



World Flora Online

Guidelines for Content Data Contributors

Version 2.01 – March 31, 2017

Updated by William Ulate Apr 2018

Version 2.02 - Updated by Technical Working Group June 2018

Version 2.03 – Updated by Chuck Miller & Walter Berendsohn June 2018

Version 2.04 – Updated by Chuck Miller April 2019, Adopted by WFO Council at Istanbul.

Prepared by the World Flora Online Technology Working Group

Table of Contents

1. Introduction	3
2. Content data ingestion process.....	3
STEP 1: Authorization and Registration	3
STEP 2: Name Matching and WFO ID Assignment to Names	3
NAME MATCHING INPUT FILE	3
NAME MATCHING RESPONSE FILE	4
STEP 3: Content data contribution	4
Classification orTaxon table	4
Description table	5
Reference table	6
Measurement_Or_Fact table	6
Vernacular_Name table.....	7
Images table	7
Distribution table	8
3. Further definitions and guidance on data elements	10
4. Typical issues with Content data contributions	14
When creating organization:.....	14
When creating the DwCA file:	14

1. Introduction

This document sets out how to contribute **Content data** to WFO. *Content*¹ includes information describing a taxon and its properties along with the scientific name accepted by the data provider (no synonymy or other classification data are needed). These accepted names are used to link *Content* to names in the *Taxonomic Backbone* via WFO's unique identifiers for scientific names (WFO-IDs). Updates to the nomenclature and classification used in the *WFO Taxonomic Backbone* are made using a separate process detailed in the WFO Guidelines for Taxonomic Data Contributors.

2. Content data ingestion process

Please refer to the General Guidelines for Data Contributors for how to be authorized and registered as a Data Provider for WFO. All Data Providers must first be authorized and registered before submitting a dataset.

Data Providers must commit to working with the *WFO Gatekeeper* to correct errors and missing data in their dataset so that it can be included in the World Flora Online *WFO Taxonomic Backbone*.

Please note that for *Content* data provision, only accepted names (taxon names) with their associated *Content* information are to be submitted.

STEP 1: Authorization and Registration

Before submitting data to the WFO, a Data Provider must be authorized and registered. The process for authorization and registration is provided in the General Guidelines.

STEP 2: Name Matching and WFO ID Assignment to Names

For the time being, name matching service is being provided by the WFO Gatekeeper. Please contact [WFO Gatekeeper](#) for information about submitting your Name Matching Input File for matching against the WFO Backbone. In future, a self-service name matching tool will be offered.

NAME MATCHING INPUT FILE

Field ²	Simplified Description
WFO ID	Provide WFO ID if the provider already has one for the name.
localID	The provider's local identifier for the name.
scientificName	Scientific name, with rank abbreviations for trinomials, but not with author citation
scientificNameAuthorship	Author citation of the names following IPNI standard author abbreviations
taxonRank	The rank of the name using a controlled vocabulary (see below)

¹ For explanations of terms in italics, see the General Guidelines for Data Contributors

² ^B eMonocot term, ^C Darwin Core term, ^D Dublin Core term

Field ²	Simplified Description
family	Scientific name of the family in which the name is/was classified (for accepted names / synonyms), generally according to APG 4

An Excel or file will be returned that provides you with the WFO-IDs you need to fill the obligatory “taxonID” field in the Darwin Core Archive Taxon table. The Comments field is used to point out possible inconsistencies in your data. Additional Excel tabs are sometimes used to list particular name issues.

