CESP 2020-006 :
"National Portals Workshop: taxonomic databases, species data information and visualization"

Webinar 3 "International biodiversity taxonomic databases"

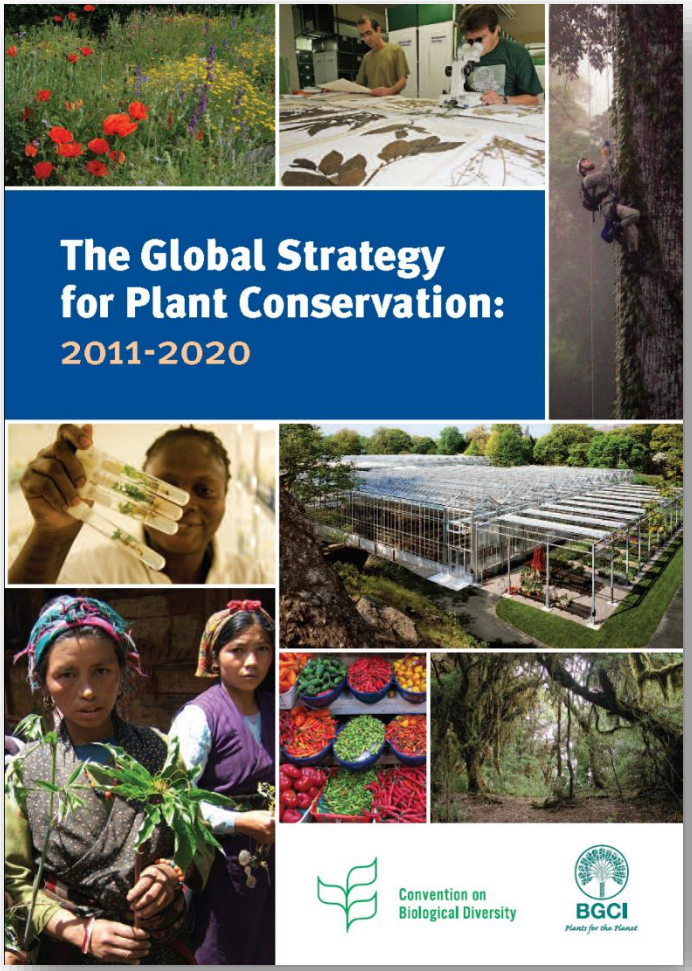
Buenas practicas y herramientas para la construcción y mantenimiento de la Columna Taxonómica de la Flora Mundial en Línea

Presentado por: William Ulate Rodríguez

Centro de Informática de la Biodiversidad

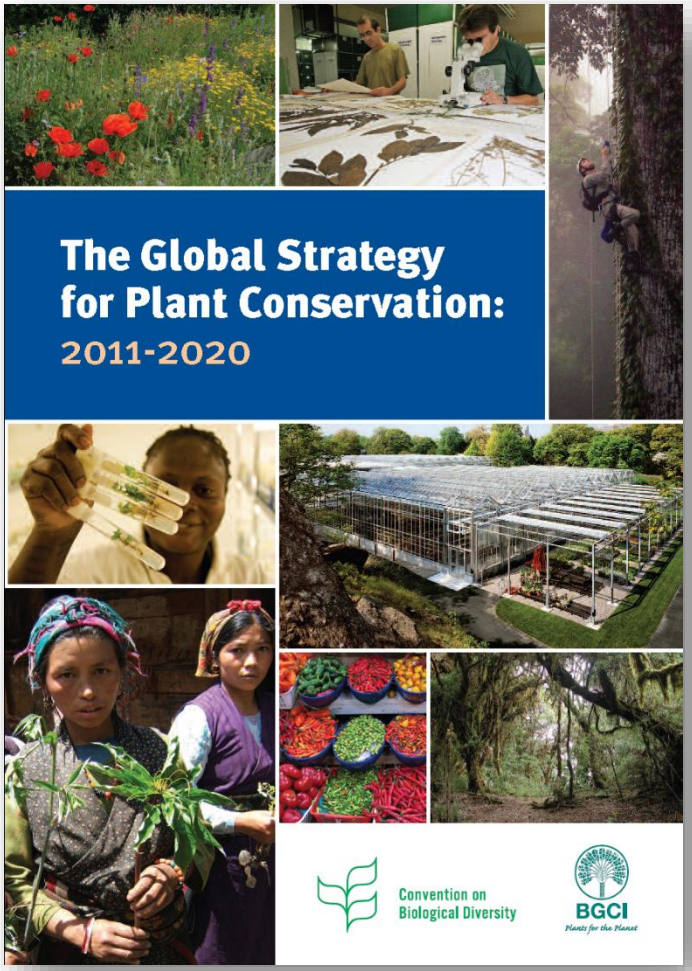
Depto. de TI, Jardín Botánico de Missouri, EEUU

27 de Setiembre de 2021



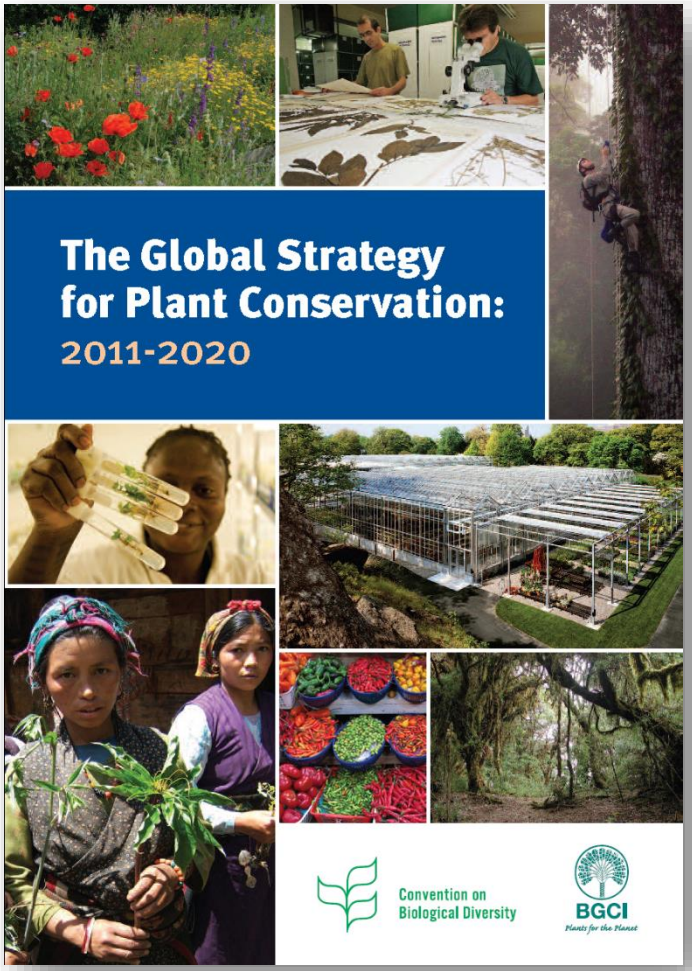
Estrategia Mundial para la Conservación de las Especies Vegetales 2011-2020

GSPC



- Adoptado en 2002
- Actualizado para el 2011-2020
- **Reducir la pérdida de especies vegetales para el 2020**

GSPC



Objetivo 1: La diversidad de plantas está bien entendida, documentada y reconocida

Meta 1: Una flora en línea de todas las plantas conocidas

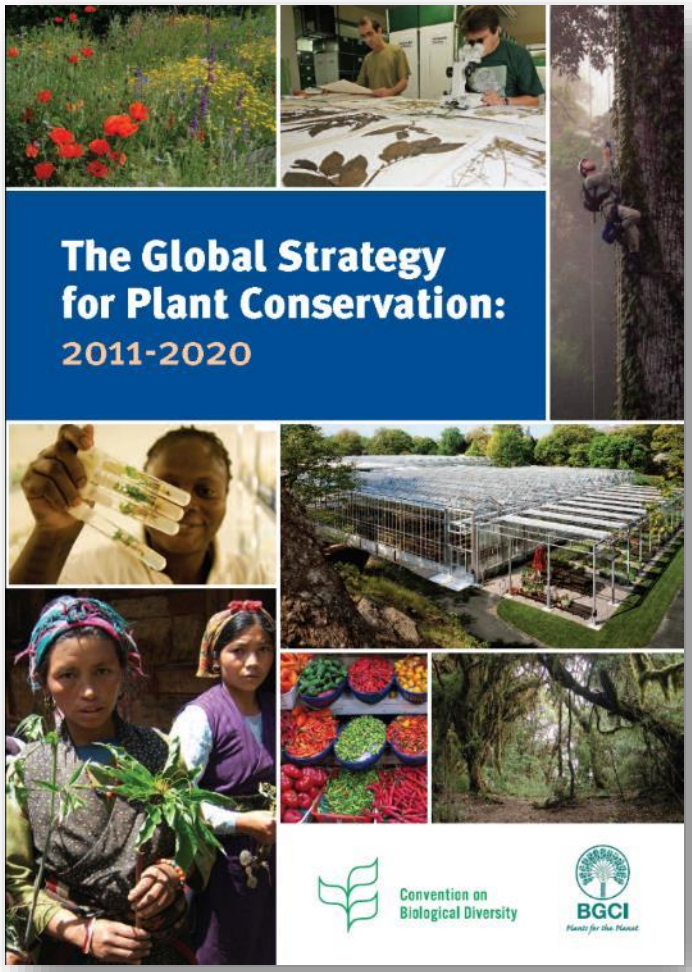
Meta 1 - 2020

Una Flora en línea de todas las plantas conocidas incluye:

- Nombres Aceptados
- Una sinonimia comprensiva

La Meta 1 – 2010 provee la base para la creación de una Flora Global

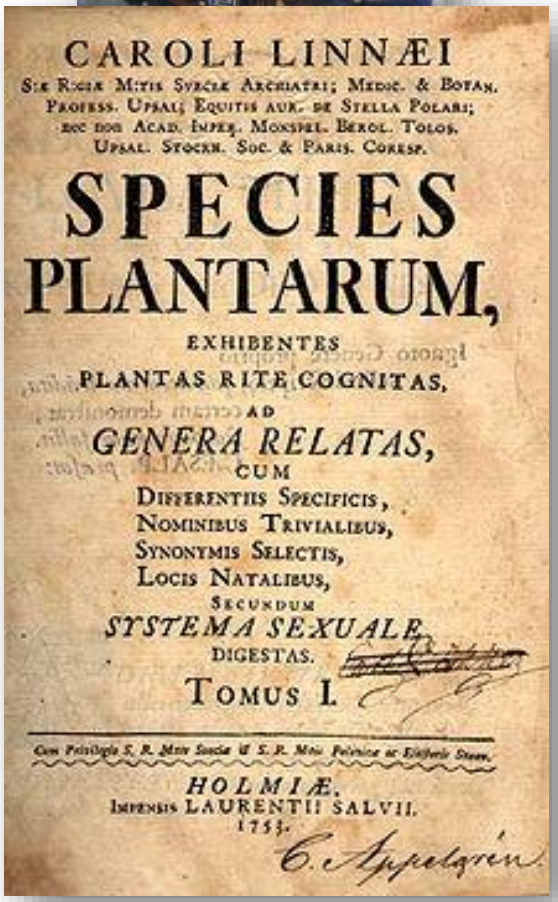
Muchas de las Partes están implementando proyectos de Floras digitales a nivel nacional



Esfuerzos previos para catalogar todas las Plantas del Mundo



Carl Linnaeus (1707–1778)


CAROLI LINNÆI
SIE REGIA MITIS SVECIÆ ARCHIATRI; MEDIC. & BOTAN. PROFESS. UPSAL; EQUITIS AUR. DE STELLA POLARI; NEC NON ACAD. IMPER. MONSPEL. BEROL. TOLOS. UPSAL. STOCKH. SOC. & PARIS. CORESP.

