Accelerating biodiversity research through DNA barcodes, collection and observation data

GBIF Secretariat

Version 2, April 2023

Table of Contents

Course description

Aim

This course will teach you how to use DNA barcodes, collection and observation data to resolve research questions in biodiversity. The programme uses a combination of lectures, tutorials, and hands-on exercises. You will learn to handle biodiversity data including DNA barcoding. You will gain practical experience in using open and digitally documented biodiversity data through GBIF and BOLD to answer biodiversity research questions. You will understand and practice capturing observation, collection, and genetic data from analogue and digital sources. Finally, this course provides basic skills in data publishing through GBIF and BOLD.

Scope

Data management skills for accessing and publishing data through biodiversity data platforms. This is an observation/specimen → published record course that does not include wet lab steps.

This course is a collaboration between the Diku-funded University of Oslo project BioDATA, the BMBF-funded project Caucasus Barcode of Life (CaBOL), and GBIF - Global Biodiversity Information Facility as has been developed by Dag Endresen, Dmitry Schigel, Helena Wirta, Hugo de Boer, Laura Russell, and Stefaniya Kamenova.

Audience

The course is suitable for MSc and PhD students in biology and other professionals in relevant fields.

Prerequisites

Participants should have an affinity or professional interest in biodiversity. Participants need to have the motivation and interest to handle DNA barcodes, museum collection data, and observation data. A good understanding of English is necessary to follow the course, carry out the exercises, and receive support during the teaching.

Intended learning outcomes

- Understand and be able to explain the concept of species delimitation.
- Learn to use genetic sequence data as a DNA barcode to identify a species.
- Learn to publish and retrieve data from GBIF and BOLD.
- Learn the basics of data capture, cleaning, storage, geo-referencing, and citation.
- Critically assess the quality of own and external data and their fitness for purpose.
- Practice the key tools and approaches to maximize data quality, data linking and data reuse.
- Explore the benefits of FAIR and open data principles in biodiversity research and collaboration.
- Understand the value of data management as a research-enabling tool.
- Broadly understand the importance of international biodiversity infrastructures, and how these can contribute to biodiversity assessments, monitoring, conservation, and red- listing.

Course preparation

Required

Registrations and accounts

- Create an account at ORCID.
- Create a personal (user) account at GBIF.org, top right corner. You can login to GBIF using ORCID.
- Create a personal (user) account at BOLD. You can login to BOLD using ORCID.
- Create an account at iNaturalist.

Foundations videos

To prepare for the course lectures, please watch the Foundations video playlist. These materials are terms-heavy, but will ensure that all participants have the same basis of knowledge to begin and will prepare you for the course lectures; please capture questions as you go through and ask them during the course.

Foundations videos - 6 videos, 67 mins.

Recommended

The course will include a practical part where you will design a research project and its data elements. This project and its "data journey" will be on a plant-pollinator system. To prepare for this part, please read the following. Skip if you are working on pollinators yourself already.

- Honey detective work raises fears for bees
- Loss of bees causes shortage of key food crops, study finds

Optional

The following are optional activities if you have completed all of the above. Many aspects of these will be introduced during the course days.

- What is GBIF? - video, 8 minutes
- What is BOLD? - video, 9.5 minutes
- Introduction to GBIF - online course
- More than 75 percent decline over 27 years in total flying insect biomass in protected areas - paper

Files for download

All files for the course may be downloaded from this page. Or if you prefer, all files are linked individually throughout the course as they occur in the curriculum. The video files are embedded throughout the course, as well, and play from YouTube. Subtitles are available when playing from YouTube for most videos. If you have difficulty accessing the embedded videos, please download the mp4 files to play them locally on your computer.

Videos

The videos are narrated in English. Subtitles are not available for the downloaded videos.

Foundations1.zip (73.7 MB)

Foundations2.zip (90.2 MB)

Capture.zip (63.1 MB)

Management.zip (30.2 MB)

Publishing.zip (77.9 MB)

Exercise data

Data journey step 4: sequences.zip file (ZIP 5 KB)

Data journey step 6:

1. Excel template
2. Vicia.zip file (ZIP 51 MB)

Data journey step 7:

1. ViciaForCleaning.txt file (ZIP 66 KB)
2. UC1-3c-open-refine.csv. (207.5 KB)

Software installation

Install OpenRefine

Install software required for activities later in the course

OpenRefine is a tool with a set of features for working with tabular data that improves the overall quality of a dataset. It is an application that runs on your own computer as a small web server, and in order to use it your web browser should point at that web server. So, think of OpenRefine as a personal and private web application.

We will use OpenRefine during the data mobilization portion of the course, especially during the practical exercises. It will be necessary to install OpenRefine on your laptop. If you are a skilled computer user, you can follow these steps to install the software on your computer. If you are not confident, please ask for help. Refer to the OpenRefine download page for more details.

Administrative passwords may be required to install software.

Installation Requirements

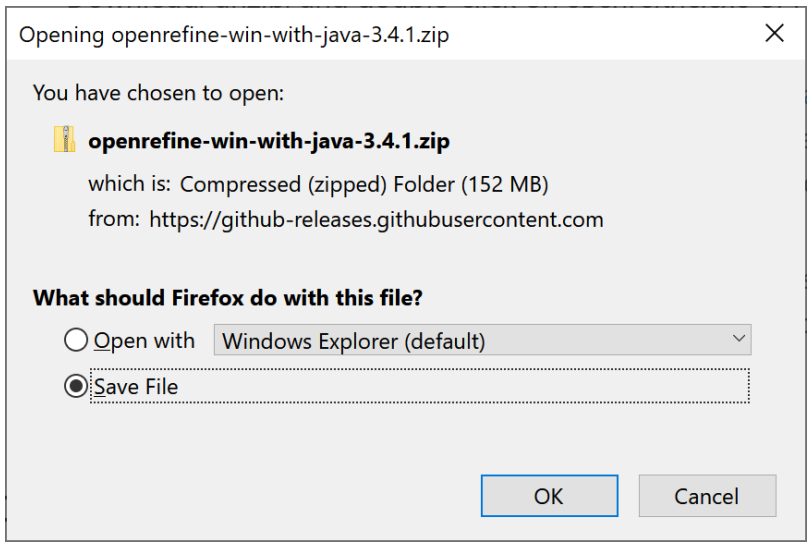
1. Linux users only: Java JRE installed.
2. Google Chrome, Microsoft Edge or Mozilla Firefox installed. Internet Explorer is not supported.

The latest stable release is OpenRefine 3.4.1, released on September 24, 2020. Detailed installation instructions are available at https://docs.openrefine.org/manual/installing/.

