



Practical demo on how to prepare and map different types of data

Christian Svindseth, GBIF Norway

INDEX

- GBIF Integrated data Publishing Toolkit, IPT
- Install IPT
- Add and map data source



Steps to publish biodiversity data with IPT

- Step 1: Get access to the resources (IPT)
- Step 2: Create a new resource in the IPT
- Step 3: Import data files
- Step 4: Darwin Core mappings
- Step 5: Metadata authoring
- Step 6: Publishing
- Step 7: Configuring visibility
- Step 8: Managers

Demo hands-on IPT

- Workshop demo IPT:
- <http://data.gbif.no/ipt-sandbox/>
- Login: <removed from public slides>
- Password: <removed from public slides>
- GBIF Demo portal (after you publish and register):
- <http://www.gbif-uat.org/installation/...>



GBIF Integrated data Publishing Toolkit - IPT

The GBIF Integrated Publishing Toolkit: Vision

A single platform allowing the sharing of



Primary biodiversity data



Species name information



Dataset descriptions (metadata)

The ability to register with GBIF



Technical contact information
E.g. Internet URLs



Physical contact information
E.g. telephone details



Institutional affiliations
Accurate attribution

Connect



Databases



Upload text files

Lower the technical threshold for participation



Flexibility to accommodate data extensions



Support efficient and simple transfer of content



An open source project

The GBIF Integrated data Publishing Toolkit (IPT)

is a free open source software tool written in Java that is used to publish and share biodiversity datasets through the GBIF network.

<http://www.gbif.org/ipt>

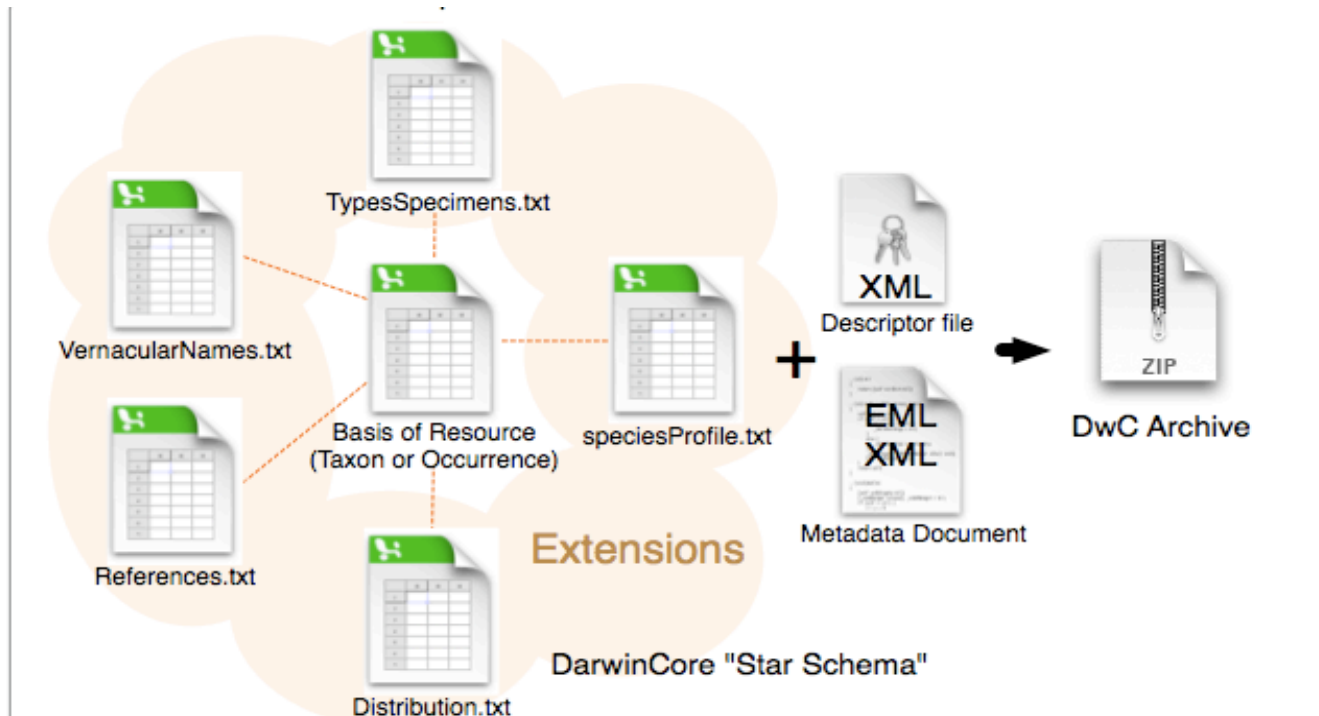
IPT User Manual:

<https://github.com/gbif/ipt/wiki/IPT2ManualNotes.wiki>

Robertson T, Döring M, Guralnick R, Bloom D, Wieczorek J, Braak K, Otegui J, Russell L, Desmet P (2014). The GBIF integrated publishing toolkit: Facilitating the efficient publishing of biodiversity data on the internet. PLoS One 9(8). [doi:10.1371/journal.pone.0102623](https://doi.org/10.1371/journal.pone.0102623)

The screenshot shows the GBIF Integrated Publishing Toolkit (IPT) user interface. At the top, the GBIF logo is on the left, and the text "GBIF INTEGRATED PUBLISHING TOOLKIT (IPT) free and open access to biodiversity data" is on the right. There is a search bar with "email" and "login" buttons, and a language selector set to "ENGLISH". Below the header, there are two tabs: "Home" (active) and "About". The main content area features a vertical list of menu items, each with a circular icon: "IPT settings" (with a gear icon and a "Publish all resources" button), "Users accounts" (with a person icon), "GBIF registration options" (with a leaf icon), "Organisations" (with a network icon and a note: "The IPT must be registered with GBIF before associating organisations"), "Core Types and Extensions" (with a puzzle piece icon), and "Logs" (with a magnifying glass icon). At the bottom, there is a footer with the text "The most recently updated resources are also available as an [RSS feed](#)." and a navigation bar with links for "IPT Version 2.3.2-rea67259", "About the IPT", "User manual", "Report a bug", and "Request new feature". The copyright notice "©2015 Global Biodiversity Information Facility." is at the very bottom.

Darwin Core Archive structure





IPT - add new user account

IPT add user account



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

TEST MODE

ENGLISH

Home

About

Login as Admin

Hosted resources available through this IPT

Filter:

Logo	Name	Organisation	Type	Subtype	Records	Last modified	Last publication	Next publication
------	------	--------------	------	---------	---------	---------------	------------------	------------------

No resources are currently available

Showing 0 to 0 of 0

◀ previous next ▶

No public resources exist.

IPT Version 2.3.2-rea67259

[About the IPT](#)

[User manual](#)

[Report a bug](#)

[Request new feature](#)

©2015 Global Biodiversity Information Facility.



Home

Manage Resources

Administration

About



IPT settings 

[Publish all resources](#)



Users accounts



GBIF registration options



Organisations

The IPT must be registered with GBIF before associating organisations



Core Types and Extensions



Logs

New User Account

Email addresses identify users and cannot be changed later on. Please select them carefully when creating a new account. If an email address needs to be changed you will have to create a new account and delete the old one.

Users having a manager role are entitled to create new resources, but cannot manage resources created by other users unless they have been invited to do so in the configuration for that resource. You can create any number of users having the Admin role, but at least one must remain in the system.

Email

First Name

Last Name

Role

Password

Verify Password

IPT add user account (4)



Home

Manage Resources

Administration

About

Users accounts

Filter:

Name	Email	Role	Last Login
Dag Endresen	dag.endresen@nhm.uio.no	Admin	never
GBIF Drift	gbif-drift@nhm.uio.no	Admin	7/4/16 5:50 AM

Showing 1 to 2 of 2

◀ previous next ▶

Create

Cancel

IPT Version 2.3.2-rea67259

[About the IPT](#)

[User manual](#)

[Report a bug](#)

[Request new feature](#)



IPT - add new data source

IPT Source Data – File Upload



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

TEST MODE

Logged in as larussell@vertnet.org [Account](#) [Logout](#) **ENGLISH**

Home

Manage Resources

Administration

About

Overview: *ttrs-ferns-specimens*

This is the overview page for the *ttrs-ferns-specimens* resource. Please start by uploading and mapping source data to generate and publish a Darwin Core archive. Don't forget to fill in the mandatory metadata. Published resources can be registered with the GBIF Network. Registered resources can be removed, but they cannot be reverted to a private state.