SPECIES PLANTARUM,

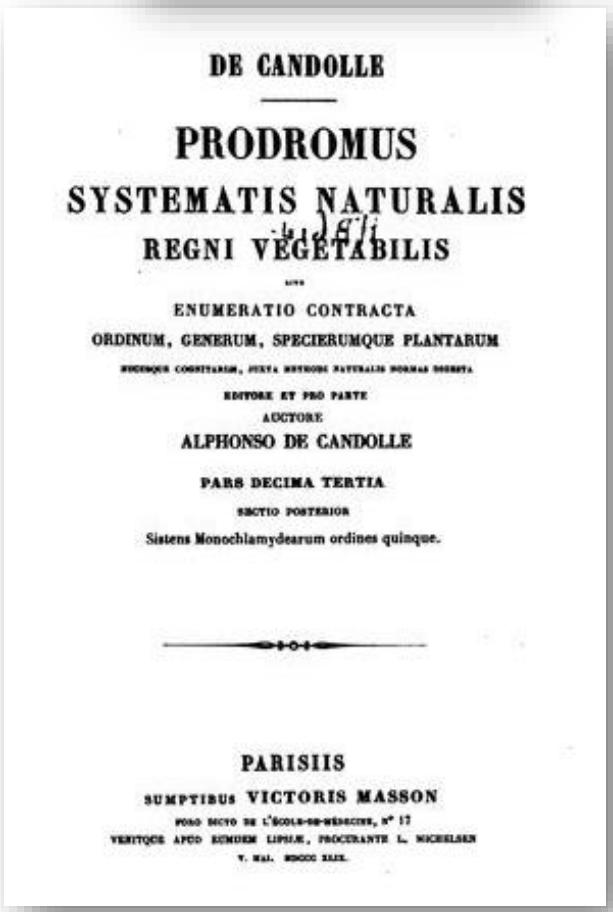
EXHIBENTES
PLANTAS RITE COGNITAS,
AD
GENERA RELATAS,
CUM
DIFFERENTIIS SPECIFICIS,
NOMINIBUS TRIVIALIBUS,
SYNONYMIS SELECTIS,
LOCIS NATALIBUS,
SECUNDUM
SYSTEMA SEXUALE
DIGESTAS.

TOMUS I.

HOLMIÆ,
IMPENSIS LAURENTII SALVII.
1753.

- *Species Plantarum* (1753) incluye **6 000** especies en ca. 1 000 géneros
- Linnaeus expandió su conocimiento para cubrir alrededor de 7 700 especies de plantas
- Él consideraba que el mundo no tenia más de 10 000 especies de plantas en total.




DE CANDOLLE

PRODROMUS
SYSTEMATIS NATURALIS
REGNI VEGETABILIS

ENUMERATIO CONTRACTA
ORDINUM, GENERUM, SPECIERUMQUE PLANTARUM

EDITORE ET PRO PARTE
AUCTORE
ALPHONSO DE CANDOLLE

PARS DECIMA TERTIA
SECTIO POSTERIOR
Sistens Monochlamydearum ordines quinque.

PARISIIS
SUMPTIBUS VICTORIS MASSON

Esfuerzos previos para catalogar todas las Plantas del Mundo

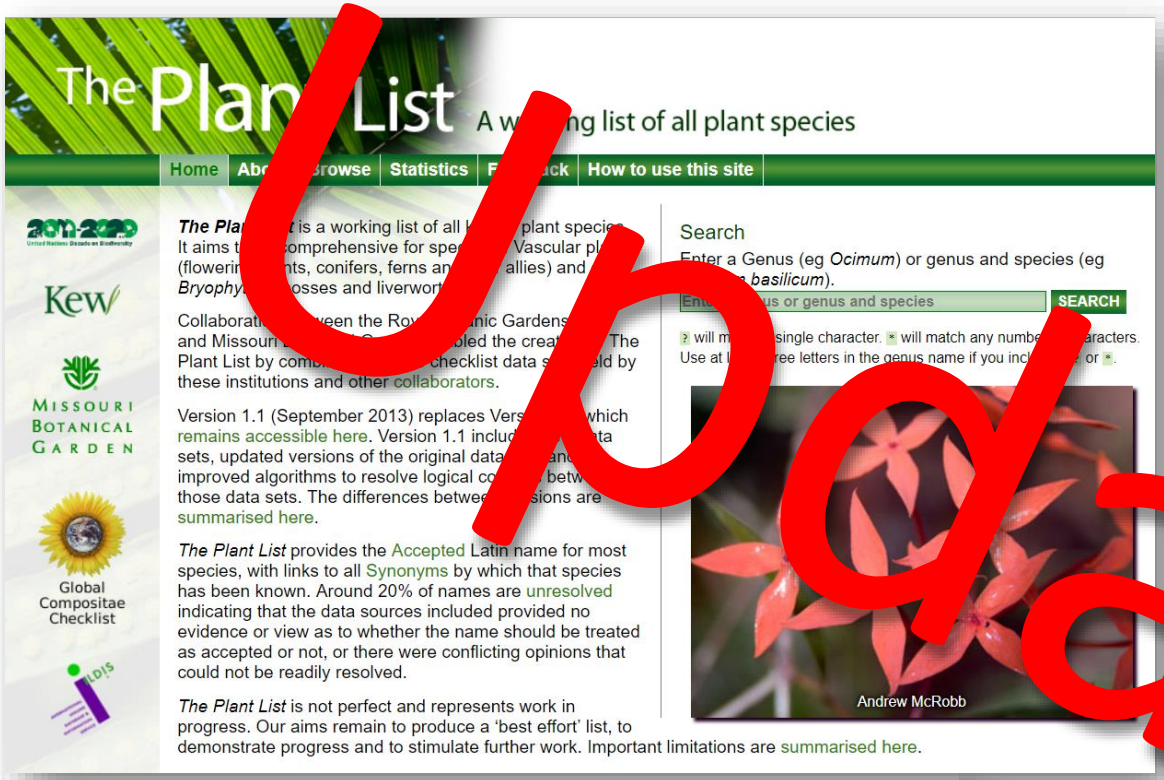
Augustin-Pyramus de Candolle (1778–1841)

- Compilación del *Prodromus Systematis Naturalis Regni Vegetabilis* inició en 1824 y, eventualmente abarcó 17 volúmenes
- Colectivamente, el *Prodromus* provee un tratamiento para **58 975** especies de Dicotiledoneas y Gymnospermas
- Las monocotiledóneas y los *Ficus* nunca fueron completados

World Flora Online
www.worldfloraonline.org

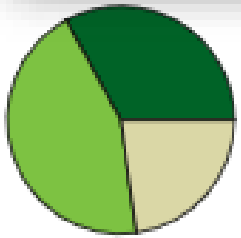
The Plant List

- Versión 1.1
- Enero 2014



Update

Rango Taxonómico	Cantidad
Familias	642
Genera	17 020
Especies	1 064 035



Status	Total	
Accepted	350,699	33.0%
Synonym	470,624	44.2%
Unresolved	242,712	22.8%

World Flora Online
www.worldfloraonline.org

http://www.worldfloraonline.org/



Ratificación de un Proyecto Internacional

Ratificaciones:

Asociación Mundial para la Conservación de las Especies Vegetales (GPPC), St Louis, 2011



Congreso Botánico Internacional (IBC), Melbourne, 2011

IBC2011
XVIII international botanical congress
MELBOURNE AUSTRALIA | 23-30 JULY 2011

Un Consorcio Internacional

- **Reunión Inicial**
  St Louis, Febrero 2012
- **Evento Alterno SBSTTA**
  Montreal, Mayo 2012
- **Planeamiento**
  St Louis, Julio 2012
- **Lanzamiento Proyecto**
  India, October 2012



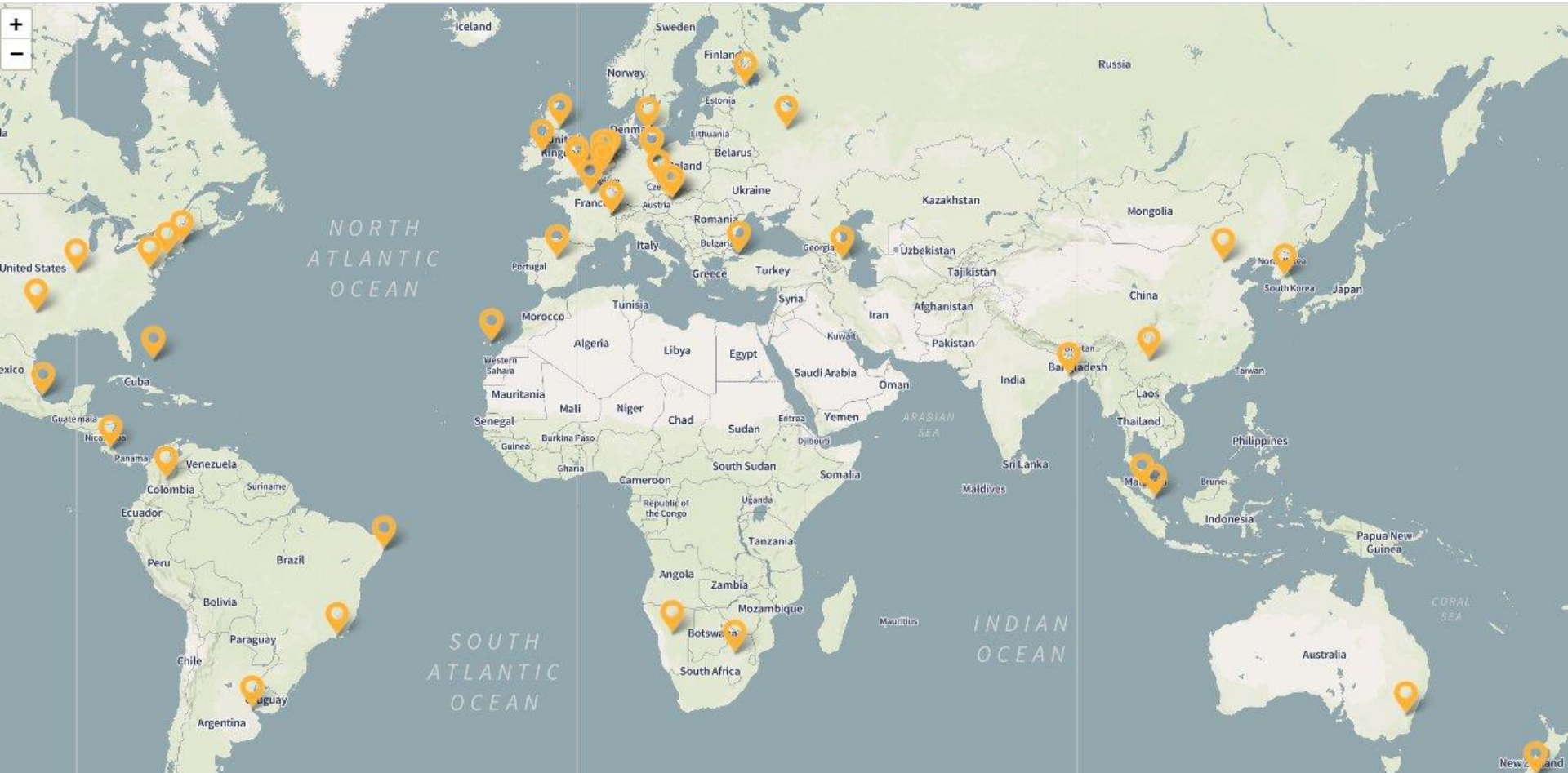




Un Consorcio Internacional

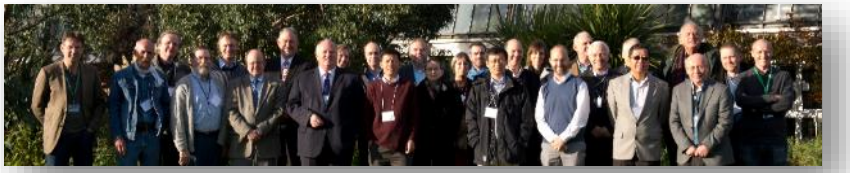
Memorandum de Entendimiento abierto para la firma en Enero 2013

47 firmantes al 2021…



Un Consorcio Internacional

- **1er Consejo WFO**
  Edinburgo
  10-14 Noviembre 2013



- **2do Consejo WFO**
  San Petersburgo
  23-27 Junio 2014



- **3er Consejo WFO**
  Ginebra
  25-29 Enero 2015



Un Consorcio Internacional

- **4to Consejo WFO**
  Rio de Janeiro
  19-24 Octubre 2015
- **5to Consejo WFO**
  New York
  25-30 Abril 2016
- **6to Consejo WFO**
  Pretoria
  8-11 Noviembre 2016



Un Consorcio Internacional

- **7mo Consejo WFO**
  Berlín
  27-31 Marzo, 2017
- **8vo Consejo WFO**
  París
  13-17 Noviembre, 2017
- **9no Consejo WFO**
  Bogotá
  18-22 Junio, 2018







Un Consorcio Internacional

- **10mo Consejo WFO**
  Dublín
  4-7 Diciembre, 2018
- **11mo Consejo WFO**
  Istambul
  23-26 Abril, 2019
- **12mo Consejo WFO**
  St. Louis
  18-22 Noviembre, 2019







Un Consorcio Internacional



- **13vo Consejo WFO**
  -Virtual-
  1 Julio, 2020
- **14vo Consejo WFO**
  -Virtual-
  29 Setiembre, 2020
- **15vo Consejo WFO**
  -Virtual-
  17 Noviembre, 2020

Un Consorcio Internacional



- **16vo Consejo WFO**
  -Virtual-
  17 Marzo, 2021
- **17vo Consejo WFO**
  -Virtual-
  6 Julio, 2021
- ***18vo Consejo WFO***
  -¿Meise?-
  ¿2022?

