



Module 12

Exporting Data





Export the Entire Database



Backup the Database

⚙️ User Tools (spadmin)

User Account

- 🔑 Change Password
- 🔙 Log Out

Customization

- ⚙️ Preferences
- 📐 Schema Config

Administrative Tools

- 📁 App Resources
- 📄 **Backup Database**
- 🔑 Generate Master Key
- ✅ Repair Tree
- 🔒 Security and Accounts

Export

- 📤 Create Dwc Archive
- 📡 Update RSS Feed

Import

- 🌐 Locality Update Tool

Documentation

- 📖 About Specify 7
- 💬 Community Forum
- 📖 Documentation

Developer Resources

- 🗑️ Clear Browser Cache
- 🗄️ Database Schema
- ⚙️ Operati
- 📁 Tables



Backup Database

No previous backup was found. Start a new one?

New

Close



Backup Database

46%

Close



Backup Database

Backup completed successfully.

Download

Close



Backup Database

A previous backup was found: **sp7demofish_2025_02_10.sql.gz** (21.96 MB)
This backup was created on 2025-08-25

Download

New

Close



Export Query Results



Export Query Results

Using the Query Builder, you can export any query built on any table to CSV.

Results: (1181)

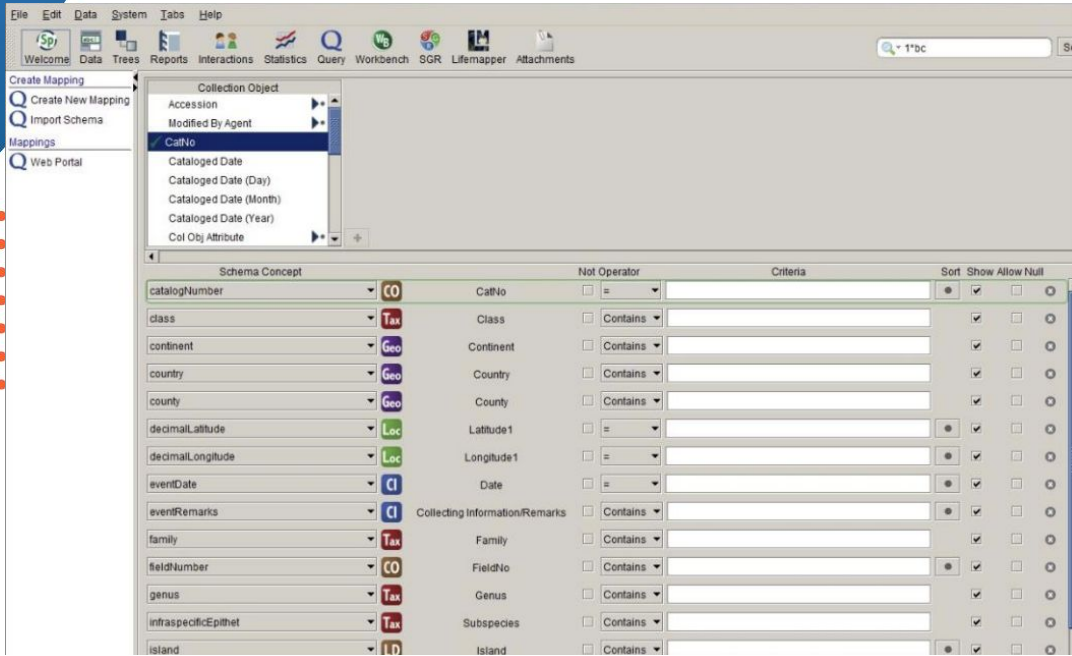
Batch Edit Create CSV Create KML Create Record Set GeoMap Browse In Forms

