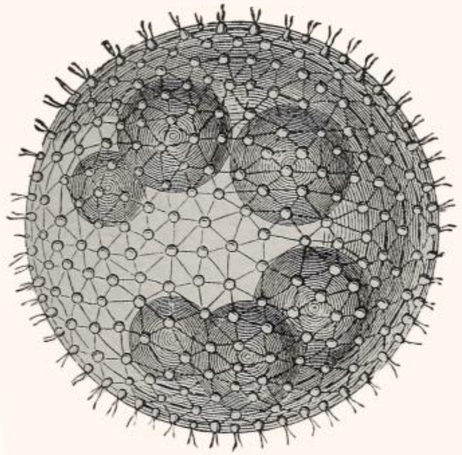
Data Use Club



Accessing and downloading species information

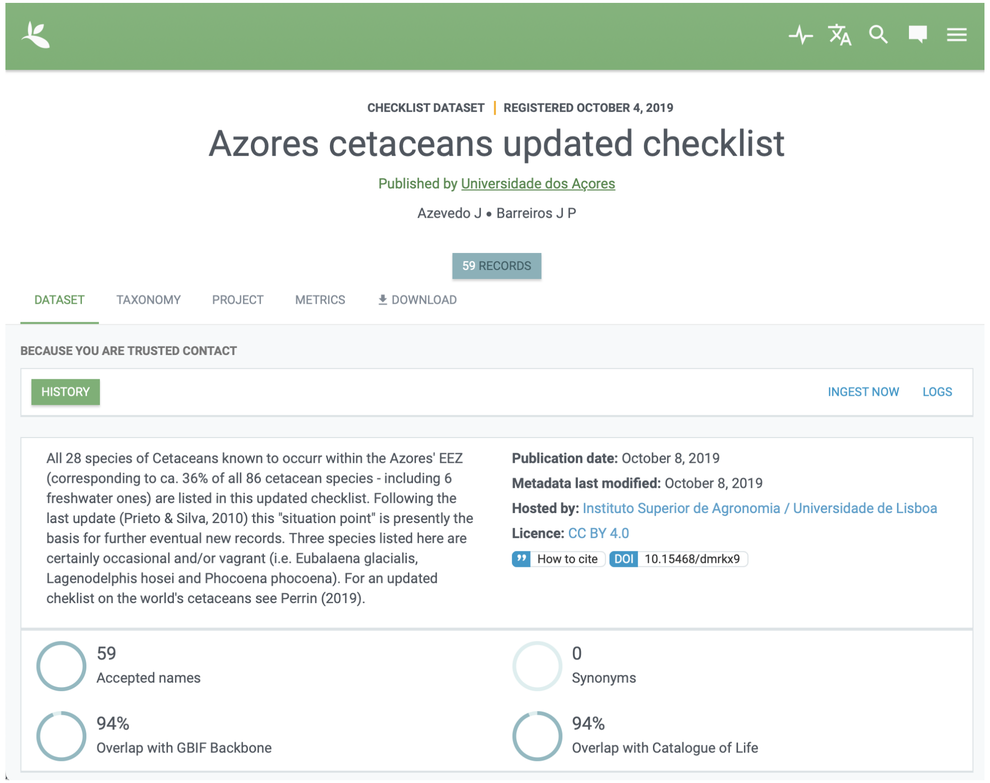
Practical sessions

Accessing species information from GBIF

GBIF.org and GBIF Species API

Ceratodon purpureus (Hedw.) Brid. observed in Belgium by Frederik V. Holsbeeck (licensed under http://creativecommons.org/licenses/by-nc/4.0/)

Checklists on GBIF



- Scientific names of organisms sharing a common theme or feature
- For example:
  - Taxonomy of a group: Sphagnopsida
  - Same geographic scope: Kenya
  - Theme: medicinal use

Checklists on GBIF – structure

Checklists on GBIF – structure – taxon core

Taxon

taxonID | scientificNameID | acceptedNameUsageID | parentNameUsageID | originalNameUsageID
nameAccordingToID | namePublishedInID | taxonConceptID | scientificName | acceptedNameUsage
parentNameUsage | originalNameUsage | nameAccordingTo | namePublishedIn | namePublishedInYear
higherClassification | kingdom | phylum | class | order | superfamily | family | subfamily | tribe | subtribe
genus | genericName | subgenus | infragenericEpithet | specificEpithet | infraspecificEpithet | cultivarEpithet
taxonRank | verbatimTaxonRank | scientificNameAuthorship | vernacularName | nomenclaturalCode
taxonomicStatus | nomenclaturalStatus | taxonRemarks

Taxon	Class
Identifier	http://rs.tdwg.org/dwc/terms/Taxon
Definition	A group of organisms (sensu http://purl.obolibrary.org/obo/OBI_0100026) considered by taxonomists to form a homogeneous unit.
Comments	
Examples	`the genus Truncorotaloides as published by Brönnimann et al. in 1953 in the Journal of Paleontology Vol. 27(6) p. 817–820`

https://dwc.tdwg.org/terms/#taxon

On this page

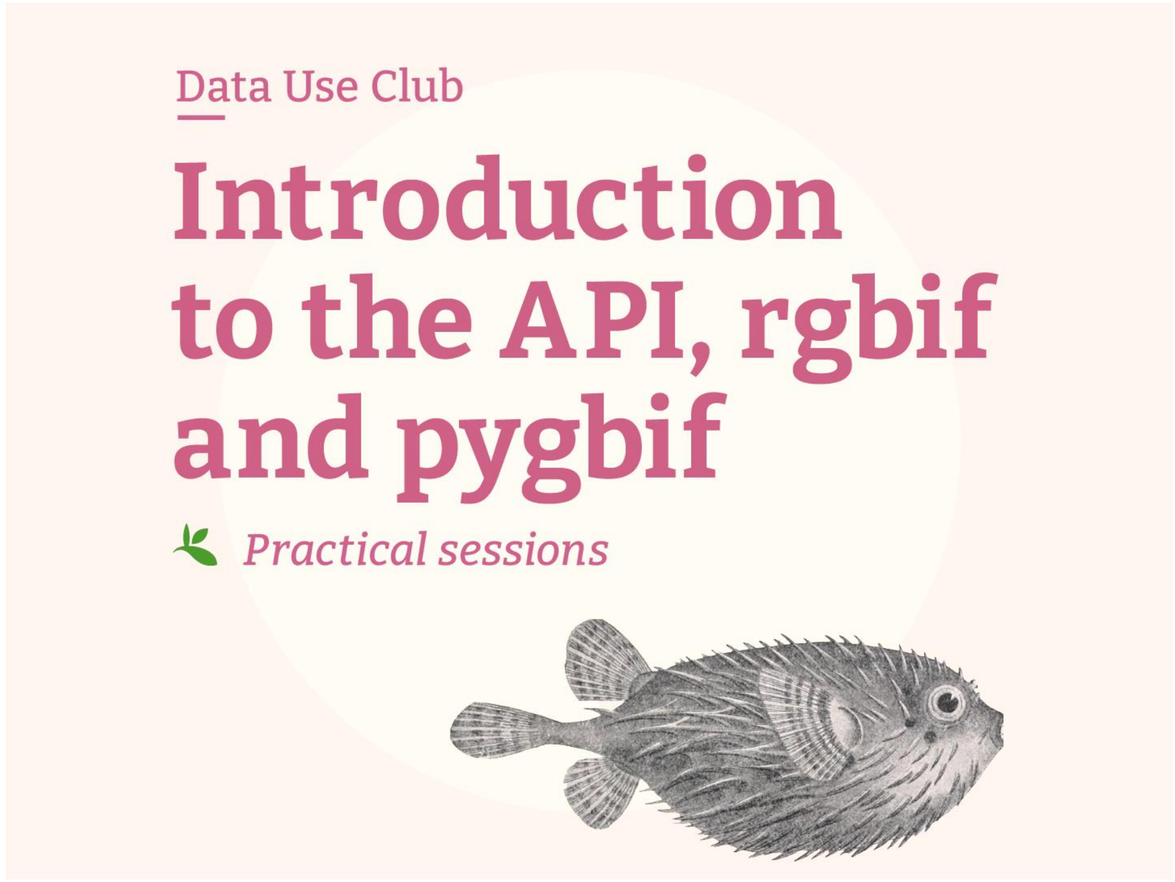
Checklists on GBIF – structure - possible extensions

GBIF Backbone vs other checklists

- GBIF Backbone Taxonomy:
  - Specific checklist
  - Assembled from selected checklists
  - All names organized and matched to it

- Other checklists:
  - All names are matched to the Backbone
  - Source of additional information

GBIF API



- Access programmatically: occurrences, taxa, registry, maps, literature
- Limitation (everything except occurrence downloads):
  - Up to 10,000 records in total
  - 300 records per page

GBIF Species API – functions and output

species/match

- Matches to scientific names
- GBIF Backbone taxonomy
- Returns taxonomy and **usageKey**

https://data-blog.gbif.org/post/gbif-species-api/

GBIF Species API – functions and output

- Matches to scientific names, description and vernacular name
- Any GBIF checklist
- Additional filters: threat, habitat, etc.
- Returns taxonomy, key and **nubKey**

https://data-blog.gbif.org/post/gbif-species-api/

GBIF Species API – functions and output

species/match

species/search

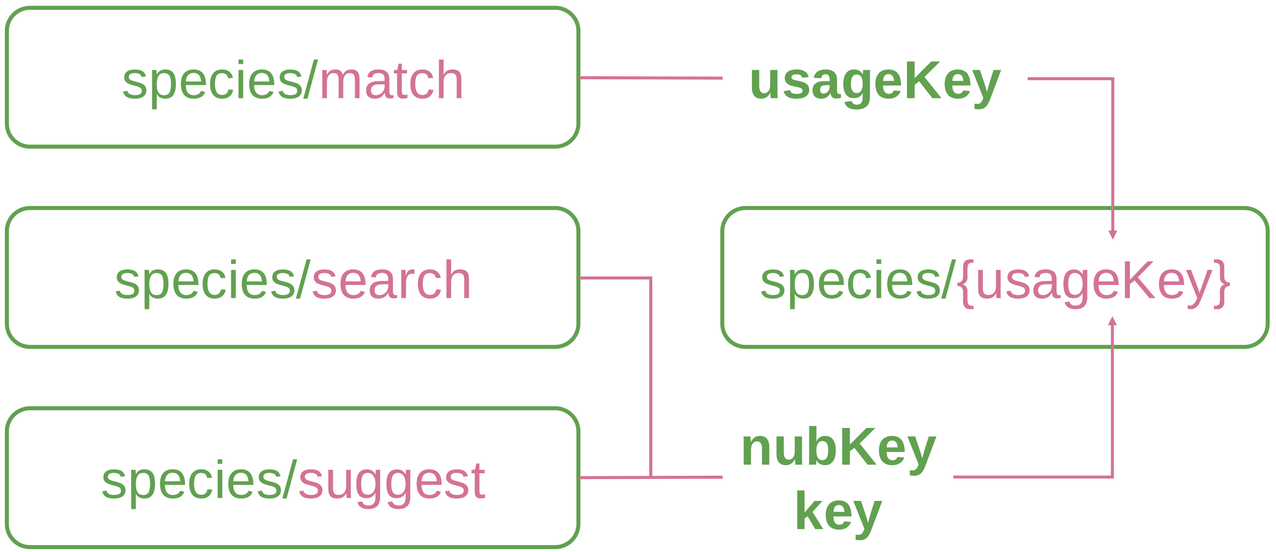
species/suggest

species/{usageKey}

- Specific taxon record
- From any checklist
- Vernacular names, synonyms, metrics, etc.