¿Qué es WFO?

Un compendio en Web de acceso abierto

que incluya:

- Todas las plantas conocidas
  - Briófitas y Plantas Vasculares
- Todas las regiones
  - Países como la unidad básica
  - Incorporando Regiones menores y mayores
- Disponible libremente en Internet



¿Qué es WFO?

Un compendio en Web de acceso abierto

que incluya:

- Línea de base única, una autoridad, orientada al usuario, precisa y coordinada de todas las plantas del mundo
- Obtenga información de las Floras publicadas, checklists y revisiones taxonómicas y de otro tipo.



Algunos públicos meta

Consumidores

- **Conservacionistas**
  Especialmente aquellos que trabajan con las metas del GSPC y otras áreas del CDB
  - Científicos Conservacionistas
  - Planificadores de Conservación/ Formuladores de Políticas
  - Investigadores de especies invasoras
  - Investigadores en Usos Sustentables
  - Botánicos Economicos/Etnobotánicos
- **Taxónomos de Plantas**
- **Otros científicos**
  Ecologistas, Antropólogos, Arqueólogos, Farmacólogos
- **Grupos de interés general**
  Historiadores naturales, científicos ciudadanos, etc..

Contribuyentes

- **Proveedores de datos primarios**
- **Administradores de Información**
- **Curadores Taxonómicos**
- **Revisores taxonómicos expertos**
- **Administrador Técnico de Sistemas y Datos**

La Organización de WFO



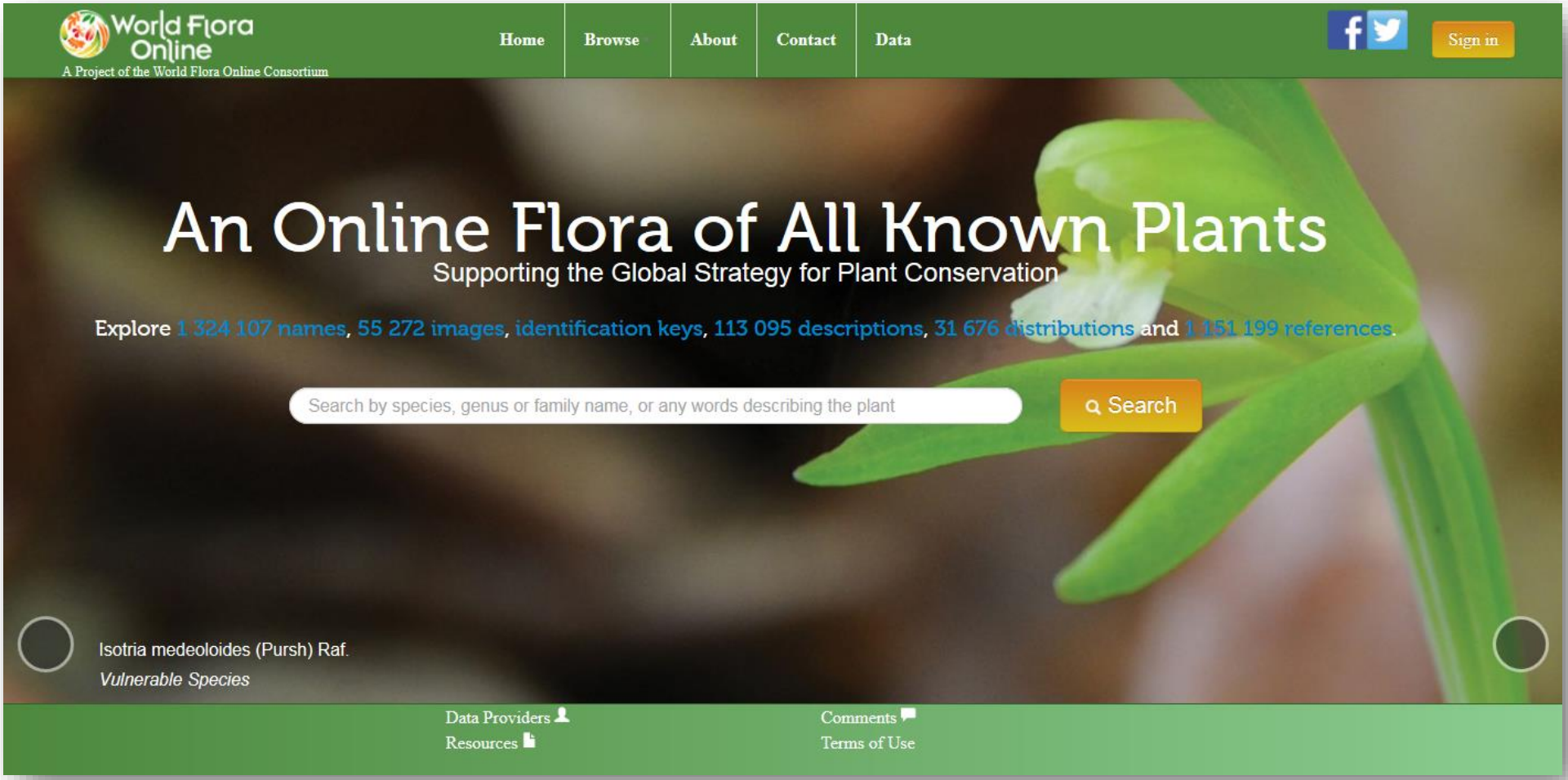
Portales Públicos

http://about.worldfloraonline.org/



Portales Públicos

http://www.worldfloraonline.org/



Interfaz de la Página de Taxones

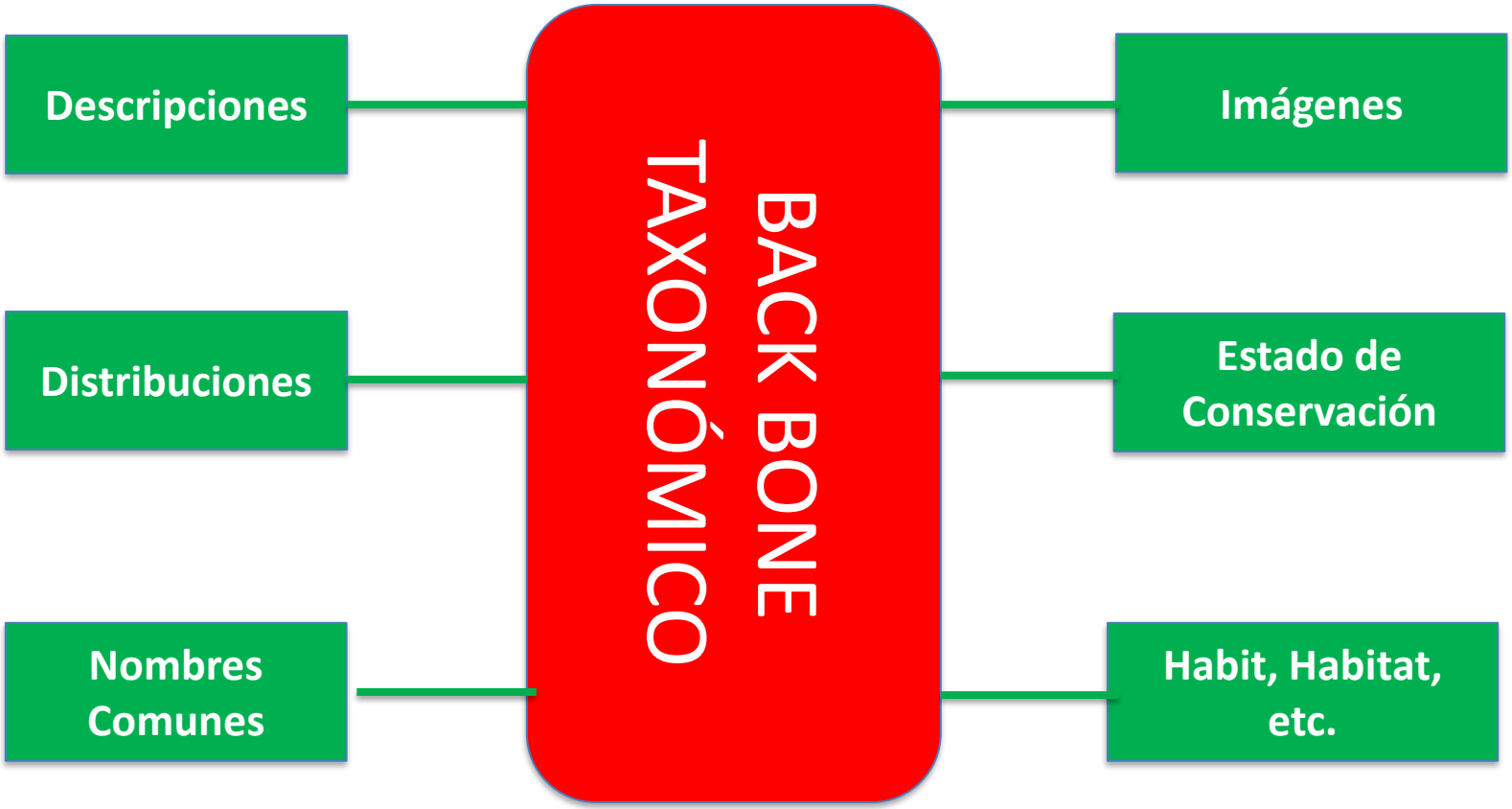


Dos Partes del Portal eMonocot / WFO

Nomenclatura & Taxonomía (Backbone)

