

OBIS-Wan t'Knowbi

(OBIS Guide)

Contents

An Introduction to OBIS	2
Helpful external references	2
Accessing data from Open Government.....	4
Navigating the Open Government website	4
Data Translation.....	7
Translating datasets into the Darwin Core (DwC) standard.	7
Required column headers.....	8
Sample tables.....	9
WoRMS taxonomic match	11
WoRMS taxa match using the web tool.....	11
WoRMS Taxa match using R.....	14
Navigating the Integrated Publishing Toolkit (IPT) website.....	15
Requesting an IPT account.....	17
Creating and uploading a new dataset to Integrated Publishing Toolkit (IPT)	17
Creating new resources and uploading datasets to the Integrated Publishing Toolkit (IPT)	17
Uploading your data to IPT	20
Source Data:.....	20
Source Data Page:	20
Darwin Core Mappings:	21
Value Translation:	22
Metadata:.....	23
Published Versions:.....	23
Visibility:.....	24
Resource Managers:	24
Review your entry	25
Removing datasets on IPT	26
IPT Tips and Tricks.....	27

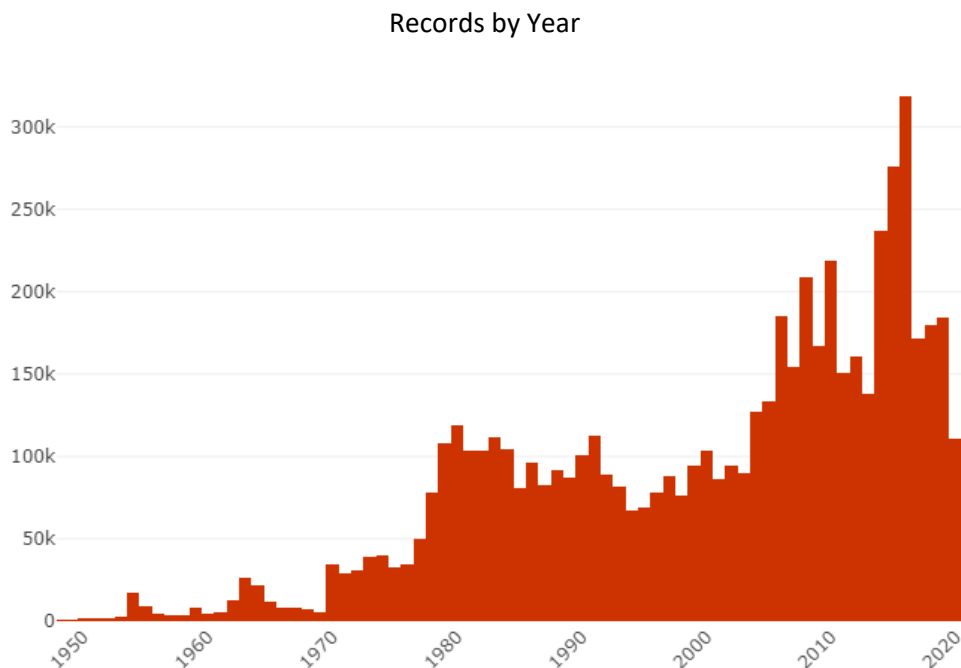
About this guide

This OBIS guide was created to support the process in transforming biological data from the source to the Darwin Core Archive (DwC-A) standard and publishing to OBIS via IPT. The guide is tailored to DFO dataholders and is NOT a complete and exhaustive guide for DwC-A, IPT, or OBIS. Please see the [Helpful external references](#) section for existing complete guides.

An Introduction to OBIS

Ocean Biodiversity Information System (OBIS) is an online information system and repository where scientists can contribute and query relational data in Darwin Core Archive (DwC-A) format. The OBIS mission is to “build and maintain a global alliance that collaborates with scientific communities to facilitate free and open access to, and application of, biodiversity and biogeographic data and information on marine life”. There are currently over 7.4 million occurrence records published to the OBIS Canada node from 164 datasets. Globally, more than 3,900 datasets have been shared through OBIS, representing more than 71 million occurrence records.

According to [OBIS Canada](#), March 2021.



Helpful external references

OBIS Website

[OBIS Manual](#)

[OBIS R guide](#)

[OBIS Canada Node](#)

Open Government (Open Canada)

Integrated Publishing Toolkit (IPT)

OBIS · GitHub

[OBIS R Resources - iobis/robis](#)

[OBIS R tools - iobis/obistools](#)

[IPT Manual on Github](#)

[Bio data guide](#)

[OBIS workshop by OTN \(English\)](#)

[OBIS workshop by OTN \(French\)](#)

Darwin Core (DwC) Quick Reference Guide

Standardized vocabulary – P01

Institution Code Search

[OceanTeacher OBIS Course](#) (requires a free account registration)

OBIS Slack group

Accessing data from Open Government

Data published on the [Open Government](http://www.open.canada.ca) website can readily be translated to DwC-A format for publication to OBIS Canada. Please make sure that the dataset holder/author is aware of your intent to publish their data and that they support the publication of their data to OBIS Canada. Ideally the dataset holder/author will conduct a final review of the OBIS-formatted dataset before their data is published and released to the public through OBIS Canada.

Navigating the Open Government website

Navigate to the Open Government (www.open.canada.ca). Publications to OBIS Canada MUST be published on Open Canada or have been approved to be published through Open Canada.

- Datasets vary greatly by structure and content, so one must be familiar with the OBIS standard and Darwin Core structure before translating the datasets to publish onto OBIS Canada.
- Most or all of the metadata required for OBIS Canada can be found on the Open Canada page for a specific dataset; if there are missing information, please contact the data author (data holders' contact information are listed on Open Canada).

Search for the datasets in the search bar below highlighted in the **red box**.

Open Government

Open Government is about making government more accessible to everyone. Participate in conversations, find data and digital records, and learn about open government.

Search data and information

Browse the collection of more than 80,000 open data and information assets

 Find

Proactive disclosure

Browse the proactive disclosure of financial and human resources-related information by federal departments and agencies.

Government Contracts ▼

Go

Access to information

Browse completed Access to Information summaries, make an ATIP request, and learn more about Access to Information in Canada.

Search completed Access to Inform. ▼

Go

A list of databases will generate when a dataset match with the keywords highlighted in yellow. Click on the [underlined title](#) to go into the dataset page.

Mortality events at British Columbia marine finfish aquaculture sites

Federal

Mortality at salmon aquaculture facilities is closely monitored. If the amount of dead **fish** at a farm exceeds thresholds outlined in conditions of licence, a **mortality** event is said to have occurred and must be reported to DFO within 24 hours of discovery. Facility managers and veterinarian(s) must determine the probable cause of the event and develop a plan to mitigate ongoing harm to the farmed **fish** and reduce any risk to wild **fish**. Companies must continue to update Fisheries and Oceans Canada (DFO) every 10 days for the duration of the **mortality** event, including daily **mortality** counts, mitigation applied, determination of the cause(s) of the event and any updated plan.