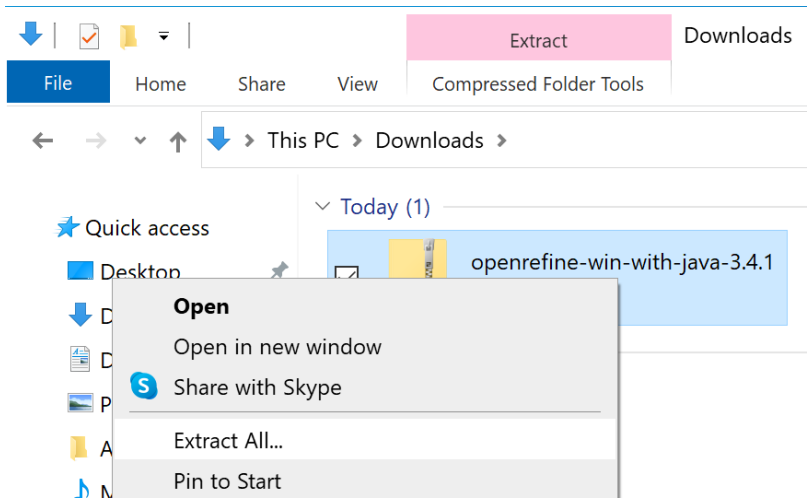
Installation on MS Windows

1. Download the Windows kit **with embedded Java**. Choose to save the file rather than open it.
2. Find the downloaded file. Right click it, and choose "Extract all…". Unzip, and double-click on openrefine.exe or refine.bat if the former does not work.
3. A command window will appear (don't close it) and soon after a new web browser window will show the application.

▼ *Detailed instructions for MS Windows (click to expand)*

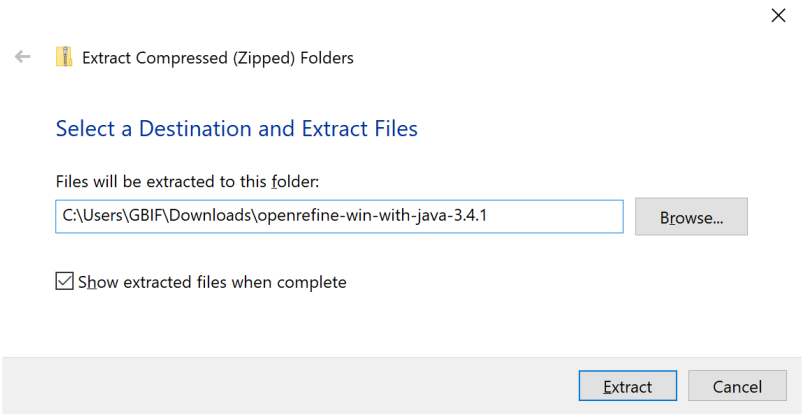
1. 

   Download the Windows kit **with embedded Java**. Choose to save the file rather than open it.

2. 

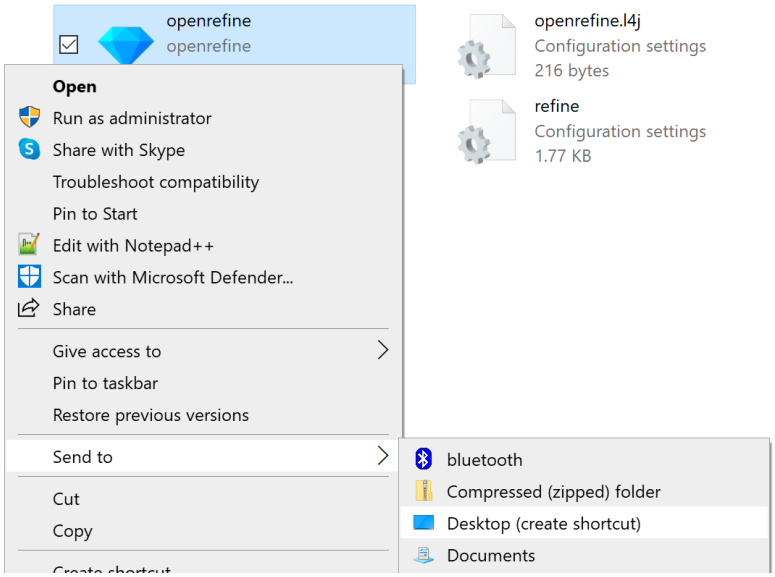
Find the file you downloaded. Right click it, and choose "Extract All..."

3.



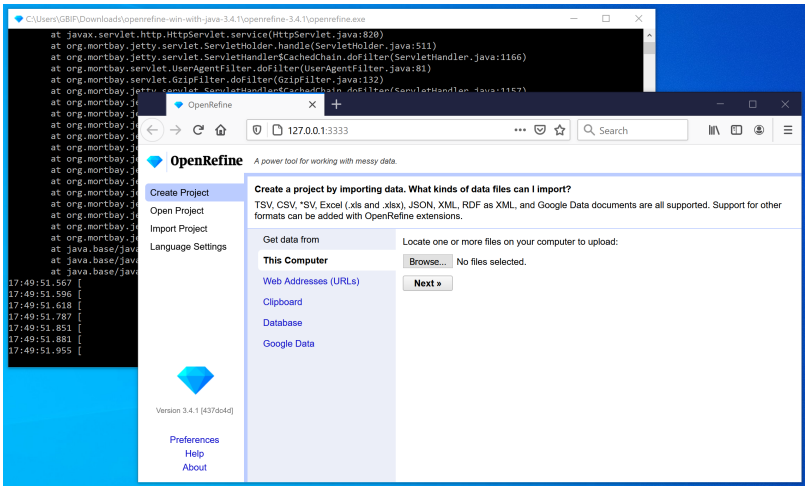
Click "Extract"

4.



Find the extracted files. Optionally, right click "openrefine" and choose "Send to → Desktop (create shortcut)" to create a shortcut on your desktop. Then double click "openrefine"

5.



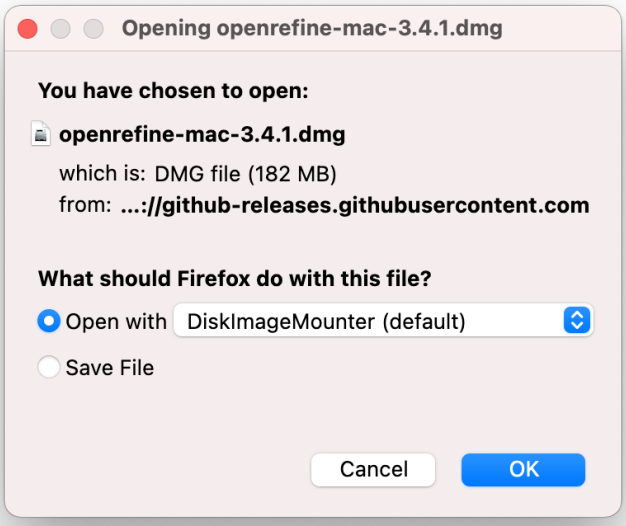
A black console window opens, and a short time later the browser opens. OpenRefine is now ready to use.

Installation on Mac

1. Download the Mac kit.
2. Download, open, drag icon into the Applications folder. You do not need to install Java separately.
3. Double click on it and a new web browser window will show the application.

▼ *Detailed instructions for Mac (click to expand)*

1.



Download the Mac kit, and choose to open it.