Contenido

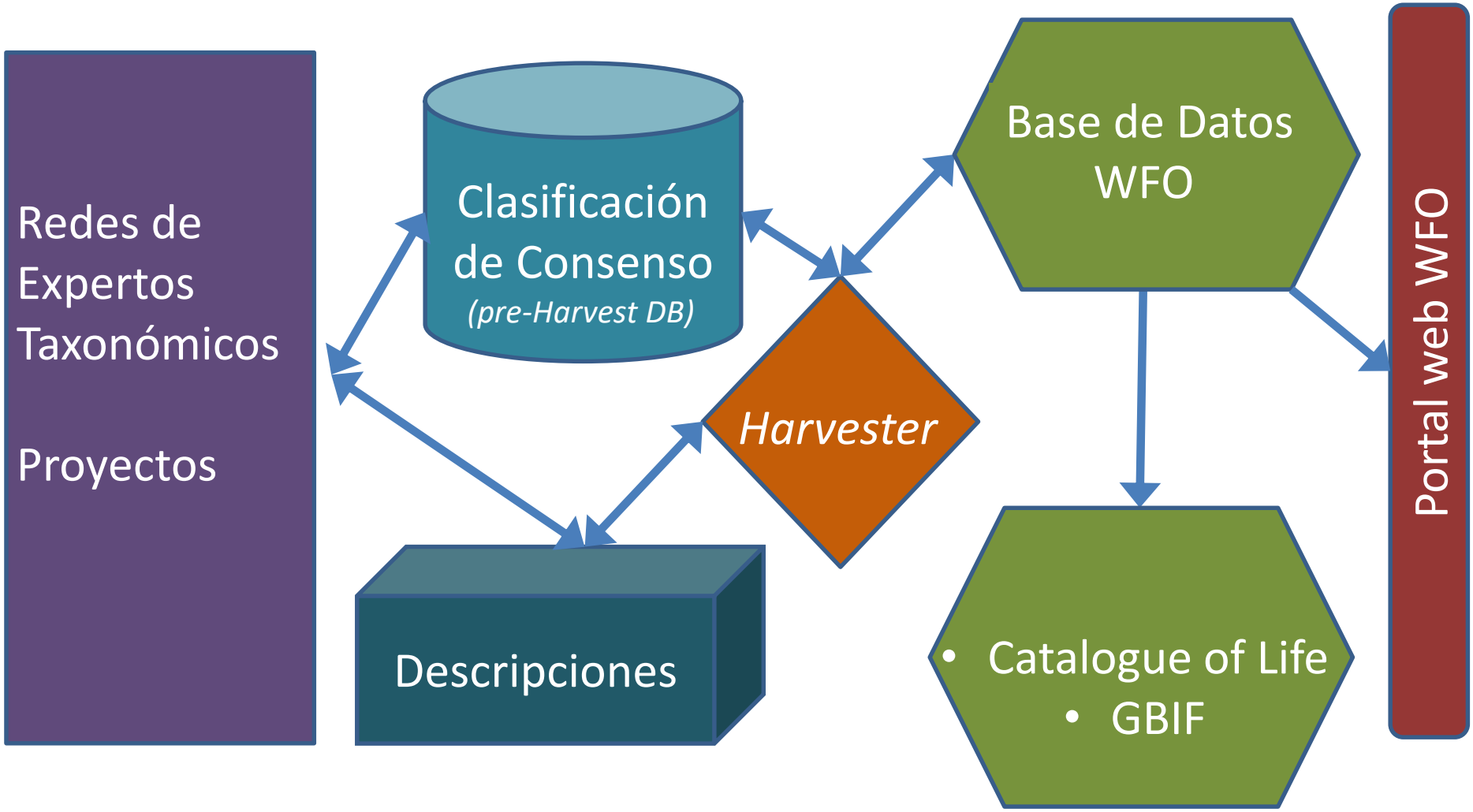
“Backbone” Taxonómico del WFO



“Backbone” Taxonómico del WFO

- Un checklist de consenso global
  - Un único concepto taxonómico
- Curado por la Comunidad Botánica
  - Redes de Expertos Taxonómicos
  - Especialistas
- Coordinado por el Consejo del WFO
  - Identificando brechas
  - Encontrando especialistas

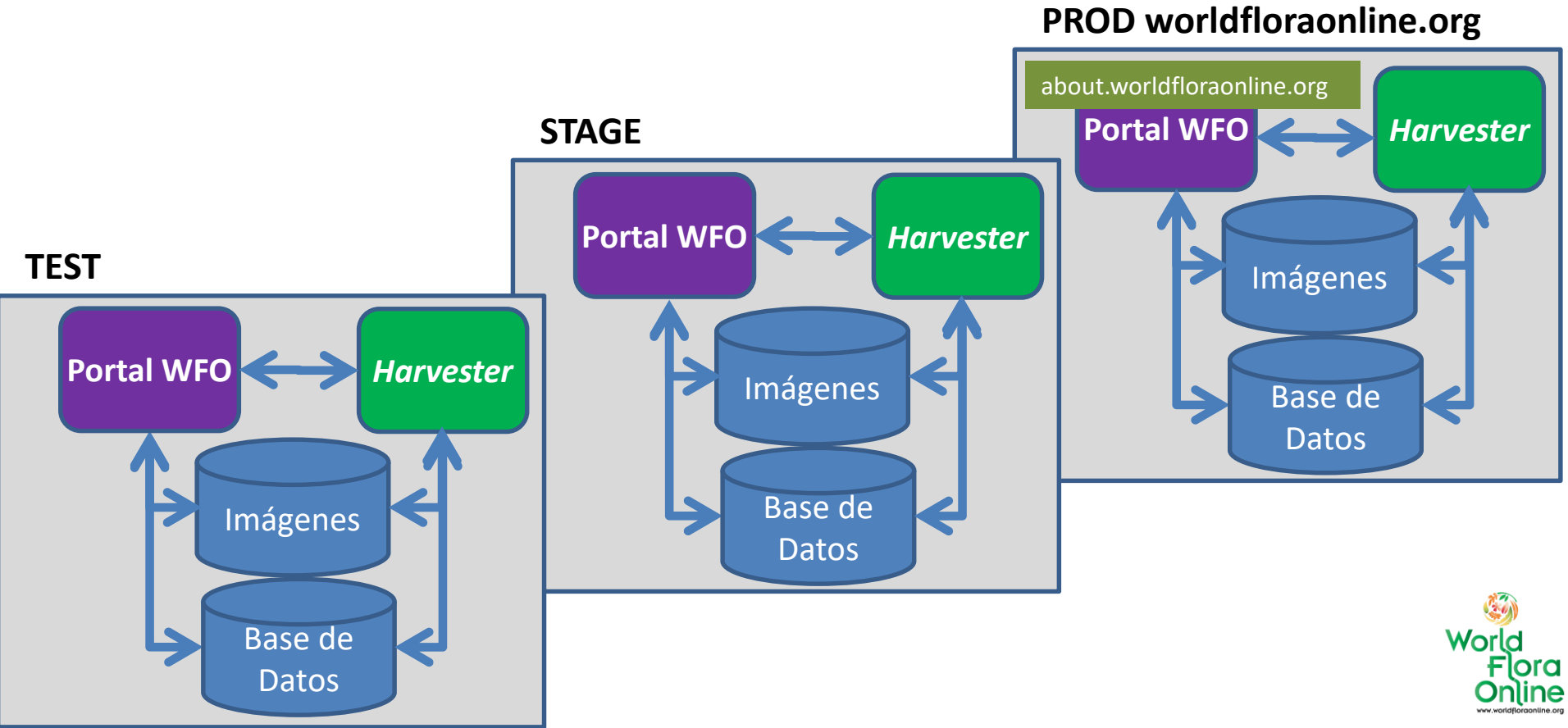
El Sistema de WFO



Infraestructura

Google Cloud Platform

Courtesía de NYBG



Un Proyecto Colaborativo

- Organización general, consejo, apoyo, etc. (Miembros del Proyecto)
- Plataforma eMonocot (Kew)
- Desarrollo del Portal Público basado en el de eMonocot (Missouri)
- Uso de la Nube de Google (New York)
- Trabajo en la Interfaz de Usuario y las herramientas de comunicación (Rio de Janeiro, Missouri y Edimburgo)

Un Proyecto Colaborativo

- Diseño gráfico general (Missouri y Edimburgo)
- Trabajo en la columna vertebral (backbone) taxonómica, TPL (Kew & Edimburgo)
- Administración Taxonómica del backbone taxonómico a través del módulo nomenclatural de Botalista (Ginebra)
- Lista de especialistas taxonómicos (all members)

Redes de Expertos Taxonómicos

List of Approved Taxonomic Expert Networks (TENs)

Bryophytes
- Bryophytes Group

Pteridophytes

Gymnosperms
- Conifers
- Cycads

Angiosperms
- Alooideae
- Annonaceae
- Apocynaceae
- Monographia Aquifoliacearum
- Arecaceae
- Begonia Resource Centre
- Boraginaceae (subtribe Amsinckiinae)
- Bruniaceae
- Cabombaceae
- Caryophyllales.org
- Commelinales
- Cordiaceae
- Dioscoreaceae
- Forestry Research Programme Dipterocarpaceae Database
- Ericaceae Resource Centre
- Fagaceae
- Gesneriaceae
- Global Conservation Consortium Acer (Subfamily Hippocastanoideae)
- Haloragaceae and related Families
- Hypericum
- Irvingiaceae
- Legume Phylogeny Working Group: Taxonomy
- Mayacaceae
- Meconopsis Group (Meconopsis & Cathcartia)
- Melastomataceae.net
- Musaceae
- Picramniaceae
- Poaceae
- Quiinoideae
- SolanaceaeSource.org
- Styracaceae
- Sapotaceae Resource Centre
- The International Compositae Alliance
- Zingiberaceae Resource Centre

Regional Endemic Family based TENs
- Endemic Plant Families of Madagascar

Identificador Único WFO-ID

wfo-999####### Dominium
wfo-997####### Kingdom
wfo-995####### SubKingdom
wfo-993####### Division
wfo-992####### SubDivision
wfo-987####### SuperClass
wfo-985####### Class
wfo-980####### Subclass
wfo-950####### SuperOrder
wfo-900####### Order
wfo-850####### Suborder
wfo-700####### Family

wfo-650####### Subfamily
wfo-550####### SuperTribe
wfo-500####### Tribe
wfo-450####### Subtribe
wfo-400####### Genus
wfo-350####### SUBGENUS
wfo-340####### Section
wfo-300####### Subsection
wfo-280####### Series
wfo-270####### Subseries
wfo-000####### Species
Subspecies
Variety
Subvariety
Form
Subform
InfraspecificName

Los identificadores WFO-ID son de la forma :

'WFO-<número único de 10-digitos>'

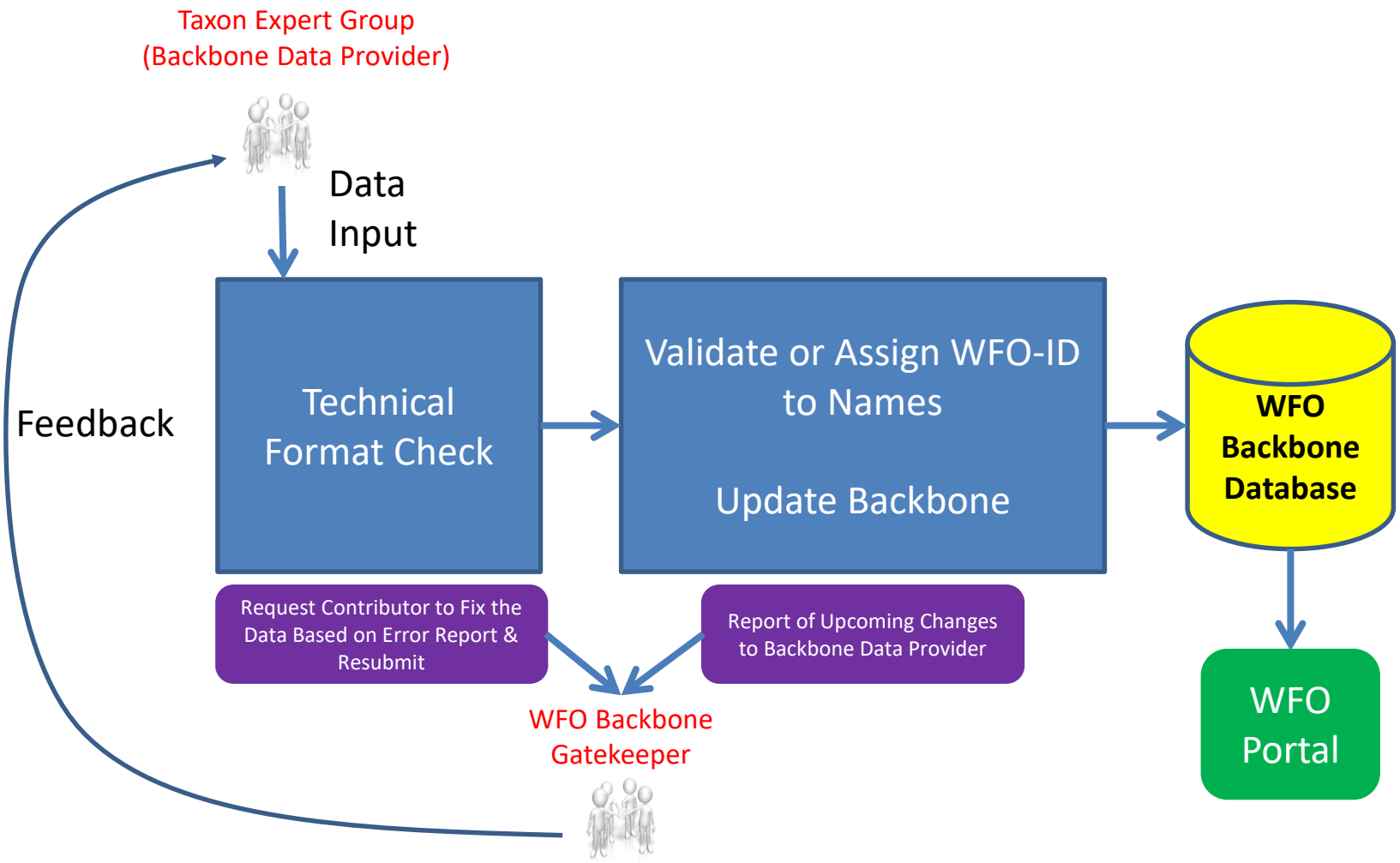
Nota: La numeración es enteramente arbitraria y no debe ser interpretada.