https://data-blog.gbif.org/post/gbif-species-api/

GBIF Species API – functions and output



https://data-blog.gbif.org/post/gbif-species-api/

GBIF Species API – functions and output

species/match

species/search

species/suggest

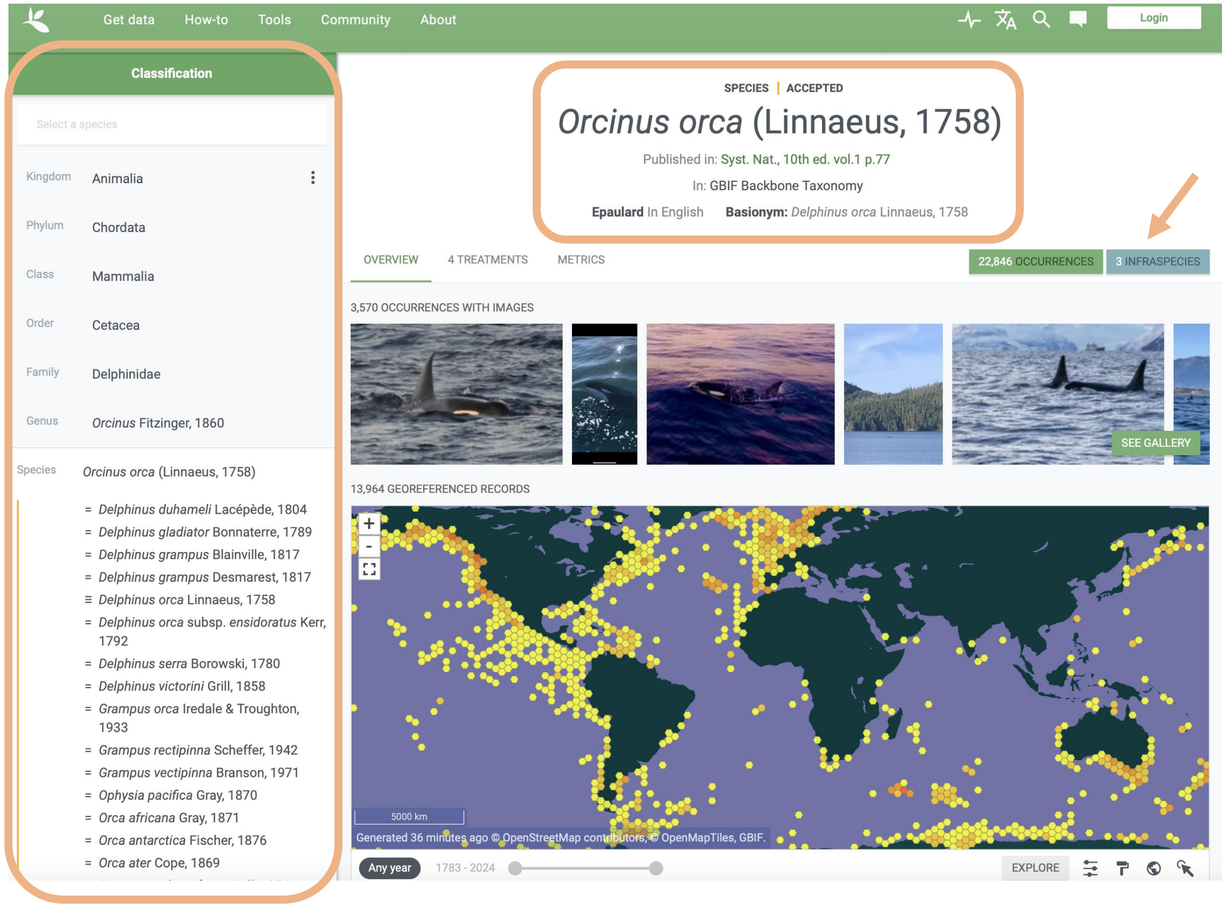
species/{usageKey}

parser/name

https://data-blog.gbif.org/post/gbif-species-api/

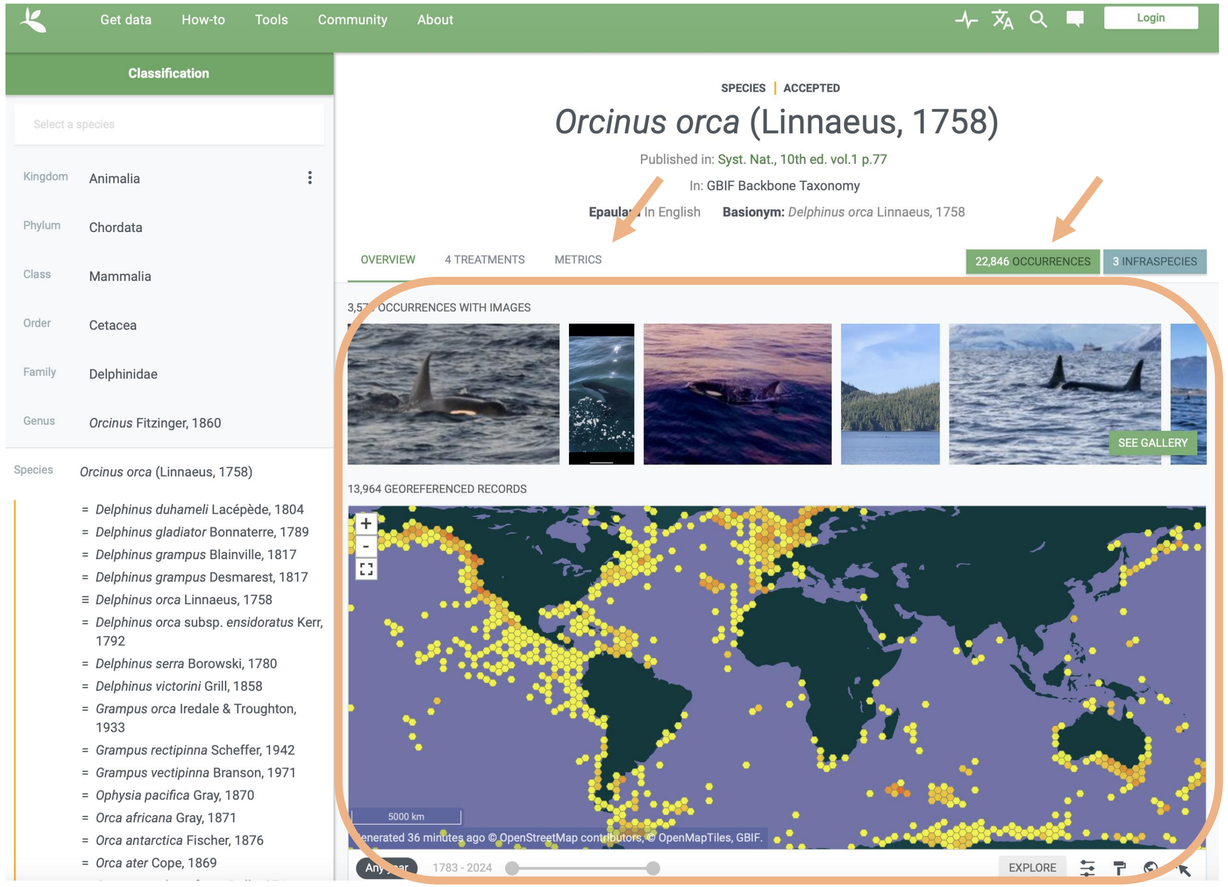
Taxon page – backbone taxonomy

species/{usageKey}/



- Name and authorship
- Taxonomic status
- Reference publication
- Source
- Classification
- Child taxa

Taxon page – occurrences aggregated

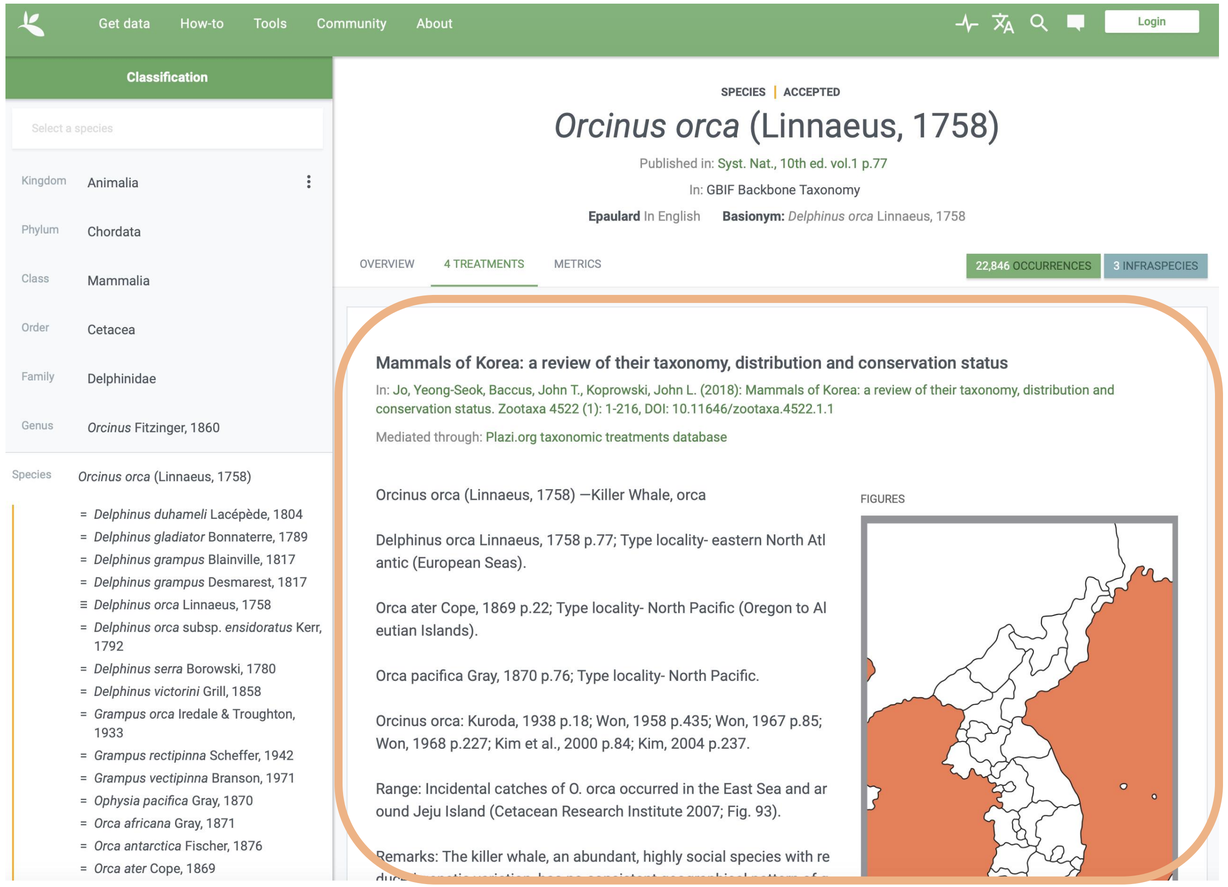


occurrence/search

map/occurrence/

- Metrics
- Gallery
- Map
- Type specimens
- Which occurrence datasets the name appears in

