# Guidelines for Mobilizing National Botanical Checklists and Red Lists to GBIF

Authors: Lycka Kamoen[1,2,3]; Gentle Langezaal[2]; Jeroen Creuwels[1,4]; Niels Raes[1,5]

1. Naturalis Biodiversity Center
2. Hortus botanicus Leiden
3. https://orcid.org/0000-0003-3682-7037
4. https://orcid.org/0000-0001-6131-7026
5. https://orcid.org/0000-0002-4329-4892

Date: 12 October 2026

Version: 1.0

DOI: 10.5281/zenodo.23160658




Cover Illustration: *Parnassia palustris* from Prof. Dr. Otto Wilhelm Thomé *Flora von Deutschland, Österreich und der Schweiz* 1885.

Part of MetaPlantCode project deliverable D4.2: Mobilization of regional, national and international botanical checklists, red lists, floras of Europe.

This research was funded by Biodiversa+, the European Biodiversity Partnership, in the context of the MetaPlantCode project under the 2022-2023 BiodivMon joint call. It was co-funded by the European Commission (GA No. 101052342).

# Contents


1. Introduction 3
2. Scope 4
    2.1 Checklists 4
    2.2 Floras 4
    2.3 Red lists 4
3. Checklist Metadata requirements 4
4. Checklist Data Requirements 5
5. Publishing through the GBIF IPT 6
    5.1 Obtaining login credentials 6
    5.2 Accessing the IPT 6
    5.3 Adding Source Data 7
6. Darwin Core Mappings 7
    6.1 Core Mapping: Darwin Core Taxon 7
7. Checklist Extensions 9
    7.1 Species Distribution Extension 10
    7.2 Vernacular Names Extension 10
8. Become a GBIF data publisher 11
9. Metadata Requirements 11
    9.1 Basic Metadata 12
    9.2 Contacts 13
    9.3 Geographic Coverage 13
    9.4 Temporal Coverage 14
    9.5 Project Data 14
    9.6 Sampling Methods 15
    9.7 Citations 16
10. Downloading and Publishing 16
    10.1 Transferring the dataset from Tryout-IPT to the real IPT and GBIF 17
11. References 21

# 1. Introduction

Effective biodiversity monitoring is essential for understanding changes in ecosystems, assessing the impacts of environmental pressures, and supporting evidence-based conservation and management decisions. In recent years, DNA metabarcoding has emerged as a powerful tool for biodiversity assessment, enabling the rapid identification of multiple species from environmental samples such as soil, water, pollen, or bulk organismal material. This approach greatly increases the efficiency and scalability of biodiversity monitoring compared with traditional survey methods.

The MetaPlantCode project aims to develop harmonized best practices for collecting, analysing, and publishing plant metabarcoding data, with a strong emphasis on ensuring that these data become Findable, Accessible, Interoperable, and Reusable (FAIR). By promoting standardized workflows and data-sharing practices, the project seeks to improve the quality, reproducibility, and long-term usability of plant metabarcoding datasets.

A key challenge in interpreting plant metabarcoding results is that many plant species cannot be reliably distinguished based solely on their DNA barcode sequences. Closely related taxa often share identical or near-identical barcode regions, making it difficult to assign environmental DNA sequences to a single species with confidence. To resolve these ambiguities, the integration of additional contextual information is required, particularly the known geographic distributions of potential candidate species. Such information is commonly available through national species checklists, floras, and red lists, which provide authoritative knowledge on species occurrence and status within a given region.

However, these resources are often scattered across institutions, published in different formats, and not readily accessible for automated analyses. To overcome these limitations, species checklist and red list data can be published through the Global Biodiversity Information Facility (GBIF), an international network and data infrastructure that provides open access to biodiversity data. By using GBIF standards and publication tools, checklist data become more discoverable, machine-readable, interoperable, and reusable. Making these resources FAIR not only facilitates their integration into metabarcoding workflows but also increases their value for biodiversity research, conservation planning, ecological modelling, and environmental policy.

This manual provides practical guidance for preparing, structuring, and publishing botanical checklists and red lists through the GBIF Integrated Publishing Toolkit (IPT). The guidelines are based on GBIF's established best practices and offer step-by-step instructions for data preparation, Darwin Core mapping, metadata creation, extension use, publication, and quality control. Although developed within the context of plant metabarcoding, the workflow described here is broadly applicable to biodiversity data managers, herbarium staff, conservation organizations, and other stakeholders involved in the publication and management of species checklist data.

# 2. Scope

This manual encompasses botanical reference data in the form of checklists, red lists, and floras. Although these resources differ in their purpose and content, they are all published through GBIF as checklist datasets, using a common data structure and publication workflow.

## 2.1 Checklists

Simple lists of species names with no explicit indication of taxonomic status but generally implied to consist of accepted names of taxa. Such lists are generally intended to identify a set of taxa included within some regional or thematic context. In some cases, checklists also include a complete higher-level taxonomy, implemented through parent–child relationships that define the hierarchical placement of each taxon.

## 2.2 Floras

These are typically books that provide detailed species accounts for a particular region. Details may include many of the data types included in annotated checklists but include specific data types, such as detailed descriptions and illustrations, specimen references and other details that explicitly define (circumscribe) the taxon within the scope of the region which is not necessarily global. Here, flora data will follow the same workflow as checklists.

## 2.3 Red lists

Red lists are authoritative assessments of the conservation status of species within a defined geographic or thematic scope. They evaluate extinction risk using standardized criteria–preferably aligned with the IUCN Red List Categories and Criteria–and provide essential information for conservation planning, policy development, and biodiversity monitoring

# 3. Checklist Metadata requirements

All checklists published through the GBIF IPT must include complete and accurate metadata following the GBIF Metadata Profile which is aligned with the Ecological Metadata Language (EML) metadata standard: https://ipt.gbif.org/manual/en/ipt/latest/gbif-metadata-profile.

Metadata must clearly describe:

- Geographic scope
- Taxonomic scope
- Purpose of the checklist
- Data sources and compilation methods
- Contributors, institutions, and licensing
- Versioning and update procedures

High-quality metadata ensures transparency, reproducibility and FAIR (findable, accessible, interoperable and reusable) data.

## 4. Checklist Data Requirements

Note that all files should use UTF-8 character encoding, as many scientific names and author citations include diacritics or special characters that may become corrupted when non standard encodings are used. When working with numerical fields, ensure that decimal separators follow international conventions, using periods rather than commas. For CSV files, be aware that some regional settings use a semicolon (;) instead of a comma (,) as the list separator. This is a common source of errors and should be corrected before uploading the file to the IPT.

Before preparing a checklist or red list for publication, ensure that the dataset is complete, internally consistent, and formatted according to international data standards. At minimum, the source file (Excel or CSV) must contain the required fields listed below. If these are absent, manually add a column to the dataset containing this value for each row, for example the year of the list.

- Scientific name (including author name, as recommended by GBIF best practices)
- Taxon ID (a unique identifier for each taxon; manually add to dataset if absent)
- Country code corresponding to the origin of the list (ISO 3166-1 alpha-2; manually add to dataset if absent)
- Date or year of list publication (at least year; manually add to dataset if absent)

For red lists, the dataset must also include:

- Threat status, preferably using the standardized IUCN Red List Categories and Criteria: https://www.iucnredlist.org/resources/categories-and-criteria. In addition to global IUCN categories, many countries apply national or regional conservation statuses, such as Regionally Extinct (RE).

Table 1. IUCN categories

| Code | Category Name |
|---|---|
| EX | Extinct |
| EW | Extinct in the Wild |
| RE | Regionally Extinct (national category) |
| CR | Critically Endangered |
| EN | Endangered |
| VU | Vulnerable |
| NT | Near Threatened |
| LC | Least Concern |
| DD | Data Deficient |
| NE | Not Evaluated |

Additional fields–such as taxonomic rank, higher classification, remarks, or bibliographic references–should be included whenever available.