Source Data

Choose File No file chosen

Connect to database

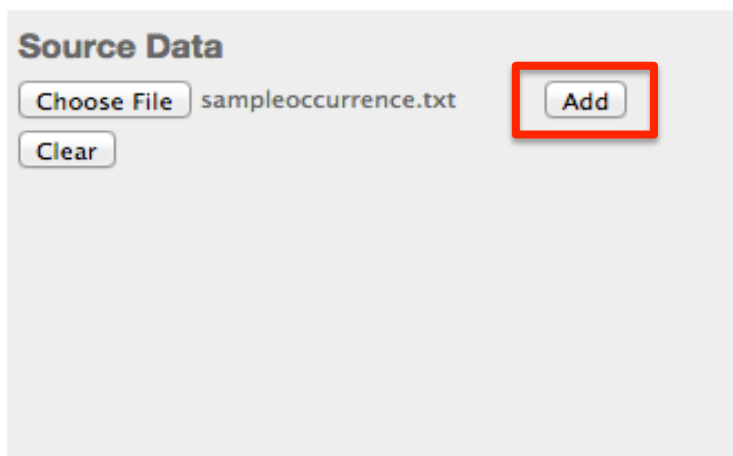
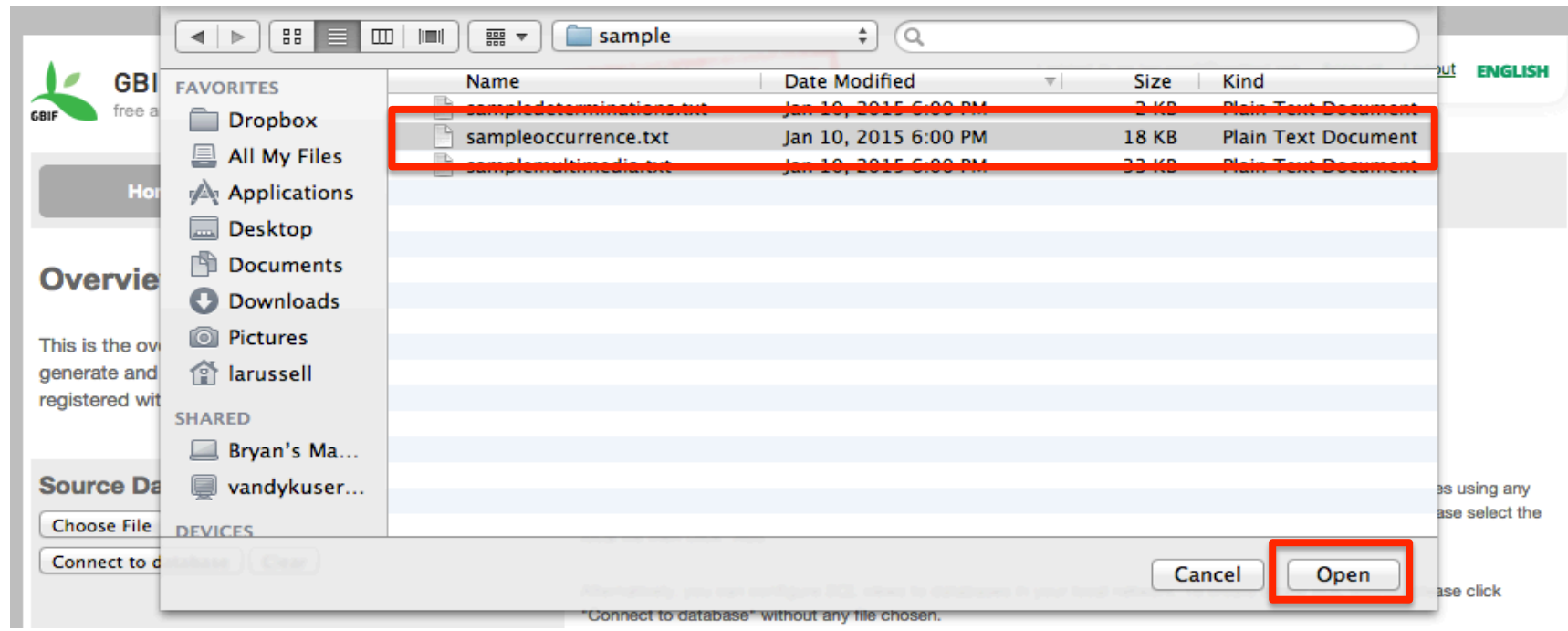
Clear

Your source data files for generating a Darwin Core Archive. You can upload delimited text files (csv, tab, and files using any other delimiter) either directly or compressed (zip or gzip). Excel files are also supported. To (re)upload a file, please select the local file then click "Add".