NAME MATCHING RESPONSE FILE

Field ³	Simplified Description
WFO ID	The WFO ID assigned to this name. Will be empty if no match found.
localID	The provider's identifier for the name
scientificName	Scientific name, with rank abbreviations for trinomials, but not with author citation
scientificNameAuthorship	Author citation of the names following IPNI standard author abbreviations
taxonRank	The rank of the name using a controlled vocabulary (see below)
family	Scientific name of the family in which the name is/was classified (for accepted names / synonyms), generally according to APG 4
Comment	Explanation of matching results and any issues during name matching. Kinds of match are: Match, Near Match, May Match, Not Match or Empty

For names that did not match, (this is common)the Data Provider should review the response file and make adjustments as necessary. Commonly, the response file may be modified and resubmitted to the WFO Gatekeeper in an iterative fashion. Unmatched names that are new names to the WFO Backbone that are discovered during this process may be added to the Taxonomic Backbone.

After the completion of Step 2, all names provided should have WFO IDs assigned but often many names are not matched.

STEP 3: Content data contribution

WFO *Content* Data Provider are expected to contribute the following tables and data elements using a *Darwin Core (DwC) Archive* (see WFO General Guidelines for Contributors).

Classification or Taxon file

(**Bold** = required field)

Field ⁴	Simplified Description
--------------------	------------------------

³ B eMonocot term, ^C Darwin Core term, ^D Dublin Core term

⁴ B eMonocot term, ^C Darwin Core term, ^D Dublin Core term

Field ⁴	Simplified Description
taxonID ^c	The <i>WFO-ID</i> of the taxon name
scientificNameID ^c	Nomenclator ID (IPNI, etc)
scientificName ^c	Scientific name, with rank abbreviations for trinomials, but not with author citation
taxonRank ^c	The rank of the name using a controlled vocabulary (see below)
parentNameUsageID ^c	taxonID (<i>WFO-ID</i>) of name of taxon at next higher rank (accepted names only; see explanation below)
scientificNameAuthorship ^c	Author citation of the names following IPNI standard author abbreviations
family	Scientific name of the family in which the name is/was classified (for accepted names / synonyms), generally according to APG 4
genus ^c	The generic name (for a genus or as part of a species or infraspecies name) w/o authors. Obligatory for names at or below the rank of genus.
specificEpithet ^c	The species epithet (as part of a species or infraspecies name) w/o authors. Obligatory for names at or below the rank of species.
infraspecificEpithet ^c	The infraspecific epithet (as part of a species or infraspecies name) w/o authors. Obligatory for names below the rank of species.
namePublishedIn ^c	Nomenclatorial citation of the name. The abbreviation should follow established community standards (e.g. "Fl. Cameroun 34: 210. 1998")

Description file

These data to be provided for each descriptive record. In essence this is a combination of a descriptive text and the type of description (e.g. "distribution", "morphology", "ecology").

(**Bold** = required field)

Field ⁵	Simplified Description
taxonID ^c	The <i>WFO-ID</i> of the taxon name the description refers to
type ^d	Kind of description (shown as the title of the section in the portal). See list of permitted values below.
language ^d	Language of description. See list of permitted values below (not shown in the portal).
description ^d	Any descriptive text matching the kind of description defined by the "type". <i>Note: Authors, editors, etc. should not be included in the description field, rather in the "source" field.</i>
contributor ^d	An entity (e.g. a person) having been an author or made a contribution to the textual information given in the "description" field (not shown in the portal). N.B.: also see "source".
audience ^d	For whom the description is intended (not shown in the portal)

⁵ B eMonocot term, ^c Darwin Core term, ^d Dublin Core term

Field ⁵	Simplified Description
rightsHolder ^D	Text describing the person or organisation holding the rights (not shown in the portal)
created ^D	Date the record was created at the provider's side. (Not displayed in the portal, format: YYYYMMDD or YYYY-MM-DD.)
creator ^D	The person or organisation who created the record (not displayed in the portal).
source ^D	The identifier (from the Reference.txt file) of a reference that is to be cited as the source of this descriptive record (shown in the portal under Bibliography). <i>Note: all attribution of authors, editors, etc. should be included in the bibliographicCitation of the reference record for this source and not embedded within the "description" field.</i> In case of multiple references, use a comma-separated list of identifiers.
rights ^D	Text describing the rights being given (shown in the portal preceding the licence information)
license ^D	URL to the Creative Commons license (shown in the portal under the section "Information from" for the respective footnote)