Last Updated: Jan. 11, 2021 **Date Published:** Mar. 12, 2019

The dataset page displays various information about the dataset in a format similar to below:

- Contact information, keywords, and metadata.
- Summary of dataset, publishing, and rights details

Mortality events at British Columbia marine finfish aquaculture sites

Mortality at salmon aquaculture facilities is closely monitored. If the amount of dead fish at a farm exceeds thresholds outlined in conditions of licence, a mortality event is said to have occurred and must be reported to DFO within 24 hours of discovery. Facility managers and veterinarian(s) must determine the probable cause of the event and develop a plan to mitigate ongoing harm to the farmed fish and reduce any risk to wild fish. Companies must continue to update Fisheries and Oceans Canada (DFO) every 10 days for the duration of the mortality event, including daily mortality counts, mitigation applied, determination of the cause(s) of the event and any updated plan.

This report provides a summary of mortality events reported by aquaculture companies to Fisheries and Oceans Canada.

Historical data are available from 2011 to the present. Explanation of the terms used in the report's column headings can be found in the terminology file below.

Publisher - Current Organization Name: Fisheries and Oceans Canada
Publisher - Organization Section Name: Aquaculture, Communications
Licence: Open Government Licence - Canada

Contact Email:
webmasterpacrhq@dfo-mpo.gc.ca

Keywords:
Fisheries and Oceans Canada
Pacific region | Aquaculture
fish farming | salmon farming
fish health monitoring
fish health events
British Columbia | BC

Subject:
Nature and Environment

Place of Publication: British Columbia

Homepage: <http://www.dfo-mpo.gc.ca/aquaculture/protect-protege/reduce-disease-reduire-maladie-eng.html>

Series Title: Public Reporting on Aquaculture - Pacific Region

Maintenance and Update Frequency: As Needed

Date Published: 2019-03-12

Date Modified: 2020-07-10

Temporal Coverage: 2011-06-21 to 2018-12-02

Openness Rating: ★★☆☆

Resources

Show entries

Resource Name	Resource Type	Format	Language	Links
---------------	---------------	--------	----------	-------

Each dataset contains a list of resources similar to the example provided below that contain the dataset files and/or supplemental files associated with the dataset. Common file formats include CSV, XLS, TXT, and possible external links to where the data and supplemental files are held. An English and French version will be available to download for most datasets. As of now, we only publish datasets in *ENGLISH* onto OBIS Canada. Datasets published in French will require translation prior to OBIS publication.

Resources

Show entries

Resource Name 	Resource Type 	Format 	Language 	Links
Mortality events at BC marine finfish aquaculture sites	Terminology	XLSX	English	Access
Mortality events at BC marine finfish aquaculture sites	Terminology	XLSX	French	Access
Mortality events at BC marine finfish aquaculture sites	Terminology	CSV	English	Access
Mortality events at BC marine finfish aquaculture sites	Terminology	CSV	French	Access
Mortality events at BC marine finfish aquaculture sites 2011 and ongoing	Dataset	XLSX	English	Access
Mortality events at BC marine finfish aquaculture sites 2011 and ongoing	Dataset	XLSX	French	Access
Mortality events at BC marine finfish aquaculture sites 2011 and ongoing	Dataset	CSV	English	Access
Mortality events at BC marine finfish aquaculture sites 2011 and ongoing	Dataset	CSV	French	Access

Data Translation

Dataset translation can be done with various tools, including Microsoft Excel, Microsoft Access, R, etc. This guide will provide a basic visual representation of the tables in a sample dataset. This process would vary based on which tool the user choose to use for the translation process.

Follow the links below to the full guides:

[OBIS Darwin Core](#)

[OBIS R guide](#)

[Template Table](#) from Github

Translating datasets into the Darwin Core (DwC) standard.

This step requires knowledge of the dataset and the DwC-A standards, more information can be found [here](#):

<https://obis.org/manual/darwincore/>

<https://dwc.tdwg.org/terms/>

For taxonomic matching, please use [WoRMs](#) ([jump to taxon matching](#))

A Darwin Core Archive (DwC-A) contains up to 4 components, which include the following:

- Core files (an Event Core and an Occurrence Core are standard for most datasets)
- Extension files (Extended Measurement or Fact or 'EMoF') - optional
- eml.xml (Ecological Metadata Language)
- meta.xml (describes the linkage between the archives)

* IPT contains a metadata editor for the datasets that is capable of compiling the eml.xml and the meta.xml files automatically from the editor so these files are not required to be created separately. Alternatively, datasets can be linked to a database provided there is another platform and infrastructure to host the database. For more information on entering metadata on IPT, [jump to the metadata section](#).

Required column headers

The following column headers are the minimum required for the DwC-A format to be uploaded onto IPT. The absence of any one of these columns will prevent a dataset from being published on IPT.

Refer to the [DwC-A Quick Reference guide](#) for a complete list of terms by usage and file type (the navigation menu on the right will allow you to readily access definitions and examples for terms used in Event Core, Occurrence Core and EMOF files).

Column Header	Example	Details
occurrenceID	PBS-15-2	An unique identifier and Primary Key (PK) for each occurrence (row). Typically the occurrenceID is concatenated from institutionCode + collectionCode + catalogNumber. However, it can be a concatenated with other identifiers as well.
eventDate	2021-02-11	The date, in ISO 8601 standard
decimalLatitude	49.2354	Decimal latitude as spatial reference system EPSG:4326
decimalLongitude	-112.5342	Decimal longitude as spatial reference system EPSG:4326
scientificName	<i>Oncorhynchus tshawytscha</i>	The scientific name of the sample with the highest precision possible. Use originally recorded name.
scientificNameID	urn:lsid:marinespecies.org:taxname:158075	Matched from WoRMS, enter ID match of the originally recorded scientificName even if the term is no longer accepted.
occurrenceStatus	Present or Absent	Presence or absence of the taxon
basisOfRecord	HumanObservation	Based on the method of data collection; HumanObservation, PreservedSpecimen, LivingSpecimen, etc.

Sample tables

An example of each of the Event Core, Occurrence Core, and EMOF files are shown below with some common column headers. The dataset is that of a bottom trawl, where trawl catch data along with a subset of biodata are captured.

*some headers are optional depending on the dataset.

*The use of eventRemarks or occurrenceRemarks columns can be used in each of the respective event core or occurrence core files to provide additional remarks or information.