# 5. Publishing through the GBIF IPT

*For a complete IPT user manual by GBIF see: [https://ipt.gbif.org/manual/en/ipt/latest/](https://ipt.gbif.org/manual/en/ipt/latest/)*

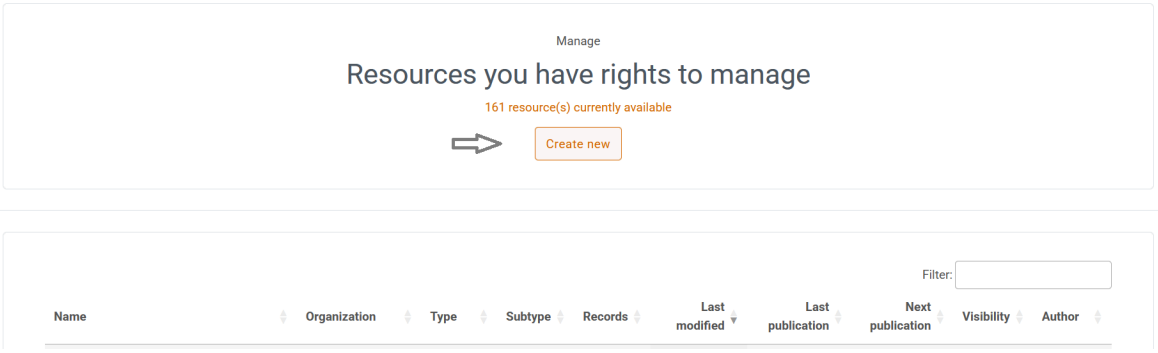
*Screen dumps are provided for visualization using a vascular plant red list of the Netherlands as an example.*

## 5.1 Obtaining login credentials

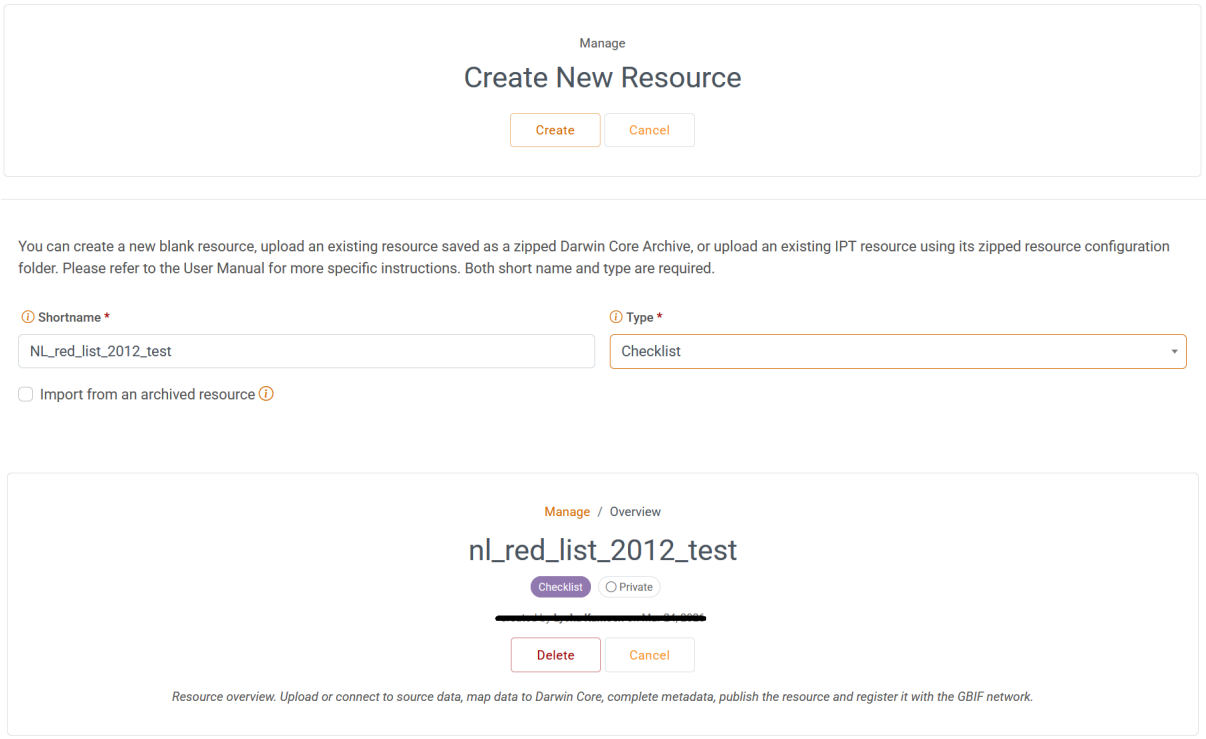
To gain access to the IPT, users must request an account and obtain login credentials. Account creation requires affiliation with an existing and endorsed GBIF data-publishing organization, ensuring that all published datasets are managed within a trusted institutional framework.

## 5.2 Accessing the IPT

1. Log in to the IPT instance.
2. Navigate to **Manage Resources**.
3. Select **Create New**.



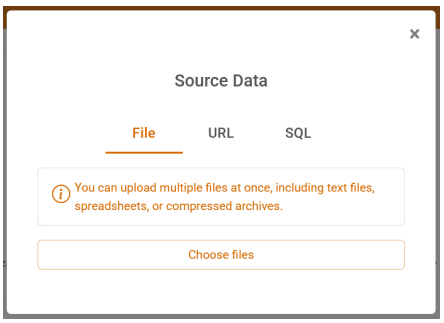

4. Provide a meaningful **Shortname**, including the two-letter country code for clarity.
5. Set **Type** to *Checklist*.
6. Click **Create**.




## 5.3 Adding Source Data

1. Under **Source Data**, click **+ Add**.
2. Upload the checklist or red list file (Excel or CSV).
3. If the dataset contains multiple sheets or supplementary files, upload them as additional sources.

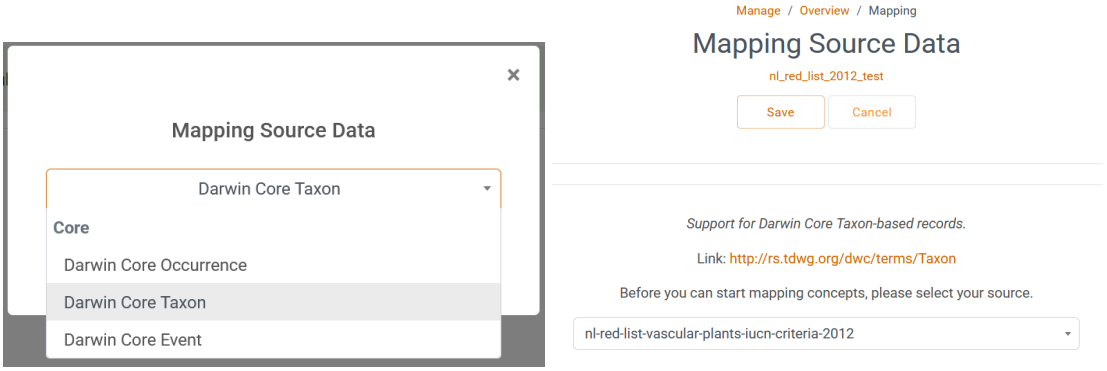



# 6. Darwin Core Mappings

## 6.1 Core Mapping: Darwin Core Taxon

1. Go to **Darwin Core Mappings** and click **+ Add**.
2. Select **Darwin Core Taxon** as the core and click **Add**.
3. Choose the appropriate source file.
4. Map all relevant fields.



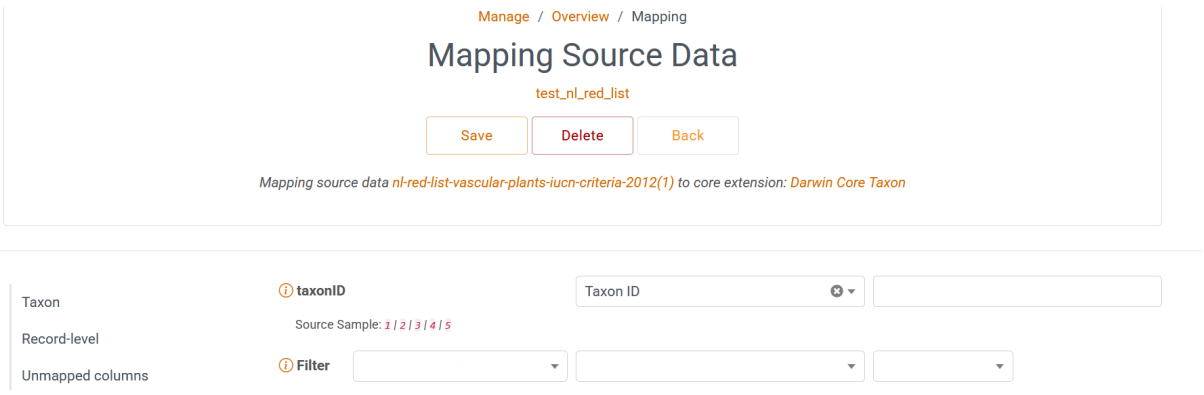

Mandatory mappings:

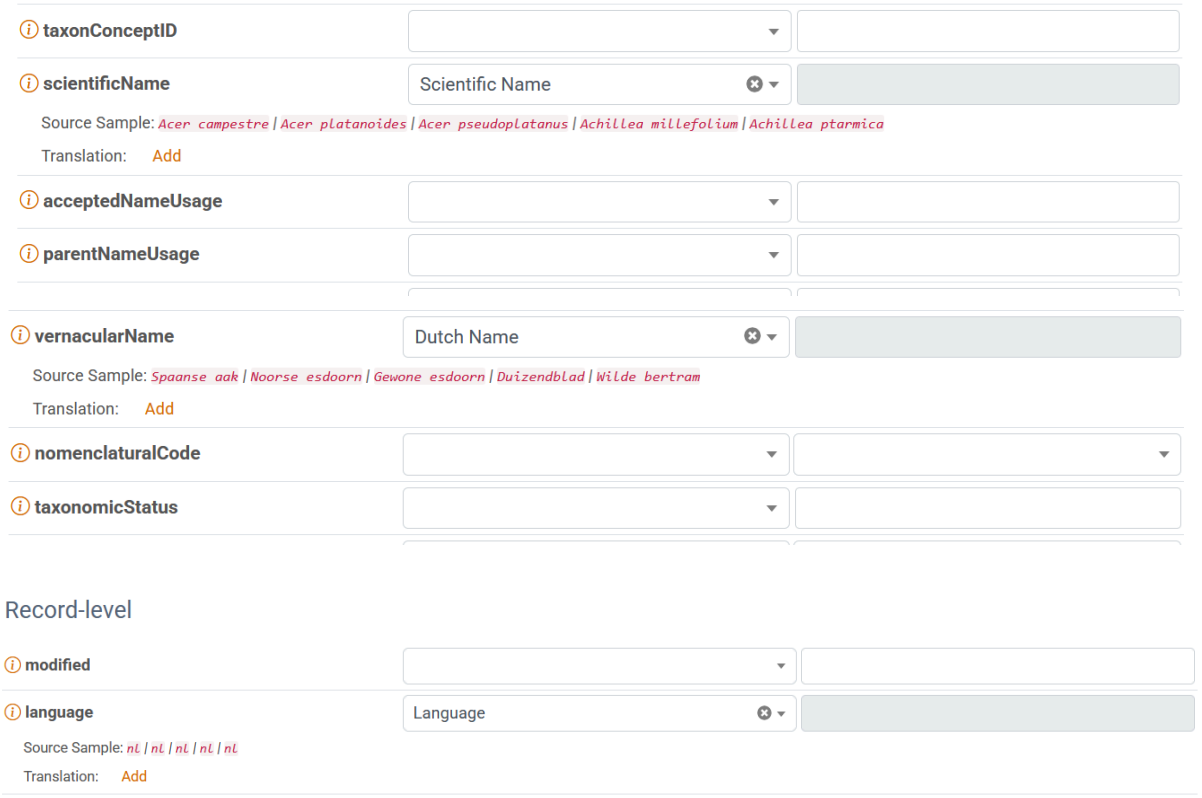
- **taxonID**
- **scientificName**

Recommended mappings (when available):

- **scientificNameAuthorship**
- **taxonRank**
- **kingdom**, **phylum**, **class**, **order**, **family**, **genus**
- **taxonomicStatus**
- **language** (original language of the list)

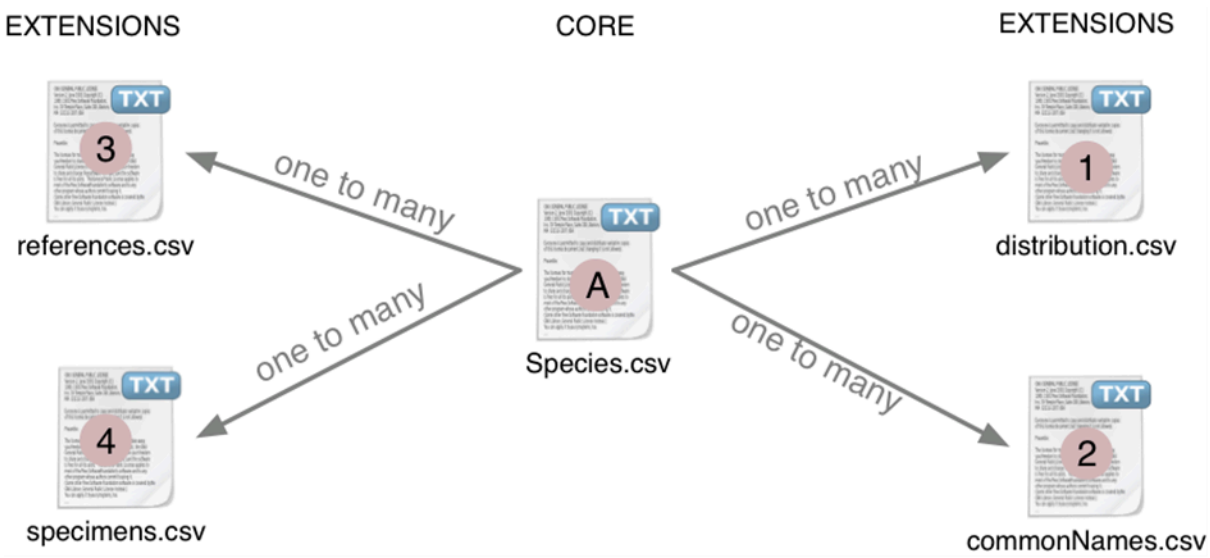
Mapping as many fields as possible increases data quality and interoperability.






# 7. Checklist Extensions

GBIF provides several extensions to enrich checklist data:
https://ipt.gbif.org/manual/en/ipt/latest/best-practices-checklists#gbif-checklist-extensions



To add a Darwin Core extension:

1. Go to Darwin Core Mappings and click **+ Add**.
2. Select the desired extension and click **Add**.
3. Choose the appropriate source file.
4. Map all relevant fields.

The source file can be the same file that is used for the checklist core mapping if all information is contained in a single file. However, extension data are often stored in separate files.

### 7.1 Species Distribution Extension

Latest version: 2022-02-02

https://rs.gbif.org/extension/gbif/1.0/distribution_2022-02-02.xml

Use this extension to share information on one or more distribution references for a taxon. One or more locality records may be linked to the same taxon. For example multiple localities, regions, or countries may be listed. Use this extension to describe the threat status for a taxon, seasonal distribution changes, and other properties linked to a taxon in a particular region.

Mandatory fields:

- **taxonID**
- **countryCode**
- **eventDate**

For red lists:

- **threatStatus**

### 7.2 Vernacular Names Extension

Latest version: 2015-02-13

https://rs.gbif.org/extension/gbif/1.0/vernacularname.xml

This extension provides the means to share information related to common (vernacular) names linked to taxa in the core data file. Multiple vernacular names can be linked to the same taxon via the taxonID.

Mandatory fields:

- **taxonID**
- **vernacularName**

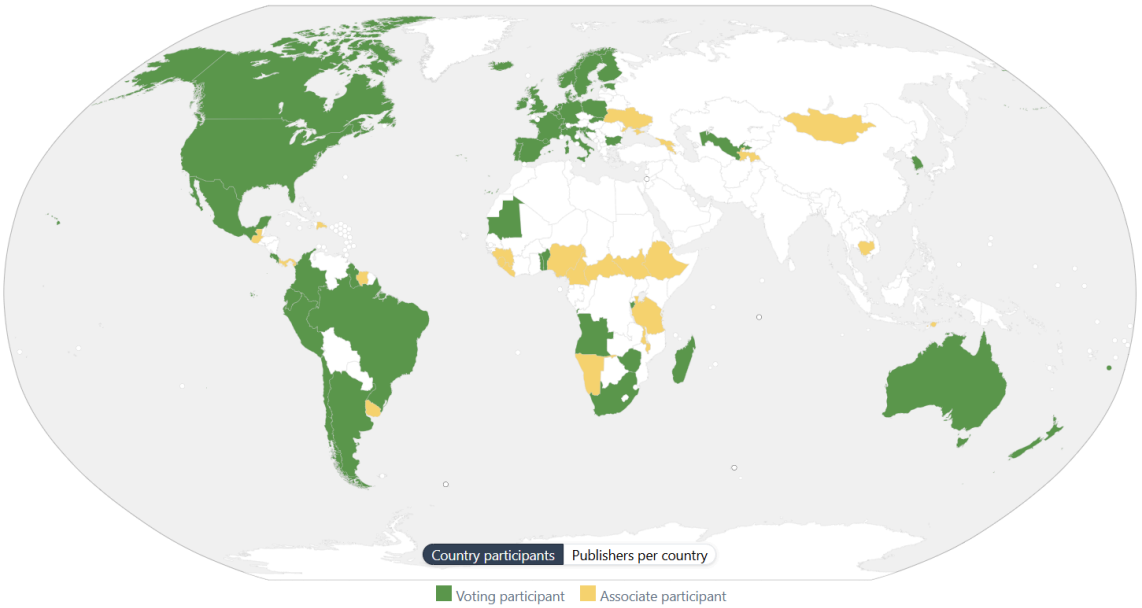
Recommended fields:

- **language** (to specify the language of the name)

If you wish to publish vernacular names in multiple languages, it is recommended to create a separate source file for each language. Upload each file as a separate extension and map the fields accordingly.