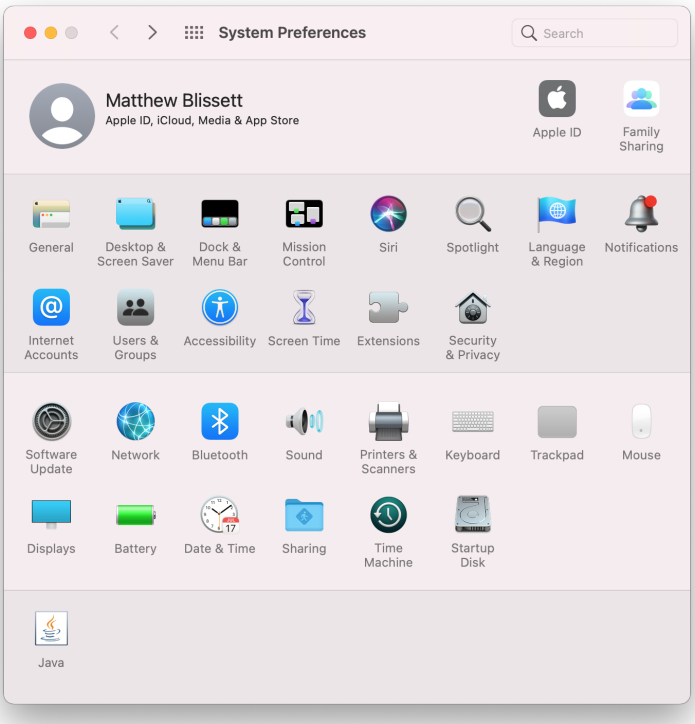
2.

A warning is shown. Click "OK".

3. 

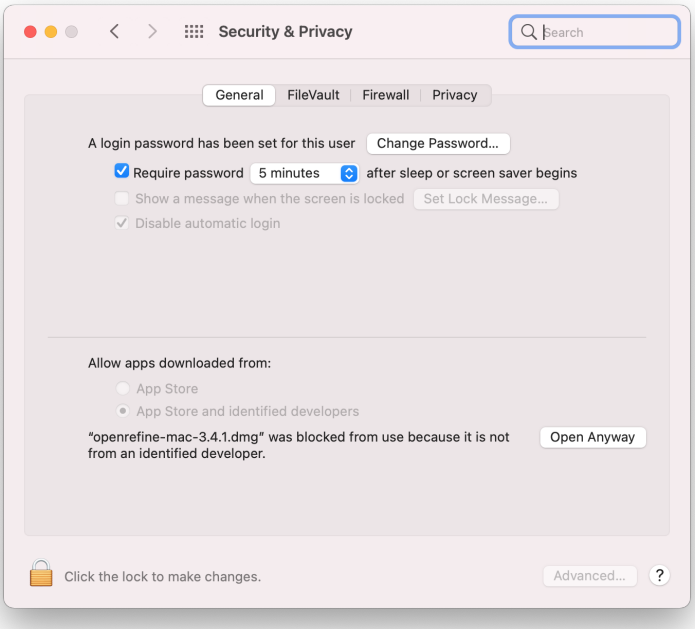
System Preferences

Open System Preferences.

4. 

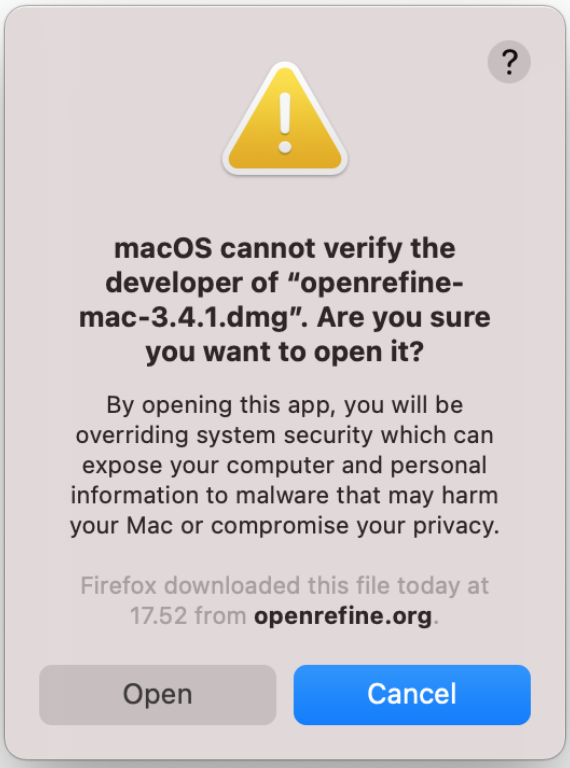
Open Security & Privacy

5.



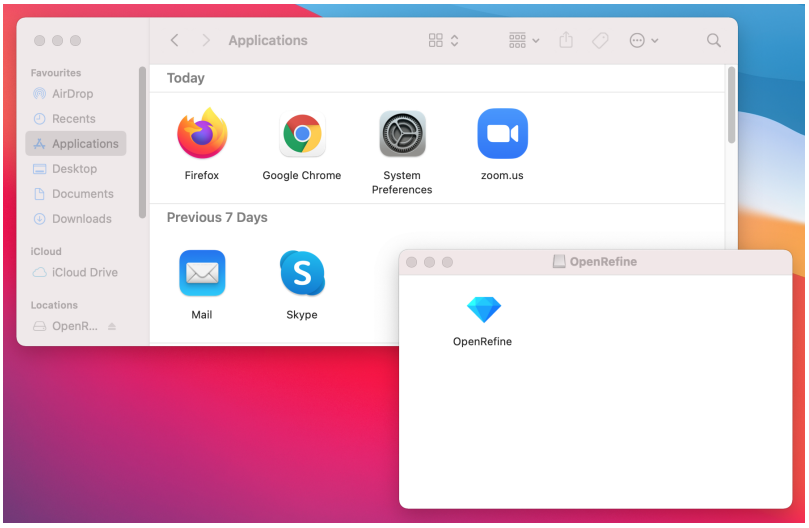
Choose "Open Anyway" at the bottom.

6.



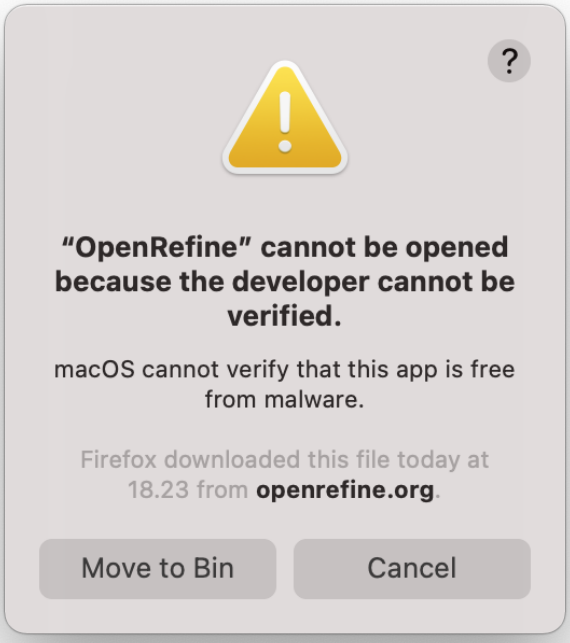
Choose "Open"

7.



