iDigBio
Integrated Digitized Biocollections

CYWG (CyberInfrastructure Working Group)
October 2014

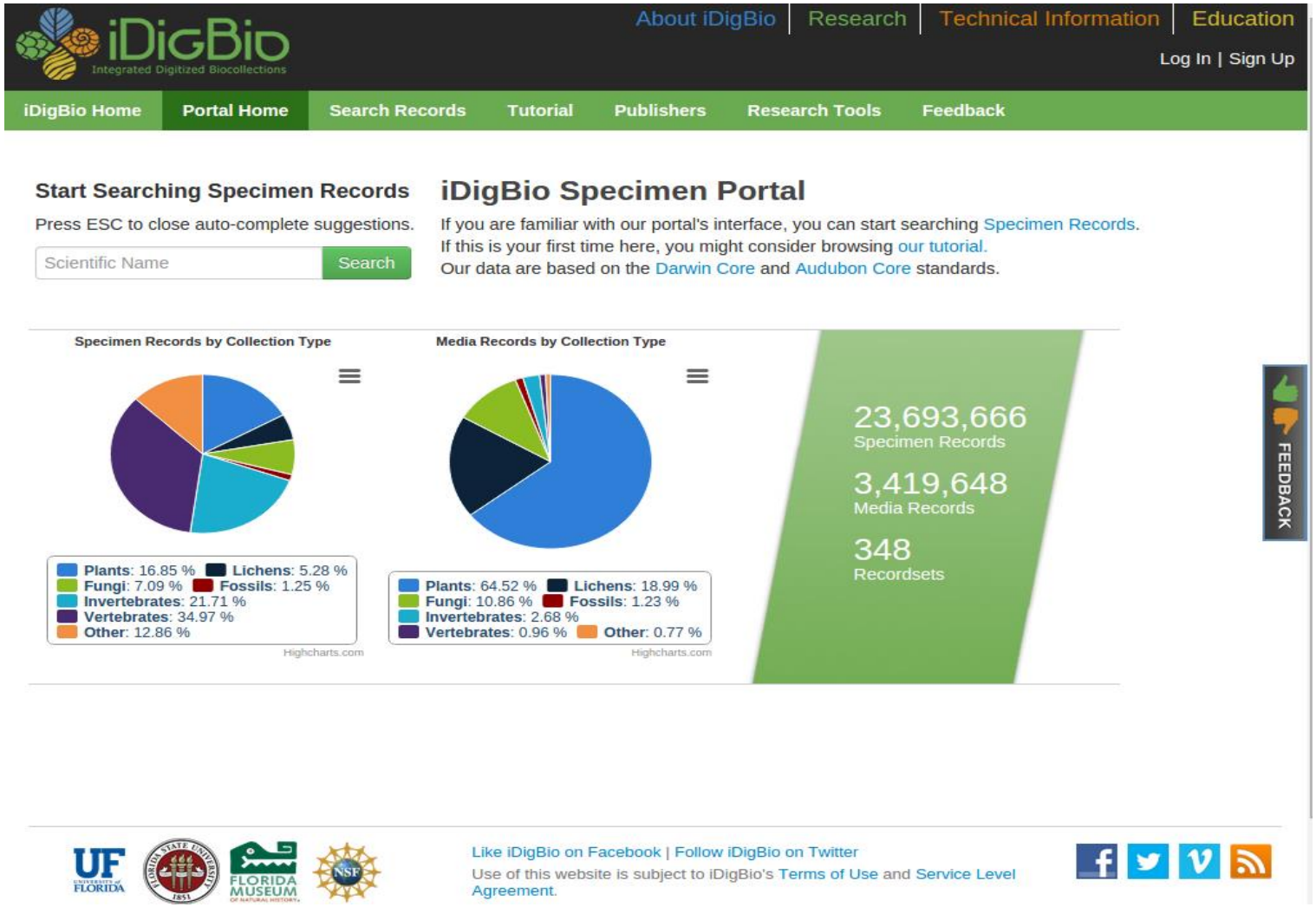
iDigBio Search Portal

Greg Traub
Advanced Computing and Information Systems Laboratory (ACIS)
University of Florida

gtraub@acis.ufl.edu

iDigBio is funded by a grant from the National Science Foundation's Advancing Digitization of Biodiversity Collections Program (Cooperative Agreement EF-1115210). 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. All images used with permission or are free from copyright.