Asociar un Nombre con su WFO-ID

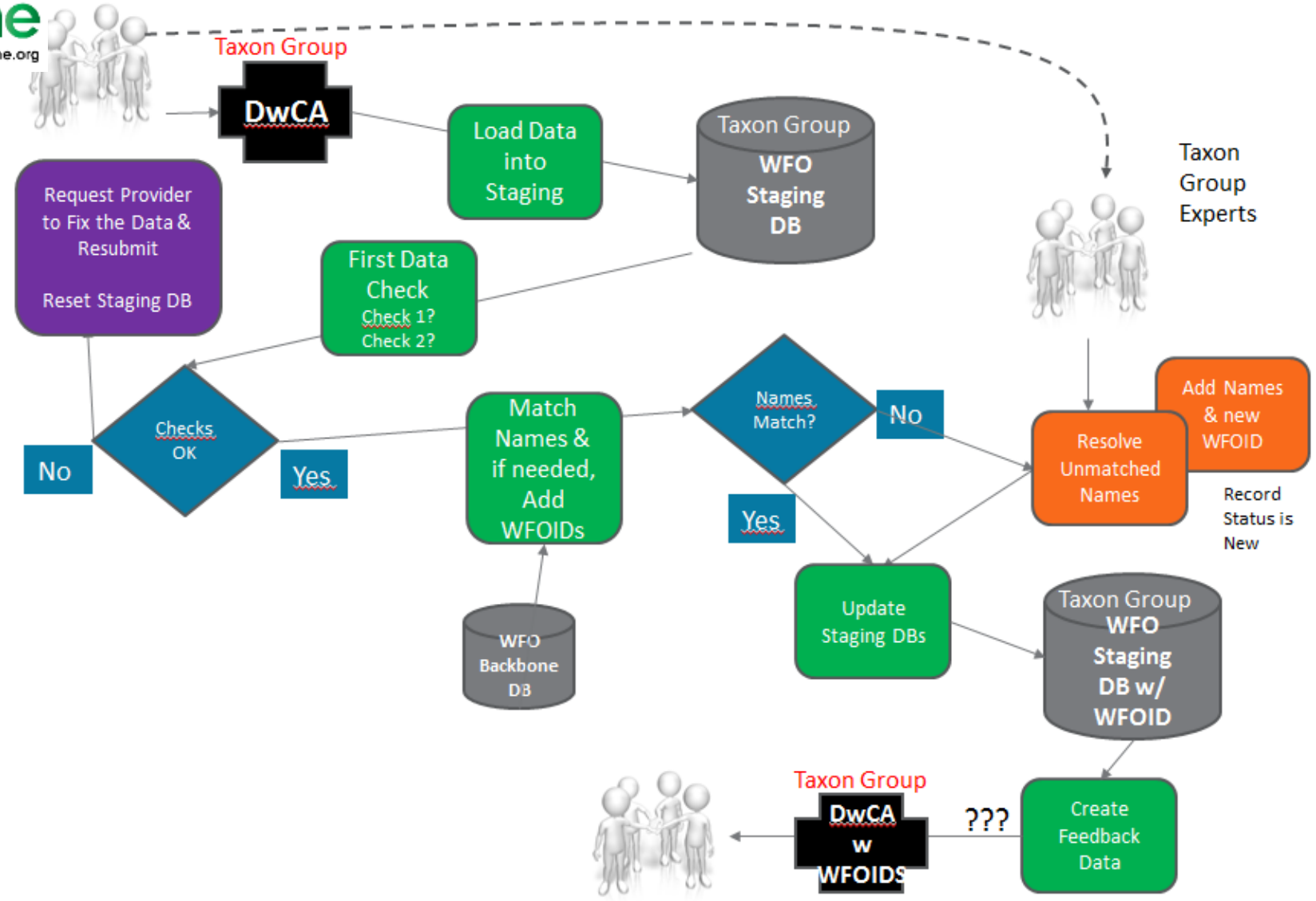
Principales pasos en el proceso de asociación de un Nombre con su WFO-ID

- Leer el archivo del usuario como insumo.
- Leer cada registro del archivo.
- Comparar los valores del registro contra el Backbone taxonómico para encontrar una asociación.
- Recuperar el WFO-ID del registro asociado.
- Generar un archivo de salida con los WFO-IDs asignados y comentarios explicativos, si no hubo asociación.

Diagrama de Flujo de Trabajo Simplificado



Proceso *Pre-Harvesting*



Need 2 Identifiers – Provider ID & WFO ID

Validador DwCA de GBIF

TOOL | 10 FEBRUARY 2015

GBIF data validator

By using the GBIF data validator, publishers can check their datasets for potential issues and preview the GBIF interpreted version before actually publishing the data

DATA VALIDATOR

Powered by Biodiversity Information Standards TDWG

What is the GBIF data validator?

The GBIF data validator is a service that allows anyone with a GBIF-relevant dataset to receive a report on the syntactical correctness and the validity of the content contained within the dataset. By submitting a dataset to the validator, you can go through the validation and interpretation procedures usually associated with publishing in GBIF and quickly determine potential issues in data - without having to publish it.

Best Practices - DwCA

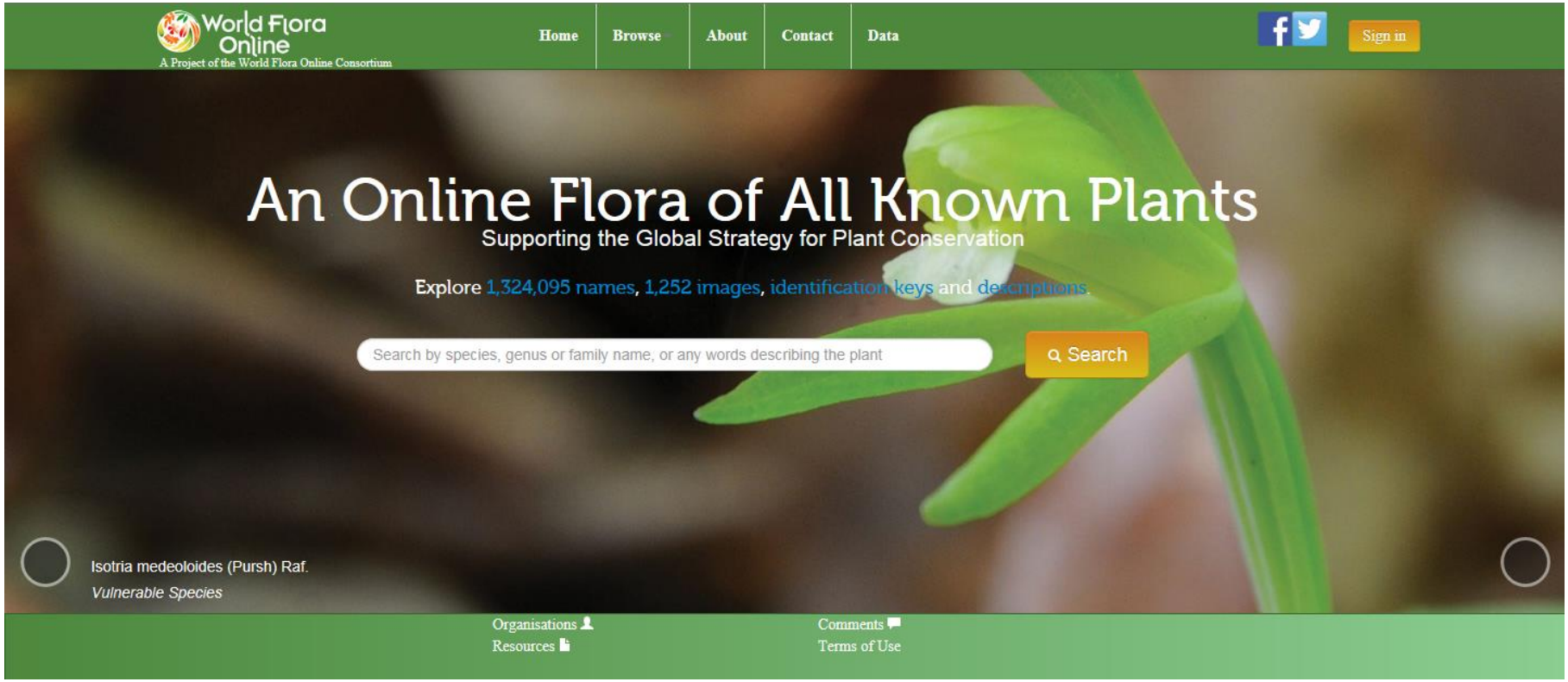
- Validate DwCA file (compressed zip) with GBIF's DwCA validator (it will catch the most basic errors)
- Check encoding, line ending and field separator
- Every taxon must have an ID and reference existing IDs (in the file or the backbone)
- In the EML FILE, LOCATION is relative path within the DwCA zip file.
- The correct format for dates is YYYYMMDDTHH:MM:SSZ: "20140516T13:03:00Z "
- Separate Description References with a comma in the source field.
- The text value in namePublishedIn can not be bigger than 255 characters long.

Biodiversity Information Standards TDWG

GBIF

Nueva interfaz de Usuario del Portal

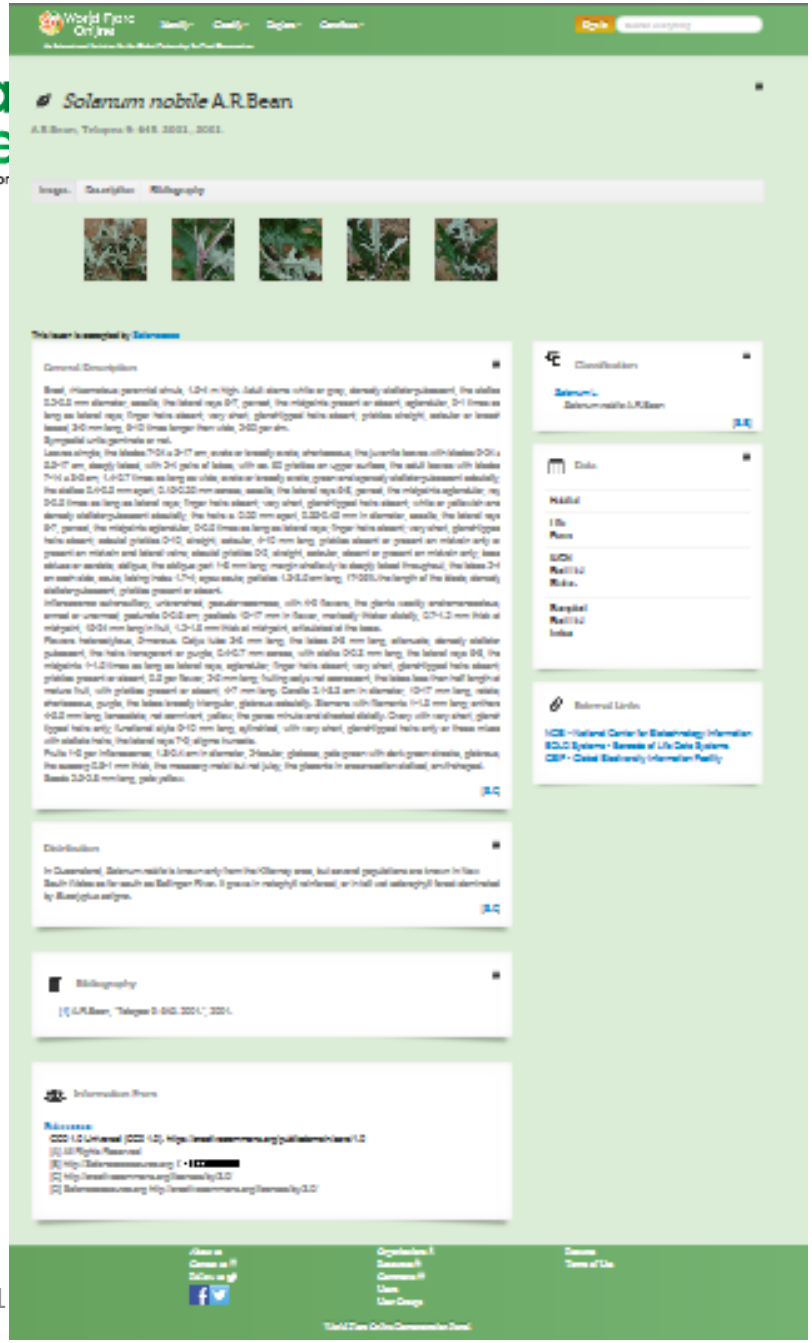
Congreso Internacional de Botánica , 2017