	Ins	Institution - Code	Col	Collection - Code	CO	Catalog Number	CO	Collection Object - GUID	CR	Left Side Rel	Tax	Order - Full Name	Tax	Class - Full Name	Tax	Family - Full Name	Tax	Taxon - Full Name	Tax	Genus - Full Name	Tax	Species - F
☐		KU		HERB		000146828		5bdf2ea3-df83-44cb-ad1d-8ea19c6fb53b		<Aggregator not defined.>		Asterales		Magnoliopsida		Asteraceae		Rudbeckia hirta pulcherrima		Rudbeckia		hirta
☐		KU		HERB		000137381		39d2bb3f-7eba-4c65-939e-abadb7de314b		<Aggregator not defined.>		Asterales		Magnoliopsida		Asteraceae		Helianthus nuttallii nuttallii		Helianthus		nuttallii
☐		KU		HERB		000258020		3e4d9f7e-1996-40f9-afb4-d35be7ce3e2d		<Aggregator not defined.>		Lamiales		Magnoliopsida		Lamiaceae		Nepeta cataria		Nepeta		cataria
☐		KU		HERB		000137424		b93afa9a-9258-4af7-b0de-9571d880808a		<Aggregator not defined.>		Asterales		Magnoliopsida		Asteraceae		Helianthus nuttallii nuttallii		Helianthus		nuttallii
☐		KU		HERB		000291685		cdb19acc-d407-47d9-93f1-9067ec6b7214		<Aggregator not defined.>		Lamiales		Magnoliopsida		Lamiaceae		Scutellaria lateriflora		Scutellaria		lateriflora

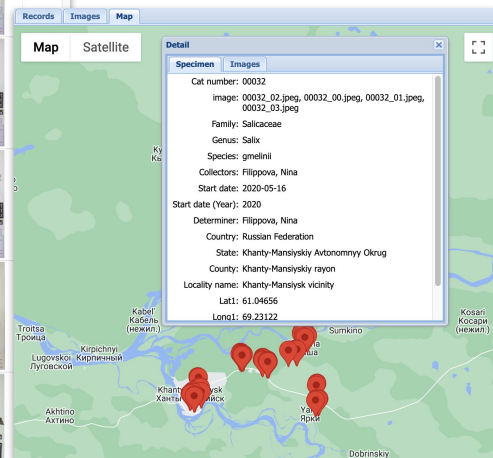
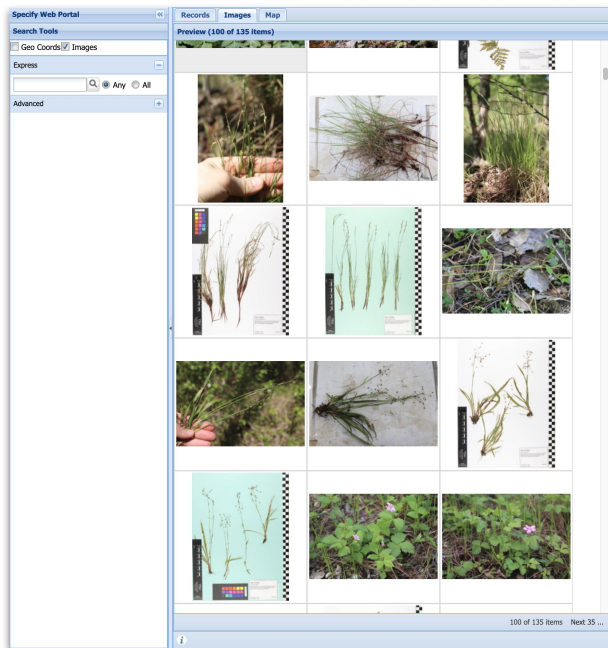


Web Portal Exporting





Create an export query, package a ZIP file, import into the Web Portal!
Using the Data Exporter tool in Specify 6. This can be automated using the command line.





Publishing Data



Why Data Publishing Matters

- **Increasing Mandates**

- Funders (e.g. NSF) and journals may require open access to specimen– and sequence–level data
- Institutional data policies emphasize transparency, reproducibility, and compliance

- **Global Visibility & Impact**

- Publishing to aggregators (GBIF, ALA, GGBN, iDigBio) multiplies discoverability
- Shared data drives research in ecology, conservation, climate change, and beyond

Why Data Publishing Matters

- **Efficiency Gains**

- Automated exports save curator time vs. manual CSV assembly
- Scheduled or on-demand feeds keep aggregator portals up to date with minimal effort