Event Core

eventID	eventDate	decimalLatitude	decimalLongitude	rightsHolder	institutionCode	institutionID	eventRemarks
R90-001	1990-07-25	48.6568	-124.9608	Her Majesty the Queen in right of	PBS	https://edmo.seadatanet.org/report/4180	Sunny

Column Header	Example	Details
eventID	R90-001	Unique identifier for the event
eventDate	1990-07-05	Date of event, in ISO 8601 YYYY-MM-DD format
decimalLatitude	48.6568	Decimal latitude
decimalLongitude	-124.9608	Decimal Longitude
rightsHolder	Her Majesty the Queen in.....	Rights holder of dataset,
institutionCode	PBS	Institution Code from EDMO
institutionID	https://edmo.seadatanet.org/report/4180	Institution ID (link) from EDMO
eventRemarks	Sunny	Remarks about the event

Occurrence

eventID	lifeStage	occurrenceID	vernacularName	scientificName	scientificNameID	Kingdom	Phylum	Class	Order	Family	Genus	specificEpithet	occurrenceStatus	basisOfRecord
R90-001	smolt	R90-001-1	SOCKEYE SALMON	Oncorhynchus nerka	urn:lsid:marinespecies.org:t	Animalia	Chordata	Actinopterygii	Salmoniformes	Salmonidae	Oncorhynchus	nerka	present	HumanObservation
R90-001	smolt	R90-001-2	CHINOOK SALMON	Oncorhynchus tshawytscha	urn:lsid:marinespecies.org:t	Animalia	Chordata	Actinopterygii	Salmoniformes	Salmonidae	Oncorhynchus	tshawytscha	present	HumanObservation
R90-001	smolt	R90-001-3	COHO SALMON	Oncorhynchus kisutch	urn:lsid:marinespecies.org:t	Animalia	Chordata	Actinopterygii	Salmoniformes	Salmonidae	Oncorhynchus	kisutch	present	HumanObservation

Column Header	Example	Details
eventID	R90-001	Reference to the eventID
occurrenceID	R90-001-1	Unique identifier for the occurrence
lifeStage	Smolt	Life stage of the sample
vernacularName	Chum Salmon	Common name of the sample
scientificName	<i>Oncorhynchus keta</i>	Scientific name of the sample
Phylum, Class, Order, etc.	Animalia, Chordata, etc.	Taxonomic rank info for the sample
specificEpithet	keta	The second half of a species' binomial name, separated from its generic counterpart
occurrenceStatus	Present	Presence or absence of a species sample
basisOfRecord	HumanObservation	Based on method of data collection

Event EMoF

eventID	measurementType	measurementTypeID	measurementValue	measurementUnit	measurementUnitID	measurementMethod
R90-001	Distance Off Shore		2	nautical miles		Port Beam Trawl
R90-001	Duration	http://vocab.nerc.ac.uk	1.0333	hours	http://vocab.nerc.ac.uk/co	Port Beam Trawl
R90-001	TowSpeed	http://vocab.nerc.ac.uk	3.7	knots	http://vocab.nerc.ac.uk/co	Port Beam Trawl

Column Header	Example	Details
eventID	R90-001	Links the measurements to the taxonomic occurrence defined in the Event Core
measurementType	Duration	Type of measurement
measurementTypeID	http://vocab.nerc.ac.uk/collection/P01/current/AZDRZZ01/2/	Search the ID on NVS *Some vocabulary terms will require the introduction of a definition to NVS and the creation of an ID — it's an ever evolving system.
measurementValue	1.0333	The value of the measurement
measurementUnit	Hours	The unit of the measurement
measurementUnitID	http://vocab.nerc.ac.uk/collection/P06/current/UHOR/	Search for the ID on NVS
measurementMethod	Port Beam Trawl	Method of measurement

*Contact the [OBIS Canada node manager](#) for information on how to submit vocabulary request to NVS.

Occurrence EMoF

occurrenceID	measurementType	measurementTypeID	measurementValue	measurementUnit	measurementUnitID
2000-7-25-10-39:1	weight	http://vocab.nerc.ac.uk/collection/	0.5	Kilograms	http://vocab.nerc.ac.uk/collection/
2000-7-25-10-39:2	fork length	http://vocab.nerc.ac.uk/collection/	12	millimetres	http://vocab.nerc.ac.uk/collection/

Column Header	Example	Details
occurrenceID	2000-7-25-10-39:1	Links the measurements to the taxonomic occurrence defined in the Occurrence Core
measurementType	Weight	Type of measurement
measurementTypeID	http://vocab.nerc.ac.uk/collection/.....	Search the ID on NVS *Some vocabulary terms will require the introduction of a definition to NVS and the creation of an ID — it's an ever evolving system.
measurementValue	0.5	The value of the measurement

measurementUnit	Kilograms	The unit of the measurement
measurementUnitID	http://vocab.nerc.ac.uk/collection/P06/current/EFKC/	Search the ID on NVS

*Contact the [OBIS Canada node manager](#) for information on how to submit vocabulary request to NVS.

WoRMS taxonomic match

WoRMS, or the World Register of Marine Species, hosts an extensive library of marine taxa and a [highly effective taxa matching tool](#).

WoRMS taxa match using the web tool

To obtain complete taxonomic profiles for the species listed in your dataset, draft a simple list of scientific names in an input file in txt, csv, or xls format. Ensure that any identification qualifiers — “sp.” for example — are removed from your list prior to upload.

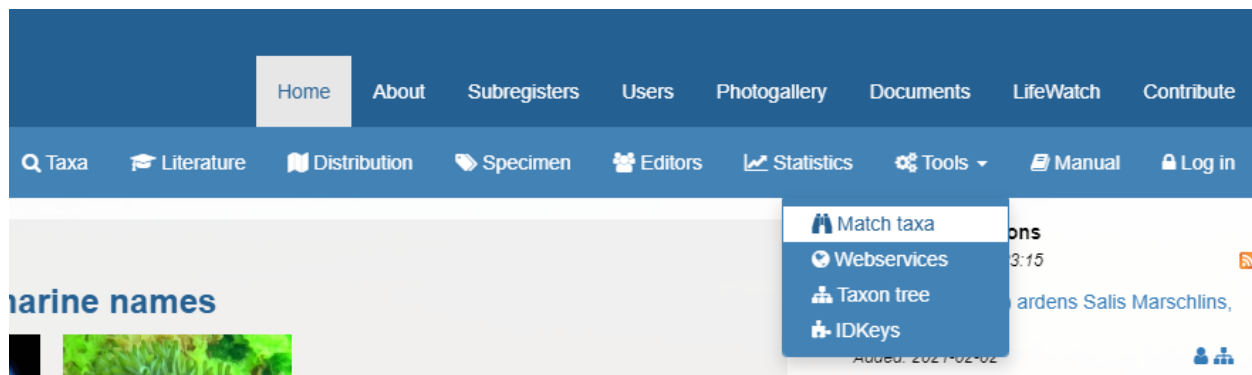
To access the WoRMS taxonomic matching tool, visit the [WoRMS website](#), select the “Tools” dropdown and click on the “Match taxa” button.

Note: If the lowest level (species) cannot be matched, move on to the next highest level (genus, family, etc.).