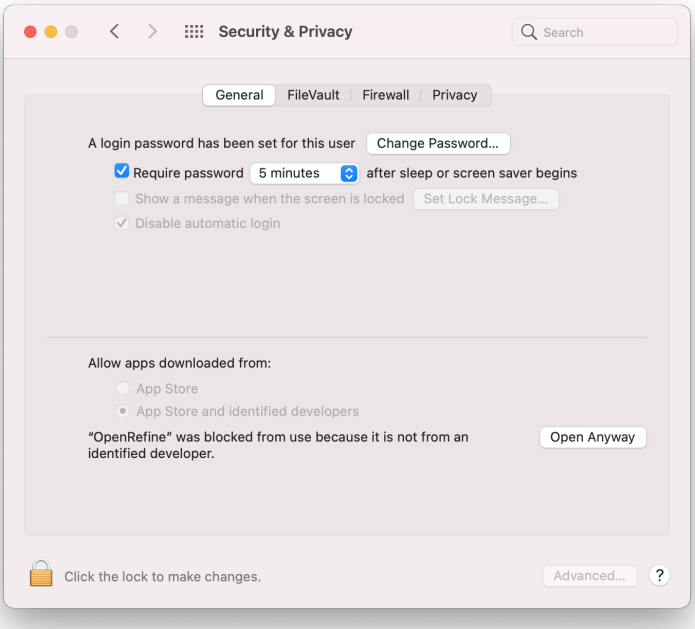
Finally, the application archive is opened! Drag it to your Applications folder.

8.



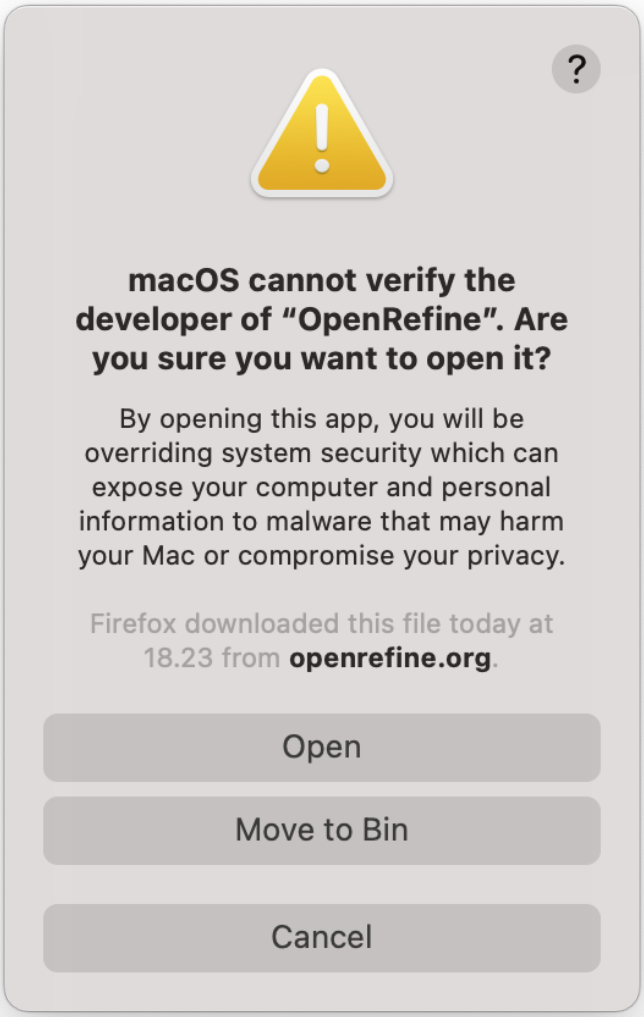
Double-click the OpenRefine icon. Another security warning appears!

9.



Go back to "Security & Privacy" and click "Open Anyway" — again.

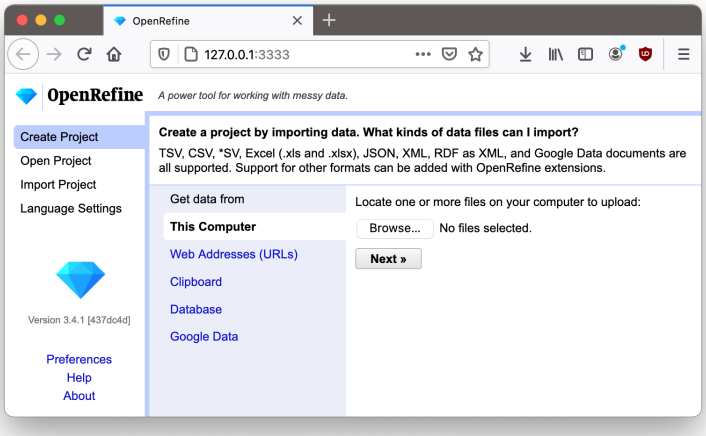
10.



(To avoid these warnings, the OpenRefine developers would need to pay Apple.)

Click "Open".

11.



Finally! The application is running.

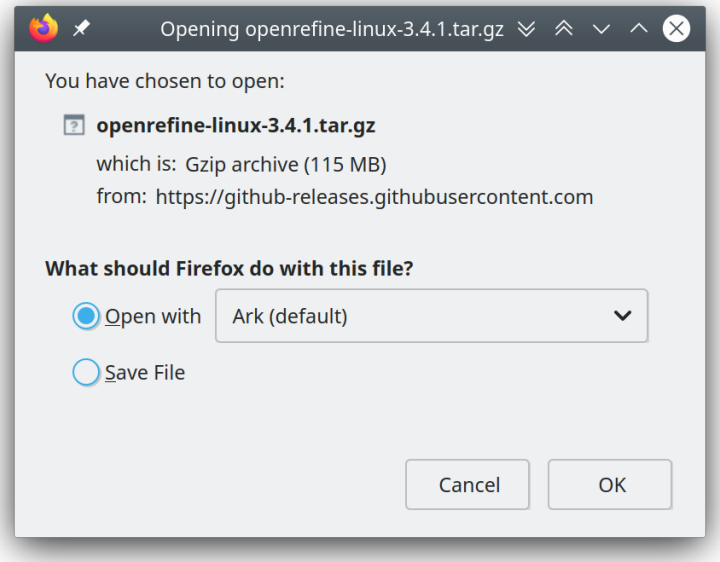
Installation on Linux

1. Download the Linux kit.
2. Download, extract, then type ./refine to start. This requires Java to be installed on your computer.