Alternatively, you can configure SQL views to databases in your local network. To create a new SQL source, please click "Connect to database" without any file chosen.

i It is not possible to update a resource by uploading a Darwin Core Archive. This feature is not implemented yet.

IPT Source Data – File Upload



1. Select an occurrence file and click the Open button.
2. Once the file is selected and the file name shows next to “Choose File,” click the Add button.

IPT Source Data – Upload Review

✓ Added new file source >>sampleoccurrence<<.

Resource Title ttrs-ferns-specimens

Source Data

Edit your source data format

Source Name

Readable
Columns 42

File </var/www/iptworkshop.idigbio.org/larussell/resources/ttrs-ferns-specimens/sources/sampleoccurrence.txt>

Size 17.9 KB

Rows 50

Modified 12/01/15

Source log [Download](#)

42 columns as read from source

Upload file location on server

Number of records as read from source

File specifications as read from source

Number of Header Rows

Field Quotes

Date Format

Field Delimiter

Character Encoding

IPT Source Data – Upload Preview

id	Kingdom	Phylum	Class	Order	Family	Genus	Common Name	Specific Epithet	Infraspecific Epithet	Scientific Name
urn:uuid:becf9816-8078-4896-a76e-39602fcf7b31	Plantae	Lycopodiophyta	Lycopodiopsida	Lycopodiales	Lycopodiaceae	Lycopodiella				Lycopodiella alopecuroides (L.) Cranfill
urn:uuid:783e9ef1-e6d1-4f37-af52-9209df61671b	Plantae	Lycopodiophyta	Lycopodiopsida	Lycopodiales	Lycopodiaceae	Lycopodiella				Lycopodiella alopecuroides (L.) Cranfill
urn:uuid:2137cde4-db0a-4aba-9293-8ad8a6e927b1	Plantae	Lycopodiophyta	Lycopodiopsida	Lycopodiales	Lycopodiaceae	Lycopodiella				Lycopodiella alopecuroides (L.) Cranfill
urn:uuid:0b401ab0-c981-489d-8448-e4b1eb8a9a10	Plantae	Lycopodiophyta	Lycopodiopsida	Lycopodiales	Lycopodiaceae	Lycopodiella				Lycopodiella alopecuroides (L.) Cranfill
urn:uuid:3c3cf351-9458-42b3-a467-7e0c80de3853	Plantae	Pteridophyta	Filicopsida	Ophioglossales	Ophioglossaceae	Botrychium				Botrychium lunarioides (Michx.) Sw.
urn:uuid:5fae0586-7fff-4de7-a118-a6a3c0395c11	Plantae	Pteridophyta	Filicopsida	Ophioglossales	Ophioglossaceae	Botrychium				Botrychium lunarioides (Michx.) Sw.

IPT Source Data – Upload Complete

Source Data

No file chosen

Your source data files for generating a Darwin Core Archive. You can upload delimited text files (csv, tab, and files using any other delimiter) either directly or compressed (zip or gzip). Excel files are also supported. To (re)upload a file, please select the local file then click "Add".

Alternatively, you can configure SQL views to databases in your local network. To create a new SQL source, please click "Connect to database" without any file chosen.

sampleoccurrence [file] 17.9 KB, 50 rows, 42 columns. 12/01/15

IPT Source Data – Database Connection



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

TEST MODE

Logged in as larussell@vertnet.org [Account](#) [Logout](#) **ENGLISH**

Home

Manage Resources

Administration

About

Overview: *ttrs-ferns-specimens*

This is the overview page for the *ttrs-ferns-specimens* resource. Please start by uploading and mapping source data to generate and publish a Darwin Core archive. Don't forget to fill in the mandatory metadata. Published resources can be registered with the GBIF Network. Registered resources can be removed, but they cannot be reverted to a private state.

Source Data

Choose File No file chosen

Connect to database Clear

Your source data files for generating a Darwin Core Archive. You can upload delimited text files (csv, tab, and files using any other delimiter) either directly or compressed (zip or gzip). Excel files are also supported. To (re)upload a file, please select the local file then click "Add".