Reference file

These References data to be provided for every reference to be associated to a descriptive data.
(**Bold** = required field)

Field	Simplified Description
taxonID ^C	The <i>WFO-ID</i> of the taxon name the reference refers to (Required)
Identifier ^D	An identifier that is unique to the datasource and that can be used as a reference in other files
bibliographicCitation ^C	Text giving the reference to be displayed in the portal
URI	URL weblink to the electronic publication. (Optional)

Measurement_Or_Fact file

This file provides

(**Bold** = required field)

Field	Simplified Description
taxonID ^C	The <i>WFO-ID</i> of the taxon name the reference refers to
MeasurementType ^C	Kind of the measurement, fact or assertion. The permitted values are: http://iucn.org/terms/threatStatus http://kew.org/wcs/terms/habitat http://kew.org/wcs/terms/lifeform
MeasurementValue ^C	The value of the measurement, fact or assertion. See list of permitted values for the respective type below. If terms are provided that are not in the list, then they will be shown as provided.

Field	Simplified Description
source ^D	Source of this measurement, fact or assertion. A URL or full publication citation (shown in the portal where the data are displayed).
rights ^D	Text describing the rights being given (shown in the portal in a footnote preceding the licence information)
license ^D	URL to the Creative Commons license (shown in the portal under the section "Information from" for the respective footnote)

Vernacular_Name files

(**Bold** = required field)

Field	Simplified Description
taxonID ^C	The <i>WFO-ID</i> of the taxon name the reference refers to
vernacularName ^C	The name, shown in the portal in the "common name" section.
language ^D	The language of the name (shown next to the name in the portal).
countryCode ^D	The ISO 3166-1-alpha-2 code for the country
rights ^D	Text describing the rights being given (shown in the portal in a footnote preceding the licence information)
license ^D	URL to the Creative Commons license (shown in the portal under the section "Information from" for the respective footnote)

Images table

(**Bold** = required field)

Field (Image table)	Simplified Description
taxonID ^C	The <i>WFO-ID</i> of the taxon name the reference refers to (calls up the taxon name in the image metadata table "image of")
title ^C	The caption of the image (shown as title, under the image itself and as "Caption" in the image metadata table)
identifier	The public URL that identifies and locates the image file directly, not the html page it might be shown on (see also "references").
rights ^C	Information about rights held in and over the image (shown in the "Information from" section under the image metadata table)/
license ^C	License for this image. Can be text or a url like creative commons uses (displayed after the rights statement in the "information from" section under the image metadata table).
format ^C	The file format, physical medium, or dimensions of the image (not shown in the portal)

Field (Image table)	Simplified Description
source ^c	If the image was derived or taken from another source this is the reference to that resource. For example a book from which an image was scanned or the original provider of a photo, such as photography agencies (not shown in the portal).
audience ^c	A description for whom the image is intended or useful (shown under "Audience" in the image metadata table)
creator ^c	A person or legal person primarily responsible for making the resource useful (shown under "Creator" in the image metadata table and under the image itself)
created ^c	Date the record was created at the provider's side (not displayed in the portal, format: YYYYMMDD or YYYY-MM-DD).
modified ^b	Date the record was modified the last time at the provider's side. (Not displayed in the portal, format: YYYYMMDD or YYYY-MM-DD.)
contributor ^c	Any contributor in addition to the creator that helped in recording the media item (shown under "Contributor" in the image metadata table)
publisher ^c	The responsible for making the image available (shown under "Publisher" in the image metadata table)
description ^c	A textual description of the content of the media item (shown under "Description" in the image metadata table).
locality ^b	A description of the locality where the image was taken (shown under "Locality" in the image metadata table)
subject ^b	A list of terms describing the image (shown under "Keywords" in the image metadata table).
rightsHolder ^c	A person or organization owning or managing rights over the media item (not displayed in the portal).
latitude ^b	The latitude at which the image was taken (is shown under "Location" in the image metadata table). Use decimal degrees.
longitude ^b	The longitude at which the image was taken (is shown under "Location" in the image metadata table). Use decimal degrees.
references ^c	An html webpage that shows the image or its metadata. This link should be used for html based image viewers like FSI and should be used as a source link wherever the media item is shown. It is recommended to provide this url even if a media file exists as it will be used for linking out (not shown in the portal?)