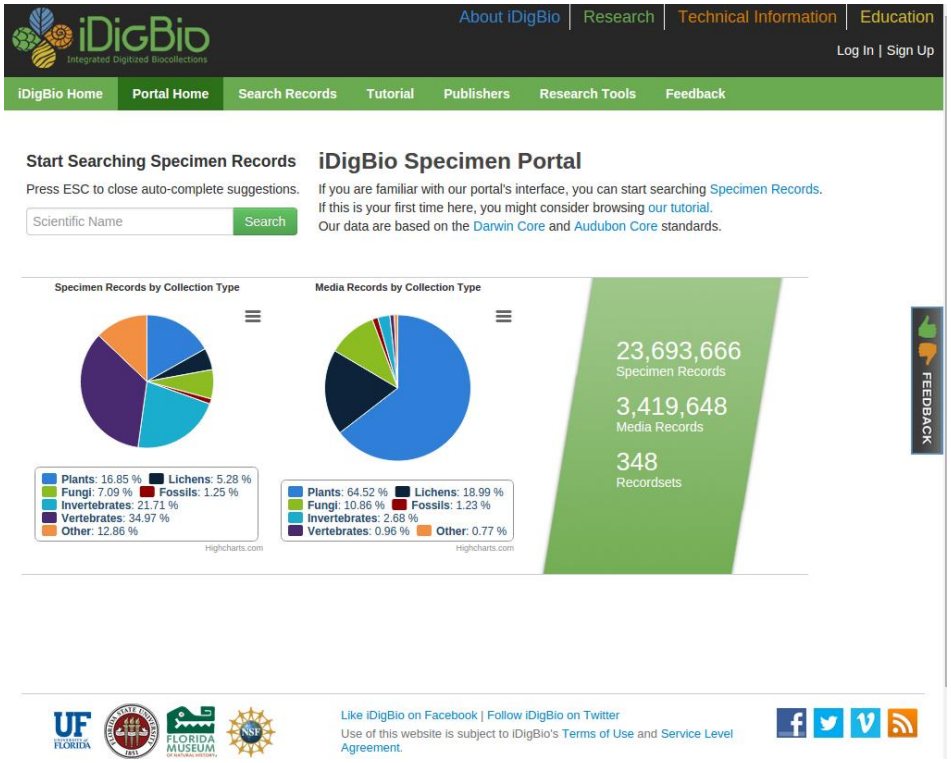
iDigBio Search Portal



http://portal.idigbio.org

Portal Home Page

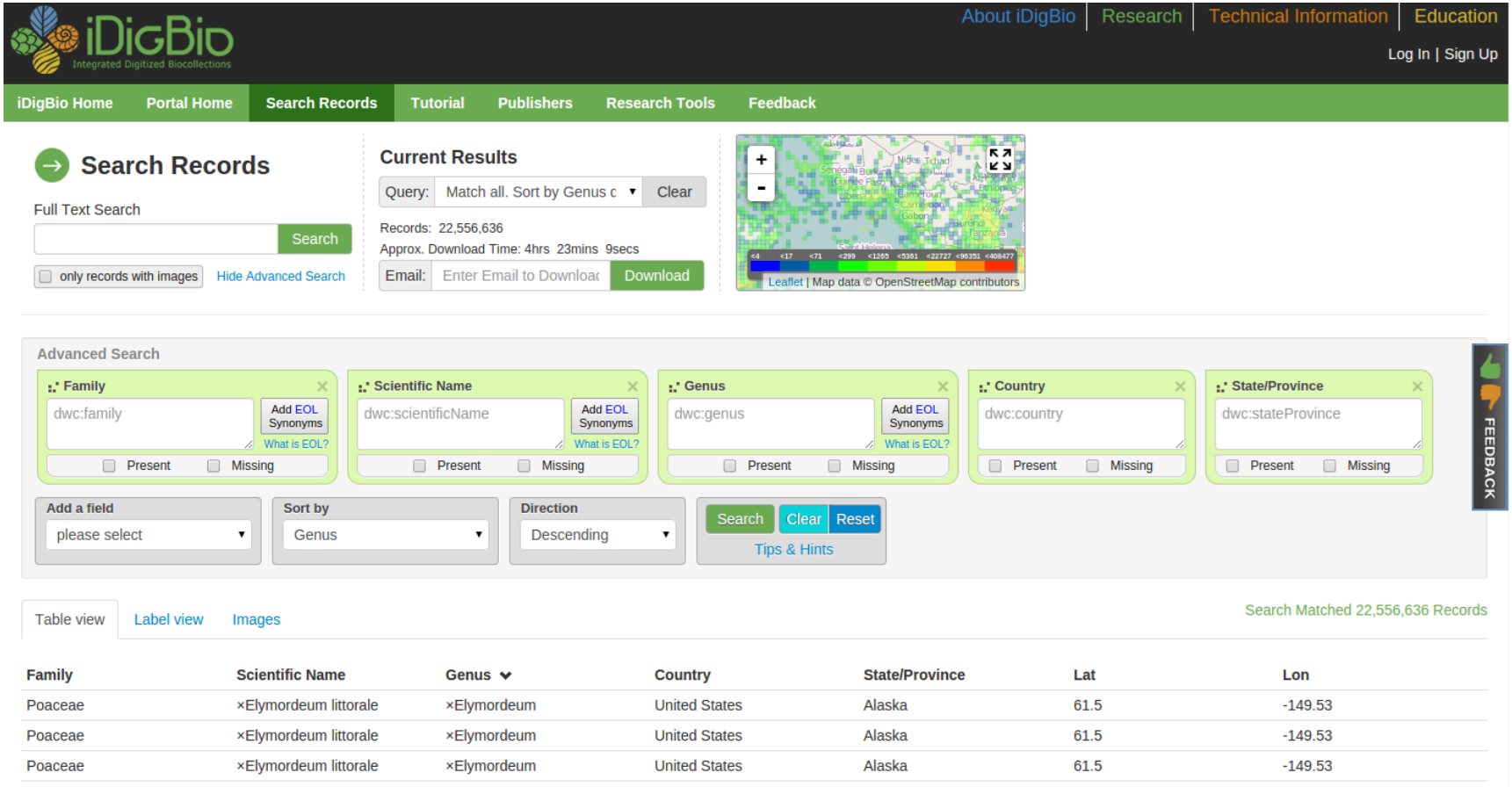
- Provides a quick search box with search suggestions for Scientific Name
- Provides graphs for the specimen and media record count breakdowns by collection category
- Provides up-to-date record counts of Specimen, Media and Recordset counts