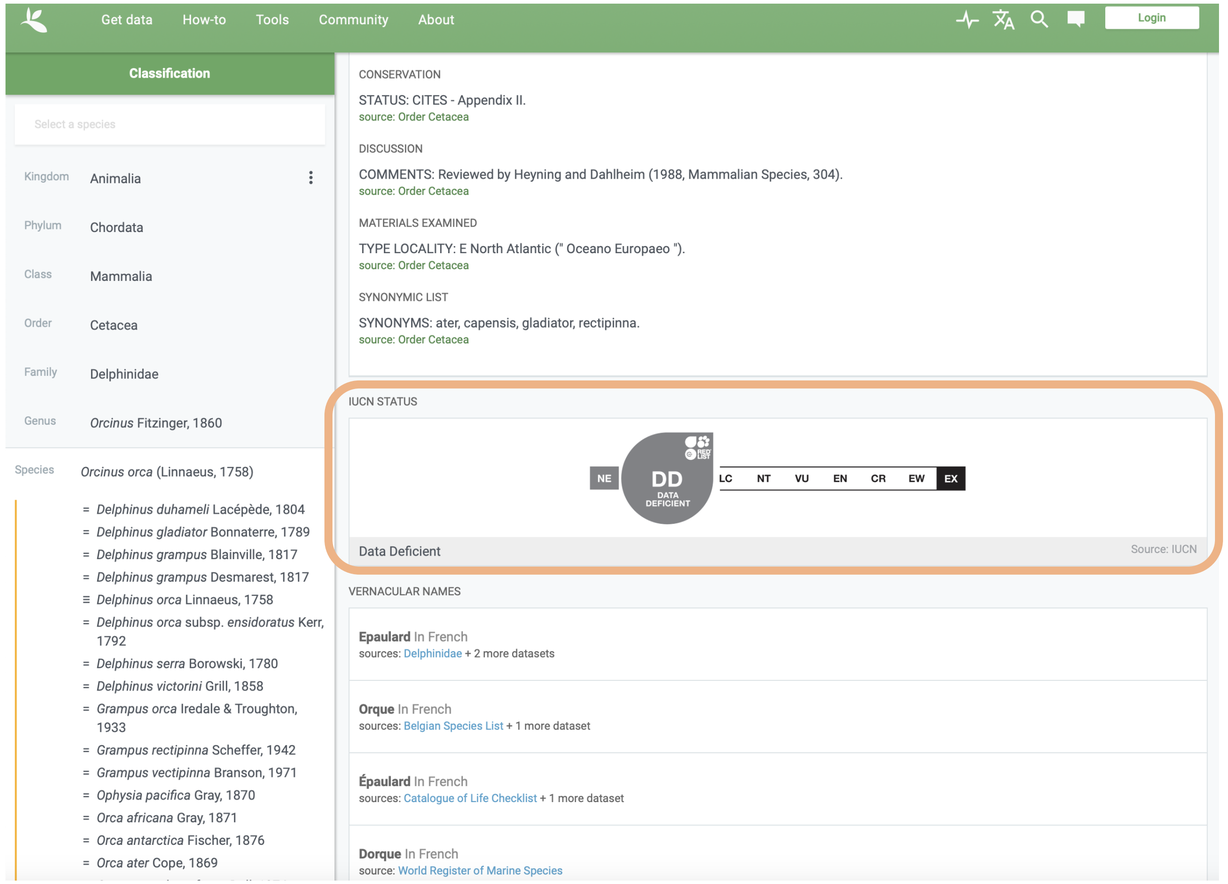
Taxon pages – treatments



- Digitization of publications published by PLAZI.org
- May contain images

https://github.com/gbif/portal16/issues/1385

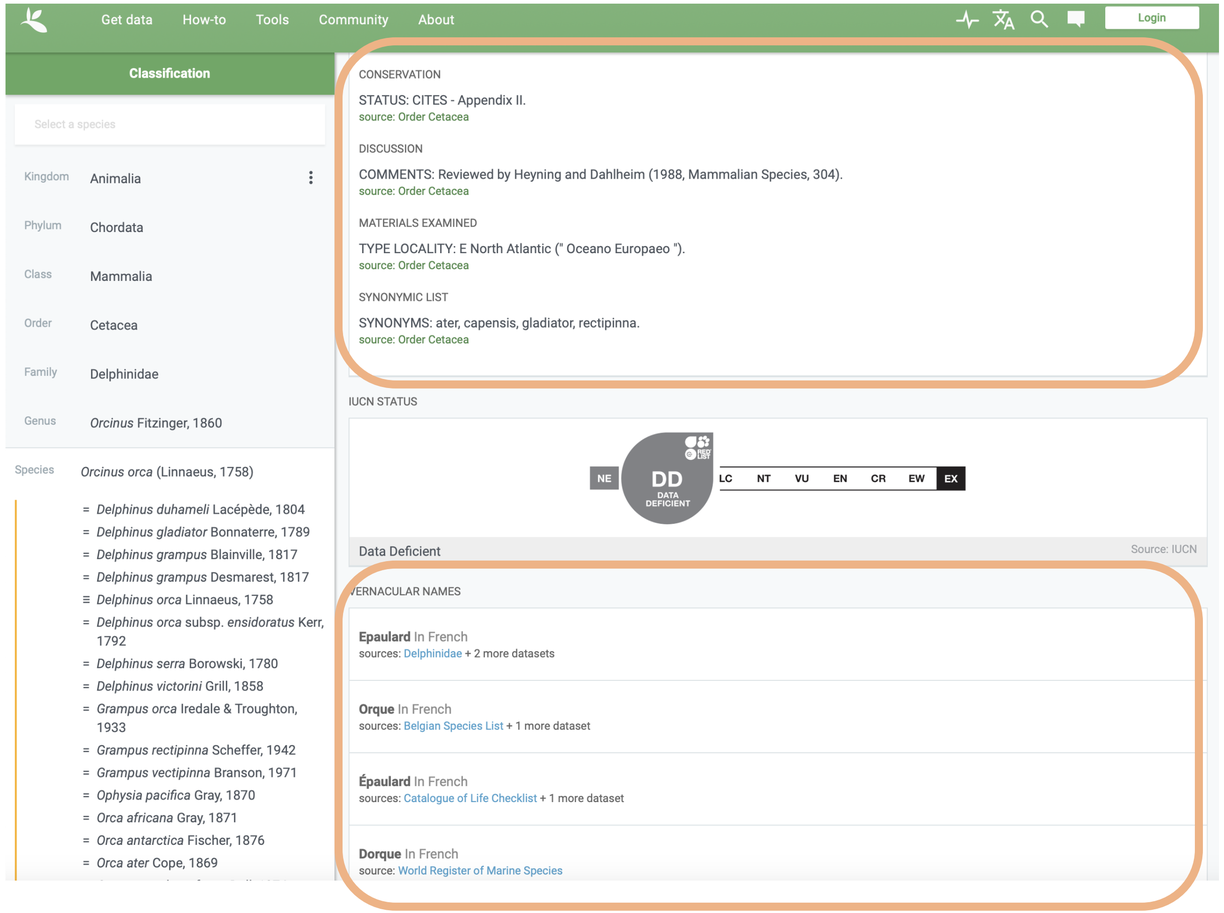
Taxon pages – IUCN Red List of Threatened Species



- The IUCN Red List of Threatened Species (https://doi.org/10.15468/0qnb58)
- IUCN Threat status

/species/{usageKey}/iucnRedListCategory

Taxon pages – any checklist

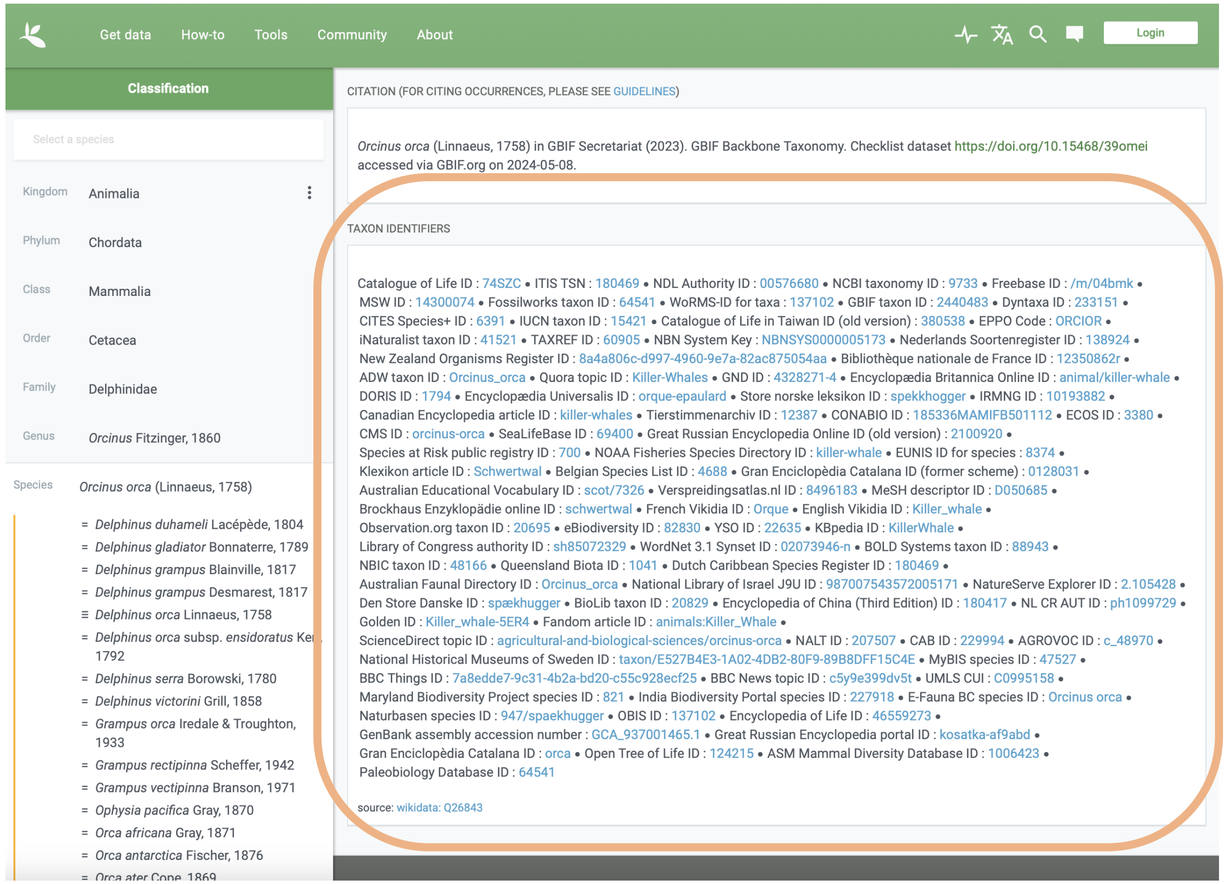


- Any checklist containing name matched to the taxon:
  - Description from taxon description extension
  - Vernacular name from vernacular name extension

