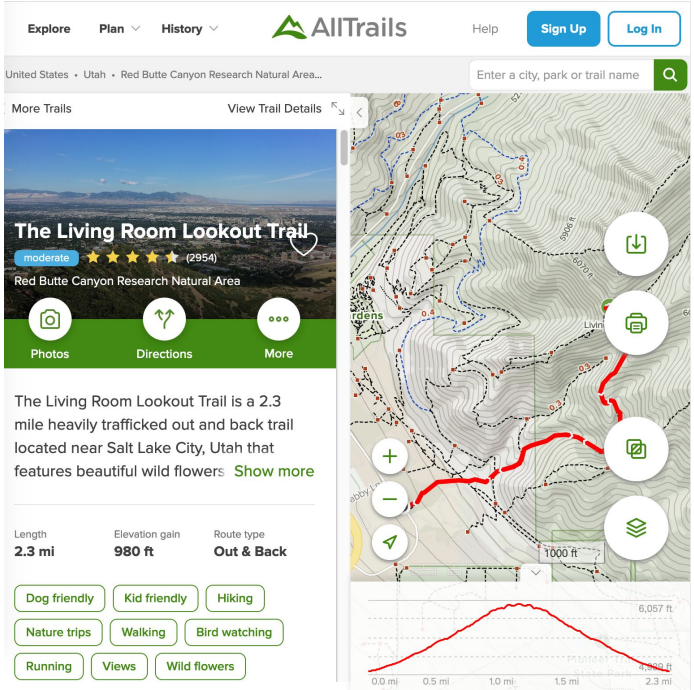
Molly Phillips
EODI
mphillips@floridamuseum.ufl.edu

iDigBio is funded by grants from the National Science Foundation's Advancing Digitization of Biodiversity Collections Program [DBI-1115210 (2011-2018) and DBI-1547229 (2016-2021)]. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

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?

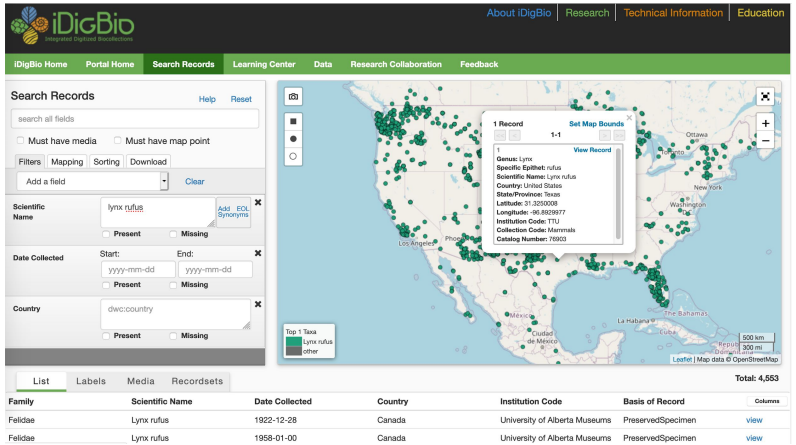
Application Programming Interface

iDigBio
Integrated Digitized Biocollections

1,832 recordsets

124.9 million specimen records

35.5 million specimen images



An organization has a system 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!

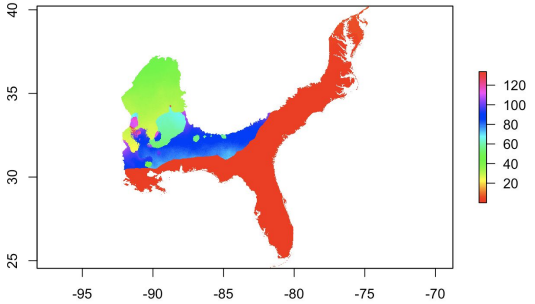
For example…

Retrieve the data → Clean the data → Bring in more data → Do the analysis → Visualize the analysis