- **Community & Collaboration**

- Shared datasets enable large-scale analyses (e.g., species range modeling, phylogeography)
- Contributing institutions gain credit and visibility for their collections

- **Future-Proofing Collections**

- Published, standardized data remain accessible and reusable long after systems evolve

Data Publishing

- Specify 7 can function as an Integrated Publishing Toolkit (IPT) to publish collections data to GBIF and other data aggregators
- Built on top of the Query Builder
- Data sets can be shared via
 - ZIP archive
 - RSS feed





Data Publishing Pipeline



Step 1: Build an export query

Query: BI portal KUIT

	Determinations	→	✓ Current	→	+		True or Empty	→					
	Collection	→	✓ Code	→	+		Contains	→					
	Collection	→	✓ Description	→	+		Contains	→					
	Collection	→	Discipline	→	Division	→	Institution	→	✓ Copyright				
	Collection	→	Discipline	→	Division	→	Institution	→	✓ Terms Of Use				
	✓ Cat #	→	+		Equal	→		→					
	✓ GUID	→	+		Contains	→		→					
	Determinations	→	Preferred Taxon	→	Phylum	→	✓ Full Name	→	+		Contains	→	
	Determinations	→	Preferred Taxon	→	Class	→	✓ Full Name	→	+		Contains	→	
	Determinations	→	Preferred Taxon	→	Order	→	✓ Full Name	→	+		Contains	→	
	Determinations	→	Preferred Taxon	→	Family	→	✓ Full Name	→	+		Contains	→	
	Determinations	→	Preferred Taxon	→	Genus	→	✓ Full Name	→	+		Contains	→	
	Determinations	→	Preferred Taxon	→	Species	→	✓ Full Name	→	+		Contains	→	
	Determinations	→	Preferred Taxon	→	Subspecies	→	✓ Full Name	→	+		Contains	→	
	Determinations	→	Taxon	→	(any rank)	→	✓ Common Name	→	+		Contains	→	
	Determinations	→	✓ Type Status	→	+		Equal	→		→			
	Field No: Locality	→	✓ Field Number	→	+		Contains	→		→			
	Field No: Locality	→	✓ Date	→	+		Equal	→	mm / dd / yyyy	→			
	Field No: Locality	→	Collectors	→	✓ (aggregated)	→		→		→			
	Field No: Locality	→	Locality	→	Geography	→	Continent	→	✓ Full Name	→	+		
	Field No: Locality	→	Locality	→	Geography	→	Country	→	✓ Full Name	→	+		
	Field No: Locality	→	Locality	→	Geography	→	State	→	✓ Full Name	→	+		