/species/{usageKey}/vernacularNames

/species/{usageKey}/descriptions

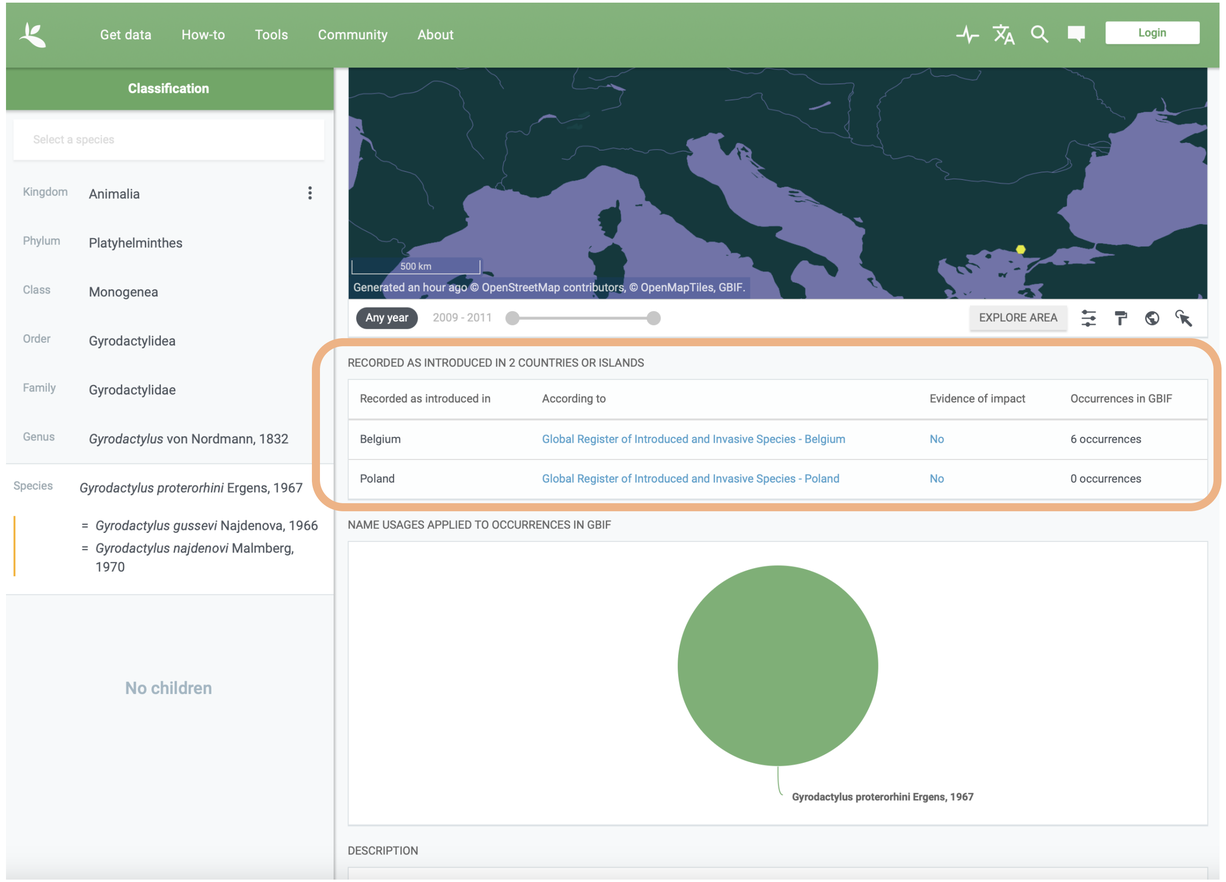
Taxon pages – Wikidata



- Taxon identifiers
- https://www.wikidata.org/
- Directly queried from the wikidata API: https://query.wikidata.org/

Taxon pages – Global Register of Introduced and Invasive Species checklists

- Introduction and invasive status
- Checklists published on GBIF by the Invasive Species Specialist Group ISSG
- One Checklist per country/area
- Combined with occurrences available on GBIF



https://github.com/gbif/portal16/issues/883#issuecomment-784536216

Example

How to access names of extinct species

Corynactis californica Carlgren, 1936observed in United States of America by thetidepooler (licensed under http://creativecommons.org/licenses/by-nc/4.0/)

Working with extinct species on GBIF

Meganeura selysii (Brongniart, 1894) collected in France (licensed under http://creativecommons.org/licenses/by/4.0/)

GBIF users are often interested in getting occurrences of only living species or “**no extinct**” species.

This seemingly straightforward request is often difficult to fulfill in practice.

isExtinct

GBIF has a species search term called **isExtinct**.

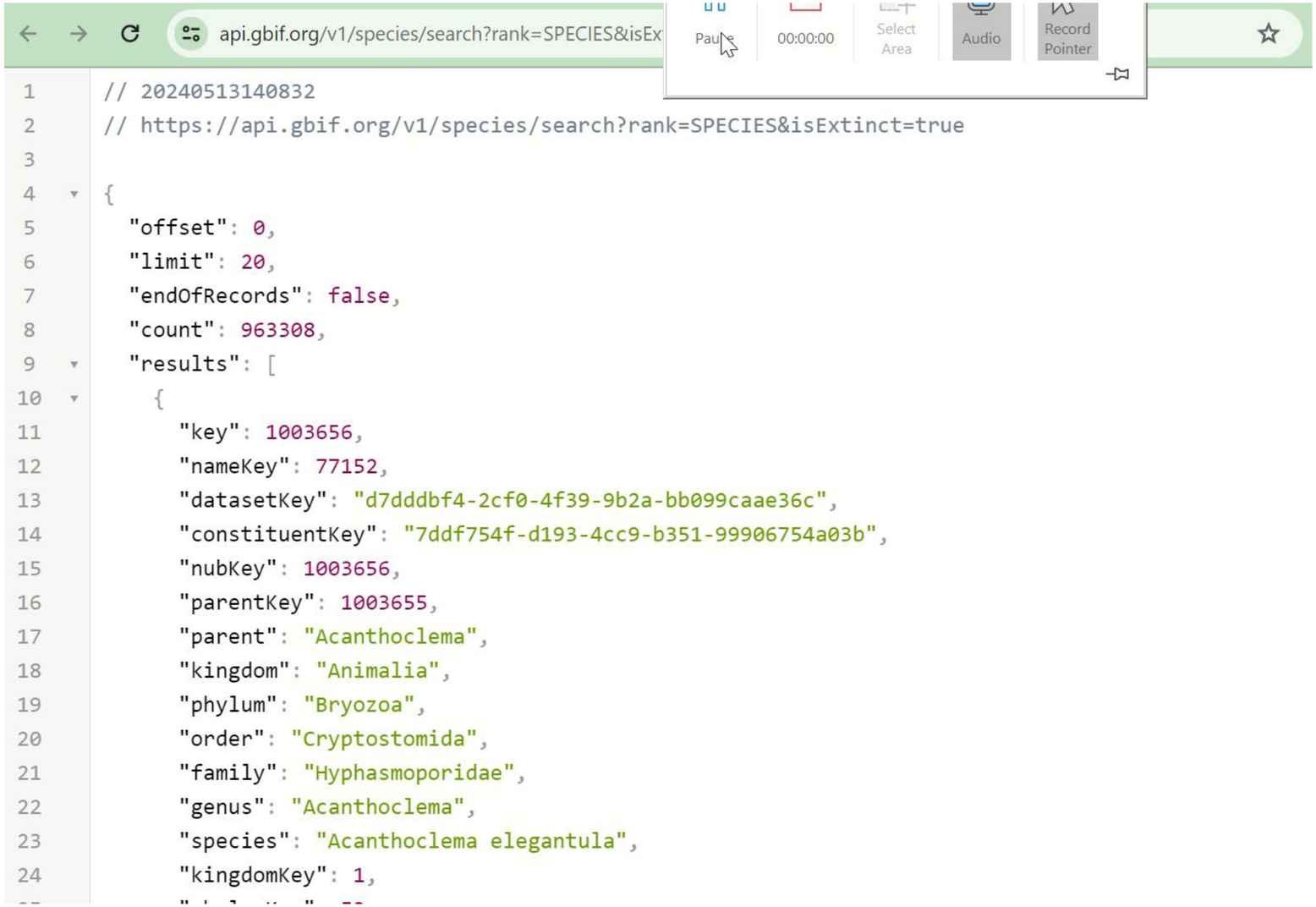
This term is is not visible on the web UI, but it is available when using the **API** or rgbif/pygbif.