Distribution file

(**Bold** = required field)

Field	Simplified Description

Field	Simplified Description
taxonID ^c	The <i>WFO-ID</i> of the taxon name the distribution record refers to.
locality ^c	The verbatim name of the area this distribution record is about.
locationID ^c	The TDWG code for the locality. See https://www.kew.org/tdwg-world-geographical-scheme-recording-plant-distributions for the valid TDWG codes.
countryCode ^c	ISO3166 alpha 2 (3 is permissible) country codes the area belongs to or as an alternative for a locationID if the area is a country. For multiple countries separate values with a comma ","
establishmentMeans ^c	Term describing the status of the organism in the given area based on how frequent the species occurs
source ^c	Source reference for this distribution record.
occurrenceRemarks ^c	Comments or notes about the distribution

3. Further definitions and guidance on data elements.

For DwC terms, some of the definitions given here focus on their use in the botanical context of WFO. They thus represent special cases of the general definitions given in the 'DwC Terms' – <http://rs.tdvwg.org/dwc/terms/index.htm>. The 'Type value' must be all lower-case letters.

type (Description file) – list of permitted values

Type value	URI with the description of the type
general	http://rs.gbif.org/vocabulary/gbif/descriptionType/general
diagnostic	http://rs.gbif.org/vocabulary/gbif/descriptionType/diagnostic
morphology	http://rs.gbif.org/vocabulary/gbif/descriptionType/morphology
habit	http://rs.gbif.org/vocabulary/gbif/descriptionType/habit
cytology	http://rs.gbif.org/vocabulary/gbif/descriptionType/cytology
physiology	http://rs.gbif.org/vocabulary/gbif/descriptionType/physiology
size	http://rs.gbif.org/vocabulary/gbif/descriptionType/size
lifespan	http://rs.gbif.org/vocabulary/gbif/descriptionType/lifespan
lifetime	http://rs.gbif.org/vocabulary/gbif/descriptionType/lifetime
biology	http://rs.gbif.org/vocabulary/gbif/descriptionType/biology
ecology	http://rs.gbif.org/vocabulary/gbif/descriptionType/ecology
habitat	http://rs.gbif.org/vocabulary/gbif/descriptionType/habitat
distribution	http://rs.gbif.org/vocabulary/gbif/descriptionType/distribution
reproduction	http://rs.gbif.org/vocabulary/gbif/descriptionType/reproduction
conservation	http://rs.gbif.org/vocabulary/gbif/descriptionType/conservation
use	http://rs.gbif.org/vocabulary/gbif/descriptionType/use
dispersal	http://rs.gbif.org/vocabulary/gbif/descriptionType/dispersal
lifecycle	http://rs.gbif.org/vocabulary/gbif/descriptionType/lifecycle
growth	http://rs.gbif.org/vocabulary/gbif/descriptionType/growth
genetics	http://rs.gbif.org/vocabulary/gbif/descriptionType/genetics
chemistry	http://rs.gbif.org/vocabulary/gbif/descriptionType/chemistry
associations	http://rs.gbif.org/vocabulary/gbif/descriptionType/associations
population	http://rs.gbif.org/vocabulary/gbif/descriptionType/population
management	http://rs.gbif.org/vocabulary/gbif/descriptionType/management
legislation	http://rs.gbif.org/vocabulary/gbif/descriptionType/legislation
threats	http://rs.gbif.org/vocabulary/gbif/descriptionType/threats
typematerial	http://rs.gbif.org/vocabulary/gbif/descriptionType/typematerial
typelocality	http://rs.gbif.org/vocabulary/gbif/descriptionType/typelocality
phylogeny	http://rs.gbif.org/vocabulary/gbif/descriptionType/phylogeny
hybrids	http://rs.gbif.org/vocabulary/gbif/descriptionType/hybrids
literature	http://rs.gbif.org/vocabulary/gbif/descriptionType/literature
culture	http://rs.gbif.org/vocabulary/gbif/descriptionType/culture
vernacular	http://rs.gbif.org/vocabulary/gbif/descriptionType/vernacular