Alternatively, you can configure SQL views to databases in your local network. To create a new SQL source, please click "Connect to database" without any file chosen.

i It is not possible to update a resource by uploading a Darwin Core Archive. This feature is not implemented yet.

IPT Source Data – Database Connection

Resource Title ttrs-ferns-specimens

Source Data

Edit your source data format

Source Name

Readable

Columns 0

Database System

Host

Database

Database user

Database password

SQL Statement

Character Encoding

Date Format



IPT mapping data

IPT Darwin Core Mappings

Source Data

No file chosen

Your source data files for generating a Darwin Core Archive. You can upload delimited text files (csv, tab, and files using any other delimiter) either directly or compressed (zip or gzip). Excel files are also supported. To (re)upload a file, please select the local file then click "Add".

Alternatively, you can configure *SQL views* to databases in your local network. To create a new SQL source, please click "Connect to database" without any file chosen.

sampleoccurrence [file] 17.9 KB, 50 rows, 42 columns. 12/01/15

Darwin Core Mappings

⌵

Your mapping between the source data and Darwin Core terms.

IPT Darwin Core Mappings – Selecting Source

Darwin Core Occurrence


The category of information pertaining to evidence of an occurrence in nature, in a collection, or in a dataset (specimen, observation, etc.). Updated 2 Apr 2014 with materialSampleID.

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

Source data

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

Source data

Save

Cancel

IPT Darwin Core Mappings – Automapping

Automapped 20 columns based on header names. If source names match DwC, automapping occurs

 basisOfRecord is required Alerts users to missing required field

Resource Title [ttrs-ferns-specimens](#)

Mapping Source Data: [sampleoccurrence](#)

For each property you can assign a dynamic value taken from the specified column of your source data file *sampleoccurrence* or provide a fixed, static value that is used for all records.

Shortcut to Data Source

[Hide Unmapped Fields](#)

Darwin Core Occurrence

The category of information pertaining to evidence of an occurrence in nature, in a collection, or in a dataset (specimen, observation, etc.). Updated 2 Apr 2014 with materialSampleID.

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

Please remember occurrenceID

occurrenceID

 No ID

The column in your source that holds a unique key for every record/row, also known as the "core id". If you want to map other sources, this identifier is required and will be used to link your source rows together.

Filter
Before Translation



The filter allows you to include only records that match a criterion set for one of the source data fields.

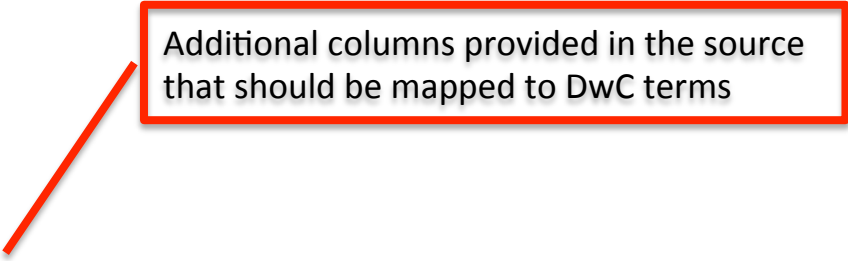
Save Delete Back

IPT Darwin Core Mappings – Unmapped Fields

Unmapped columns

The following columns of the source are not assigned to the properties of this mapping:







- id
- Common Name
- Rank
- Type of Preservation
- Coden
- Occurrence Remarks
- Catalog
- Determiner
- Collection Date
- Collection Date2
- Collector
- Nomenclatural Notes
- Preparation
- State
- Habitat Description
- Longitude
- Latitude
- Datum
- Dynamic Property
- Morphbank URL
- License



Additional columns provided in the source that should be mapped to DwC terms

IPT Darwin Core Mappings – Dropdowns

Record Level

type	  <input type="text" value="Collection"/>
modified	 <input type="text"/>
language	 <input type="text"/>
rights	 Rights <input type="text"/>
	Source Sample: Translation: <input type="button" value="Add"/>
rightsHolder	 <input type="text"/>

- Collection ✓
- Dataset
- Event
- Image
- Interactive Resource
- Moving Image
- Physical Object
- Service
- Software
- Sound
- Still Image
- Text

Some terms, like type, have recommended vocabularies, that you can use if the data is not present in your source. basisOfRecord is another term with a recommended vocabulary.