The API call to get all extinct taxa across any checklist on GBIF

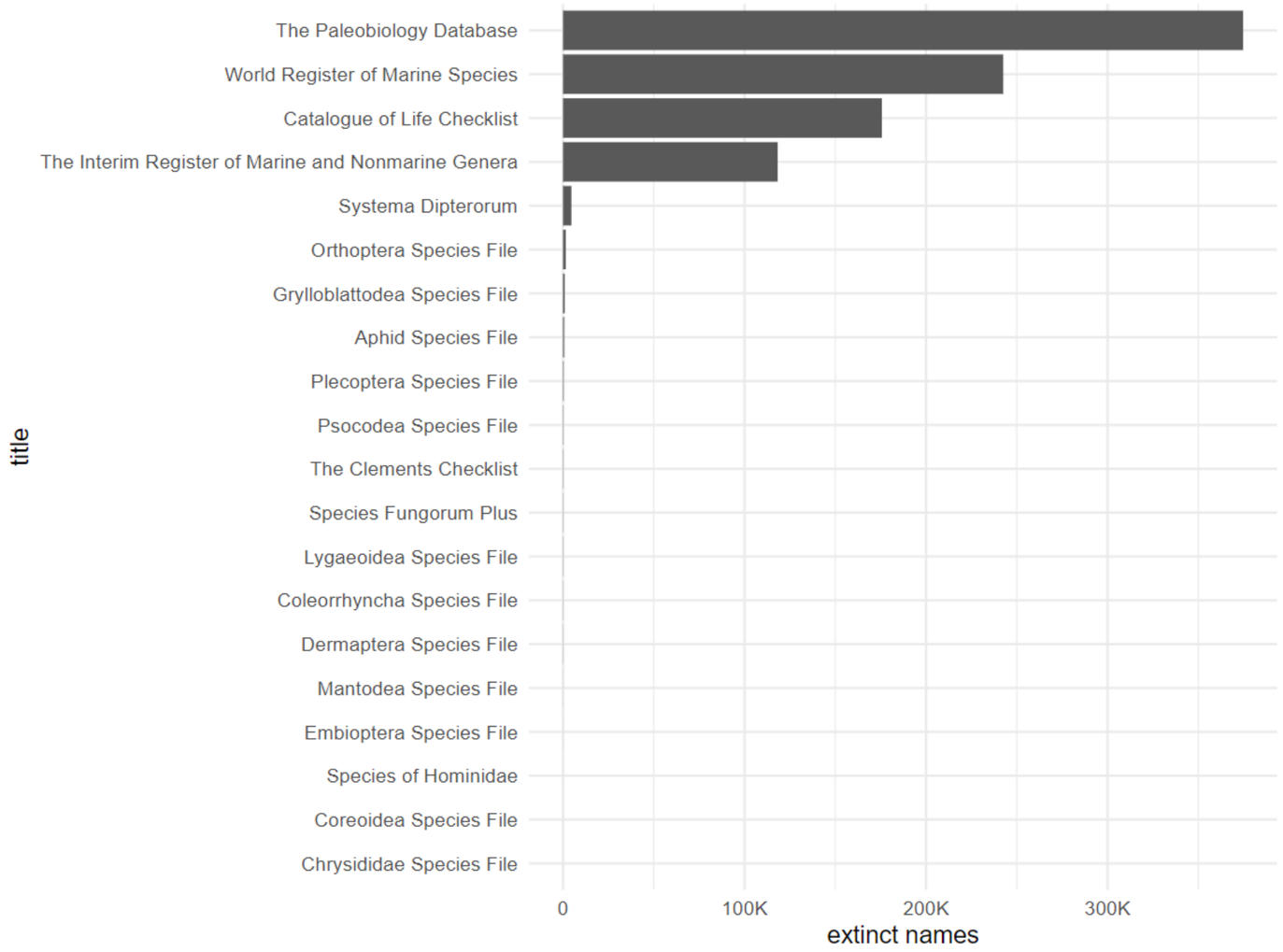
https://api.gbif.org/v1/species/search?rank=SPECIES&isExtinct=true

Limited to the GBIF backbone

https://api.gbif.org/v1/species/search?rank=SPECIES&isExtinct=true&datasetKey=d7dddbf4-2cf0-4f39-9b2a-bb099caae36c


api.gbif.org/v1/species/search?rank=SPECIES&isEx
Pau
00:00:00
Select Area
Audio
Record Pointer
// 20240513140832
// https://api.gbif.org/v1/species/search?rank=SPECIES&isExtinct=true
{
"offset": 0,
"limit": 20,
"endOfRecords": false,
"count": 963308,
"results": [
{
"key": 1003656,
"nameKey": 77152,
"datasetKey": "d7dddbf4-2cf0-4f39-9b2a-bb099caae36c",
"constituentKey": "7ddf754f-d193-4cc9-b351-99906754a03b",
"nubKey": 1003656,
"parentKey": 1003655,
"parent": "Acanthoclema",
"kingdom": "Animalia",
"phylum": "Bryozoa",
"order": "Cryptostomida",
"family": "Hyphasmoporidae",
"genus": "Acanthoclema",
"species": "Acanthoclema elegantula",
"kingdomKey": 1,

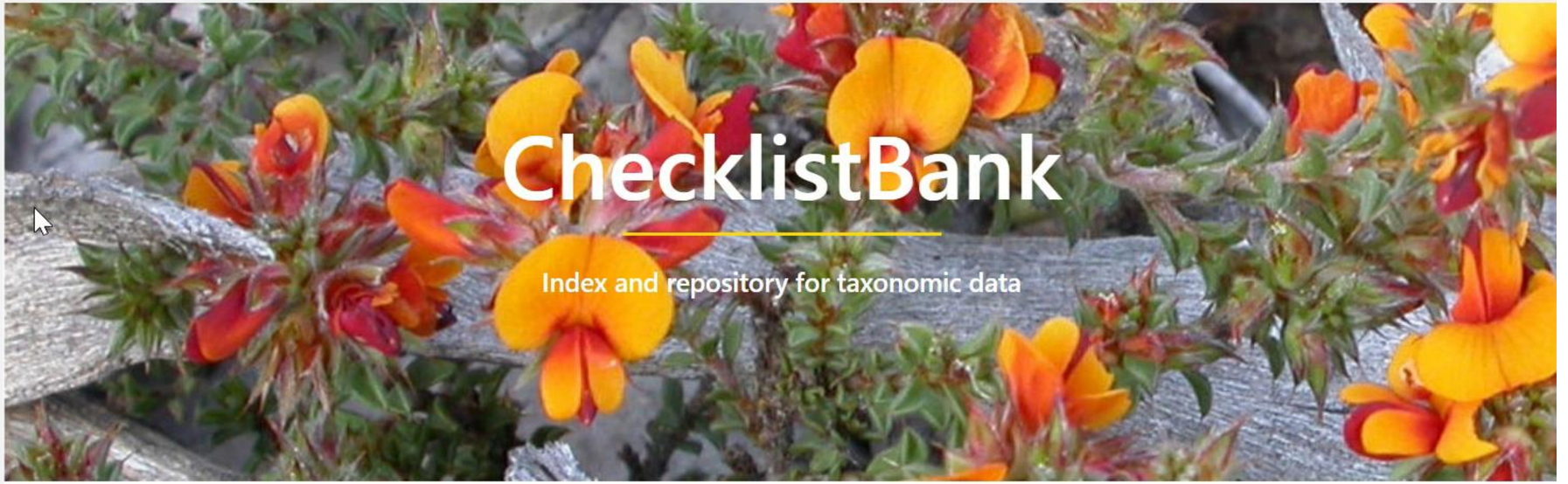
Number of isExtinct names in GBIF backbone



ChecklistBank

- Datasets
- Tools
- COL
  - Metadata
  - Browse
  - Search
  - Download
  - References
  - Sources
  - Releases

jwaller



Species in Catalogue of Life	Name Usages in ChecklistBank	Datasets in ChecklistBank	Latest COL Checklist
2,158,135	306,446,084	52,146	2024-04-26

The Catalogue of Life (COL) aims to support the publication and curation of checklists and to provide a platform for their consistent discovery, use and citation. GBIF has for some time maintained a checklist index and supported the network of repositories for its community to share checklist data. COL and GBIF have united their capabilities to make ChecklistBank a

Latest datasets added

- First record of Corydalus diasi Navás, 1915 (Megaloptera: Corydalidae) from Rio de Janeiro state, Brazil (May 14th 2024)

Checklistbank API

Export all extinct names from COL

http://api.checklistbank.org/dataset/3LR/export.txt?extinct=true

Only from a certain group “Trilobites taxonID=TRL”

http://api.checklistbank.org/dataset/3LR/export.txt?taxonID=TRL&extinct=true

Catalogue of Life

The most complete authoritative list of the world's species - maintained by hundreds of global taxonomists - Learn more