Example of an input taxa file below:

```
Taxa4.txt - Notepad
File Edit Format View Help
Prionace glauca
Galeorhinus galeus
Squalus acanthias Linnaeus
Sebastes Cuvier
Oncorhynchus gorbuscha
Oncorhynchus keta
Oncorhynchus kisutch
Oncorhynchus mykiss
Loligo Lamarck
Clupea pallasii Valenciennes
Scomber colias Gmelin
Oncorhynchus tshawytscha
Anoplopoma fimbria
Ophiodon elongatus Girard
Gadus chalcogrammus Pallas
Oncorhynchus nerka
Merluccius productus
Trachurus symmetricus
Sebastes melanops Girard
Sebastes paucispinis Ayres
Sebastes ruberrimus
Scorpaenichthys marmoratus
```

To access the WoRMS taxonomic match tool, visit the [WoRMS website](#) and select the “Tools” dropdown and click on the “Match taxa” button.



On the Match taxa page of WoRMS, select the “Choose File” button to upload the taxonomic input file. Select LSID (the IPT-compatible “scientificNameID”), Scientific Name, Authority (“scientificNameAuthorship”), Accepted name, Taxon status and Environment to include this information in your output file then click “Next >”

WoRMS Taxon match

You can use the WoRMS Taxon Match Tool ([credits](#)) to automatically match your species list or taxon list with WoRMS. After matching, the tool will return your file with the AphiaID's, va
For performance reasons, the limit is set to 1,500 rows. For matching larger files, non-marine or multiple datasources, please use the [Lifewatch Species Information Backbone](#).

File * No file chosen

Allowed filetypes: Plain text [TXT], Comma Separated [CSV] & Excel sheet [XLS, XLSX]

Row delimiter ☐ First row contains column names

Column delimiter

Match authority ☐

Match upto Higher taxa only possible if a full classification is given in additional columns

Limit to

Limit to taxa belonging to

Output ☒ AphiaID ☒ LSID ☐ TSN ☒ ScientificName ☐ Authority ☒ Accepted name ☐ Classification ☐ Qualitystatus ☐ Taxon status ☐ Environment ☐ Citation

The WoRMS Taxon Match will read the file and prompt a column header selection for a list of scientific names.

WoRMS Taxon match

Preview for the file 'taxa.txt' (first 20 records) [\[new match\]](#)
Please select a WoRMS term that corresponds to your column and click 'Match'.

ScientificName

- Pseudopleuronectes Americanus
- Tautoglabrus Adspersus
- Homarus Americanus
- Mytilidae
- Ophiuroidea
- Anthozoa
- Osmerus Mordax
- Scyphozoa
- Limanda Ferruginea
- Gasterosteus Aculeatus
- Strongylocentrotus
- Ammodytes Dubius
- Cancer Irroratus
- Mollusca
- Myoxocephalus Octodecemspinosus
- Crangon

[\[options\]](#) [\[preview\]](#)

[< Back](#) [Match](#)

Choose “ScientificName” from the dropdown menu and click “Match.”

WoRMS Taxon match

Match preview for the file 'taxa.txt' - matching: 94.54% [\[new match\]](#)
If available, please select the [WoRMS](#) taxon that corresponds to your taxon. Then click 'Download'.

Cancer Irroratus	Cancer Linnaeus, 1758
Mollusca	Mollusca
Myoxocephalus Octodecemspinosus	Myoxocephalus Tilesius, 1811
Crangon	(ambiguous - select below)
Gadus Morhua	(ambiguous - select below)
Aspidophoroides Monopterygius	Aspidophoroides Lacepède, 1801
Chionoecetes Opilio	Chionoecetes Krøyer, 1838
Asterias	Asterias Linnaeus, 1758
Urophycis Tenuis	Urophycis Gill, 1863
Hemitripterus Americanus	Hemitripterus Cuvier, 1829
Pectinidae	Pectinidae Rafinesque, 1815
Scophthalmus Aquosus	Scophthalmus Rafinesque, 1810
Leucoraja Ocellata	Leucoraja Malm, 1877
Myoxocephalus Scorpis	Myoxocephalus Tilesius, 1811
Clypeasteroida	Clypeasteroida A. Agassiz, 1872
Holothuroidea	Holothuroidea

☐ Excel sheet (XLS)
☐ Excel sheet (XLSX)
☒ Text file
☐ SGML

[< Back](#) [Download](#)

Once the match is completed, toggle your choice of output format (XLS, XLSX, TXT or SGML) and click “Download.” If an ambiguous match occurs, a drop down will appear and a selection can be made manually for each taxa.

WoRMs Taxa match using R

There are two R packages that you can use to conduct taxa matches through WoRMS: obistools and worrms.

Make sure to have a list of unique taxa for your dataset before using the functions. The function `match_taxa()` in the obistools package only returns the LSID but can work with larger data sets. The `worms_records_names()` and `worms_records_name()` functions in the worrms package returns LSID, all taxonomy, authorship and taxon rank from the scientific name.

OBIS Tools R package:

```
1 #worms taxonomy search using worrms package
2 #***only works if no ambiguous matches and max 50 rows***
3 #returns all taxonomy, authority and taxon rank
4 install.packages("worms")
5 library("worms")
6
7 #***only works if no ambiguous matches and max 50 rows***
8 #returns all taxonomy, authority and taxon rank
9 worms_taxonomy <- wm_records_names(name = uniquetaxa$scientificName)
10
11 #convert the list of tibbles to one data frame
12 worms_taxa_df <- dplyr::bind_rows(worms_taxonomy)
13
14 #add specific epithet
15 worms_taxa_df <- worms_taxa_df %>% dplyr::mutate(specificEpithet = stringr::word(scientificname, 2))
16
17 #can also search just one name
18 salmon_taxonomy <- wm_records_name(name = "Salmo salar")
19
20
21
```

Worrms R package:

```
1 #worms taxonomy search using worrms package
2 #***only works if no ambiguous matches and max 50 rows***
3 #returns all taxonomy, authority and taxon rank
4 install.packages("worms")
5 library("worms")
6
7 #***only works if no ambiguous matches and max 50 rows***
8 #returns all taxonomy, authority and taxon rank
9 worms_taxonomy <- wm_records_names(name = uniquetaxa$scientificName)
10
11 #convert the list of tibbles to one data frame
12 worms_taxa_df <- dplyr::bind_rows(worms_taxonomy)
13
14 #add specific epithet
15 worms_taxa_df <- worms_taxa_df %>% dplyr::mutate(specificEpithet = stringr::word(scientificname, 2))
16
17 #can also search just one name
18 salmon_taxonomy <- wm_records_name(name = "Salmo salar")
19
20
21
```

*The worrms package can only handle maximum 50 taxa at a time.

Navigating the Integrated Publishing Toolkit (IPT) website

This will be a high level introduction to IPT for the purpose of uploading datasets, please visit the [full length IPT Manual](#) for more information.

To navigate to the IPT website for the Canadian node, please visit: <http://ipt.iobis.org/obiscanada/>

Sign in with your email and your password on the top right hand corner of the IPT home page.