## 8. Become a GBIF data publisher

Before an organisation, or person that is affiliated to an organisation, can publish data to GBIF it is required to become a GBIF data publisher. Endorsement ensures the quality of data that is shared with GBIF. New publishers are endorsed by the national GBIF nodes. If a publisher is from a country that is not yet a GBIF member, the endorsement request is processed by the regional GBIF nodes, or at the level of the GBIF secretariat. You can find information about participating countries on the 'GBIF Participant network' page and the figure below.




**GBIF Participant network of voting and associate participants. Date: 2 October 2026.**

Details on how to become a GBIF data publisher can be found at https://www.gbif.org/become-a-publisher. Your national GBIF node or the GBIF secretariat can assist if there are any issues or questions regarding the process.

## 9. Metadata Requirements

Data without metadata can not be published at GBIF. Mandatory metadata includes information on the publishing organisation, contact details, geographic and temporal coverage of the dataset, etcetera. Which are described below.

Open the metadata editor by selecting **Metadata** and clicking **edit**, or upload a pre-prepared metadata file if available. All checklist and red list datasets must include complete and accurate metadata following the GBIF Metadata Profile.

**Click *Save* after every step, otherwise progress could be lost.**



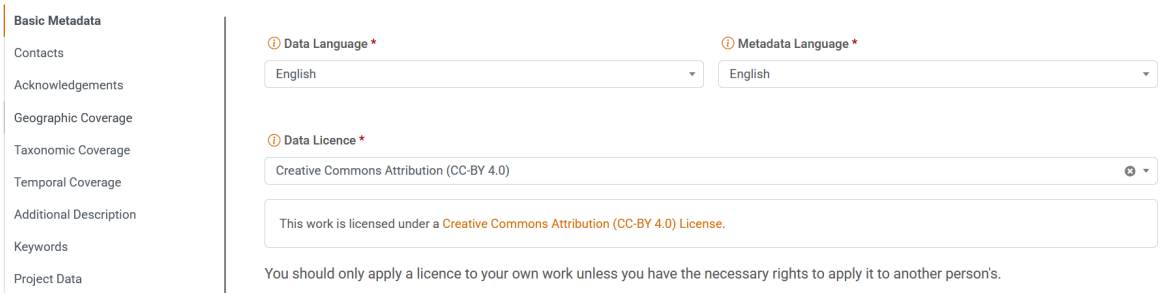

At minimum, the following fields must be filled in:

## 9.1 Basic Metadata

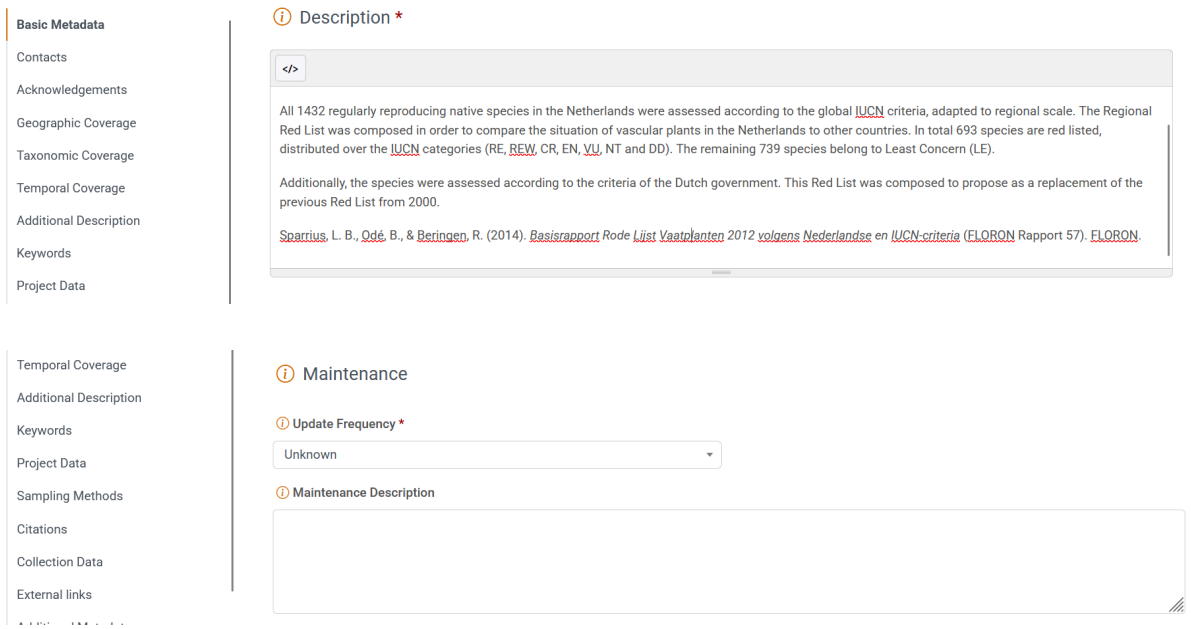
- **Title** – Provide the official title of the list, preferably in both English and the original language. The title must include at least the year of publication and the country name.
- **Publishing Organization** – Select your own organisation. If your organisation is not registered you should check section '8. Become a GBIF data publisher'.
- **Type** – Set to *Checklist*.




- **Data Language** – English.
- **Metadata Language** – English.
- **Data Licence** – Creative Commons Attribution (CC-BY 4.0).

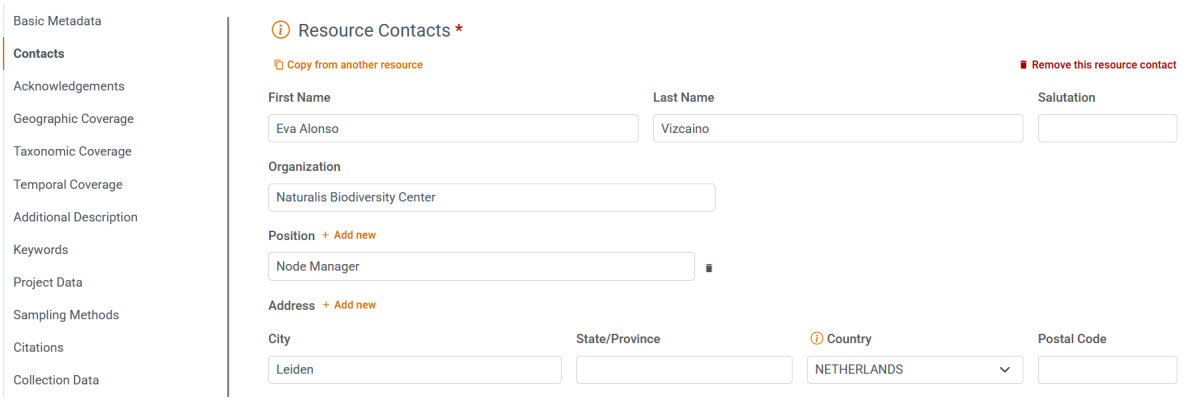



- **Description** – Provide a clear description of the checklist or red list, including its purpose, scope, and methodology. Always include the original citation of the source list.
- **Update Frequency** – Set to *Unknown*, as update cycles for these lists are typically not predetermined.