Search Interface

- Portal provides an easy search interface for all indexed data in iDigBio
- A user can search across all fields (full-text search)
- A user can search individual fields (advanced search fields)
- Advanced search fields provide progressive autocompletion (meaning any used fields are taken into account when providing suggestions)
- Lower taxonomy search fields (Family, Genus, Specific Epithet) provide an “Add EOL Synonyms” button for adding synonyms to a users entries utilizing the Encyclopedia Of Life API.

Search Interface



Search Results

- Three options to view search results
- A table view with columns that correspond to the advanced search fields (or default columns). Clicking on a row opens a Specimen Record view
- A label view that displays the results with a summary of the record information as well as a thumbnail of any media related to the record. Clicking on a label opens a Specimen Record view
- An Image gallery that displays all images associated with results in a contact sheet view. Clicking on an image opens the associated Media Record view

Search Results

Table results

Table view | Label view | Images

Search Matched 65 Records

Family	Scientific Name ▲	Genus	Country	State/Province
Endodontidae	Aaadonta	Aaadonta	Palau	Ulebsechel Island
Endodontidae	Aaadonta	Aaadonta	Palau	
Endodontidae	Aaadonta	Aaadonta	Palau	

Label results

Table view | Label view | Images

Search Matched 3,512,348 Records

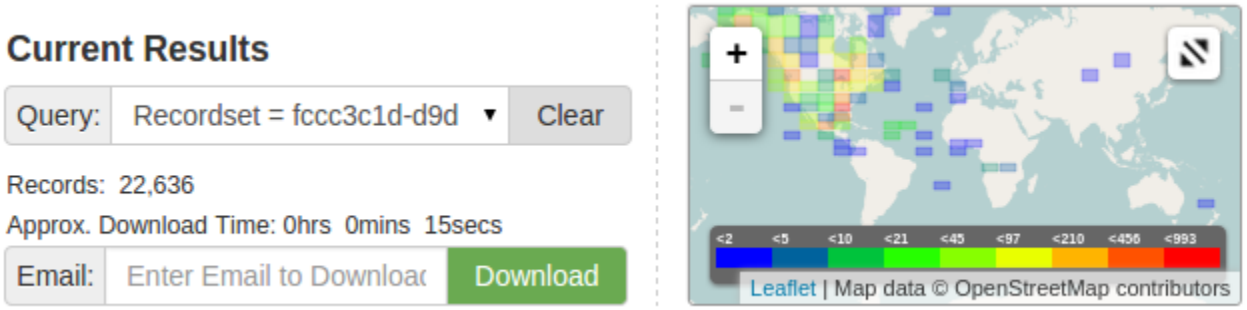
Image results

Table view | Label view | Images

Search Matched 749,373 Records

Search Results

- Previous results can be viewed by using the Query select box to access the users search history
- Any results that contain valid geopoint data are placed in the map box. If results return more than 5,000 geopoints a heat layer is rendered on the map. Less than 5,000 and individual points are rendered. Clustering points also appear to help distinguish points that are close to each other.
- Every result set can be downloaded with our email to download feature. This features provides an estimation of the time it takes to build the download. The user is notified by an email with a download link when the dataset is ready.



Specimen Record Page

- Specimen records provide all details of the specimen record.
- Record data is separated into a tabbed view for each category of data (Taxonomy, Specimen, Locality, Collection Event, Other)
- All associated media are displayed in thumbnails if available. Each thumbnail links to its own Media Record page
- Valid georeference point (decimal degree lat/lon) is displayed on a map
- Data provider (institution) summary provided with contact information
- Button for viewing raw data
- Button for going to associated Recordset Page

Specimen Record Page

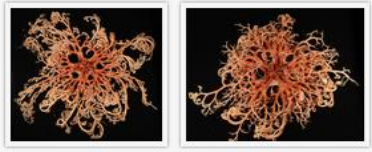
Astrophyton FLMNH, Invertebrate Zoology, 11905-Echinodermata

Taxonomy | Specimen | Collection Event | Locality | Other

Class	Ophiuroidea