Step 2: Map each field to the corresponding identifier via XML

</> DwCA_voucher

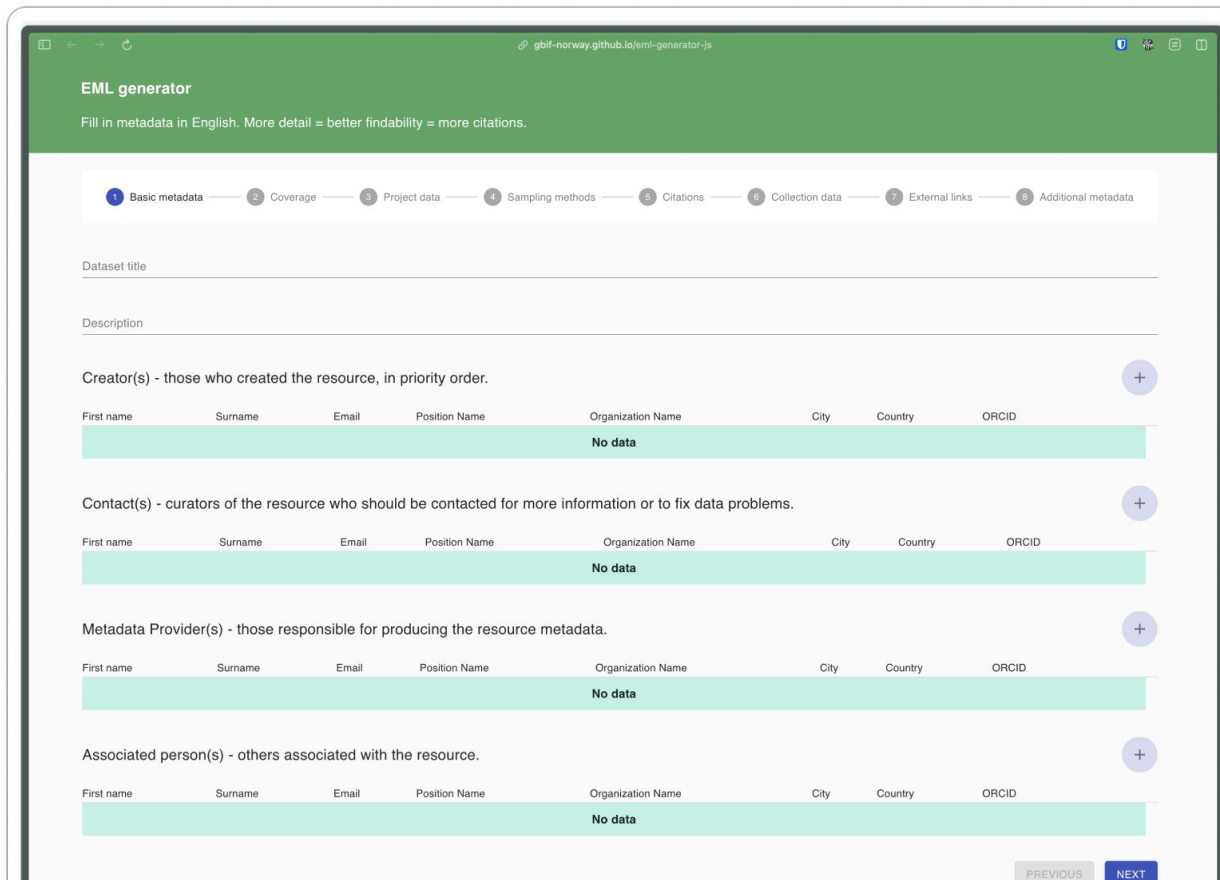


Load File

Download

```
1 <?xml version="1.0" encoding="utf-8"?>
2 <archive>
3   <core rowType="http://rs.tdwg.org/dwc/terms/Occurrence">
4     <queries>
5       <query name="occurrence.csv" contextTableId="1">
6         <id term="http://rs.tdwg.org/dwc/terms/occurrenceID" isNot="false" isRelFld="false" oper="11" stringId="1.collectionobject.guid" value=""/>
7         <field term="http://rs.tdwg.org/dwc/terms/institutionCode" isNot="false" isRelFld="false" oper="11" stringId="1,23,26,96,94.institution.code" value=""/>
8         <field term="http://rs.tdwg.org/dwc/terms/institutionID" isNot="false" isRelFld="false" oper="11" stringId="1,23,26,96,94.institution.altName" value=""/>
9         <field term="http://purl.org/dc/terms/license" isNot="false" isRelFld="false" oper="11" stringId="1,23,26,96,94.institution.copyright" value=""/>
10        <field term="http://purl.org/dc/terms/accessRights" isNot="false" isRelFld="false" oper="11" stringId="1,23,26,96,94.institution.termsOfUse" value=""/>
11        <field term="http://rs.tdwg.org/dwc/terms/collectionCode" isNot="false" isRelFld="false" oper="11" stringId="1,23.collection.code" value=""/>
12        <field term="http://rs.tdwg.org/dwc/terms/basisOfRecord" isNot="false" isRelFld="false" oper="11" stringId="1,23.collection.collectionType" value=""/>
13        <field term="http://rs.tdwg.org/dwc/terms/datasetName" isNot="false" isRelFld="false" oper="11" stringId="1,23.collection.description" value=""/>
14        <field term="http://rs.tdwg.org/dwc/terms/datasetID" isNot="false" isRelFld="false" oper="11" stringId="1,23.collection.guid" value=""/>
15        <field term="http://rs.tdwg.org/dwc/terms/catalogNumber" isNot="false" isRelFld="false" oper="11" stringId="1.collectionobject.catalogNumber" value=""/>
16        <field term="http://rs.tdwg.org/dwc/terms/otherCatalogNumbers" isNot="false" isRelFld="false" oper="11" stringId="1.collectionobject.altCatalogNumber" value=""/>
17        <field term="http://purl.org/dc/terms/modified" isNot="false" isRelFld="false" oper="11" stringId="1.collectionobject.timestampModified" value=""/>
18        <field term="http://rs.tdwg.org/dwc/terms/kingdom" isNot="false" isRelFld="false" oper="11" stringId="1,9-determinations,4-preferredTaxon.taxon.Kingdom" value=""/>
19        <field term="http://rs.tdwg.org/dwc/terms/phylum" isNot="false" isRelFld="false" oper="11" stringId="1,9-determinations,4-preferredTaxon.taxon.Phylum" value=""/>
20        <field term="http://rs.tdwg.org/dwc/terms/class" isNot="false" isRelFld="false" oper="11" stringId="1,9-determinations,4-preferredTaxon.taxon.Class" value=""/>