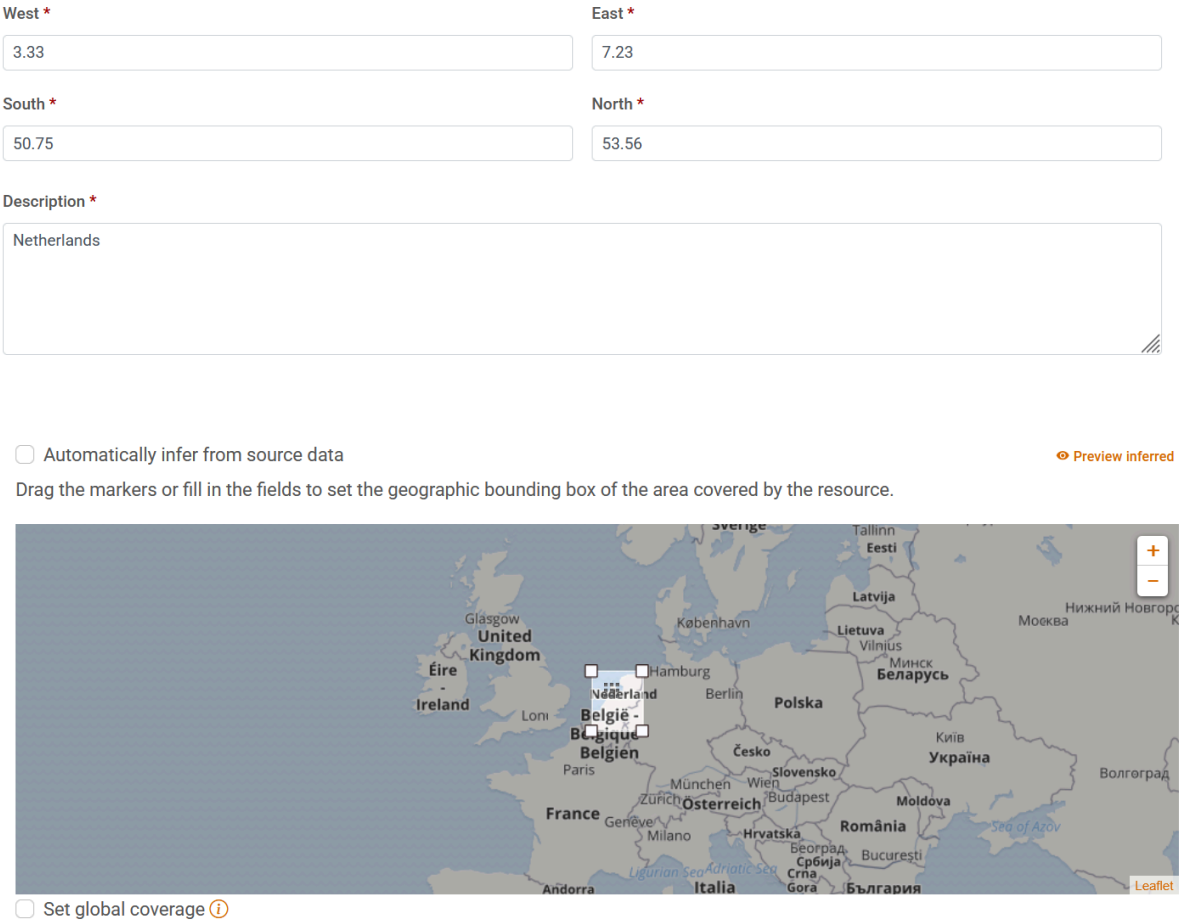

## 9.2 Contacts

- **Resource Contacts** – The contact details of the GBIF publishing organisation.
- **Resource Creators** – The authors of the original list.
- **Metadata Providers** – The person preparing the IPT resource (your name).



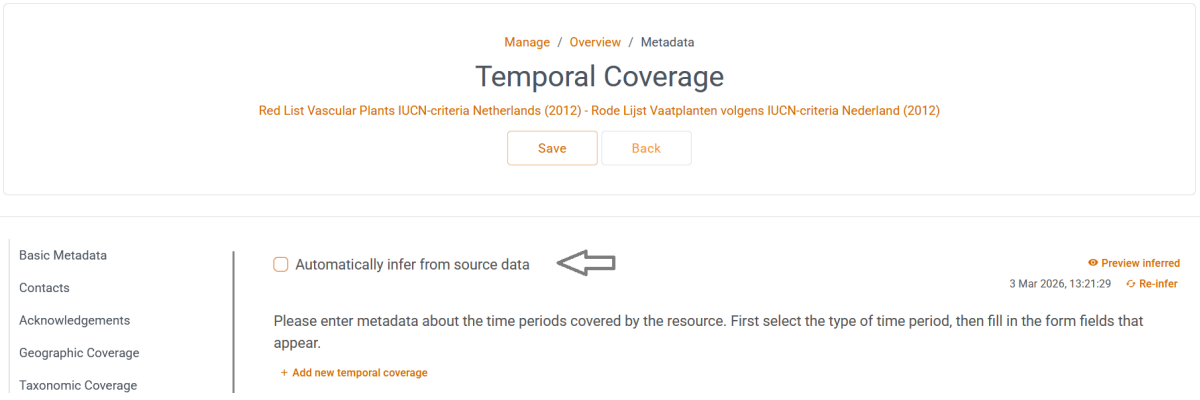

## 9.3 Geographic Coverage

- **West / East / South / North** – Provide the bounding coordinates of the country.
- **Description** – State the country name.



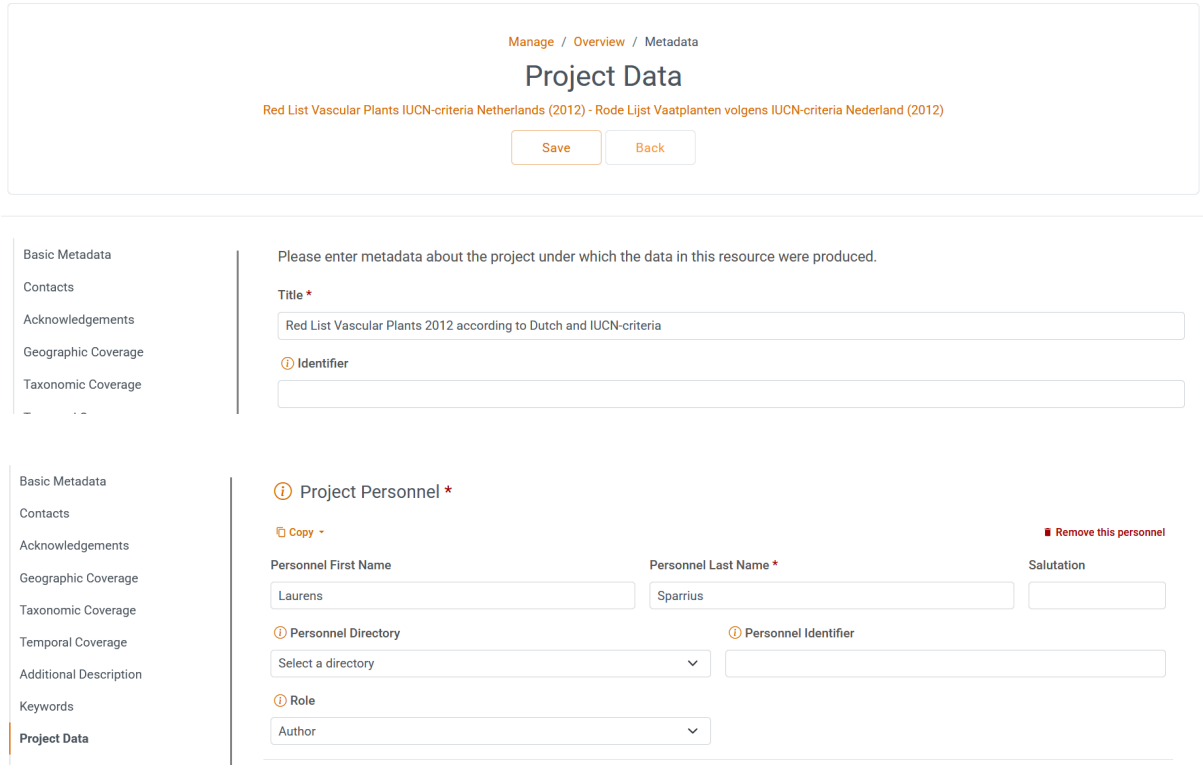

## 9.4 Temporal Coverage

- Set to **Automatically infer from source data**, unless specific assessment periods are known and should be manually entered.