Order	Euryalida
Family	Gorgonocephalidae
Genus	Astrophyton

Go To Recordset | View Raw Data

Associated Media



Specimen Georeference

Record Provided By

invertebratezoology

The UF Invertebrate collection holds ~510,000 databased lots of mollusks and marine invertebrates. It began as a Malacology collection almost 100 years ago and ~85% of the holdings are still mollusks. Since 2000 the collection was expanded to cover all invertebrate phyla, focusing on marine taxa. Today it holds >40,000 species from 28 phyla.

Contacts

Gustav Paulay
Curator of Invertebrate Zoology
paulay@flmnh.ufl.edu

William Paine
IT Director
netadmin@flmnh.ufl.edu

FEEDBACK

Media Record Page

- Provides a large format view of the media (if its an image) along with metadata associated with image
- Displays the associated specimen record title
- Displays thumbnails of other media associated with its parent specimen record for easy navigation of other media belonging to the same specimen
- Data provider (institution) summary is provided with contact information
- Button for going to associated Specimen Record
- Button for viewing raw data
- Button for going to associated Recordset Page

Media Record Page

Media Record: *Agrilus planipennis*

Download Media File

Other Media

Go To Specimen Record | Go To Recordset | View Raw Data

Media Metadata

Resource Creation Technique	AutoMontage extended-focus imaging
Creator	ELIJAH
Identifier	http://bioguid.osu.edu/osuc_occurrence_images/2828
Modified	2014-08-07T11:14:45Z
Rights	Copyright Norman Johnson 2014
Title	Image for occurrence record OSUC 211414
Type of Resource	StillImage
Subtype	Photograph
Credit	Norman Johnson
License Terms	CC-BY-NC-SA
License URL	http://creativecommons.org/licenses/by-nc-sa/3.0/
Published Source	http://specimage.osu.edu/getImageInfo.html?image_id=2828
Caption	head, lateral view
Access URI	http://osuc.osu.edu/spm_images/OSUC211414/2828_compress.jpg
Format	image/jpeg

Record Provided By

C.A. Triplehorn Insect Collection (OSUC), Ohio State University

http://osuc.osu.edu

Vouchered occurrence records for insects from the C.A. Triplehorn Insect Collection at the Ohio State University.

Contacts

Publishers Page

- Provides record count details of all data in the ingestion pipeline per publisher
- Details include record digest counts, api counts and search index counts
- Each individual recordset can be clicked to view the details of the provider information (Recordset page)

Publishers Page

iDigBio Data Publishers

This page shows all iDigBio data contributors. If you are interested in providing data, consult the data ingestion guide for more information.

	Record Count	Media Record Count
Total from Providers	22,556,733	3,167,330
Total in API	22,606,002	3,232,444
Total Published (all data incorporated in new workflow)	22,605,241	3,231,947
Total Indexed (all data) *	22,605,241	3,231,947

* Data that is marked deleted in iDigBio remains indexed until a cleanup is run.