21        <field term="http://rs.tdwg.org/dwc/terms/order" isNot="false" isRelFld="false" oper="11" stringId="1,9-determinations,4-preferredTaxon.taxon.Order" value=""/>
22        <field term="http://rs.tdwg.org/dwc/terms/family" isNot="false" isRelFld="false" oper="11" stringId="1,9-determinations,4-preferredTaxon.taxon.Family" value=""/>
23        <field term="http://rs.tdwg.org/dwc/terms/genus" isNot="false" isRelFld="false" oper="11" stringId="1,9-determinations,4-preferredTaxon.taxon.Genus" value=""/>
24        <field term="http://rs.tdwg.org/dwc/terms/specificEpithet" isNot="false" isRelFld="false" oper="11" stringId="1,9-determinations,4-preferredTaxon.taxon.Species" value=""/>
25        <field term="http://rs.tdwg.org/dwc/terms/infraspecificEpithet" isNot="false" isRelFld="false" oper="11" stringId="1,9-determinations,4-preferredTaxon.taxon.Subspecies" value=""/>
26        <field term="http://rs.tdwg.org/dwc/terms/scientificName" isNot="false" isRelFld="false" oper="11" stringId="1,9-determinations,4-preferredTaxon.taxon.fullName" value=""/>
27        <field term="http://rs.tdwg.org/dwc/terms/scientificNameAuthorship" isNot="false" isRelFld="false" oper="11" stringId="1,9-determinations,4.taxon.Species Author" value=""/>
28        <field term="http://rs.tdwg.org/dwc/terms/taxonRank" isNot="false" isRelFld="false" oper="11" stringId="1,9-determinations,4,77-definitionItem.taxontreedefitem.name" value=""/>
29        <field term="http://rs.tdwg.org/dwc/terms/identifiedBy" isNot="false" isRelFld="true" oper="11" stringId="1,9-determinations,5-determiner.agent.determiner" value=""/>
30        <field isNot="false" isRelFld="false" oper="6" stringId="1,9-determinations.determination.isCurrent" value=""/>
31        <field term="http://rs.tdwg.org/dwc/terms/dateIdentified" isNot="false" isRelFld="false" oper="11" stringId="1,9-determinations.determination.determinedDate" value=""/>
32        <field term="http://rs.tdwg.org/dwc/terms/typeStatus" isNot="false" isRelFld="false" oper="11" stringId="1,9-determinations.determination.typeStatusName" value=""/>
33        <field term="http://rs.tdwg.org/dwc/terms/eventID" isNot="false" isRelFld="false" oper="11" stringId="1,10.collectingevent.guid" value=""/>
34        <field term="http://rs.tdwg.org/dwc/terms/fieldNumber" isNot="false" isRelFld="false" oper="11" stringId="1,10.collectingevent.stationFieldNumber" value=""/>
35        <field term="http://rs.tdwg.org/dwc/terms/eventDate" isNot="false" isRelFld="false" oper="11" stringId="1,10.collectingevent.startDate" value=""/>
36        <field term="http://rs.tdwg.org/dwc/terms/dw" isNot="false" isRelFld="false" oper="11" stringId="1,10.collectingevent.startDataNumber" value=""/>
```