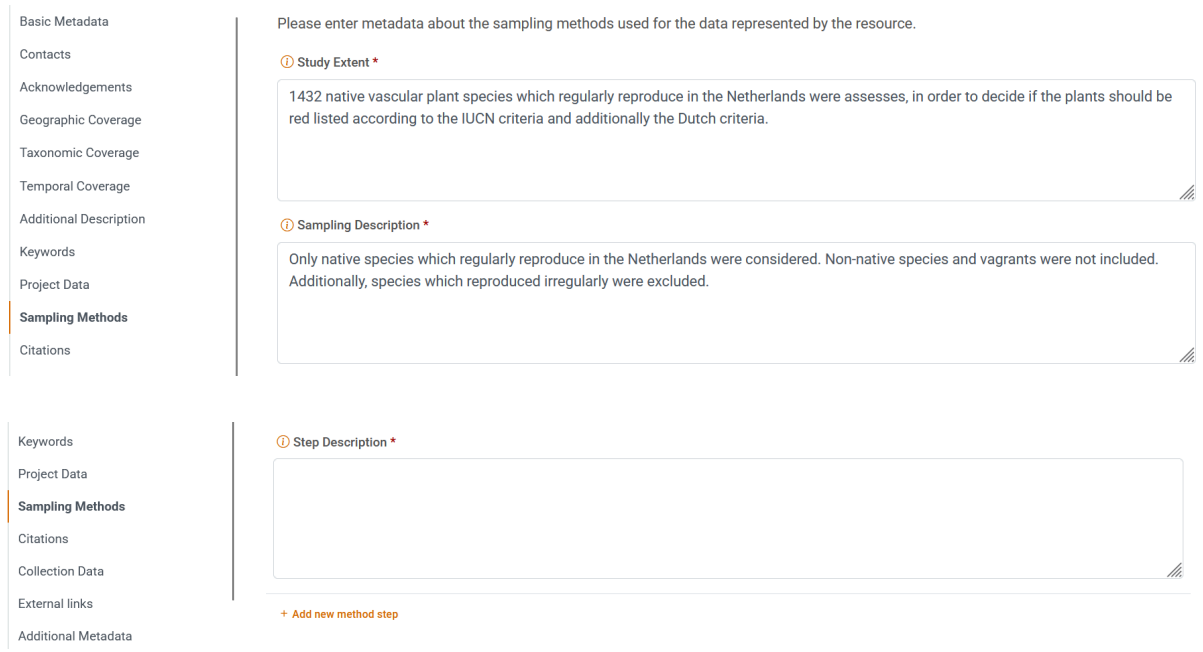

## 9.5 Project Data

- **Title** – The original title of the list or assessment project.
- **Project Personnel** – The authors or compilers of the list.



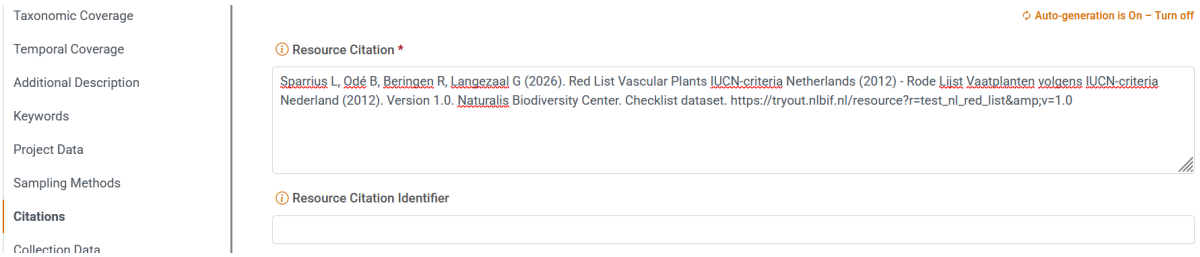

## 9.6 Sampling Methods

- **Study Extent** – Describe the geographic and thematic scope of the study, as documented in the accompanying report or publication.
- **Sampling Description** – Provide methodological details if available.
- **Step Description** – Outline any procedural steps relevant to the compilation or assessment process.



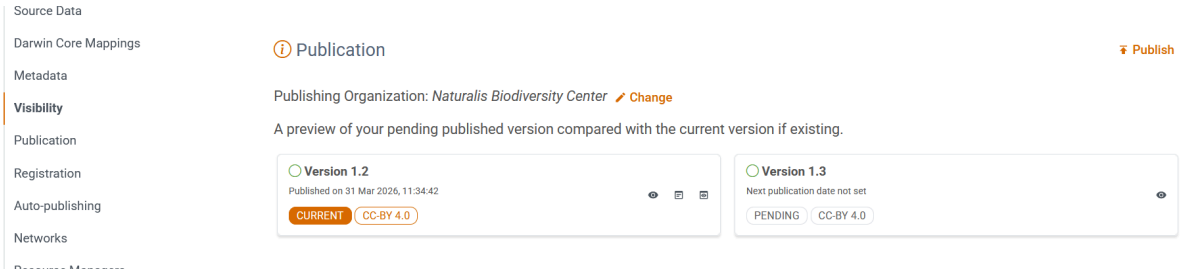
### 9.7 Citations

- **Resource Citation** – Use the IPT's auto-generated citation.
- **Resource Citation Identifier** – A DOI will be assigned automatically upon publication.

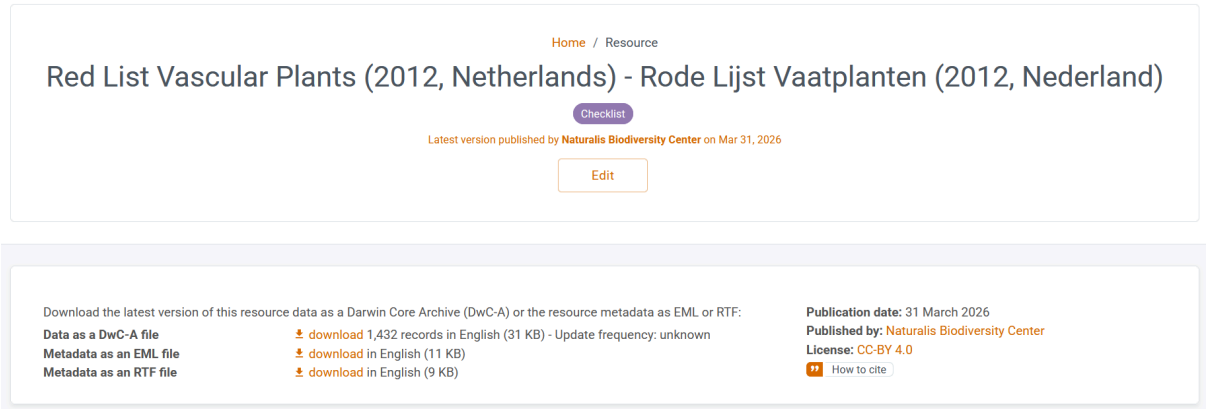


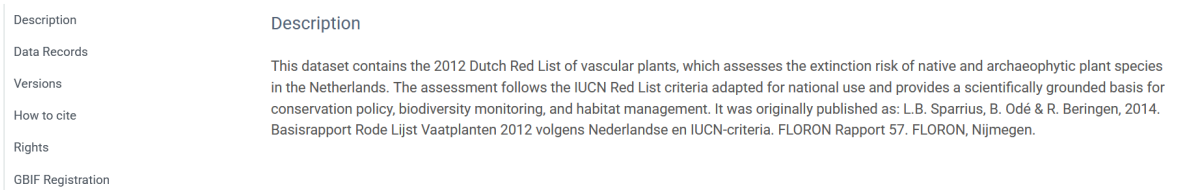

## 10. Downloading and Publishing

In order to view the checklist as they would appear on GBIF, it first needs to be published. Select **Publication**. This section shows the current version and the publishing organisation. To publish the current version, simply press *Publish* at the top right.

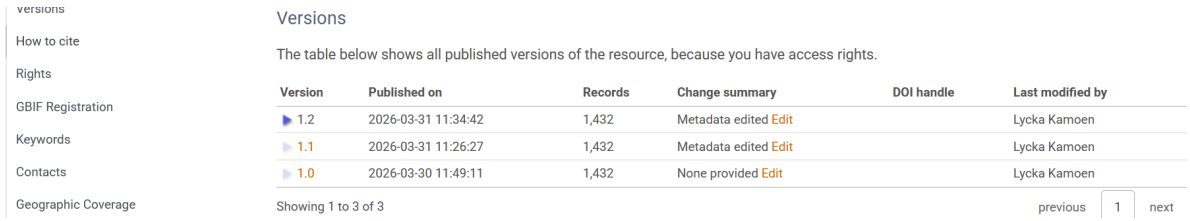



When pressing the button ***Overview*** followed by ***View*** right beneath the title of the checklist, a new page appears with the metadata of the checklist and a download link to the Darwin Core Archive (DwC-A).



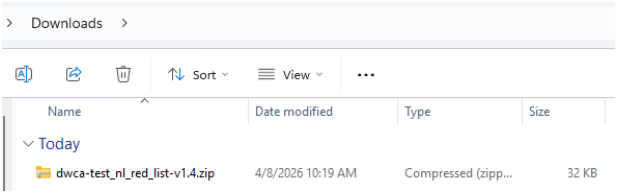



It is good practice to provide a brief summary of the changes made each time a new version is uploaded.