Fungi

154,534 species in COL

COL Version: 2024-04-26

Advanced search

Find taxon

Info Include extinct

- unranked: *Biota*
  - kingdom: Animalia
  - kingdom: Archaea

APRIL 26, 2024

MONTHLY RELEASE APRIL 2024

Monthly release April 2024 of the Catalogue of Life Checklist

MARCH 25, 2024

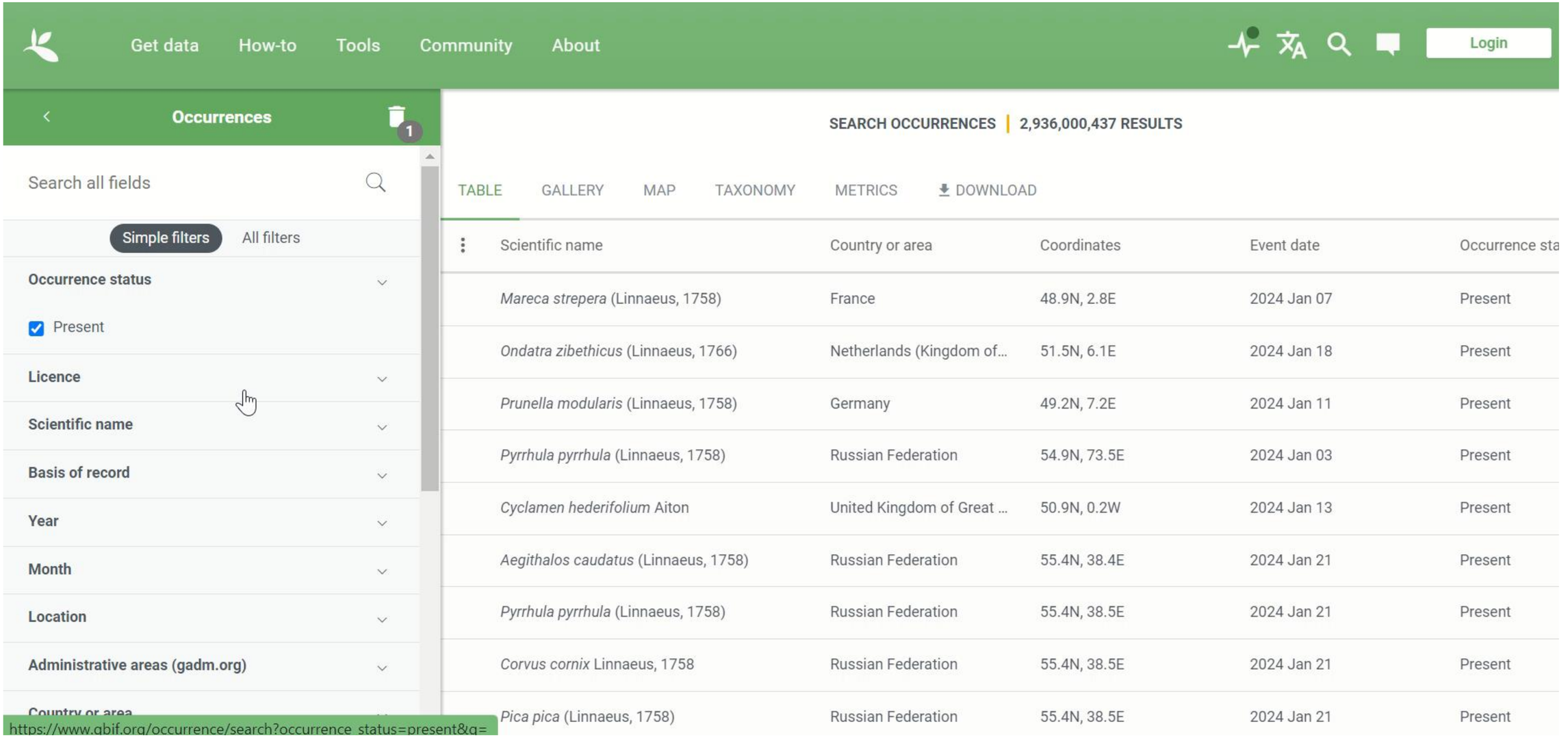
MONTHLY RELEASE MARCH 2024

Monthly release March 2024 of the Catalogue of Life Checklist

FEBRUARY 22, 2024

MONTHLY RELEASE FEBRUARY 2024

Monthly release February 2024 of the Catalogue of Life Checklist

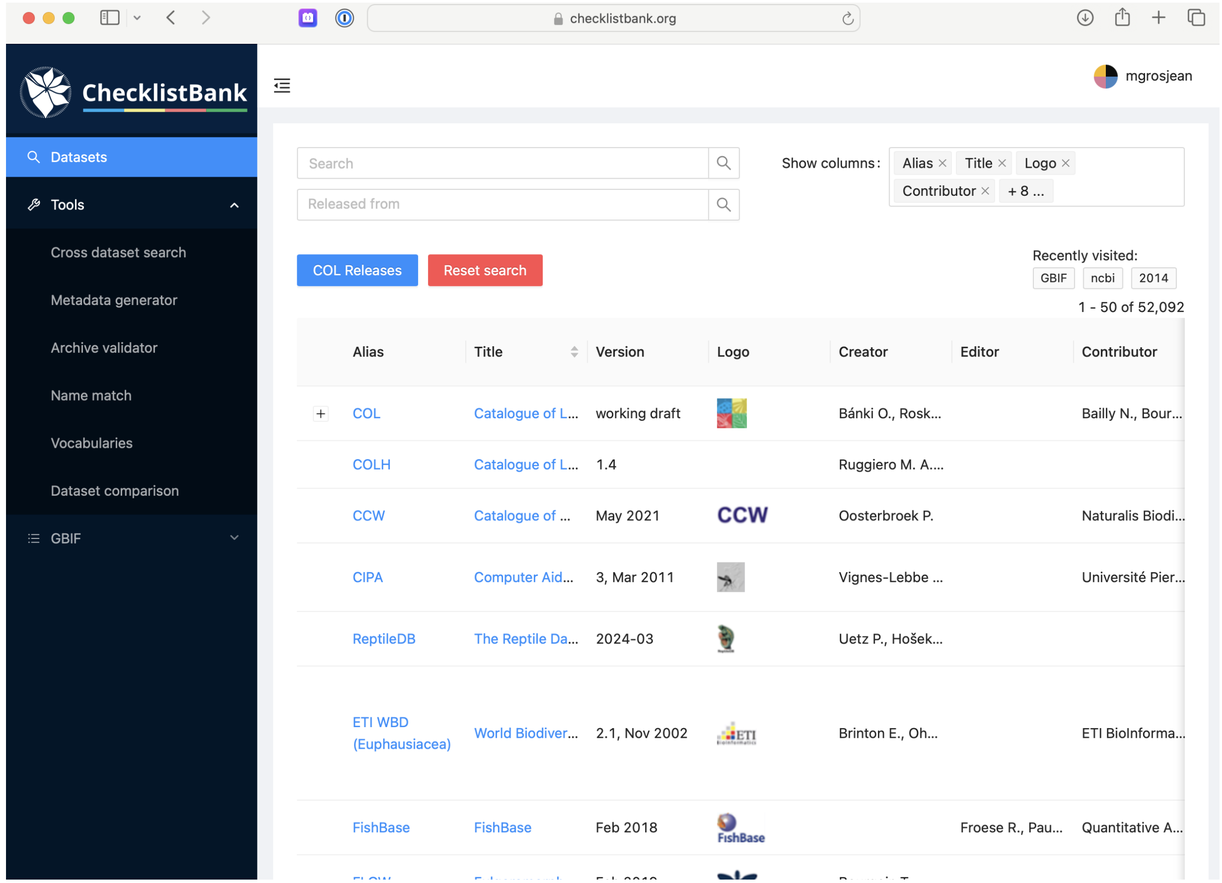

Get data
How-to
Tools
Community
About
Login
Occurrences
1
Search all fields
Simple filters
All filters
Occurrence status
Present
Licence
Scientific name
Basis of record
Year
Month
Location
Administrative areas (gadm.org)
Country or area
https://www.gbif.org/occurrence/search?occurrence_status=present&q=
SEARCH OCCURRENCES | 2,936,000,437 RESULTS
TABLE
GALLERY
MAP
TAXONOMY
METRICS
DOWNLOAD
Scientific name
Country or area
Coordinates
Event date
Occurrence sta
Mareca strepera (Linnaeus, 1758)
France
48.9N, 2.8E
2024 Jan 07
Present
Ondatra zibethicus (Linnaeus, 1766)
Netherlands (Kingdom of...
51.5N, 6.1E
2024 Jan 18
Present
Prunella modularis (Linnaeus, 1758)
Germany
49.2N, 7.2E
2024 Jan 11
Present
Pyrrhula pyrrhula (Linnaeus, 1758)
Russian Federation
54.9N, 73.5E
2024 Jan 03
Present
Cyclamen hederifolium Aiton
United Kingdom of Great ...
50.9N, 0.2W
2024 Jan 13
Present
Aegithalos caudatus (Linnaeus, 1758)
Russian Federation
55.4N, 38.4E
2024 Jan 21
Present
Pyrrhula pyrrhula (Linnaeus, 1758)
Russian Federation
55.4N, 38.5E
2024 Jan 21
Present
Corvus cornix Linnaeus, 1758
Russian Federation
55.4N, 38.5E
2024 Jan 21
Present
Pica pica (Linnaeus, 1758)
Russian Federation
55.4N, 38.5E
2024 Jan 21
Present

Accessing species information from Checklistbank

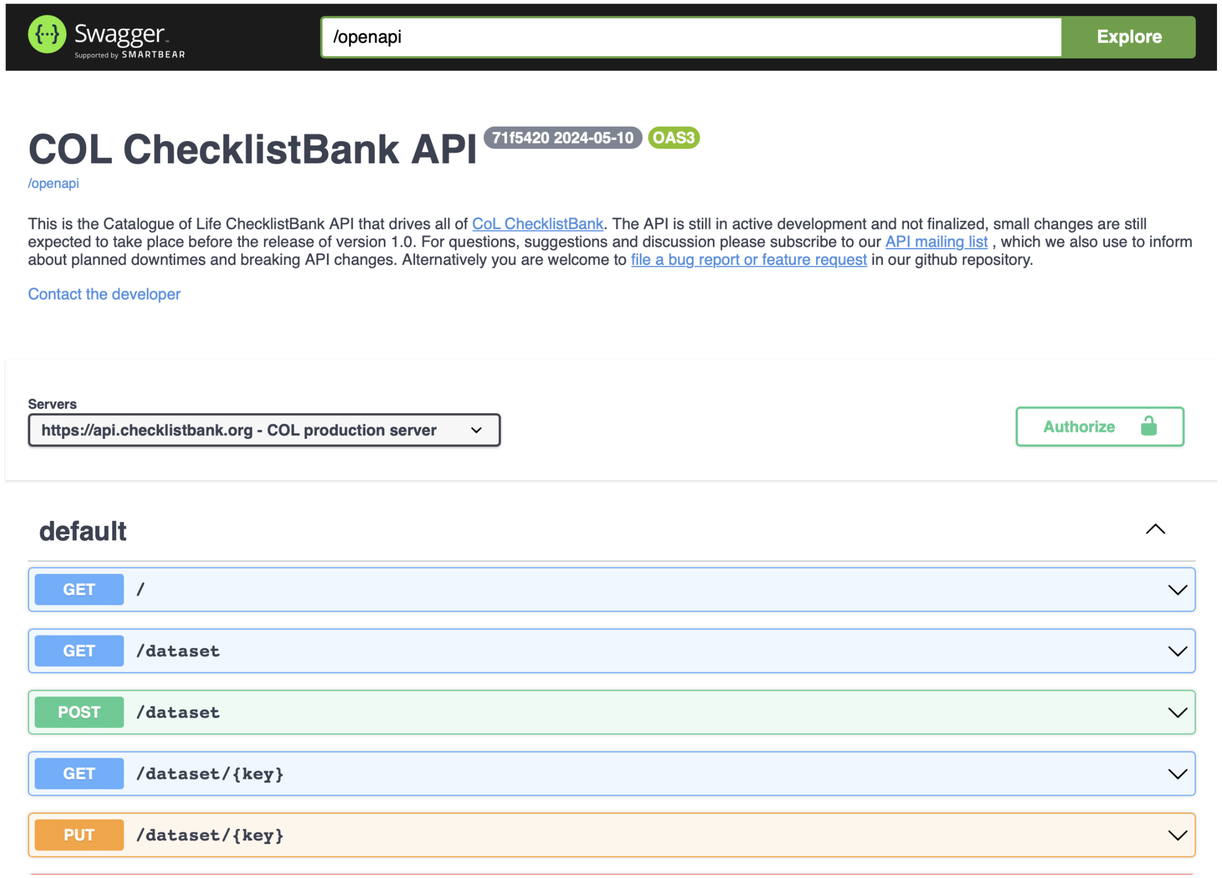
Checklistbank.org and Checklistbank API

What is Checklistbank

- Platform for checklists to:
  - Publish, edit, build
  - Find, access
- Contains all checklists in GBIF and more
- Developed by GBIF and the Catalogue of Life
- Same login as GBIF