Step 3: Create or grab Ecological Metadata Language (EML)



The screenshot shows the 'EML generator' web application. At the top, a green header contains the title 'EML generator' and the instruction 'Fill in metadata in English. More detail = better findability = more citations.' Below the header is a progress bar with eight steps: 1. Basic metadata (active), 2. Coverage, 3. Project data, 4. Sampling methods, 5. Citations, 6. Collection data, 7. External links, and 8. Additional metadata. The main form area includes a 'Dataset title' field, a 'Description' field, and four sections for adding metadata providers: 'Creator(s)', 'Contact(s)', 'Metadata Provider(s)', and 'Associated person(s)'. Each section has a '+' button and a table with columns for First name, Surname, Email, Position Name, Organization Name, City, Country, and ORCID. All tables currently display 'No data'. At the bottom right, there are 'PREVIOUS' and 'NEXT' buttons.

EML generator

Fill in metadata in English. More detail = better findability = more citations.

1 Basic metadata — 2 Coverage — 3 Project data — 4 Sampling methods — 5 Citations — 6 Collection data — 7 External links — 8 Additional metadata

Dataset title

Description

Creator(s) - those who created the resource, in priority order. +

First name	Surname	Email	Position Name	Organization Name	City	Country	ORCID
No data							

Contact(s) - curators of the resource who should be contacted for more information or to fix data problems. +

First name	Surname	Email	Position Name	Organization Name	City	Country	ORCID
No data							

Metadata Provider(s) - those responsible for producing the resource metadata. +

First name	Surname	Email	Position Name	Organization Name	City	Country	ORCID
No data							

Associated person(s) - others associated with the resource. +

First name	Surname	Email	Position Name	Organization Name	City	Country	ORCID
No data							

PREVIOUS NEXT

**EML Generator
by GBIF Norway**

<https://github.com/gbif-norway/eml-generator-js>

Step 4: Upload the data set's ecological metadata (EML)

</> DwCA_voucher_metadata 



Load File

Download

```
1 <eml:eml xmlns:eml="eml://ecoinformatics.org/eml-2.1.1"
2   xmlns:dc="http://purl.org/dc/terms/"
3   xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
4   xsi:schemaLocation="eml://ecoinformatics.org/eml-2.1.1 http://rs.gbif.org/schema/eml-gbif-profile/1.0.2/eml.xsd"
5   packageId="8f79c802-a58c-447f-99aa-1d6a0790825a/v17" system="http://gbif.org" scope="system"
6   xml:lang="eng">
7 <dataset>
8   <alternateIdentifier>8f79c802-a58c-447f-99aa-1d6a0790825a</alternateIdentifier>
9   <alternateIdentifier>http://ipt.nhm.ku.edu/ipt/resource.do?r=kubi_ichthyology</alternateIdentifier>
10  <title xml:lang="eng">KUBI Ichthyology Collection</title>
11  <creator>
12    <individualName>
13      <givenName>Andrew</givenName>
14      <surName>Bentley</surName>
15    </individualName>
16    <organizationName>KU Biodiversity Institute</organizationName>
17    <positionName>Collection Manager</positionName>
18    <address>
19      <deliveryPoint>1345 Jayhawk Blvd.</deliveryPoint>
20      <city>Lawrence</city>
21      <administrativeArea>KS</administrativeArea>
22      <postalCode>66045</postalCode>
23      <country>US</country>
24    </address>
25    <phone>+01 785-864-3863</phone>
26    <electronicMailAddress>abentley@ku.edu</electronicMailAddress>
27    <onlineUrl>http://ichthyology.biodiversity.ku.edu/</onlineUrl>
28  </creator>
29  <metadataProvider>
30    <individualName>
31      <givenName>Andrew</givenName>
32      <surName>Bentley</surName>
33    </individualName>
34    <organizationName>KU Biodiversity Institute</organizationName>
```

Step 4½: Validate data set with GBIF

Using the GBIF API, users can validate that the export data set meets the requirements for publishing.