OBIS Canada INTEGRATED PUBLISHING TOOLKIT (IPT)
free and open access to biodiversity data

email login **ENGLISH**

Home **About**


Hosted resources available through this IPT

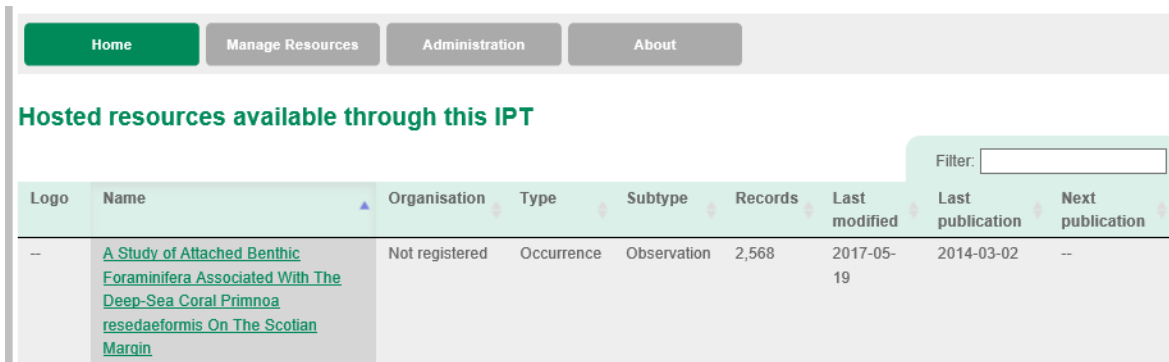
Filter:

Logo	Name	Organisation	Type	Subtype	Records	Last modified	Last publication	Next publication
--	A Study of Attached Benthic Foraminifera Associated With The Deep-Sea Coral <i>Primnoa resedaeformis</i> On The Scotian Margin	Not registered	Occurrence	Observation	2,568	2017-05-19	2014-03-02	--
--	Abundance of hydroids in a mangrove ecosystem at Twin Cays, Belize, Central America	Not registered	Occurrence	Observation	49	2016-03-29	2016-03-29	--
--	Abundance of intertidal algae and invertebrates on the Atlantic coast of Nova Scotia	Not registered	Occurrence	Observation	100,620	2016-04-06	2016-04-06	--
--	Acadia University: Ectoparasites on Atlantic sturgeon (<i>Acipenser oxyrinchus</i> Mitchill, 1814) in the Minas Basin	Not registered	Occurrence	Observation	50	2017-09-03	2017-09-03	--

The navigation tab expands once an account login is established, as shown below:

Home: the home screen of the IPT website for the Canadian node displays datasets that are published and available to the public. Search for specific datasets by entering keywords in the “Filter” box.

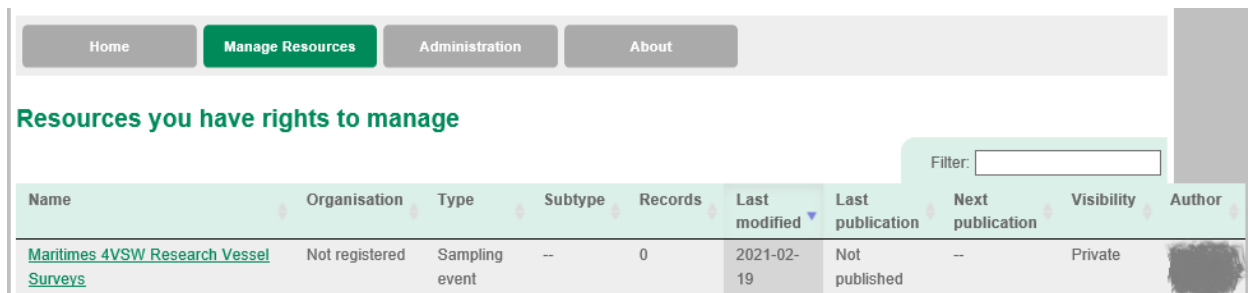
Each column (except Logo) in the dataset table can be sorted in either ascending or descending order by clicking on the  button.




The screenshot shows the 'Home' tab selected in the navigation bar. Below the navigation bar, the heading 'Hosted resources available through this IPT' is displayed. A table lists datasets with columns: Logo, Name, Organisation, Type, Subtype, Records, Last modified, Last publication, and Next publication. A filter box is located above the table. The first row of data is highlighted.

Logo	Name	Organisation	Type	Subtype	Records	Last modified	Last publication	Next publication
--	A Study of Attached Benthic Foraminifera Associated With The Deep-Sea Coral <i>Primnoa resedaeformis</i> On The Scotian Margin	Not registered	Occurrence	Observation	2,568	2017-05-19	2014-03-02	--

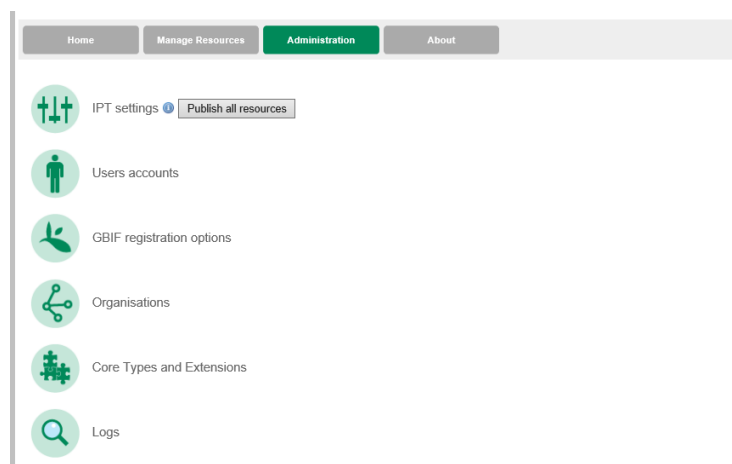
Manage Resources: the datasets that you have permission to manage will be displayed under this tab. Only datasets that you have created or that you have been assigned to will be visible to you on this tab if you have Manager privileges. Administrators will have access to a complete list of all datasets, including published resources (public and private) as well as private unpublished drafts.



The screenshot shows the 'Manage Resources' tab selected in the navigation bar. Below the navigation bar, the heading 'Resources you have rights to manage' is displayed. A table lists resources with columns: Name, Organisation, Type, Subtype, Records, Last modified, Last publication, Next publication, Visibility, and Author. A filter box is located above the table. The first row of data is highlighted.

Name	Organisation	Type	Subtype	Records	Last modified	Last publication	Next publication	Visibility	Author
Maritimes 4VSW Research Vessel Surveys	Not registered	Sampling event	--	0	2021-02-19	Not published	--	Private	

Administration: only accessible by administrator accounts, the ability to create user accounts and other administrative tasks are hosted on this page.



Requesting an IPT account

Please contact the OBIS Canada Node manager Maria Cornthwaite (Maria.Cornthwaite@dfm-mpo.gc.ca) for an IPT account.