▼ *Detailed instructions for Linux (click to expand)*

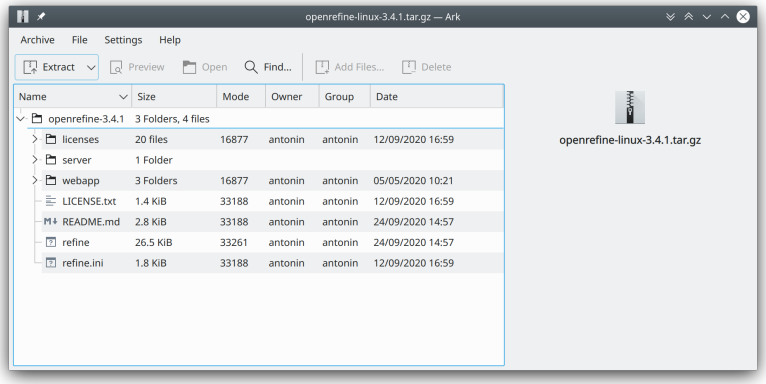
These instructions are for KDE (e.g. Kubuntu, SuSE), but the process is similar for Gnome (e.g. Ubuntu, Red Hat, CentOS).

1.



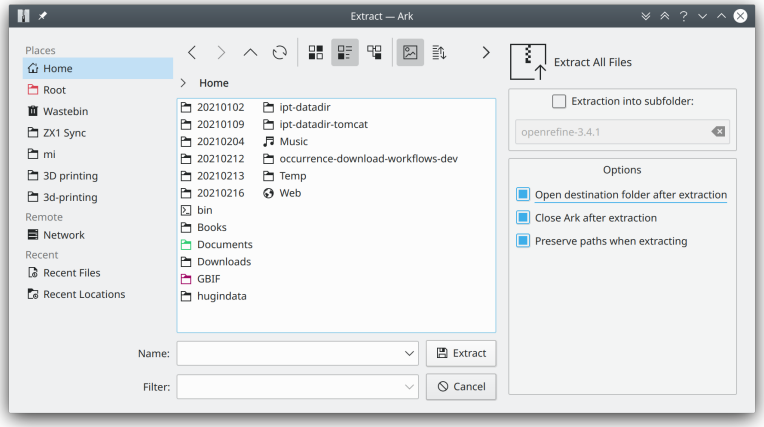
Download the Linux kit. Open it.

2.



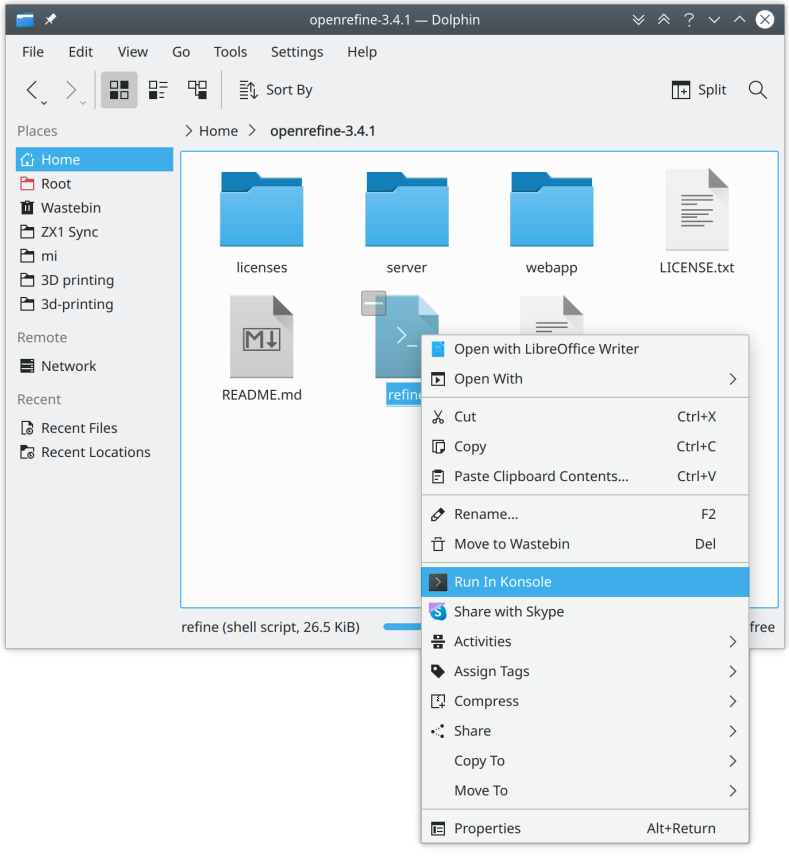
Click "Extract" to unpack the downloaded application.

3.

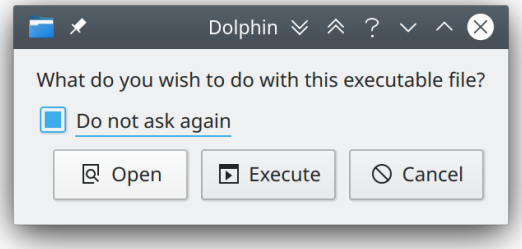


Choose a suitable place. I also selected "Open destination folder after extraction" and "Close Ark after extraction"

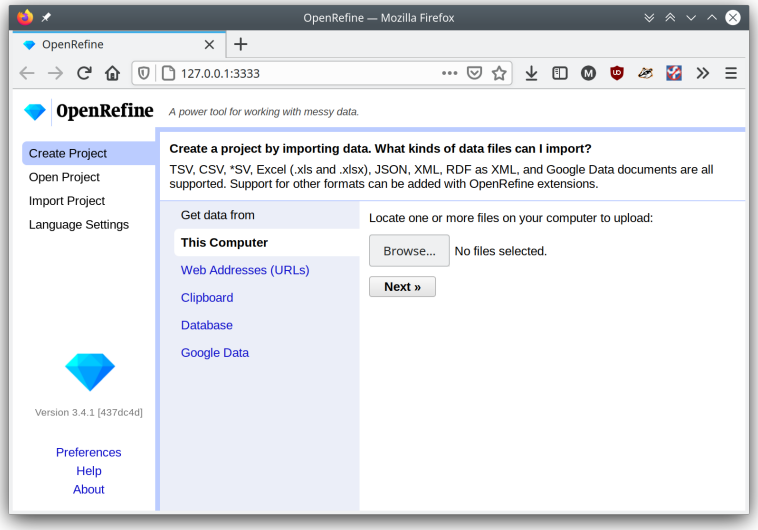
4.



Right click "refine" and choose "Run in Konsole". This is needed so you can safely exit OpenRefine later, by closing the Konsole window.

5. 

Confirm that you wish to execute the downloaded application.

6. 

OpenRefine is now running.

1. Introduction

1.1. Biodiversity of Bulgaria

The presentation can be viewed in the online version of the course.

1.2. Barcoding, data and biodiversity

The presentation can be viewed in the online version of the course.

1.3. GBIF for science and policy

The presentation can be viewed in the online version of the course.

2. Data as a first-class research citizen

This module includes information about trends to treat research data as a first-class scientific citizen, to be independently cited, and contribute to researcher career

assessment.

2.1. Open research data

In this presentation, you will review recent trends to treat research data as a first-class scientific citizen, to be independently cited, and contribute to researcher career assessment used in this course.