- Verifies data mappings to Darwin Core profiles
- Ensures all required fields are present and correctly filled
- Validates data quality (date format, geodetic datum)
- Detects internal inconsistencies (e.g., coordinates outside declared country)
- Generates detailed errors and warnings for indexing issues
- Reviews metadata structure and required content (e.g., contacts)
- Offers recommendations to improve data and metadata quality



Step 4: Define the export procedure and publish!

↑ ExportFeed ✎

Visual Editor ▾  Load File Download

Title
KUBI Ichthyology RSS Feed

Description
RSS feed export

Language

Exports



Title	DwCA Resource
KU Fish	DwCA_tissue 
Description	Metadata Resource
	DwCA_voucher_metadata 
File name	Run as user
kui-dwca.zip   	abentley
	Send completion notification to user
	abentley   
ID	Run in collection
B8B6960B-394E-4561-81C0-2A32789EFB74	KU Fish Voucher Collection ▾
GUID	☒ Publish
	Publish every N days
	7

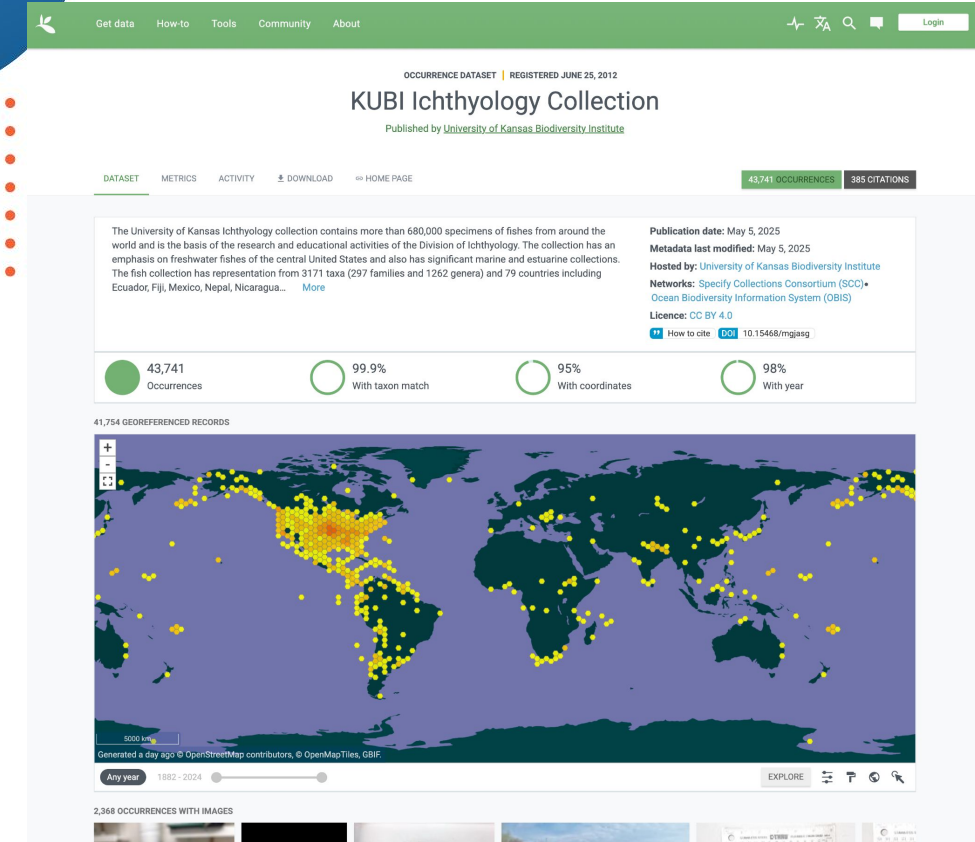


GBIF



New data
available!

Step 5: Publish & verify data set appears correctly downstream!