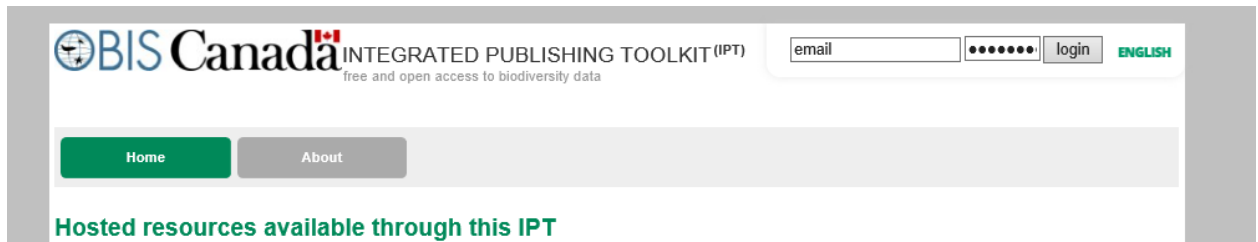
Creating and uploading a new dataset to Integrated Publishing Toolkit (IPT)

Creating new resources and uploading datasets to the Integrated Publishing Toolkit (IPT)

1. Sign onto IPT with your email & password
2. Select the Manage Resources tab
3. Scroll to the bottom of the page. In the section titled “Create New Resource” enter a “Shortname” for your dataset (a short descriptive working title free of whitespace which will become part of the URL on IPT), select the type as a “Sampling event” and select “Create” (no file uploads are required at this stage).
4. Once your resource page has been created, your overview page will be displayed. This page can be accessed by selecting your dataset from the list displayed under the “Manage Resources tab.”
 - i. **Source Data:** This section will allow you to manually upload your core and EMoF files to the IPT. Once uploaded, your file names are displayed here alongside an “Edit” button beside each file that will allow you to edit, delete or preview the mapping of a given source file. You can load all of your files at once if you zip them together. Large files (>104 mb) will also need to be zipped before uploading.
 - ii. **Darwin Core Mappings:** this dropdown menu contains the different core mapping options that you will apply to your source files. Mapping your files is a required step.
 - iii. **Metadata:** metadata will be entered via this section by clicking on the edit button. Once metadata has been input, its status will be displayed in this section. Click save before moving onto the next section.
 - iv. **Published Versions:** the publish button is located here if the data hasn’t been published yet, otherwise the status of the version published will be listed in this section.

- v. **Visibility:** datasets can be published publicly or privately. Data should be published privately initially and released to the public only after Q/C is completed. Dataset authors should be notified when data will be published on OBIS Canada.
- vi. **Resource Managers:** Information regarding the resource managers and creators is displayed here. Managers can be added or removed in this section.

1



The screenshot shows the OBIS Canada Integrated Publishing Toolkit (IPT) homepage. The header features the OBIS Canada logo with a Canadian flag, the text 'INTEGRATED PUBLISHING TOOLKIT (IPT)', and the tagline 'free and open access to biodiversity data'. To the right of the header is a login section with an 'email' input field, a password field represented by dots, a 'login' button, and a language selector set to 'ENGLISH'. Below the header is a navigation bar with two buttons: 'Home' (highlighted in green) and 'About'. Below the navigation bar is a section titled 'Hosted resources available through this IPT'.

2

Create New Resource

You can create a new blank resource, upload an existing resource saved as a zipped Darwin Core archive, or upload an existing IPT resource using its zipped resource configuration folder. Please refer to the User Manual for more specific instructions. A short name is required.

Shortname

Type

Optional archived resource to load:

3

i

Source Data

Browse...

Connect to database

Clear

Your source data files and SQL sources for generating a Darwin Core Archive.

Not modified since last publication

ep-mort-ev-2011-ongoing-rpt-pac-dfo-33.5 KB, 323 rows, 3 columns. Oct 15, 2020	Edit
mpo-aquaculture-emof [file]	
ep-mort-ev-2011-ongoing-rpt-pac-dfo-83.5 KB, 323 rows, 12 columns. Oct 15, 2020	Edit
mpo-aquaculture-event [file]	
ep-mort-ev-2011-ongoing-rpt-pac-dfo-97.6 KB, 323 rows, 14 columns. Oct 19, 2020	Edit
mpo-aquaculture-occurrence [file]	

ii

Darwin Core Mappings

Darwin Core Event

Add

Your mapping between the source data and Darwin Core terms.

Not modified since last publication

Core

Darwin Core Event	12 terms mapped to ep-mort-ev-2011-ongoing-rpt-pac-dfo-mpo-aquaculture-event. Oct 15, 2020	Edit
-------------------	--	------

Extensions

Extended Measurement Or Facts	2 terms mapped to ep-mort-ev-2011-ongoing-rpt-pac-dfo-mpo-aquaculture-emof. Oct 15, 2020	Edit
Darwin Core Occurrence	14 terms mapped to ep-mort-ev-2011-ongoing-rpt-pac-dfo-mpo-aquaculture-occurrence. Oct 19, 2020	Edit

iii

Metadata

Edit

Your resource metadata.

Not modified since last publication

iv

Published Versions

Publish

Auto-publishing

Select interval

A preview of your pending published version compared with the current version if existing.

	Current version	Pending version
Version	1.0 View	1.1 Preview
Published on	Oct 19, 2020	-
Publication log	Download	-
Publication report	Show	-

v

Visibility Private

Public

This resource is private to managers. To make this resource available to everyone, please make it public.

vi

Resource Managers

Add

Delete

Managers granted permission to modify this resource.

Creator	Melissa Nottingham, melissa.nottingham@dfo-mpo.gc.ca	
Manager	Melissa Nottingham, melissa.nottingham@dfo-mpo.gc.ca	Delete
Manager	Yeongha Jung, yeongha.jung@dfo-mpo.gc.ca	Delete
Manager	Henry Ye, henry.ye@dfo-mpo.gc.ca	Delete

Uploading your data to IPT

NOTE: SAVE YOUR WORK OFTEN AS IPT WILL LOG YOU OUT AFTER A RELATIVELY SHORT TIME

Once you've created a new resource page for your dataset, you can begin to upload your files.

Source Data: upload files onto IPT via the 'Choose File' button highlighted in the **red box** below on the dataset Overview page. Files can be uploaded individually or in a .zip compressed format (IPT will automatically unzip them). Files which have been uploaded will show up in the area highlighted by the **blue box**.

Source Data

Choose File No file chosen

Connect to database Clear

Your source data files and SQL sources for generating a Darwin Core Archive.

Last modified Feb 24, 2021

industrysealicecounts_2020ongoing_biologyemof2 [file]	711.2 KB, 6,305 rows, 7 columns. Feb 24, 2021	Edit
industrysealicecounts_2020ongoing_eventcore2 [file]	538 KB, 1,261 rows, 15 columns. Feb 24, 2021	Edit
industrysealicecounts_2020ongoing_eventemof2 [file]	145.1 KB, 1,261 rows, 7 columns. Feb 24, 2021	Edit
industrysealicecounts_2020ongoing_occurrencecore22 [file]	1 MB, 6,305 rows, 21 columns. Feb 24, 2021	Edit

Source Data Page: this page opens upon successful file upload. Confirm or edit the information pertaining to the data file and click 'Save' to verify the file. The information in the **blue box** below outlines some metadata for the data file, including the number of columns and whether it is readable or not.