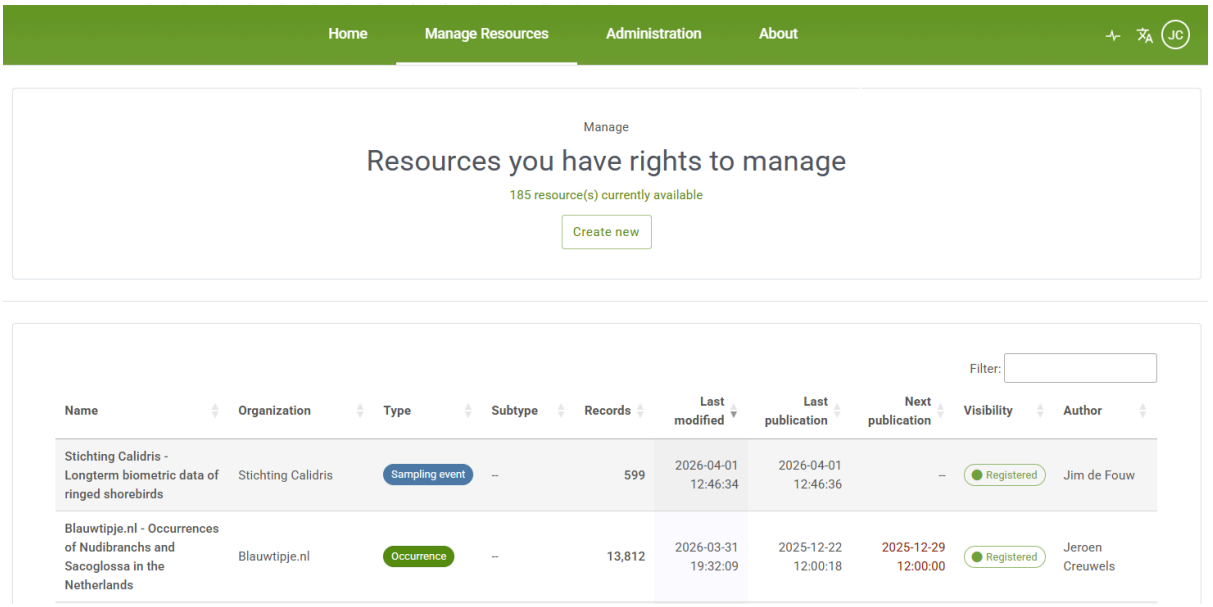



## 10.1 Transferring the dataset from Tryout-IPT to the real IPT and GBIF

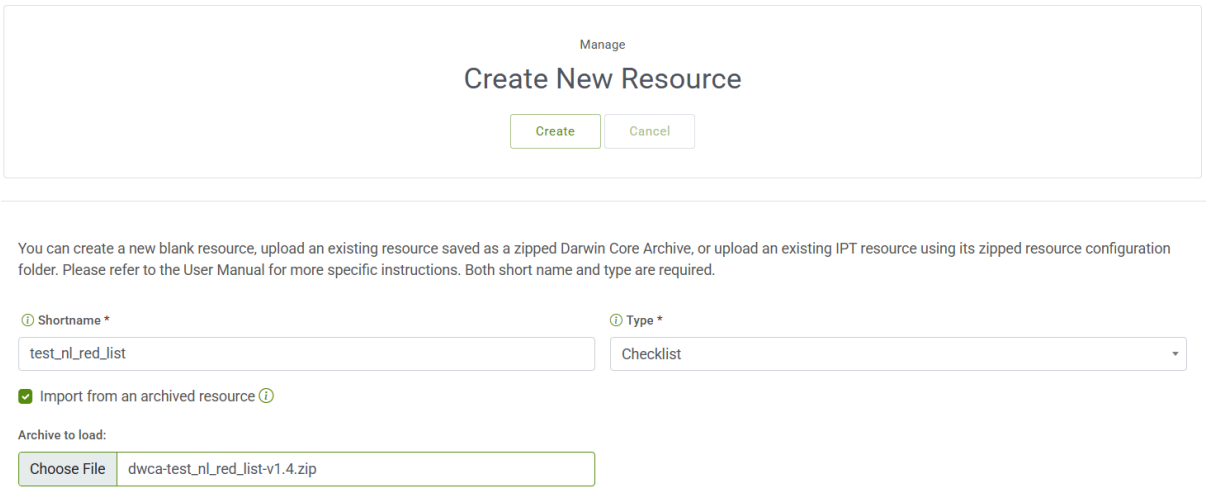


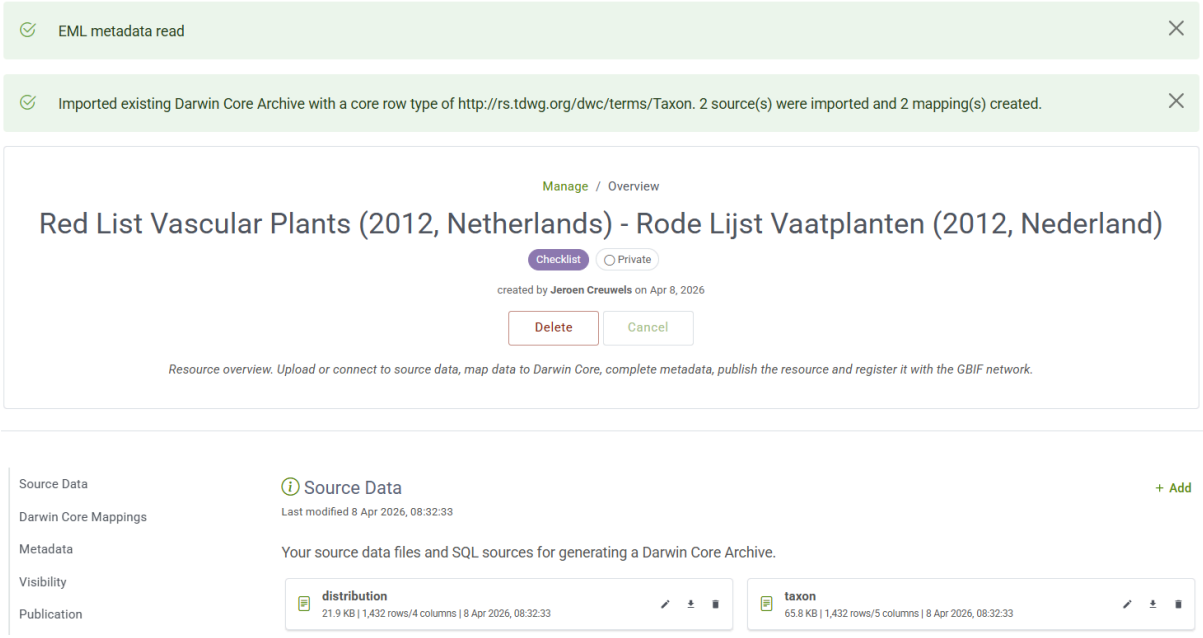

To transfer the dataset made on the Tryout-IPT, download the DwC-A file after publishing on the Tryout-IPT. Open the real IPT and navigate to Manage Resources, after logging in to your account.




Click ***Create new*** to start making a new dataset.

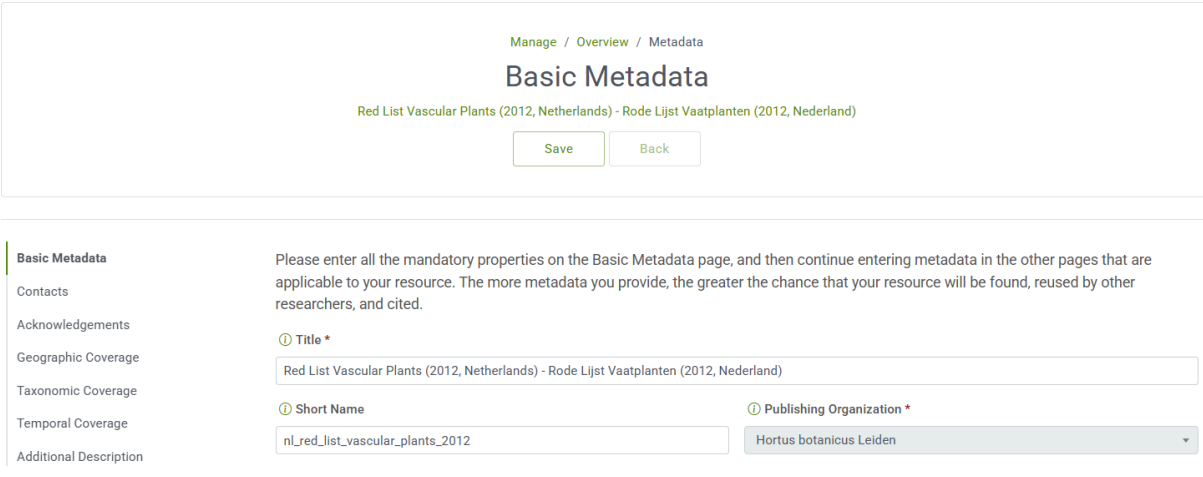



1. To import the dataset from the Tryout-IPT simply select ***Import from an archived resource*** and select your downloaded DwC-A file. The ***Shortname*** and ***Type*** will be automatically filled in.
2. Click ***Create***.