IPT Darwin Core Mappings – Preview data

kingdom	<input type="text" value="Kingdom"/>	<input type="text"/>
	<i>Source Sample:</i> Plantae Plantae Plantae Plantae Plantae	
	<i>Translation:</i> <input type="button" value="Add"/>	
phylum	<input type="text" value="Phylum"/>	<input type="text"/>
	<i>Source Sample:</i> Lycopodiophyta Lycopodiophyta Lycopodiophyta Lycopodiophyta Pteridophyta	
	<i>Translation:</i> <input type="button" value="Add"/>	<div style="border: 2px solid red; padding: 5px; display: inline-block;">Mapped fields with data in the first 5 records, show a preview of that data.</div>
class	<input type="text" value="Class"/>	<input type="text"/>
	<i>Source Sample:</i> Lycopodiopsida Lycopodiopsida Lycopodiopsida Lycopodiopsida Filicopsida	
	<i>Translation:</i> <input type="button" value="Add"/>	
order	<input type="text" value="Order"/>	<input type="text"/>
	<i>Source Sample:</i> Lycopodiales Lycopodiales Lycopodiales Lycopodiales Ophioglossales	
	<i>Translation:</i> <input type="button" value="Add"/>	
family	<input type="text" value="Family"/>	<input type="text"/>
	<i>Source Sample:</i> Lycopodiaceae Lycopodiaceae Lycopodiaceae Lycopodiaceae Ophioglossaceae	
	<i>Translation:</i> <input type="button" value="Add"/>	
genus	<input type="text" value="Genus"/>	<input type="text"/>
	<i>Source Sample:</i> Lycopodiella Lycopodiella Lycopodiella Lycopodiella Botrychium	
	<i>Translation:</i> <input type="button" value="Add"/>	

IPT Darwin Core Mappings – Fill-in Values

If no mapping is provided for a term, a value can be supplied.

institutionCode



TTRS

collectionCode



Source Sample: TROY | TROY | TROY | TROY | TROY


Translation:




IPT dataset publish and register

Publish dataset

Published Versions

Publish 

Auto-publishing

Select interval 

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

Version


Pending version

1.0 [Preview](#)

Visibility

Private

DOI

[Reserve](#) 

Data Licence

-



Published on

-

Published Versions

[Publish](#)


Auto-publishing

Select interval  

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

Version

Current version

1.4 [View](#) 

Published on

Jul 6, 2016

Publication log

[Download](#)

Publication report

[Show](#)

Pending version

1.5 [Preview](#)

-

-

-

Register dataset

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

Public

Visibility *Public* This resource is public and available to everyone.

Register ⚠️

Private

The button labeled "Register" will only be enabled if metadata complete

Visibility *Public* This resource is public and available to everyone.

Register

Private

Are you sure?

Confirm that you have read and understood the [GBIF data sharing agreement](#)

You are about to register the resource with GBIF. Do you still want to register? Be aware deletion from the GBIF Index is automatic, but cannot be undone without explicit email communication with the [GBIF Help Desk](#).

Yes No

Visibility *Registered* This public resource is registered with GBIF meaning that it is globally discoverable via the [GBIF website](#).

GBIF UUID	bcbf7a3f-9869-440a-829c-bfcd16451117
Organisation	Test Organisation #1
Organisation Contact	Kyle Braak, kbraak@gbif.org
Endorsing Node	DanBIF - Danish Biodiversity Information Facility



IPT hands-on try yourself

Demo hands-on IPT – User manual

- GBIF IPT Manual
- <https://github.com/gbif/ipt/wiki/IPT2ManualNotes.wiki>

Demo hands-on IPT

- Workshop demo IPT:
- <http://data.gbif.no/ipt-sandbox/>
- Login: <removed from public slides>
- Password: <removed from public slides>
- GBIF Demo portal (after you publish and register):
- <http://www.gbif-uat.org/installation/...>

IPT Support

- GBIF Helpdesk: helpdesk@gbif.org
- IPT GitHub site: <https://github.com/gbif/ipt/>
- Community site:
<http://community.gbif.org/pg/groups/3529/gbif-ipt-helpdesk-and-training-experts/>
- IPT mailing list: <http://lists.gbif.org/mailman/listinfo/ipt/>
- IPT user manual:
<https://github.com/gbif/ipt/wiki/IPT2ManualNotes.wiki>



Practical demo on how to prepare and map different types of data

Christian Svindseth, GBIF Norway