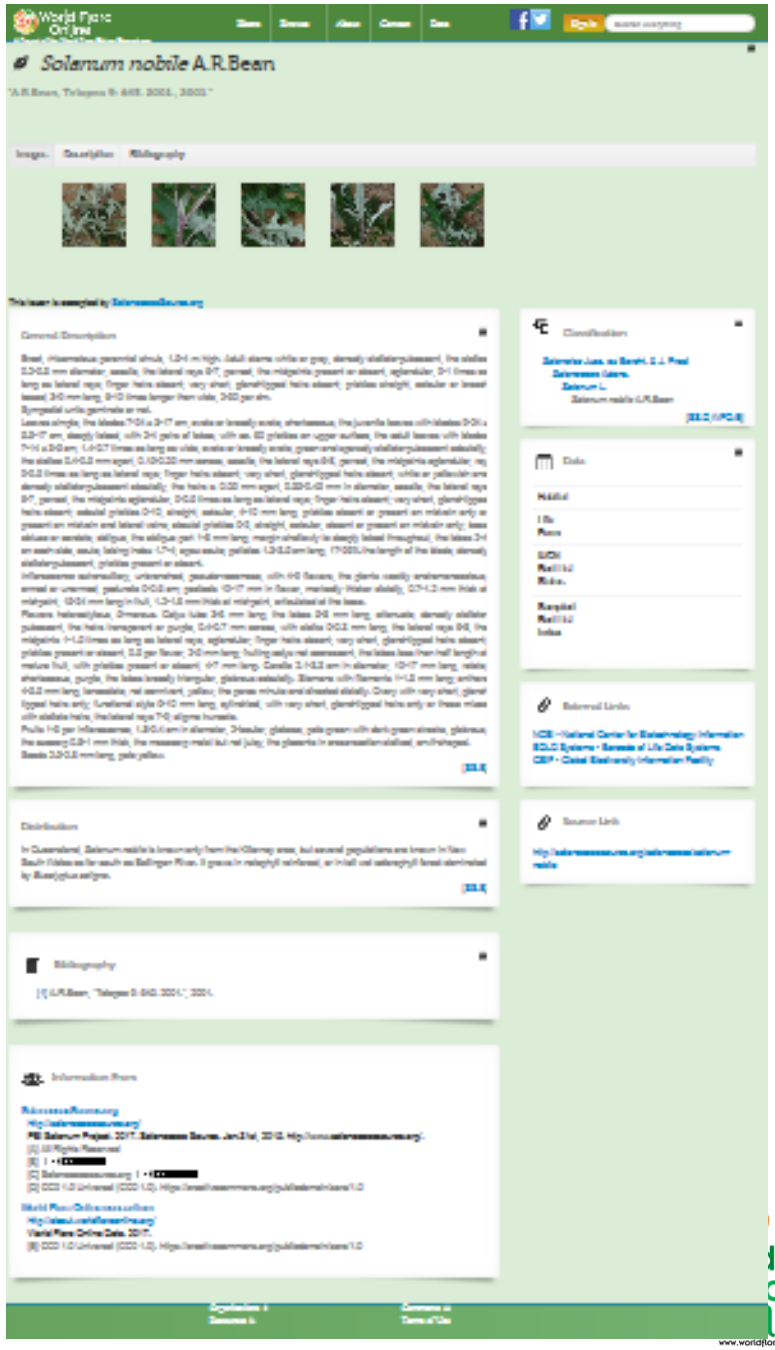


Portal de Demostración

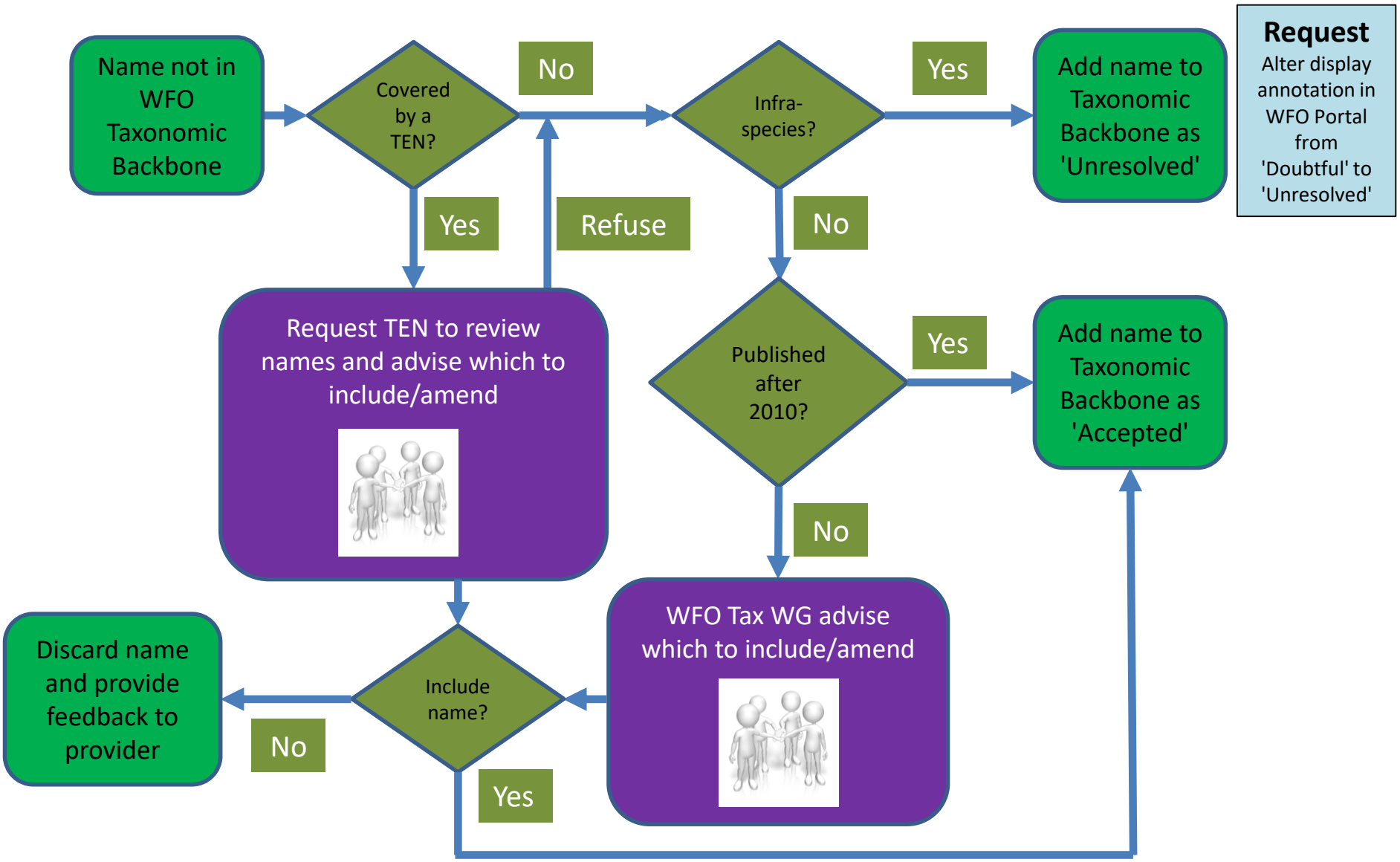
Portal Público

Página de Taxón

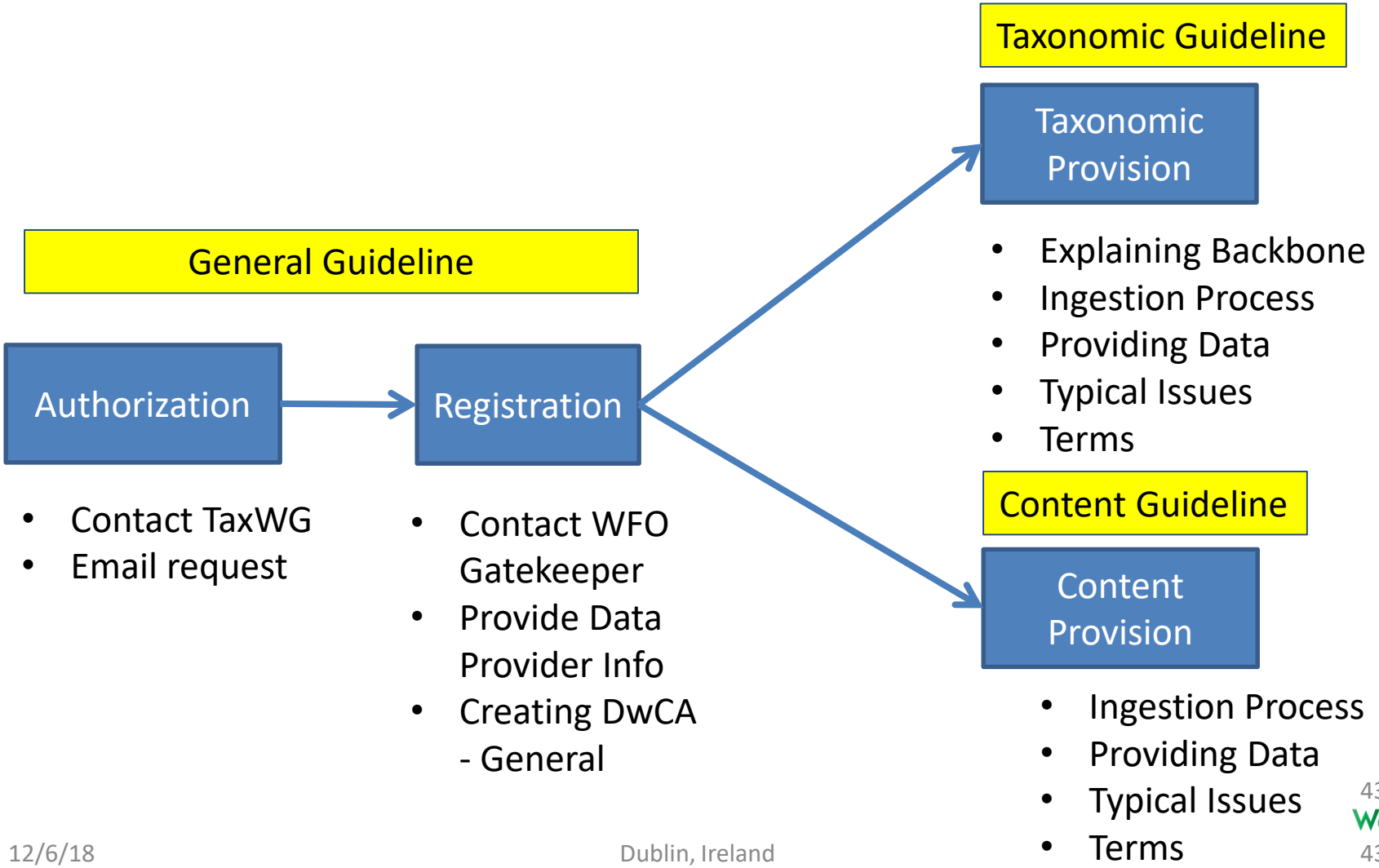




Flujo de Trabajo para manejar nuevos taxones de los Proveedores



Autorización & Registro



Revisiones en la BD *Pre-harvest*

- Synonymy, Family and MajorGroup assignment
  - Synonym of synonym: 24,888
  - Valid name with synonym as parent: 1,615
  - Excluded name as parent: 284
  - Synonym with excluded parentName: 1,912



Revisiones en la BD *Pre-harvest*

- Synonymy, Family and MajorGroup assignment
  - Valid names with no parent: 72
  - Synonyms with no accepted name: 258
  - Family mismatch of valid taxon and parent: 284
  - Same family, different Major Group: 4,759
  - Orders & families, no parent or Major Group: 319