The presentation can be viewed in the online version of the course.

3. Data journey - bumblebee pollinators

This data journey is a series of practical exercises focusing on bumblebee pollinators. Using GBIF and BOLD, you will learn: I. how to find and use already available biodiversity data in connexion with your research questions; II. How to efficiently capture and clean this data – i.e. put them in a standard format that is directly relevant and exploitable for you; III. How to generate and publish new data according to a international standards.

This data journey is comprised of nine steps. Each step (or set of steps), correlate to the different modules of the course. After a series of theoretical lectures, you will return to the Data Journey to complete the practical exercises. The practical exercises follow a path: I. the study system; II. questions and hypotheses; III. availability of data; IV. capture and cleaning of data; and V. generate and publish data.

The presentation can be viewed in the online version of the course.

4. Species identification

4.1. Why do we need to identify species

The presentation can be viewed in the online version of the course.

4.1.1. Data journey step 2

Complete step 2, tasks 4-6

4.1.2. Data journey step 3

Complete step 3, tasks 7-9

4.2. Species identification and delimitation with DNA

The presentation can be viewed in the online version of the course.

5. Barcode reference depositories

5.1. Barcode reference depositories

The presentation can be viewed in the online version of the course.

5.2. Demonstration

5.3. Data journey step 4

Complete step 4, task 10.

5.4. Data journey step 5

Complete step 5, task 11.

6. Scholarly regcognition

The presentation can be viewed in the online version of the course.

7. UNITE

The presentation can be viewed in the online version of the course.

8. Metabarcoding and eDNA applications

The presentation can be viewed in the online version of the course.

9. BOLD Systems

The presentation can be viewed in the online version of the course.

10. iNaturalist

In this activity, you will venture into the field for first-person data capture.

The presentation can be viewed in the online version of the course.

11. Data capture

In this module, you will learn the types of primary biodiversity data and how to best share that information within GBIF. You will also review principles of data quality in the context of data capture and will learn about data quality and coherence (especially on subjects such as georeferencing, dates, names and taxa cross-checking).

11.1. Data origins and types

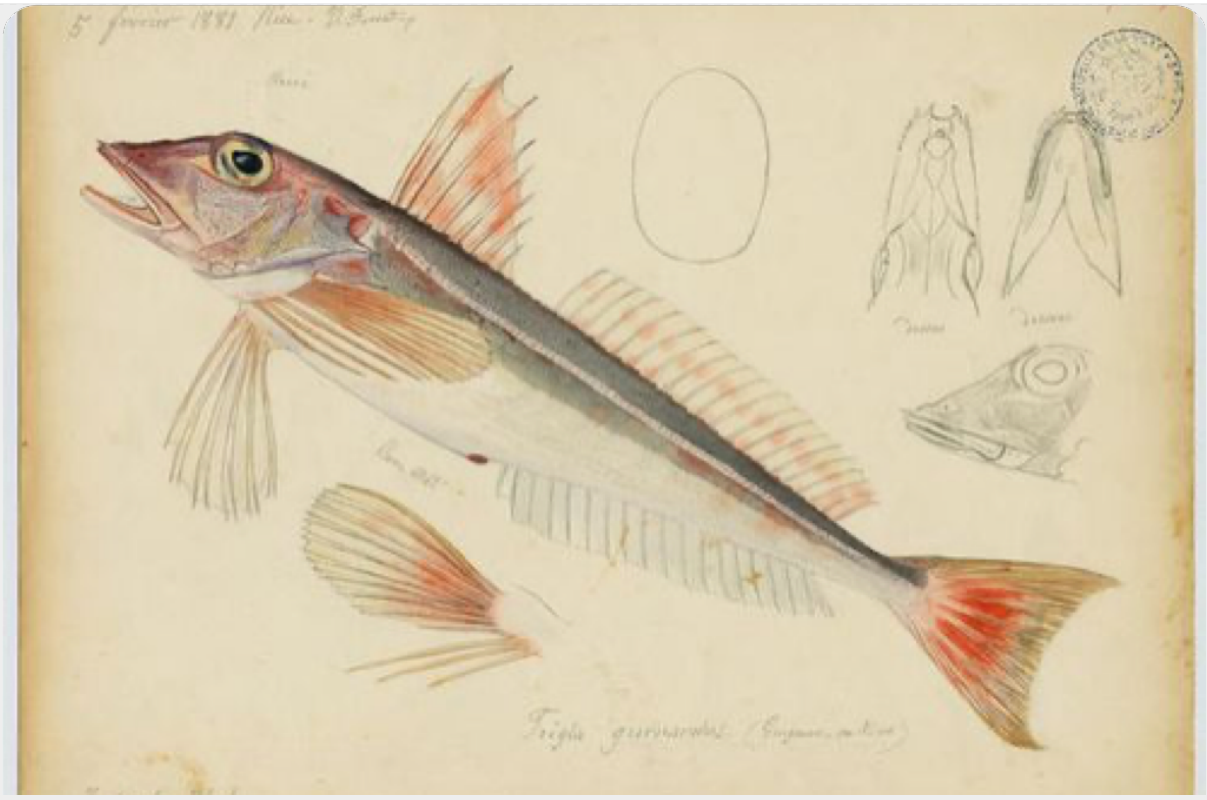
In this video (10:45), you will review primary biodiversity data that can be shared within GBIF. If you are unable to watch the embedded video, you can download it locally. (MP4 - 19 MB)

▶ https://www.youtube.com/watch?v=wKeOveydjsw *(YouTube video)*

11.1.1. Data types review

Quiz yourself on the concepts learned in this section.

1. What dataset type(s) would you choose for an ichthyology collection?



Eutrigla gurnardus (Linnaeus, 1758) | *Muséum d'histoire naturelle de Nice*

- ☐ occurrence
- ☐ checklist
- ☐ sampling event

2. What dataset type(s) would you choose for a list of invasive species?

Water hyacinth (Eichhornia crassipes) observed in Bourail, New Caledonia, where it is an introduced and invasive species by GRIIS. Photo by gérard (2016) licensed under CC BY-SA 2.0

- ☐ occurrence
- ☐ checklist
- ☐ sampling event

3. What dataset type(s) would you choose for the flora and fauna of an environmental impact study?

   Environmental impact assessment studies are done by experts in order to assess the biodiversity and biotopes of a given area, before, during and after it is affected by human activities (road works, wind turbines, mining, building construction, etc.).

Entomologist chasing butterflies *by Matthieu Gauvain (CC-BY-SA)*

- ☐ occurrence
- ☐ checklist
- ☐ sampling event

4. What dataset type(s) would you choose for bird tracking data?

   Bird-tracking data are recorded using specific devices, such as GPS trackers mounted on live birds, thus allowing scientists to track their migratory routes or breeding sites.