Publisher Summary

	Record Count			Media Record Count		
Publisher Name	Digest	API	Index	Digest	API	Index
Berkeley Natural History Museums IPT	1,860,584	1,859,985	1,859,985	0	0	0
Florida Museum of Natural History IPT Service	1,047,587	1,047,587	1,047,564	0	0	0
Northern Great Plains Herbaria Darwin Core Archive rss feed	43,012	43,012	43,012	0	0	0
MyCoPortal Darwin Core Archive rss feed	1,679,459	1,679,458	1,679,458	371,346	371,346	371,346
KU Biodiversity Institute IPT	2,010,071	2,011,170	2,011,170	0	0	0
The University of Connecticut Biological Collections	172,098	171,936	171,198	166,689	166,519	166,022
xBioD IPT in the Museum of Biological Diversity at the Ohio State University	521,710	521,782	521,782	2,593	2,593	2,593
CMC_specify	9,131	9,131	9,131	0	0	0
Consortium of North American Bryophyte Herbaria Darwin Core Archive rss feed	1,690,014	1,690,014	1,690,014	816,932	816,932	816,932
Museum of Comparative Zoology, Harvard University	1,736,357	1,736,471	1,736,471	0	0	0
CNALH Darwin Core Archive rss feed	1,232,891	1,232,891	1,232,891	649,241	649,241	649,241
SCAN Darwin Core Archive rss feed	873,024	873,160	873,160	68,696	68,718	68,718
iDigBio Feeder RSS Feed	1,316,574	1,316,574	1,316,574	19,024	19,024	19,024

Recordset Page

- Provides details about the data provider including collection information and contacts
- Displays a table of record counts and percentages of their specimen records that provide specific Darwin Core fields as they are indexed for the portal search
- Button for linking to the search interface with a search on the given recordset
- Button for viewing the raw recordset data

Recordset Page

Recordset: University of Washington Herbarium

The University of Washington Herbarium (also known as WTU) is an international resource for research into the diversity, distribution and ecology of Pacific Northwest vascular plants, non-vascular plants, fungi, lichen, and algae. With over 600,000 specimens currently in the collections and between 5,000-10,000 specimens added annually, WTU is one of the largest herbaria in the region. The University of Washington Herbarium (also known as WTU) is an international resource for research into the diversity, distribution and ecology of Pacific Northwest vascular plants, non-vascular plants, fungi, lichen, and algae. With over 600,000 specimens currently in the collections and between 5,000-10,000 specimens added annually, WTU is one of the largest herbaria in the region.

Last Update: 2014-06-24

Total Specimen Records: 74,141

Total Media Records: 23,069

Search This Recordset | View Raw Data

Website
http://www.washington.edu/burkemuseum/collections/herbarium/index.php

Contacts

egbot@asu.edu

David Giblin, Collections Manager
wtu@u.washington.edu

FEEDBACK

Specimen Fields Used for Search

Total Records: 74,141

This table represents the fields in specimen records that are used for iDigBio search. The first column represents the field name and equivalent DWC term. The last two columns represent the number and percentage of records that provide the field.

Field	Records With This Field	(%) Percent Used
Kingdom (dwc:kingdom)	0	0
Phylum (dwc:phylum)	0	0
Class (dwc:class)	0	0
Order (dwc:order)	0	0
Family (dwc:family)	66,675	89.93
Scientific Name (dwc:scientificName)	66,011	89.034
Genus (dwc:genus)	66,004	89.025
Specific Epithet (dwc:specificEpithet)	64,589	87.116
Infraspecific Epithet (dwc:infraspecificEpithet)	3,769	5.084
Higher Taxon (dwc:higherClassification)	0	0
Common Name (dwc:vernacularName)	0	0
Lat (dwc:decimalLatitude)	13,301	17.94
Lon (dwc:decimalLongitude)	13,110	17.683

When you have a moment please give the Portal a try. We welcome and would appreciate any feedback.

Feedback can be sent through our feedback page www.idigbio.org/contact

Thank You!

www.idigbio.org

facebook.com/iDigBio

twitter.com/iDigBio

vimeo.com/idigbio

idigbio.org/rss-feed.xml

webcal://www.idigbio.org/events-calendar/export.ics

iDigBio is funded by a grant from the National Science Foundation's Advancing Digitization of Biodiversity Collections Program (Cooperative Agreement EF-1115210). 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. All images used with permission or are free from copyright.