Ejemplos de modificaciones al Backbone Taxonómico de WFO durante el 2021

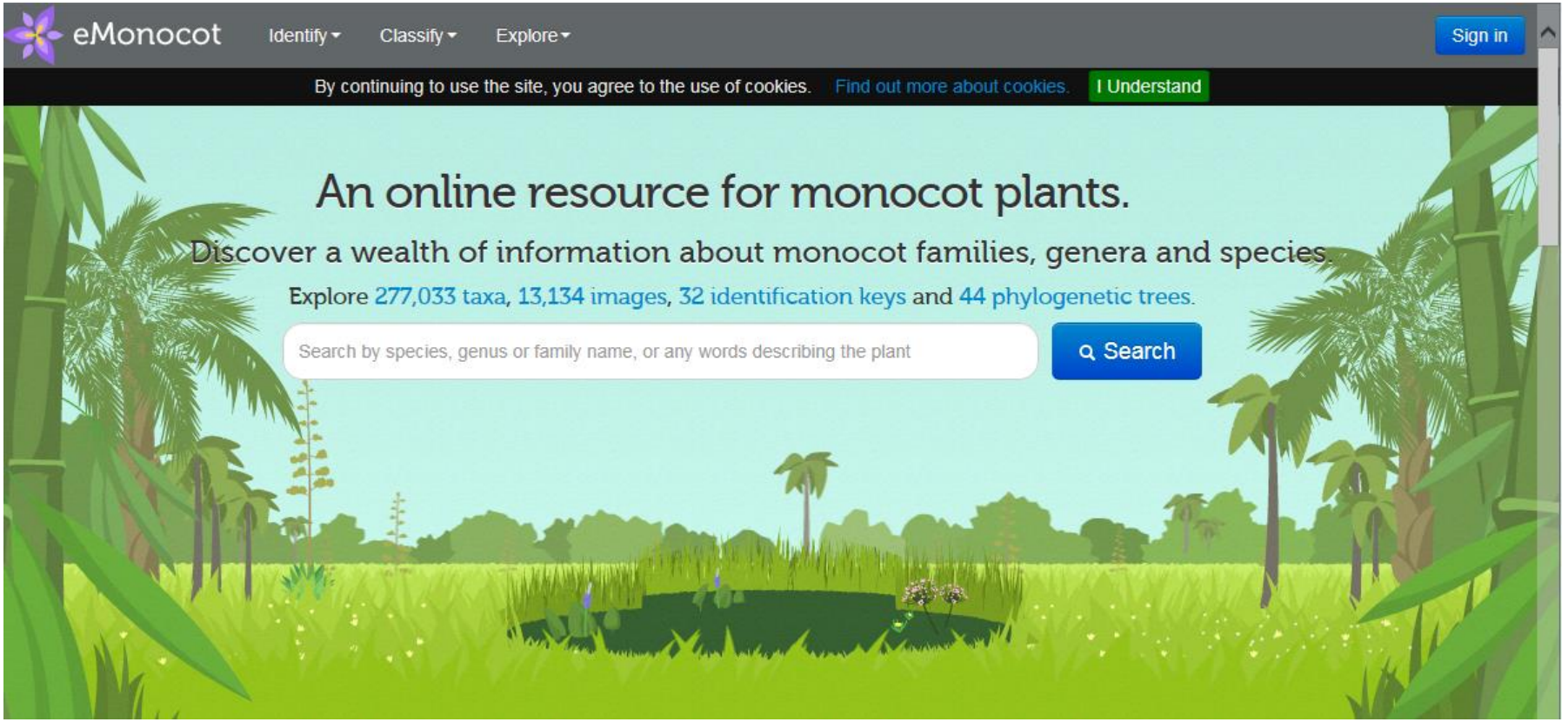
- ✓ Se corrigió una inconsistencia en Pterydophytes en los *Grupos principales*
- ✓ Irvingiaceae, se incluyeron distribuciones y el URL fuente (*taxonsource*)
- ✓ Se identificaron nombres nuevos de *Flora of Thailand*
- ✓ Se asignaron IDs a la lista de Zingiberaceae
- ✓ Clasificación de Meconopsis y Cathcartia
- ✓ SolanaceaeSource actualizado

Retos informáticos

- Manejo de información inconsistente
- Fuentes dispersas y en diferentes formatos
- Falta de Presupuesto dedicado
- Herramienta del Portal sin documentación ni soporte
- Proyecto no incluye desarrollo de la herramienta del Portal o el *Harvester*
- Poco personal técnico y en gran parte *“voluntario”*

World Flora Online

Portal basado en el Sistema de eMonocot



Kew ROYAL BOTANIC GARDENS · UNIVERSITY OF OXFORD · NATURAL HISTORY MUSEUM · NATURAL ENVIRONMENT RESEARCH COUNCIL

Nomenclatural Status

	String Value	Nomenclatural Status Code
1.	"abortivum"	Nomen_Abortivum
2.	"alternativum"	Alternative
3.	"ambigua"	Ambiguous
4.	"available"	Available
5.	"combinatio"	New_Combination
6.	"confusum"	Confused
7.	"conservandum"	Conserved
8.	"conservandumProp"	Conserved_Proposed
9.	"correctum"	<NULL>
10.	"dubimum"	Doubtful
11.	"illegitimum":	Illegitimate
12.	"invalidum"	Invalid
13.	"legitimate"	Legitimate
14.	"negatum"	Denied
15.	"novum"	Novum
16.	"nudum"	Nudum

Nomenclatural Status

FAILED

- unknown
- Unknown
- UNKNOWN
- ILLEGITIMATE
- INVALID
- CONSERVED
- illegitimate
- conserved
- nom. cons.
- nomen invalidum
- Orth. Cons. Prop.
- established
- nomen conservandum

TRY

- unknown
- Unknown
- UNKNOWN
- Illegitimate
- Invalid
- Conserved
- Illegitimate
- conserved
- nom. cons.
- nomen invalidum
- Orth. Cons. Prop.
- Established
- nomen conservandum

Taxonomic Status

- Is it case sensitive? No.

import org.gbif.ecat.voc.TaxonomicStatus;

1.	ACCEPTED	**"accepted"**
2.	DOUBTFUL	**"doubtful"**
3.	SYNONYM	**"synonym"**
4.	HETEROTYPIC_SYNONYM	**"heterotypicSynonym"**
5.	HOMOTYPIC_SYNONYM	**"homotypicSynonym"**
6.	PROPARTE_SYNONYM	**"proParteSynonym"**
7.	MISAPPLIED	**"misapplied"**
~~8.~~	~~INTERMEDIATE_RANK_SYNONYM~~ **@Deprecated**	~~**"intermediateRankSynonym"**~~
~~9.~~	~~DETERMINATION_SYNONYM~~ **@Deprecated**	~~**"determinationSynonym"**~~
	<default>	**"synonym"**

https://github.com/gbif/gbif-api/blob/master/src/main/java/org/gbif/api/vocabulary/TaxonomicStatus.java

Creative Common Licenses (1/2)

License type	Valid String
 PUBLIC_DOMAIN	**“http://creativecommons.org/licenses/publicdomain”**
 ATTRIBUTION	**“http://creativecommons.org/licenses/by/3.0”**
 ATTRIBUTION_NONCOMMERCIAL	**"http://creativecommons.org/licenses/by-nc/3.0"**
 ATTRIBUTION_SHAREALIKE	**"http://creativecommons.org/licenses/by-sa/3.0"**

Creative Common Licenses (2/2)

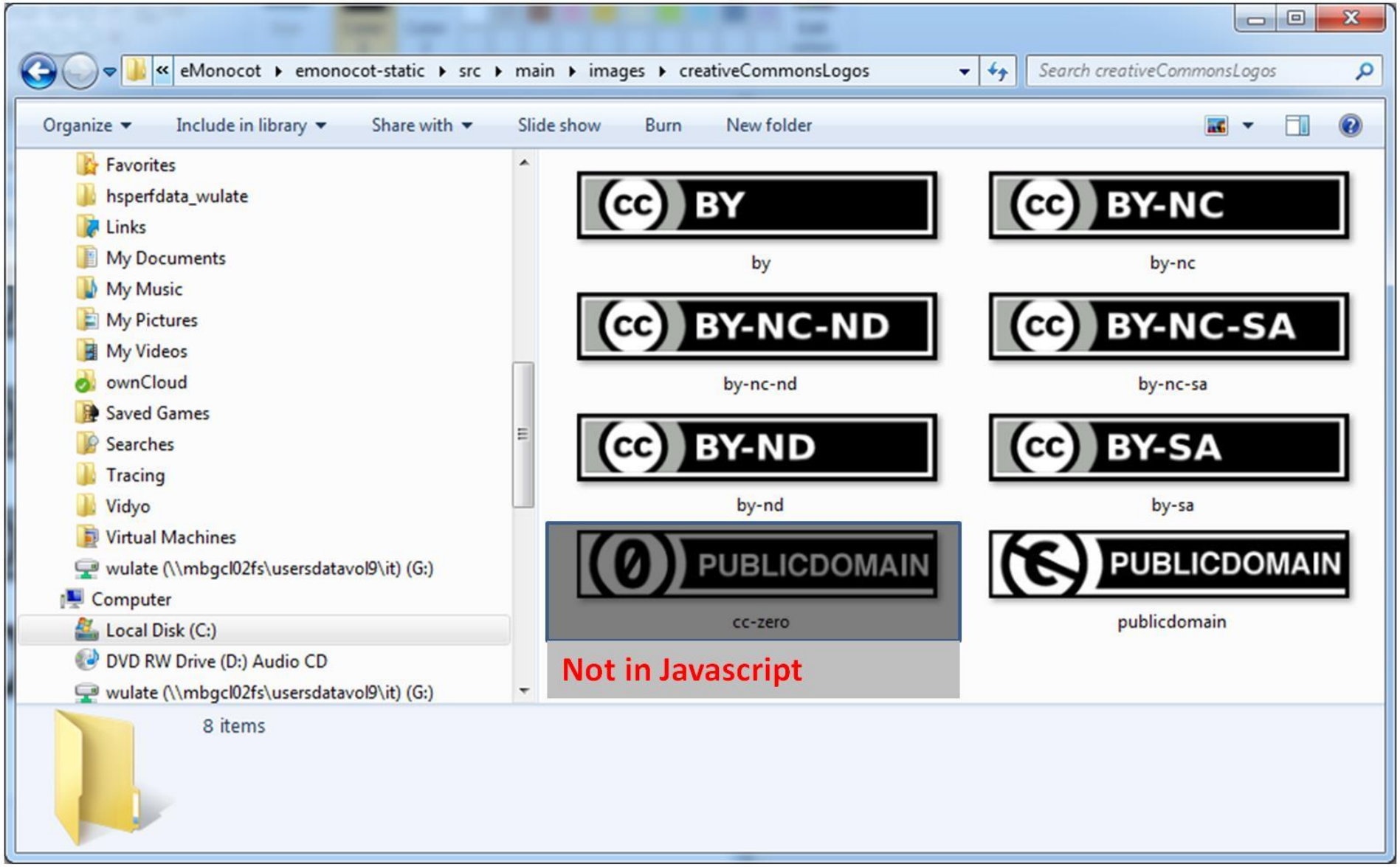
License type	Valid String
 ATTRIBUTION_NONCOMMERCIAL_SHAREALIKE	"http://creativecommons.org/licenses/by-nc-sa/3.0"
 ATTRIBUTION_NONDERIVATIVE	"http://creativecommons.org/licenses/by-nd/3.0"
 ATTRIBUTION_NONCOMMERCIAL_NONDERIVATIVE	"http://creativecommons.org/licenses/by-nc-nd/3.0"
 CC_ZERO	"http://creativecommons.org/publicdomain/zero/1.0"