Checklistbank API

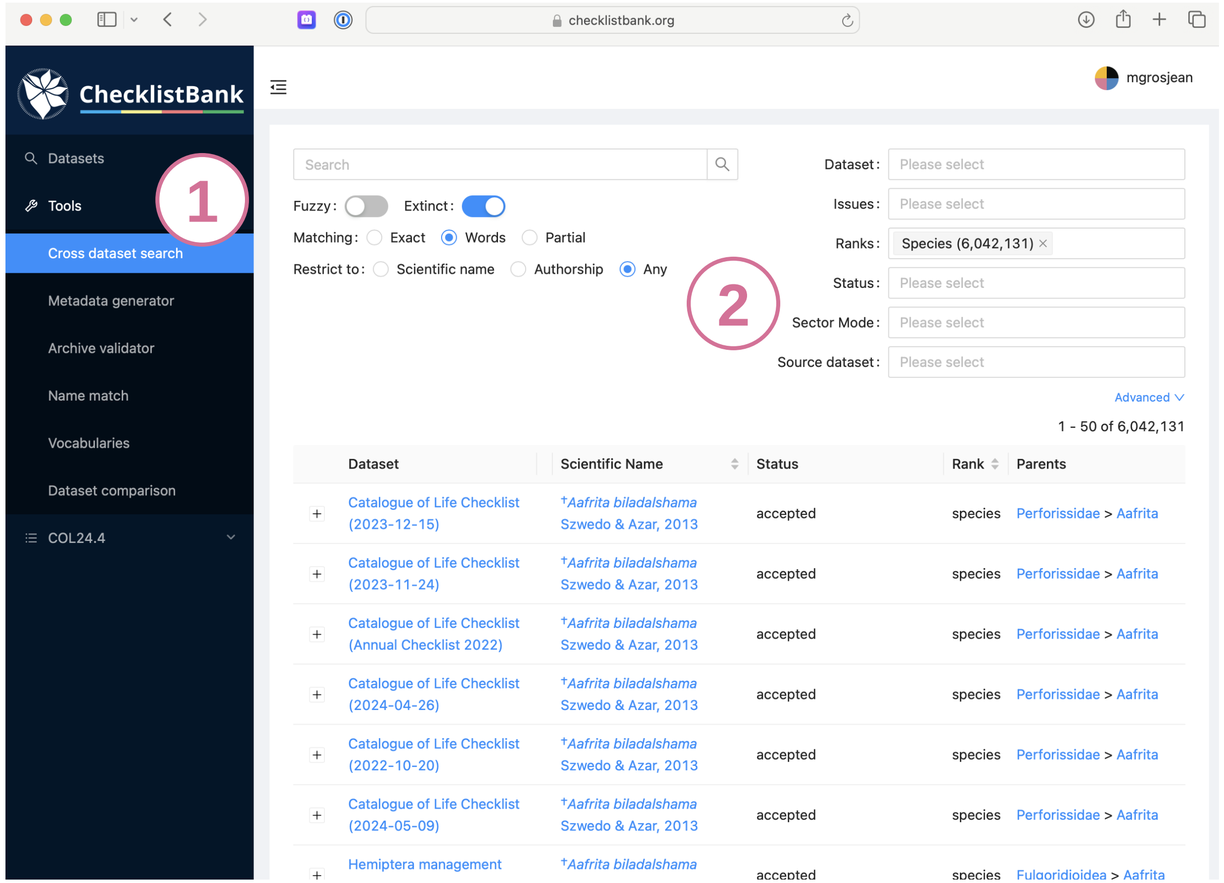


- https://api.checklistbank.org
- Centers around checklist datasets
- Functions include:
  - Name matching
  - Search
  - Access attributes like vernacular names, distribution, description, etc.

https://github.com/CatalogueOfLife/backend/blob/master/API.md

Checklistbank functions – searching – cross dataset

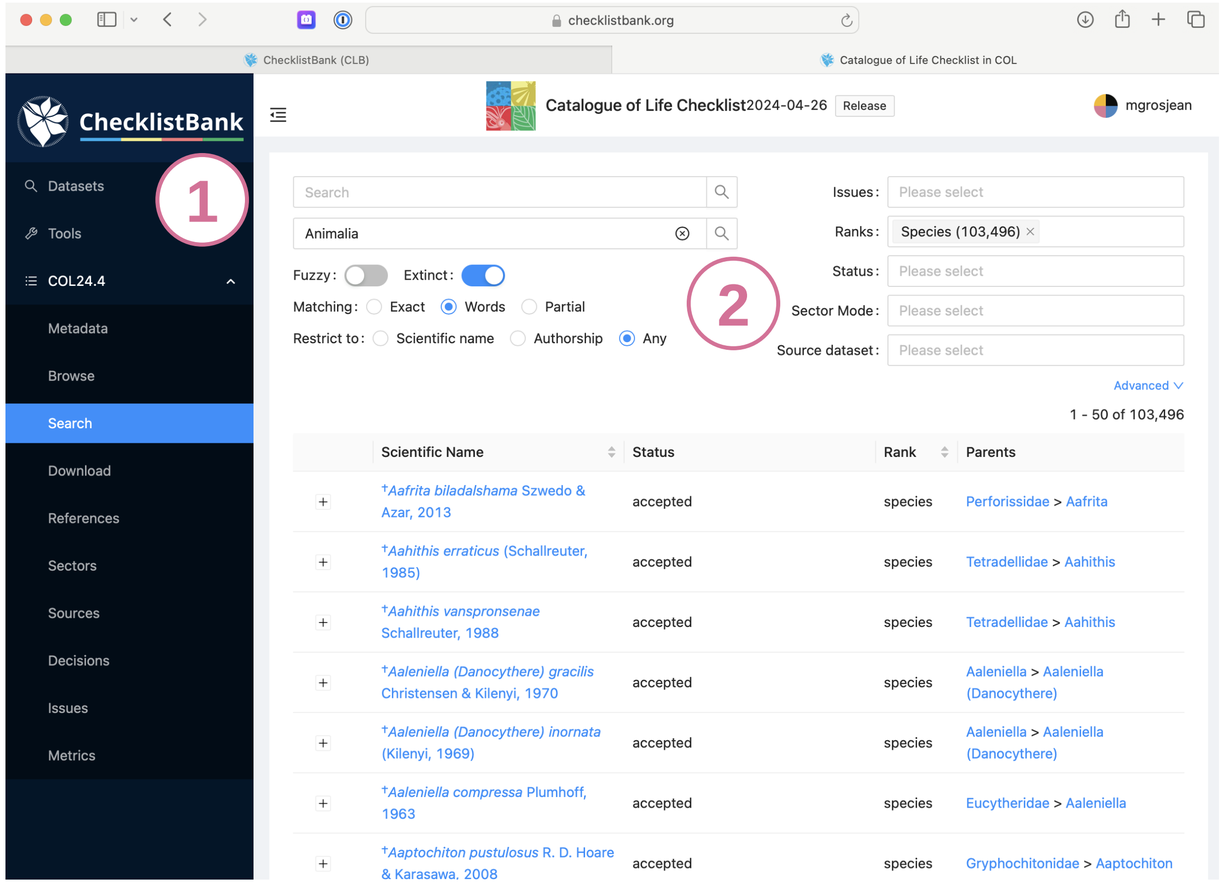
/nameusage/search



1. Cross dataset search
2. Build query

Full tutorial: https://docs.gbif.org/course-checklistbank-tutorial/en/cross-dataset-search.html

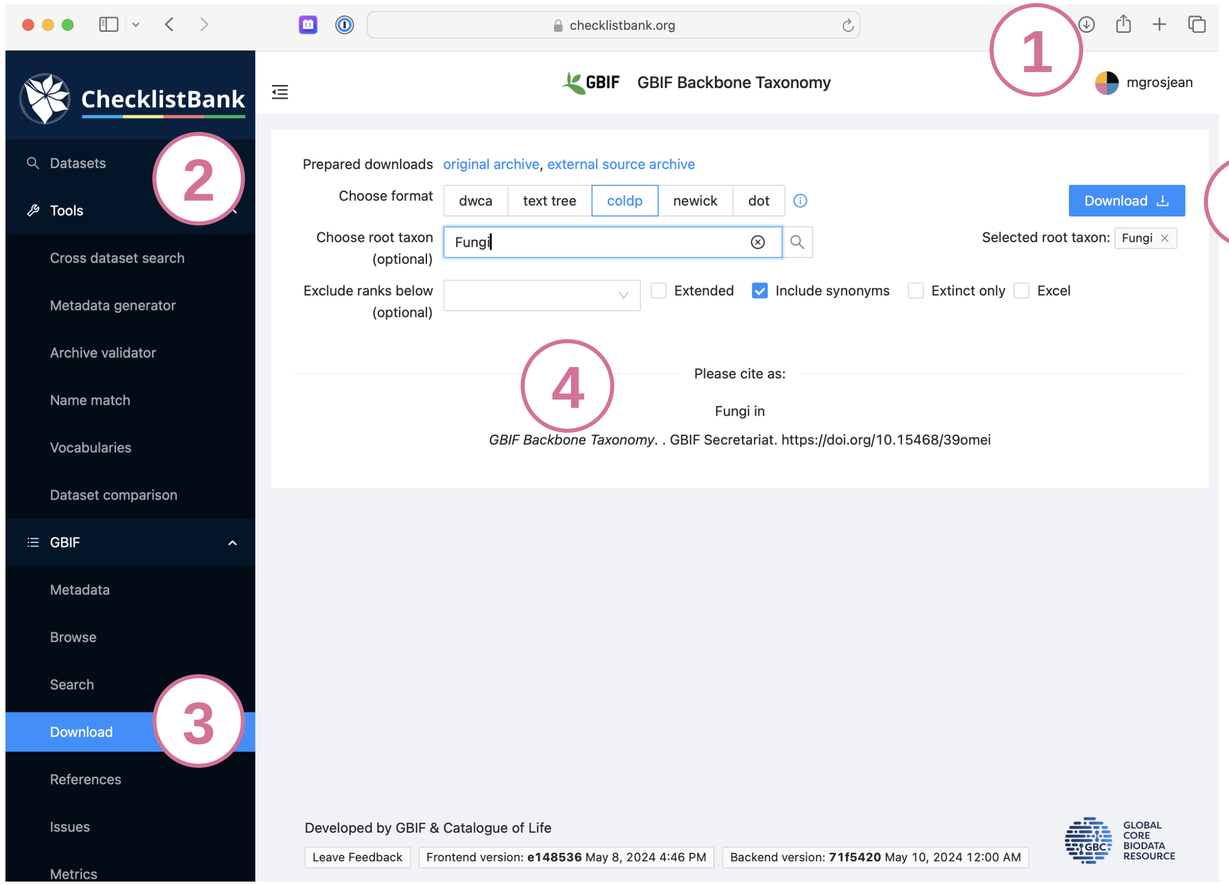
Checklistbank functions – searching – single dataset