GBIF Global Biodiversity Information Facility

USGS science for a changing world

WorldClim



iDigBio Integrated Digitized Biocollections

See a full example at https://github.com/mgaynor1/CURE-FL-Plants

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/`

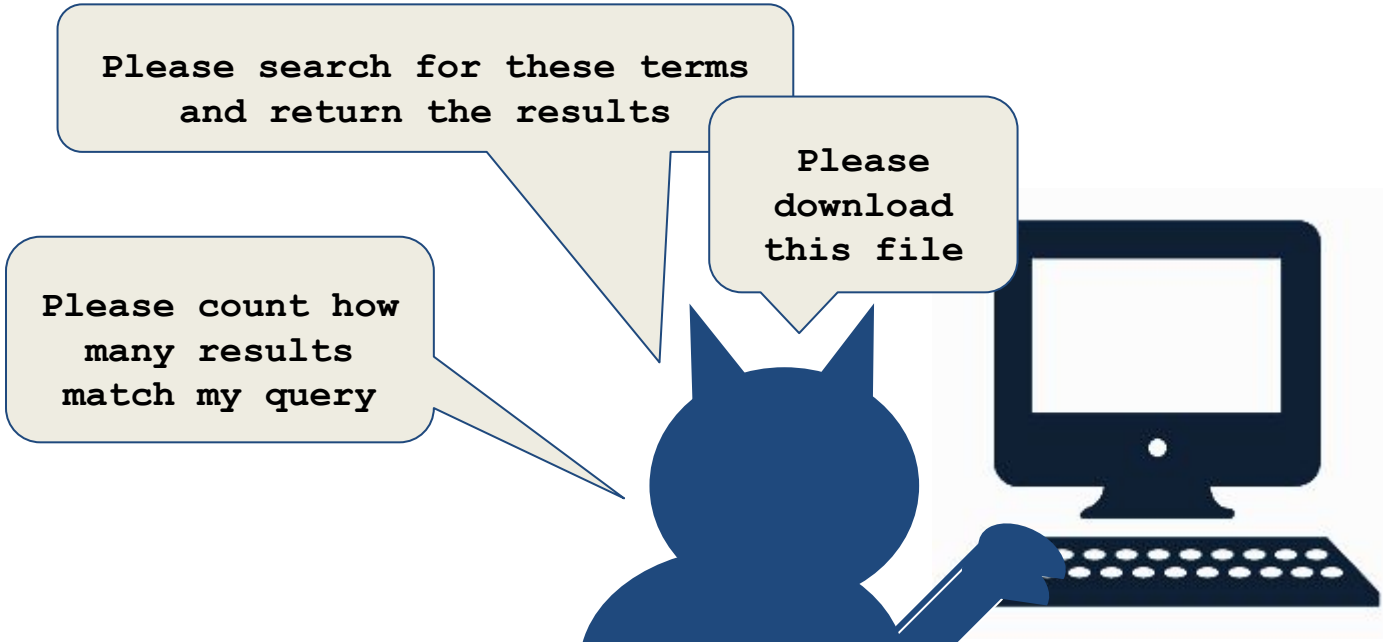
↳ Let me ask iDigBio to look for certain specimen records

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

↳ Let me ask iDigBio to download certain specimen records

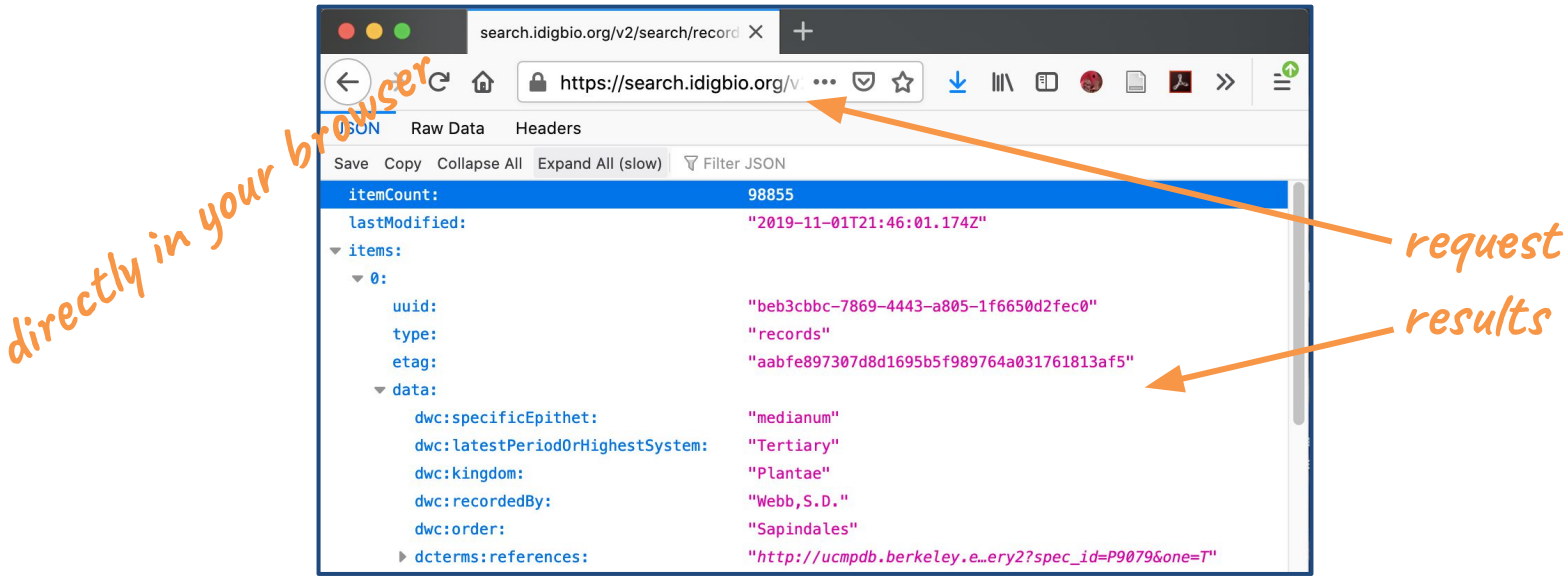
How do I talk to an API?