Recordset

[Search Recordset](#)

KUBI Ichthyology Collection

Specimen Records: 42,751

Media Records: 0

IDigBio Last Ingested Date: 2023-01-11

The University of Kansas Ichthyology collection contains more than 680,000 specimens of fishes from around the world and is the basis of the research and educational activities of the Division of Ichthyology. The collection has an emphasis on freshwater fishes of the central United States and also has significant marine and estuarine collections. The fish collection has representation from 3171 taxa (297 families and 1262 genera) and 79 countries including Ecuador, Fiji, Mexico, Nepal, Nicaragua and various marine localities. The collections and the scope of research activities in the division continue to grow due to the ongoing activities of Ichthyology staff and students. The collection is used by national and international researchers as well as by state and federal agencies. The Division of Ichthyology is designated as a Regional Center in the Midwest and Great Plains Regions (Collette & Lachner 1976, Copela 1976: 625-642; Poss and Collette 1995, Copela 1995: 48-70) and is among the top twenty ichthyological collections in the country. Almost 60% of the specimens in the collection are from the Great Plains Region. The collection is an important resource for anyone interested in the region's fishes. The data concerning these faunas are not extensively duplicated by other ichthyological collections. The Ichthyology collection comprises whole, wet, voucher specimens in 70% ethanol (88%), specimens maintained as dry skeletons (0.3%) and cleared and stained preparations in glycerine (1.7%). Other ancillary collections include a large slide and digital image collection of specimens and locations (some of which are linked to specimen records through the database), a large field note collection and a map collection. The wet specimens are housed in a state-of-the-art fluid collection facility, opened in 1996. The facility has 2,400 square feet of collection storage space on four floors and is shared with Herpetology, Mammalogy, Ornithology, Entomology and Invertebrate Zoology. The collection storage environment is maintained at 65°F year-round by an HVAC system. Specimens are housed in state of the art glassware and stainless steel tanks and protected by a sprinkler fire suppression system and UV-shielded lighting. The division has both morphological research facilities and a shared molecular systematic laboratory.

Contacts

Name	Andrew Bentley	Name	Andrew Bentley
Role	Bioinformatics Manager	Role	Collection Manager
Email	abentley@ku.edu	Email	abentley@ku.edu
Name	William Leo Smith	Name	Andrew Bentley
Role	Curator	Role	Collection Manager
Email	leosmith@ku.edu	Email	abentley@ku.edu

[Data Corrected](#) [Data Use](#) [Raw](#)

This table shows any data corrections that were performed on this recordset to improve the capabilities of IDigBio Search. The first column represents the correction performed. The last two columns represent the number and percentage of records that were corrected. A complete list of the data quality flags and their descriptions can be found [here](#). Clicking on a data flag name will take you to a search for all records with this flag in this recordset.

Flag	Records With This Flag	(%) Percent With This Flag
dwc_datasetid_added 	41958	98.145
dwc_parentnameusageid_added 	41958	98.145

Contents

[Description](#)
[Contacts](#)
[All Data](#)



Future Development



Roadmap

- **Version 1** (Planned for 2026):
 - New schema mapper & data set publishing toolkit
 - Default schema mappings for Darwin Core and popular GBIF extensions
- **Version 2** (TBD):
 - Integrate member feedback & additional defaults
 - Connect to the GBIF Validator API inside Specify
 - *Web Portal exporting*

C0 Collection Object	Det Determination
Voucher	(aggregated)
Acc Accession #	✓ Current
Acc Cataloger	Date
CO Children	Date (Day)
CoA Col Obj Attribute	Date (Month)
Col Collection	Date (Year)
CoA Collection Object Attachments	GUID
Cit Collection Object Citations	In question
Det Determinations	Remarks

→

Schema Concept	Field Name
occurrenceID	C0 Collection Object - GUID
catalogNumber	C0 Cat #
scientificName	Tax Taxon - Full Name
scientificNameAuthorship	Tax Taxon - Author
organismQuantity	Tax Order - Full Name
class	Tax Class - Full Name
family	Tax Family - Full Name
genus	Tax Genus - Full Name
specificEpithet	Tax Species - Full Name
typeStatus	Det Type Status
fieldNumber	CF Field Number

☐ Reveal Hidden Form Fields