* The green comments indicate correct import of the dataset

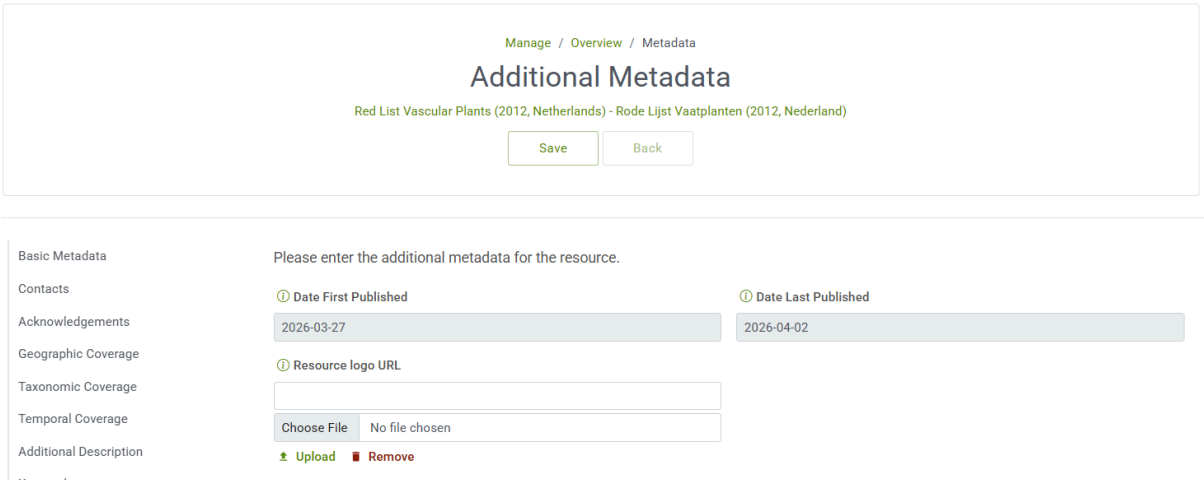
3. Navigate to ***Metadata***. This will show an “incomplete” message in read.
4. Click ***Edit***
5. Fill in the field for ***Publishing Organization*** (this will be empty when creating a new dataset from a DwC-A file)



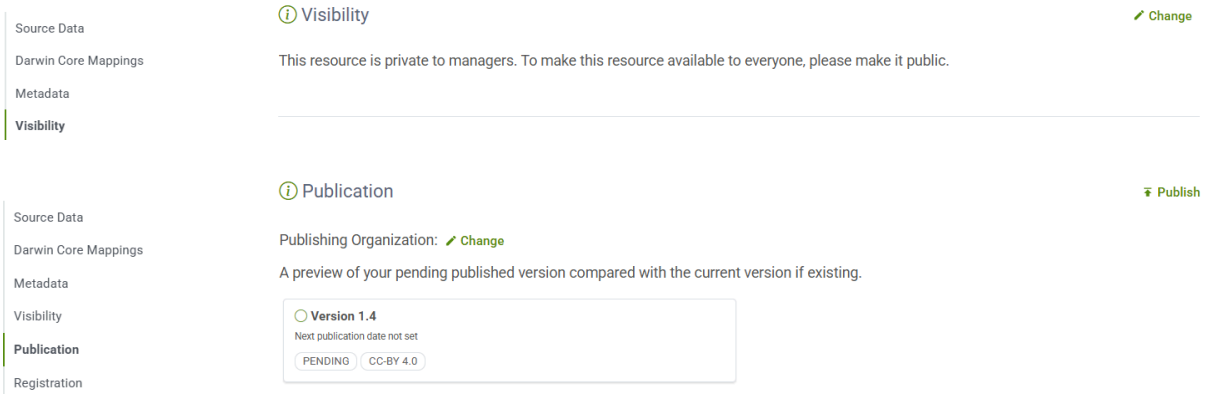

The logo of the publishing organization can be added as well (this is also possible on the Tryout-IPT).

6. Navigate to ***Additional Metadata***
7. Add the file in the ***Resource logo URL***

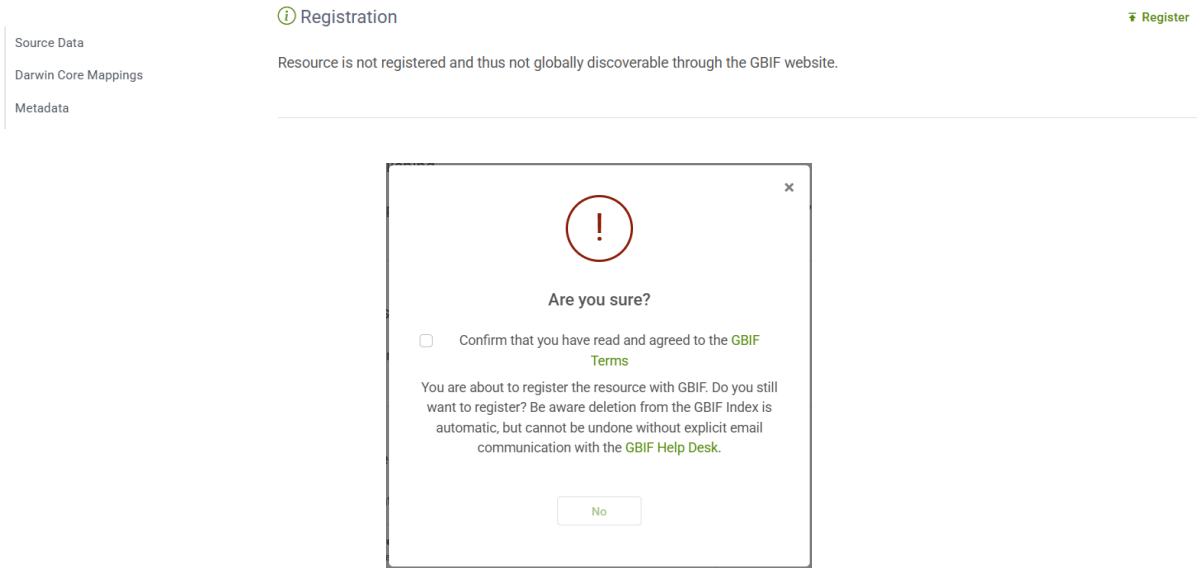
* For NLBIF many logos are available from: [http://ipt.nlbif.nl/resources/](http://ipt.nlbif.nl/resources/), simply copy the address of the logo.

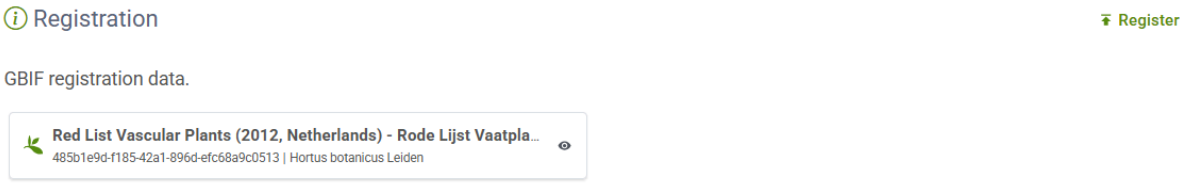



8. Navigate to ***Visibility*** and make the dataset public with ***Change***
9. Click on ***Publish*** in the ***Publication*** tab




10. Go to ***Registration*** to register the dataset on GBIF
11. Confirm in the pop-up screen and click *Yes*






Now the dataset is also visible on GBIF.




*When making edits to the dataset, always re-publish. The edits will take a while to appear on GBIF

*When altering the data itself, upload a completely new data file.

# 11. References

GBIF (2017) Best Practices in Publishing Species Checklists, version 2.1. Copenhagen: GBIF Secretariat. https://ipt.gbif.org/manual/en/ipt/3.0/best-practices-checklists.