Griffin vulture *observed at Gamla Nature Reserve by גיזונית - MinoZig (CC0)*

- ☐ occurrence
- ☐ checklist
- ☐ sampling event

5. What dataset type(s) would you choose for insect trap data?

Insect trap *by miheco (CC-BY-SA)*

- ☐ occurrence

- ☐ checklist
- ☐ sampling event

6. What dataset type(s) would you choose for national park management data?

   Data acquired in the context of protected areas management (such as national parks but also smaller nature reserves) can be diverse and have different origins: botanical surveys, tagged animals tracking, observations from rangers and guards, and even 'citizen science' data or data inferred from pictures shared on social medias.

Sri Lankan elephants *observed by pen_ash.*

- ☐ occurrence
- ☐ checklist
- ☐ sampling event

7. What dataset type(s) would you choose for a citizen science bioblitz?

   Citizen science data are often collected through thematic fieldwork days known as a "bioblitz." Volunteers typically gather in a given area and spend the day trying to observe and identify as many species as they can in this area.

   Data from each participant are captured and merged into the citizen science programme's data capture or data management tool.

Looking for birds with park staff *by US National Park Service (authorized reuse on google image search)*

- ☐ occurrence
- ☐ checklist
- ☐ sampling event

8. What dataset type(s) would you choose for a regional species list?

Black rhino *observed at the Magdeburg Zoo in Germany by Mani300*

- ☐ occurrence
- ☐ checklist
- ☐ sampling event

11.1.2. Data types review solutions

▼ *Click to expand*

What dataset type(s) would you choose for an ichthyology collection?

- occurrence
  Most of the time, specimens from collection databases are shared as occurrence data. Each occurrence (specimen or group of specimens) has its own unique identifier (sometimes derived from its catalogue number in the source collection) and the Darwin Core fields used to share them within GBIF describe each specimen: scientific name, the date it was collected on the field, who collected and/or identified it, where, etc. Each collection can have more than one specimen from a same species, as long as each specimen is identified by a unique ID.
- checklist
  It is also possible to create and share a taxonomical checklist derived from a collection database; in this case, it is recommended to share the checklist as a taxonomical dataset, with the occurrence (specimen) list associated with it by using the Occurrence core as an extension to the Taxon Core on the GBIF IPT.

What dataset type(s) would you choose for a list of invasive species?

- occurrence
  Some data publishers will share occurrence datasets coming from studies or programs tracking specimens from some specific invasive species; when the data focuses on individuals instead of the invasive species, in general, they can be shared as occurrence data.
- checklist
  Invasive species can be tracked and monitored at different scales (regional, national, thematic…); as this type of dataset focuses more on the species and their distribution across a given geographical scope, they are mainly shared as taxonomical datasets within GBIF (see GRIIS search results).

What dataset type(s) would you choose for the flora and fauna of an environmental impact study?

- occurrence
  Data are recorded by naturalists on the field and can be shared as simple occurrence datasets.
- sampling event
  They can also be shared as event datasets if standardized protocols (such as vegetation plots, transects, traps…) are used to collect the data.

What dataset type(s) would you choose for bird tracking data?

- occurrence
  These data are shared as occurrence datasets: ideally, each bird is identified with its organismID, and each occurrence (GPS ping) has its own occurrenceID, which is useful to track the different GPS locations of the same bird over the scope of the tracking programme or project. (See example)

What dataset type(s) would you choose for insect trap data?

- occurrence
  Although such data can be shared as simple occurrence datasets, it is best if they're shared as event datasets, where the location, identifier and contents of each trap can be better detailed.
- sampling event
  Insect traps (as well as other traps such as pitfall traps, malaise traps…) are typically used in monitoring programmes to check the presence (or absence) of some species and/or assess their specific abundance. Using the "eventID" field to identify each trap allows the users to get all of the specimens collected within each trap. The same logic applies to other field protocols such as transects, plots, remote cameras, etc.: by using the Event Core instead of the Occurrence core, you'll be able to share much more information about the context of the data collection, and allow users to better understand (and even replicate) your work.

What dataset type(s) would you choose for national park management data?

- occurrence
  record individuals of species
- checklist
  It is important to know how many species are present in the park/reserve perimeter and their conservation status.
- sampling event
  check and track the populations

What dataset type(s) would you choose for a citizen science bioblitz?

- occurrence
  Bioblitz datasets are mainly shared as occurrence datasets.
- sampling event
  Depending on the citizen science programme, specific sampling protocols might be used by the volunteers, in which case, the data can be shared as an event dataset.

What dataset type(s) would you choose for a regional species list?

- checklist
  Geographical or thematic species lists are often used to share information about the species present in a given area; most of the time, these lists also mention the distribution of each species as well as their conservation status in this area. Regional species lists can give a useful insight into a region's biodiversity and habitats, and need to be shared as taxonomical datasets, with or without associated occurrences.

11.2. Data capture, processing and quality

In this video (09:11), you will explore the principles of data quality applied to data capture, specifically when capturing data from collection labels, fieldwork notebooks, spreadsheets, etc. If you are unable to watch the embedded video, you can download it locally. (MP4 - 19 MB)

▶ https://www.youtube.com/watch?v=QkDJIkmwBMA *(YouTube video)*

11.3. Data journey step 6

Complete step 6, task 12.

12. Data management

In this module, you will review the main concepts, related tools and best practices for data management, particularly, data cleaning and standardization.

12.1. Principles of data management

In this video (09:49), you will review an important set of principles necessary to improve data through the processes of data cleaning. If you are unable to watch the embedded video, you can download it locally. (MP4 - 16.6 MB)

▶ https://www.youtube.com/watch?v=4ijm1cJeVHE *(YouTube video)*

12.2. Data management tools

In this video (06:42), you will learn about a variety of tools that you can use to improve the quality of your data. If you are unable to watch the embedded video, you can download it locally. (MP4 - 10.3 MB)

▶ https://www.youtube.com/watch?v=Ru3vWiYU3gw *(YouTube video)*

12.3. OpenRefine

In this video (03:27), you will learn about OpenRefine. You can use OpenRefine to standardize and improve the quality of your data. If you are unable to watch the embedded video, you can download it locally. (MP4 - 3.8 MB)

▶ https://www.youtube.com/watch?v=_YFw_bfwc3Y *(YouTube video)*

12.4. Data journey step 7

Complete step 7, tasks 13-15.

12.5. Exercise tips

12.5.1. Validation checks

Technical errors Relatively simple, often able to be automated, **checks against the integrity of the data**. These may indicate incorrect exports, data mapping, field slippage (e.g. moving 1 column to the right) or data missing at the source.

- **Completeness**: Whether all the data and metadata is available – are all fields present, are all fields filled out?
- **Bounds**: For example, are days given in the range 1-31 (depending on month)
- **Data type**: For example, does the Date field contain a date or a number?
- **Data format**: For example, are Dates provided as 01/01/2010 or 01/Jan/10?

Consistency errors

Application of real-world rules to the data. These may indicate incorrect data entry from older records, transcription errors or post processing. Some are complex to implement and **require reference data sets to check against**. E.g. a list of known collectors and collecting habits. These rules can be gathered from data users and analysts.