/dataset/{key}/nameusage/search/

1. Select Dataset: Catalogue of Life
2. Build query: extinct animal species

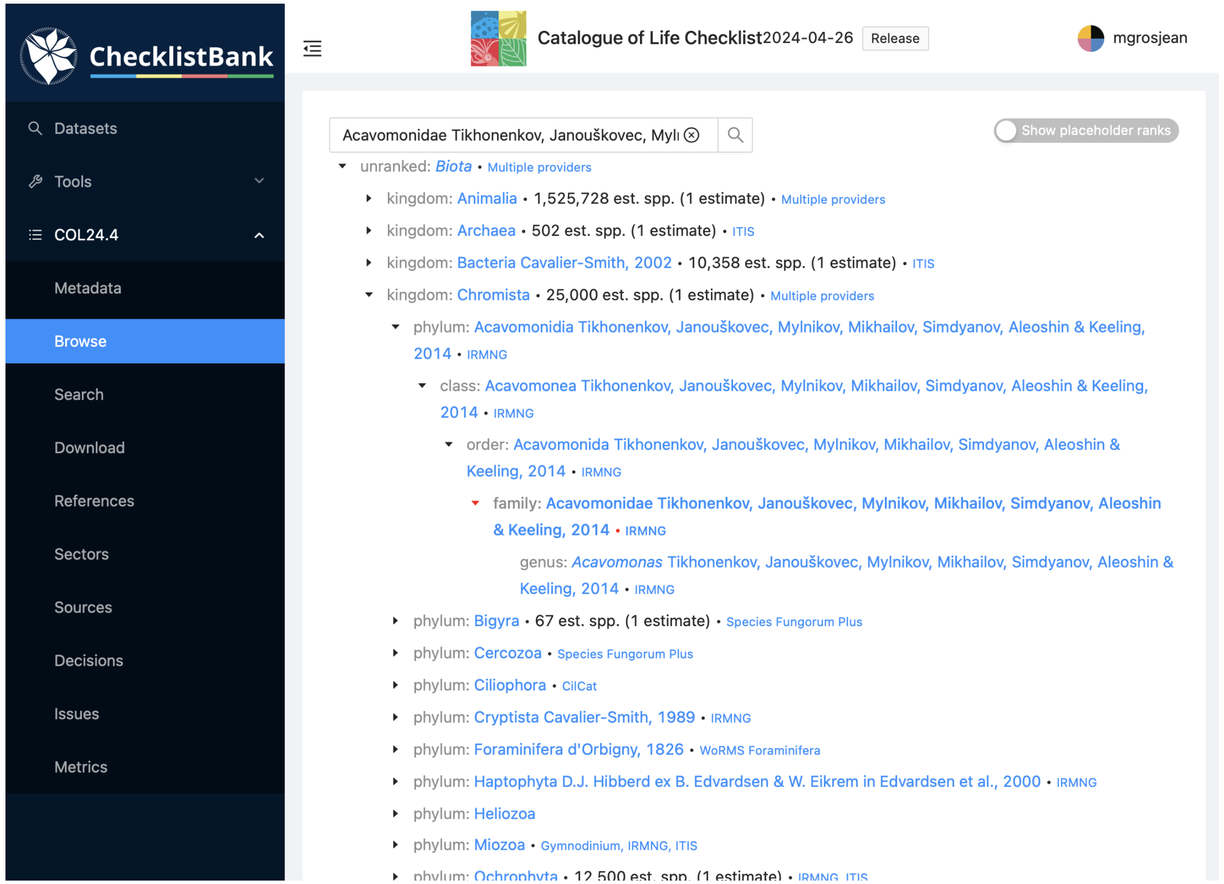
Checklistbank functions - downloading



POST /dataset/{key}/export

1. Log into Checklistbank
2. Look for dataset
3. Go to download tab
4. Choose package/subset
5. Download

Checklistbank functions – browse tree



Get the tree root taxa

dataset/{key}/tree

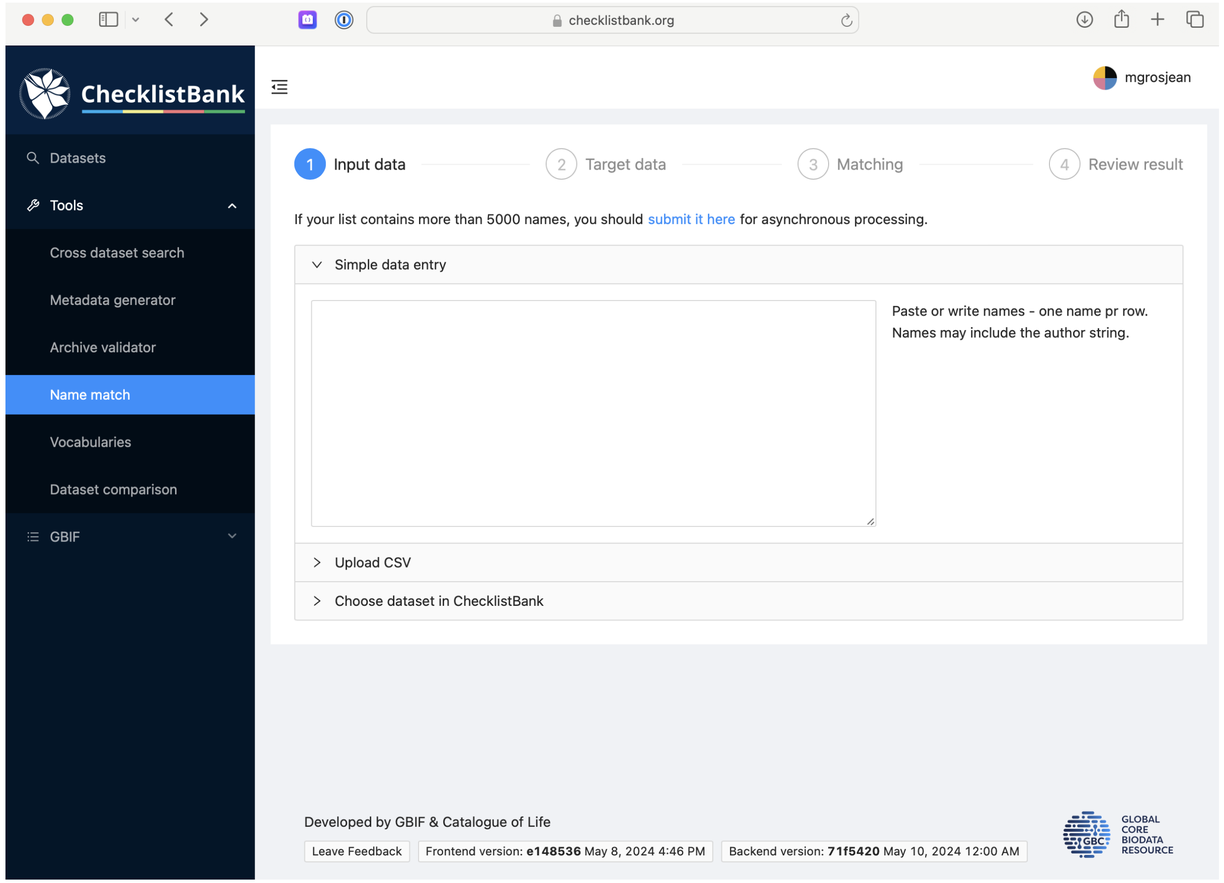
Get children

dataset/{key}/tree/{id}/children

Get parents

dataset/{key}/tree/{id}

Checklistbank functions – name matching

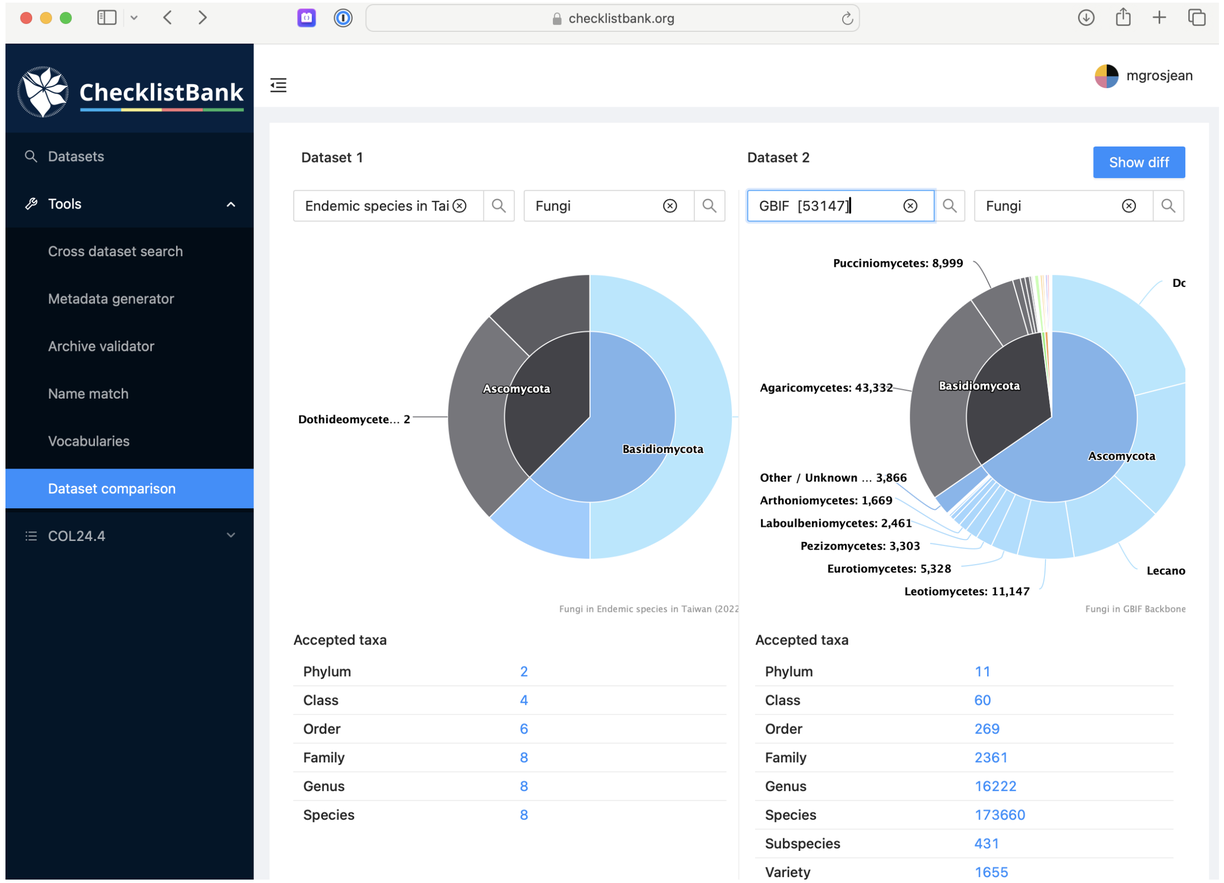


POST

/dataset/{key}/match/nameusage/job

- Match large number of names to any checklist
- Allows you to download output

Checklistbank functions – checklist comparison



/dataset/{key}/diff/{key2}

- Compare two checklists in detail
- Full tutorial: https://docs.gbif.org/course-checklistbank-tutorial/en/dataset-comparison.html

Checklistbank functions – API– others

Access taxon attributes

/dataset/{key}/taxon/{id}/

distribution	source	children
interaction	synonym	tree
media	treatment	vernacular
property	relation	classification

All in one

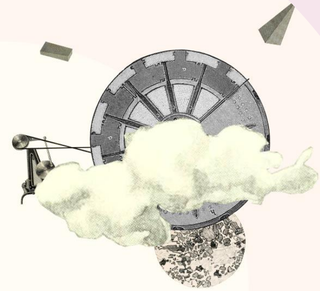
taxon/{id}/info

Search dataset per extension

dataset?
rowType=gbif:Distribution

Next session:
19 September 2024

datause@gbif.org



Data Use Club

Cloud computing services

Practical sessions