*this page can be reached by clicking the 'Edit' button beside each dataset from the previous step

✓ Added new file source >>ns_event<<

Resource Title **Northumberland Strait multi-species trawl survey dataset**

Source Data

Edit your source data format

Source Name

ns_event

Readable ✓

Columns 11

File /data/obiscanada/resources/northumberland_traw
l/sources/ns_event.txt

Size 595.4 KB

Rows 3,019

Modified Feb 18, 2021 10:38:52 PM

Source log [Download](#)

Analyze

Number of Header Rows

1

Field Quotes

"

Character Encoding

UTF-8

Field Delimiter

.

Multi-value Delimiter

Date Format

YYYY-MM-DD

Save Cancel Delete source file

Darwin Core Mappings: once all of the data files have been uploaded to IPT, the source files can be linked to the appropriate Darwin Core terms by selecting a mapping option from the dropdown menu highlighted by the **red box** and pressing the 'Add' button. Map your Event Core as a "Darwin Core Event," your Occurrence Core as a "Darwin Core Occurrence" and your EMOF files as "Extended Measurement Or Facts."

Files which have been mapped will show up in the area highlighted by the **blue box**.

Darwin Core Mappings

Your mapping between the source data and Darwin Core terms.

Not modified since last publication

Core		
Darwin Core Event	11 terms mapped to event. Feb 11, 2021	
Extensions		
Darwin Core Occurrence	17 terms mapped to occurrence. Feb 11, 2021	
Extended Measurement Or Facts	7 terms mapped to emof_occurrence. Feb 11, 2021	

The following page will pop up after clicking the 'Add' button from the screenshot above, highlighted by the **red box**. Select the dropdown highlighted in the **black box** below to select the appropriate file to map and click 'Save'.

Darwin Core Event

The category of information pertaining to a sampling event. Replaces version issued 2015-05-29 adding record-level term dwc:institutionCode

Link: <http://rs.tdwg.org/dwc/terms/index.htm#Event>

Source data

Before you can start mapping concepts, please select a source data file.

Source data

emof_occurrence ▼

Save Cancel

After each file is mapped, IPT will open the following page, this is where column headers and data cells are automatically mapped and catalogued within IPT. Column headers can be remapped with the dropdown highlighted in the **red box**. This is a required step if the data file's headers are not the same as the DwC terms — headers that do not match DwC standards will not be automatically mapped by IPT.

*this page can be reached by clicking the 'Edit' button beside each mapped core files above.

samplingEffort

eventDate

Source Sample: 2000-07-25 | 2000-07-25 | 2000-07-25 | 2000-07-25 | 2000-07-26

Translation:

eventTime

Value Translation: the following page below will open if the 'Add' button (highlighted by the **blue box** above) is clicked. From here, the source value can be translated by unique values or in bulk if all values are the same for the entire table, the translated value can be entered into the form box below, highlighted in the **black box** below.

Value Translation

You can define a translation between the values from your source to the ones saved in the generated archive. The list of distinct values found in your source is generated the first time for you, but can be manually reloaded at any time while keeping the current translations.

Property *eventDate*

"The date-time or interval during which an Event occurred. For occurrences, this is the date-time when the event was recorded. Not suitable for a time in a geological context. Recommended best practice is to use an encoding scheme, such as ISO 8601:2004(E)."

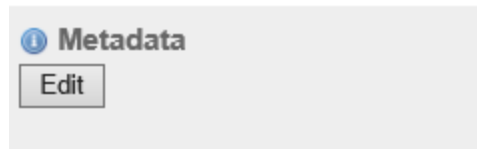
Source Value

2006-07-22

2014-07-20

Translated Value

Metadata: Metadata can be accessed by clicking on the 'Edit' button below the metadata section.



This will take you to the metadata pages where the dataset's metadata can be entered in the forms displayed on each page. The basic metadata section can now be completed. Note that most datasets should be assigned the same Data License as the example listed below, highlighted by the **red box**. All other sections of metadata can be completed by clicking the links on the right side of the page highlighted within the **blue box**.

Basic Metadata

Please enter all the mandatory properties on the Basic Metadata page, and then continue entering metadata in the other pages that are applicable to your resource. The more metadata you provide, the greater the chance that your resource will be found, reused by other researchers, and cited.

Title*

Publishing Organisation* <input type="text" value="Fisheries and Oceans Ca"/>	Type* <input type="text" value="Sampling event"/>	Metadata Language* <input type="text" value="English"/>
Update Frequency* <input type="text" value="Unknown"/>	Subtype <input type="text" value="No subtype"/>	Data Language* <input type="text" value="English"/>
Data Licence* <input type="text" value="Creative Commons Attribution (CC-BY) 4.0"/>		

Section

- Basic Metadata**
- [Geographic Coverage](#)
- [Taxonomic Coverage](#)
- [Temporal Coverage](#)
- [Keywords](#)
- [Associated Parties](#)
- [Project Data](#)
- [Sampling Methods](#)
- [Citations](#)
- [Collection Data](#)
- [External links](#)
- [Additional Metadata](#)

Published Versions: once the dataset is completely uploaded and mapped, the metadata is completed, and the dataset have been approved for release to the public; click on the 'Publish' button to publish the dataset.

Published Versions

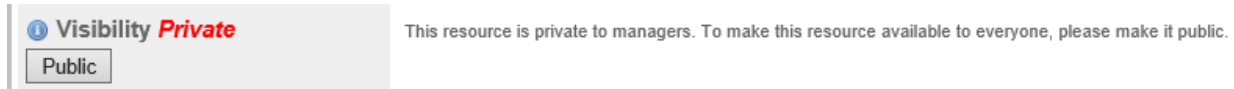
Auto-publishing

A preview of your pending published version compared with the current version if existing.

Version	Current version 1.0 <input type="button" value="View"/> ✓	Pending version 1.1 <input type="button" value="Preview"/>
Published on	Feb 11, 2021	-
Publication log	<input type="button" value="Download"/>	-
Publication report	Show	-

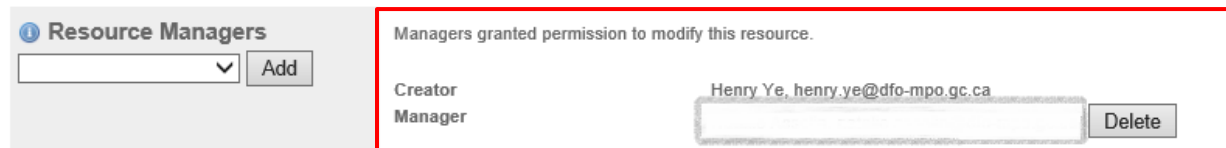
A preview/view of the published dataset page, along with the version and publication date can be found within the **blue box** above.