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



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

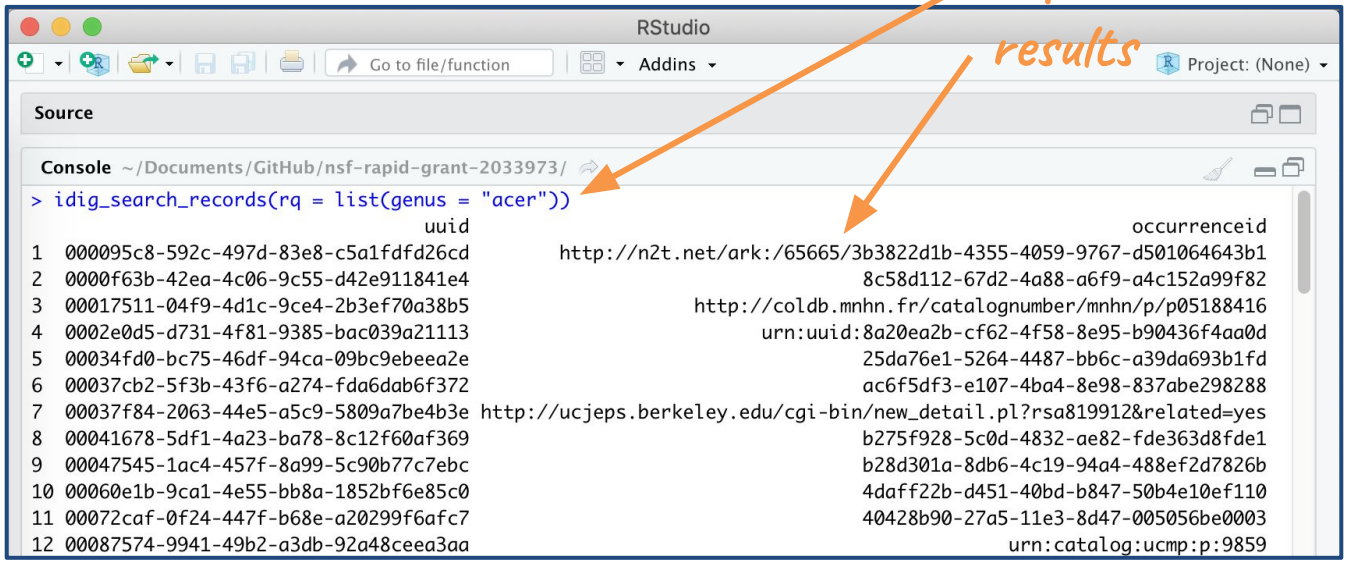
What does a request look like?

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

in a programming environment, like R

request

results



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/2Z5iYuI
- Example API code at https://bit.ly/bio-spm-data

iDigBio
Integrated Digitized Biocollections

Thank you!

www.idigbio.org

facebook.com/iDigBio

twitter.com/iDigBio

vimeo.com/idigbio

idigbio.org/rss-feed.xml

idigbio.org/events-calendar/export.ics

iDigBio is funded by grants from the National Science Foundation's Advancing Digitization of Biodiversity Collections Program [DBI-1115210 (2011-2018) and DBI-1547229 (2016-2021)]. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.