measurementType – List of Permitted Values

<http://iucn.org/terms/threatStatus>

<http://kew.org/wcs/terms/habitat>

<http://kew.org/wcs/terms/lifeform>

measurementValue – List of Translated Values

The values below will be converted to the displayed term. Other values may be provided and will be displayed exactly as provided.

Terms for measurementType from World Checklist System = http://kew.org/wcs/terms/lifeform	
Term	Displayed Term
bamboo	Bamboo
biennial	Biennial
bulb	Bulb
caudex	Caudex
cham	Chamaephyte
cl	Liana
epiphyte	Epiphyte
epiphytic	Epiphytic
geophyte	Geophyte
geopyte	Geopyte
hel	Helophyte
hemicr	Hemicryptophyte
hemimycotrophic	Hemimycotrophic
hemipar	Hemiparasite
herb	Herbaceous
holomycotroph	Holomycotroph
holomycotrophic	Holomycotrophic
holopar	Holoparasitic
hydro	Semi-aquatic
hydrocham	Hydrochamaephyte
hydrogeophyte	Hydrogeophyte
hydrohemicr	Hydrohemicryptophyte
hydrother	Hydrotherophyte
lithophyte	Lithophyte
lithophytic	Lithophytic
monocarpic	Monocarpic
nanophan	Nanophanerophyte
phan	Phanerophyte
pseudobulb	Pseudobulb
rhizome	Rhizome
rizome	Rizome

Terms for measurementType from World Checklist System = <http://kew.org/wcs/terms/lifeform>

succ	Succulent
ther	Therophyte
tuber	Tuber

**for measurementType = <http://iucn.org/terms/threatStatus>
(must be provided in lower case letters)**

cd	Conservation Dependent
cr	Critically Endangered
ex	Extinct
ew	Extinct in the wild
en	Endangered
vu	Vulnerable
nt	Near threatened
lc	Least Concern
lr	Lower Risk
dd	Data deficient
ne	Not evaluated

for measurementType = <http://kew.org/wcs/terms/habitat>

a	Tundra
dt	Desert and Xeric Shrubland
g	Temperate Mixed Forest
mt	Tropical Dry Forest
s	Mediterranean Forest, woodlands and Scrub
st	Subtropical and tropical
ta	Montane tropical
wt	Tropical Moist Forest
gc	Temperate Coniferous Forest
tc	Tropical Coniferous Forest
tg	Tropical Grassland and Savanna
ma	Mangrove
bf	Boreal Forest or Taiga
gg	Temperate Grassland and Savanna
fg	Flooded Grassland and Savanna
mg	Montane Grassland and Shrubland
ri	Rock and Ice
ma	Mangrove

aq

Aquatic

4. Typical issues with Content data contributions

When creating organization:

- Keywords must be separated with semi-colon.
- Creation date format must be YYYY-MM-DD or YYYYMMDD

When creating the DwCA file:

- Do not include a time into date fields, use YYYY-MM-DD or YYYYMMDD only
- Use only the permitted values for the nomenclatural status
- namePublishedIn is a String field whose size must be between 0 and 255
- Separation description source field references by commas (any comma in the field will produce a new record in references).