Visibility: the visibility button is displayed below. If the visibility is set to '**Private**', only the data manager and administrators can view the dataset on IPT. Select "Public" once you've completed your post-publication Q/C review, then take a moment to celebrate your contribution to global biodiversity mapping.



Resource Managers: by clicking on the drop down, one or multiple resource managers can be added to a dataset. The list of managers as well as its creator are indicated within the **red box** below.

*A dataset manager must have to have an existing IPT account at the manager level in order to be added as a resource manager.



Review your entry

Click on the 'View' button from step 6 in the previous section to see the dataset from the public's view.

Published Versions
[Publish](#)
Auto-publishing
Select interval

A preview of your pending published version compared with the current version if existing.

Current version
1.0 [View](#)
Feb 11, 2021
[Download](#)
[Show](#)

Pending version
1.1 [Preview](#)
-
-
-

A sample of a dataset's public view is illustrated below:

[Home](#) [DwC-A](#) [EML](#) [RTF](#) [Versions](#) [Rights](#) [Cite this](#)

Data Records

The data in this sampling event resource has been published as a Darwin Core Archive (DwC-A), which is a standardized format for sharing biodiversity data as a set of one or more data tables. The core data table contains 3,019 records. 2 extension data tables also exist. An extension record supplies extra information about a core record. The number of records in each extension data table is illustrated below.

Event (core)	3019
ExtendedMeasurementOrFact	24221
Occurrence	24221

This IPT archives the data and thus serves as the data repository. The data and resource metadata are available for download in the [downloads](#) section. The [versions](#) table lists other versions of the resource that have been made publicly available and allows tracking changes made to the resource over time.

Downloads

Download the latest version of this resource data as a Darwin Core Archive (DwC-A) or the resource metadata as EML or RTF:

[Data as a DwC-A file](#) [download](#) 3,019 records in English (541 KB) - Update frequency: unknown
[Metadata as an EML file](#) [download](#) in English (7 KB)
[Metadata as an RTF file](#) [download](#) in English (6 KB)

Versions

The table below shows all published versions of the resource, because you have access rights.

Version	Published on	Records	Change summary	DOI handle	Last modified by
1.0	2021-02-11	3,019	None provided Edit		Henry Ye

Showing 1 to 1 of 1 [previous](#) [next](#)

The number of records in each file (Event, Occurrence, EMoF) can be seen in the **blue box**

The different file formats and download links can be seen in the **red box**

The information for each version can be seen in the **black box**

The navigation bar on the left hand side (in the **yellow box**) will jump to the various sections on the page, most of this will be the content from the

Removing datasets on IPT

Removing datasets on IPT that have already been published and/or registered with GBIF:

- If only published to OBIS: email Pieter Provoost at p.provoost@unesco.org or helpdesk@obis.org and let them know you would like to remove the old dataset. They will let you know if the data has been cited or referenced elsewhere and whether you should archive to metadata-only or delete the page entirely.
- If archiving, delete the data files and mapping on IPT and change type to "metadata-only", link to the new IPT pages and re-publish.
- If the data is also registered to GBIF, email helpdesk@gbif.org to let them know you intend to delete/archive the data.
- Once you have had responses from OBIS and GBIF help desks, you can go ahead and delete or archive your data and let them know it has been done so the old data can be removed from the OBIS and/or GBIF databases.

IPT Tips and Tricks

Basis Of Record

When reviewing datasets some errors can be fixed without reloading data. For example HumanObservation is the required term for basisOfRecord, with no spaces in between. Changes can be made in the 'edit' section of the data file section:



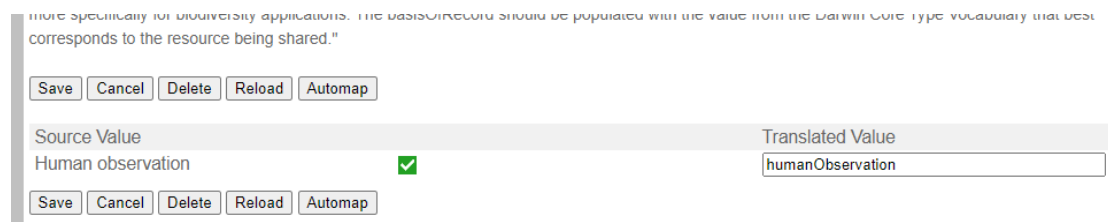
ownerInstitutionCode

basisOfRecord

Source Sample: humanObservation | humanObservation | humanObservation | humanObservation | humanOb:

Translation:

If the example above showed Human observation, it is possible to click 'add' then enter the correct value in 'translated value' and then click 'auto-map' to correct the error in all the data.



more specifically for biodiversity applications. The basisOfRecord should be populated with the value from the Darwin Core type vocabulary that best corresponds to the resource being shared."

Source Value	Translated Value
Human observation	<input type="text" value="humanObservation"/>

The icon

this icon will give you a preview of the document that is situated beside the icon. For instance, in the Darwin Core Mappings section on the dataset's home page, clicking on the icon will pop up a preview window of the data table. Any actions by mouse or keyboard will close the preview screen.

☒ Mapping preview finished successfully

Log Messages

- Start writing data file for Darwin Core Occurrence 1:10:28 AM
- Data file written for Darwin Core Occurrence with 100 records and 18 columns 1:10:28 AM

id	basisOfRecord	occurrenceID	occurrenceRemarks	organismQuantity	organismQuantityType	occurrenceStatus	eventID	scientificNameID
2000-7-25-10-39	HumanObservation	2000-7-25-10-39:1		1	individuals	present	2000-7-25-10-39	urn:lsid:marinespecies.org:taxname:158885
2000-7-25-10-39	HumanObservation	2000-7-25-10-39:2		2	individuals	present	2000-7-25-10-39	urn:lsid:marinespecies.org:taxname:159785


OBIS Canada Logo:

To add an OBIS Canada logo, navigate to the 'Additional Metadata' under the Metadata section of the dataset. Enter in the following link under the 'Resource logo URL' section in the forms box and click upload.


Additional Metadata

Please enter the additional metadata for the resource.

Date Created


 2021-01-21

Date Last Published

 2021-01-28



Resource logo URL



Alternatively if an icon image is available on the local machine, pressing 'Browse...' will pop up a window to select the logo file for the dataset to be uploaded.

Inorganic materials and unknown species

Inorganic materials such as rocks, garbage, unknown objects (non-living) should be excluded from datasets to published onto OBIS – they are not useful for most biological studies. Similarly, entries listed as “unknown species” or “unidentified” that cannot be assigned to a taxonomic kingdom should be omitted as well

Publishing Errors:

Updating or replacing occurrence or EMoF data files without altering the event core file can lead to publishing errors. In the event the error occurs, please follow the steps below to circumvent the error.

1. Delete all of the source files from IPT.
2. Remove the ghost columns from the EMoF (if applicable, ghost columns are empty columns with no headers and no data).
3. Re-zip and re-upload the event core, occurrence core, and EMoF files.
4. Remap the source files to each of the cores and EMoFs on IPT.
5. Publish!

END OF GUIDE