- **Taxonomic**: For example, if identified to species level, have a binomial scientific name and entries in genus and species fields been provided?
- **Currency**: Are dates of collection, identification, update and digitization consistent?
- **Outliers**: Detect outliers, but remember that not all outliers are necessarily errors. For example, compare against a known species range, or known environmental range (but remember that outliers may be misidentifications, rather than incorrect coordinates).
- **Geographic**: Are the coordinates within the identified locality or region? For example, are there any terrestrial occurrences in the sea or marine occurrences on land?
- **Collecting patterns**: Does the occurrence detail match the known collecting patterns of the organization or collector? Do any records appear to have been created after a collector has died (could this possibly be a different collector with a similar name)? For example, are any mammal records attributed to a bird watching group?
- **Accuracy and precision**: For example, are any georeferenced records indicating very high precision or accuracy from a pre-GPS (or pre-accurate GPS) collecting period?
- **Collecting methods**: Different survey methods (e.g. transects and area surveys) have particular characteristics. Are the records consistent with the method provided?

12.5.2. Helpful tools

- **GBIF Name Parser**: https://www.gbif.org/tools/name-parser
- **Global Names Resolver**: http://resolver.globalnames.org
- **Catalogue of Life name match**: https://data.catalogueoflife.org/tools/name-match
- **TNRS**: https://tnrs.biendata.org/
- **WoRMS**: https://www.marinespecies.org/aphia.php?p=match
- **InfoXY**: http://splink.cria.org.br/infoxy?criaLANG=en
- **Georeferencing Calculator**: http://georeferencing.org/georefcalculator/gc.html
- **Canadensys coordinate conversion**: http://data.canadensys.net/tools/coordinates
- **Canadensys date parsing**: http://data.canadensys.net/tools/dates
- **Google Maps**: https://maps.google.com/

13. Data publishing

In this module, you will learn about data publishing concepts, including the IPT, cores and extensions, and the importance of licenses, metadata, mandatory fields and hosting of datasets.

13.1. Data publishing concepts

In this video (11:45), you will learn about data publishing concepts and will receive an introduction to the Integrated Publishing Toolkit (IPT). If you are unable to watch the embedded video, you can download it locally. (MP4 - 20 MB)

▶ https://www.youtube.com/watch?v=b900d9ukjSQ *(YouTube video)*

13.2. IPT overview

In this video (06:56), you will receive an overview of the IPT data publishing interface. If you are unable to watch the embedded video, you can download it locally. (MP4 - 8.7 MB)

▶ https://www.youtube.com/watch?v=gHXsaN_JWeI *(YouTube video)*

13.2.1. Training IPT installations

If you have interest in trying the IPT, please contact training@gbif.org and you will be provided with a login and password to one of the training IPTs.

https://training-ipt-a.gbif.org/

https://training-ipt-b.gbif.org/

https://training-ipt-c.gbif.org/

13.3. IPT demonstration

In this video (24:16), you will learn how to publish an occurrence dataset using an IPT. If you are unable to watch the embedded video, you can download it locally. (MP4 - 52.6 MB)

▶ https://www.youtube.com/watch?v=eDH9IoTrMVE *(YouTube video)*

13.4. Data journey step 8

Complete step 8, task 16 using https://cloud.gbif.org/eca

14. Conclusion

Course evaluation

Complete the course evaluation

The evaluation can be completed online.

Glossary

ALA

Atlas of Living Australia. The Australian node of GBIF, who developed an open source data portal now widely used within the GBIF community & partners for biodiversity national portals.

API

Application Programming Interface. A set of clearly defined methods of communication between various software components.

BID

Biodiversity Information for Development. An EU funded project co-ordinated by GBIF whose aim is to increase data mobilization capacity in the Africa, Caribbean and Pacific regions.

BIFA

Biodiversity Fund for Asia.

CC Licences

Creative Commons. These are a series of licenses set up by the Creative Commons organization that enable sharing and reuse of creativity and knowledge through the provision of free legal tools. Three of them can be assigned to GBIF-shared datasets: CC0, CC BY and CC BY-NC.

Controlled Vocabulary

This is a restricted set of terms that are used as possible values for a given field. One can think of it as a lookup list or dropdown for a particular field. For example the DwC field basisOfRecord should only contain one of these values: "PreservedSpecimen", "FossilSpecimen", "LivingSpecimen", "HumanObservation", "MachineObservation". We would say that list of values is a controlled vocabulary for that field.

DwC

Darwin Core is a biodiversity data standard, maintained by TDWG & widely used within the GBIF community and partners. It is a set of standardized terms (or field names) and their definitions, which are used to share biodiversity information.

DOI

Digital Object Identifier. A persistent identifier or handle used to uniquely identify objects. DOIs are in wide use mainly to identify academic, professional, and government information, such as journal articles, research reports and data sets, and official publications.

DwC-A

Darwin Core Archive. A compressed (zipped) file containing all the information needed to share with GBIF, for a particular resource. Each zip contains three types of files:

1. the actual data, in one or more text files: occurrence.txt/event.txt/measurmentoffact.txt etc
2. a mapping file: rtf.xml
3. a metadata (EML) file: eml.xml When you publish using the IPT, it creates a Darwin Core Archive, which is shared with GBIF. Also, when you download data from the GBIF website you can choose a DwC-A format as well.

GUID

Globally Unique IDentifier

IPT

Integrated Publishing Toolkit. It is a free and open source web application (software) for publishing biodiversity data. The software itself lives on a server (either at your institution or elsewhere) that must have access to the internet 24/7. It is used to create and handle Darwin Core Archive files that can be shared and used by anyone including GBIF.

Loan

In the context of natural history collections, this is the procedure of lending specimens between institutions.

LSID

Life Sciences Identifier. They are persistent, globally unique identifiers for biological objects.

Data Publishing

With regards to GBIF we have a very specific definition of data publishing. It refers to making biodiversity datasets publicly accessible and discoverable, in a standardized form, via an access point, typically a web address (a URL).

Resource

A Resource is the collective term used to refer to a particular dataset and its metadata once it has been uploaded to an IPT instance.

TDWG

Taxonomic Databases Working Group, now renamed Biodiversity Information Standards.

URN

Uniform Resource Number

UUID

Universally Unique Identifier

Colophon

Suggested citation

Accelerating biodiversity research through DNA barcodes, collection and observation data. 1st edition. GBIF Secretariat: Copenhagen. https://doi.org/10.35035/DZE9-MP34. [Date of course]

Contributors

Accelerating biodiversity research through DNA barcodes, collection and observation data was originally developed as a collaboration between the Diku-funded University of Oslo project BioDATA, the BMBF-funded project Caucasus Barcode of Life (CaBOL), and GBIF - Global Biodiversity Information Facility. The original curriculum was created Dag Endresen, Dmitry Schigel, Helena Wirta, Hugo de Boer, Laura Russell, and Stefaniya Kamenova.

Acknowledgements

Icons made by Freepik from flaticon.com.

Licence

Persistent URI

https://doi.org/10.35035/DZE9-MP34

Document control

First edition, June 2021