Not in License.JAVA

Not converted as a CC License

- CC0
- http://creativecommons.org/choose/mark
- http://creativecommons.org/licenses/by/1.0
- http://creativecommons.org/licenses/by/2.0
- http://creativecommons.org/licenses/by/4.0
- http://creativecommons.org/licenses/by-sa/1.0
- http://creativecommons.org/licenses/by-sa/2.0
- http://creativecommons.org/licenses/by-sa/4.0
- http://creativecommons.org/licenses/by-nc/1.0
- http://creativecommons.org/licenses/by-nc/2.0
- http://creativecommons.org/licenses/by-nc/4.0
- http://creativecommons.org/licenses/by-nd/1.0
- http://creativecommons.org/licenses/by-nd/2.0
- http://creativecommons.org/licenses/by-nd/4.0
- http://creativecommons.org/licenses/by-nc-sa/1.0
- http://creativecommons.org/licenses/by-nc-sa/2.0
- http://creativecommons.org/licenses/by-nc-sa/4.0
- http://creativecommons.org/licenses/by-nc-nd/1.0
- http://creativecommons.org/licenses/by-nc-nd/1.0
- http://creativecommons.org/licenses/by-nc-nd/4.0
- https://creativecommons.org/licenses/publicdomain
- https://creativecommons.org/publicdomain/zero/1.0

Creative Commons Logos



Lessons learned - Portal

- The Organization identifier will be the Authority ID.
- The Acronym for the reference hyperlinks is calculated (on the fly) from the Capital letters of the Organization title.
- When defining Organizations, separate Keywords with a semi-colon
- Language is not displayed (nor used) in the Portal

Complejidad del Sistema eMonocot



MySQL Database
Ver 14.14
Distrib 5.5.62



Apache Tomcat
JavaServer Pages
web server environment
Ver. 7.0.52 (2014, deprecated)
Solr 4.0 (2014, deprecated)



Bootstrap
CSS (Cascading Style Sheets)
front-end framework
Ver. 2.3.1 (2013, deprecated)



User Interface



Linux OS
Canonical, Ubuntu 14.04 LTS
(2014, deprecated 2019)



Java
Open JDK Runtime Environment 1.7
IcedTea 2.6.17 (2007, deprecated 2016)



GeoServer
Version 2.7.6
(2016, maintenance release)
GeoTools Version13.6
GeoWebCache Version1.7.4

Complex legacy code base

- Outdated, unsupported technologies
- Programmers need to learn/relearn old coding languages
- Dependencies/software libraries no longer available – can't rebuilt from scratch without reprogramming

Casos de Uso del Portal actual

	WFO USE CASE IMPORTANCE ASSESSMENTS		
	First Phase		
	Later Phases		
	Omit		
	Refine use case definition, e.g. depending on role of Expert Networks or		
	Use means that it occurs at all. Even once.		
	USES (By type of user)	NOT INCLUDED IN CURRENT PORTAL	COMMENTS/NOTES
	A minimum set of Use Cases common to all kinds of consumer would be, I would like to...		
1	have free, open web-based access to the WFO with multiple browsers		
2	be able to print and/or download data, regardless of size	X	
	search for plant species according to one or both of:		
4	scientific plant name		
5	country (region, i.e. including range)	X	
6	be advised on a single, accepted scientific name for a species (with caveats)		[wb: depends on the record source - expert network 1, tax. backbone 4]
7	be advised on synonyms of the accepted name (with caveats)		
8	be advised on the scientific name, and synonyms, submitted by a data provider as part of a treatment	Only in a description	Not a list of all accepted/synonyms from one data provider
9	see available descriptive information of a plant (text)		
10	see available images of a plant (photos, drawings, etc.)		
11	find available identification tools (keys) to the plants in a country/region	X	
12	be advised on the most appropriate description where multiple descriptions are available (but we need to decide on what is appropriate,	X	"Most appropriate" according to who/what? [wb: "most appropriate ref. to
13	see identification tools associated with a plant (keys)	X	Access to diagnostic characters
14	be advised on the countries for which this plant is recorded within WFO (list, with caveats on differences in taxon circumscription)	Only if data is provided	[wb: if based on integrated treatment: 1, if only name-based: 4]

Portal Público



Portal Público



Report of the Taxonomic Working Group

Istanbul, 23-24 April 2019

Taxonomic Backbone

Kew WCSP/IPNI Update Project

IPNI is being updated to include new names and data from WCSP. This project will be completed by September

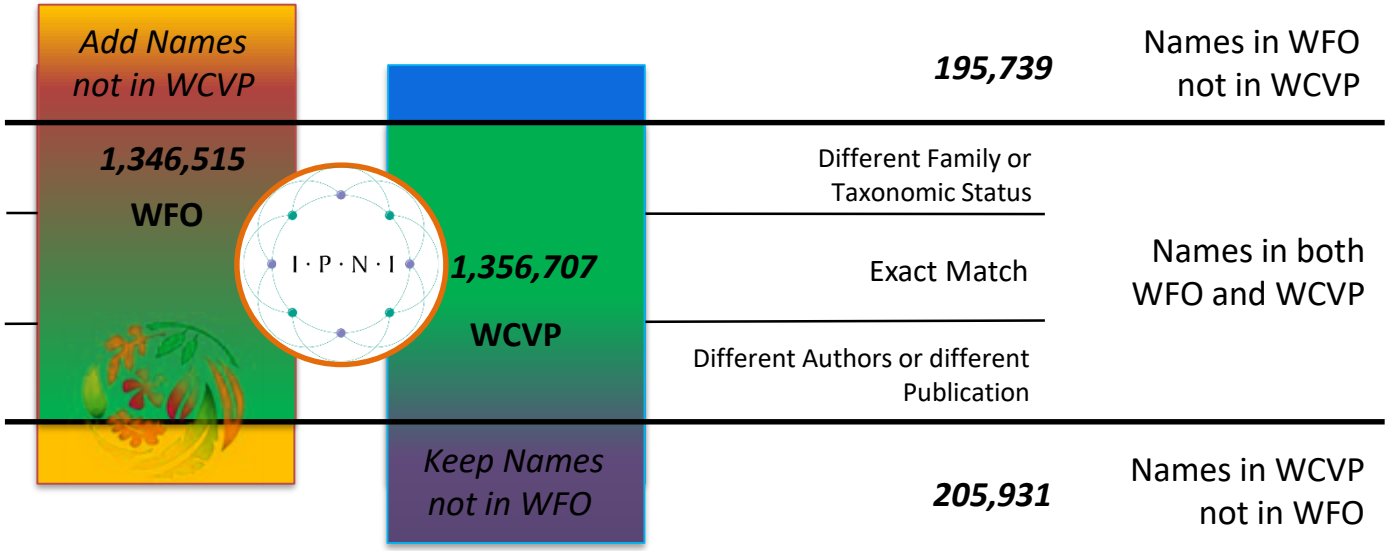
IPNI-WFO Name Matching

- 68% of names in WFO matched those in IPNI enabling assignments of IPNI-IDs to WFO Backbone names
- This should be re-run after the WCSP/IPNI Update to improve the number of matching names and an analysis of the results undertaken
- Protologue information should be updated for matched names using IPNI data – there was a problem with transfer of complete protologue info from the original TPL import

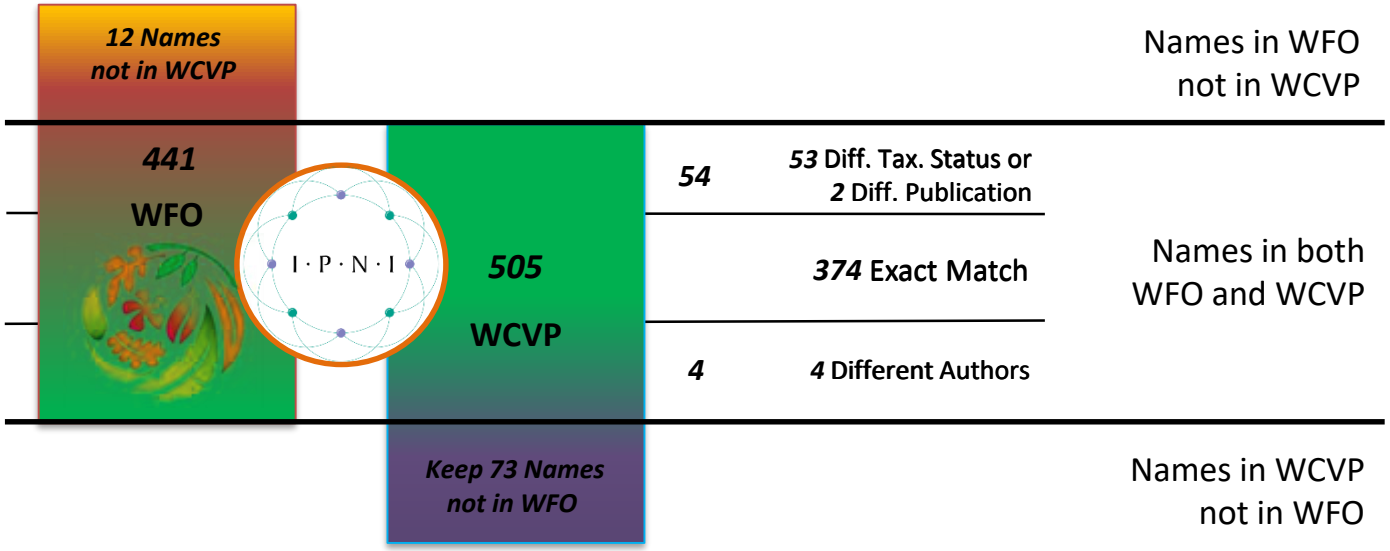
WCSP Update

WCSP to be used to update non TEN groups, to reduce 250k Unchecked Names

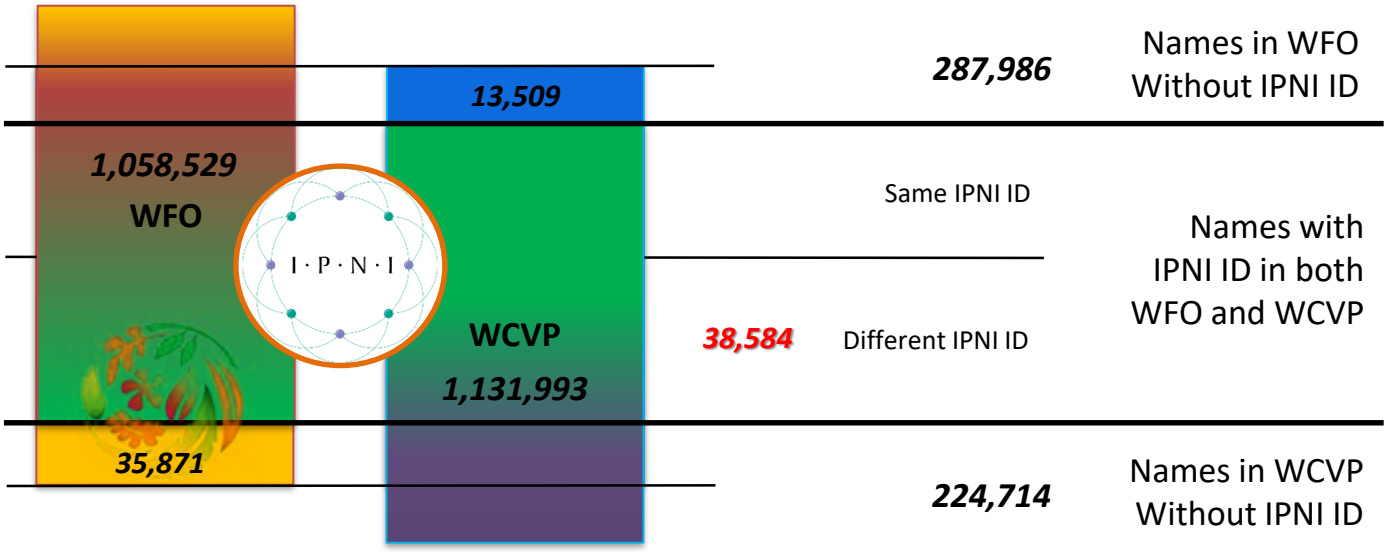
TEN: Add WFO to WCVP



Musaceae


12 Names not in WCVP
441
WFO
I · P · N · I
505
WCVP
Keep 73 Names not in WFO
54
53 Diff. Tax. Status or
2 Diff. Publication
374 Exact Match
4
4 Different Authors
Names in WFO not in WCVP
Names in both WFO and WCVP
Names in WCVP not in WFO

IPNI: Add IPNI ID to WFO



Decisions on Name and IPNI value for the TENs: Arecaceae, Musaceae, Styracaceae, Euphorbiaceae

Name		With IPNI ID in WFO			No IPNI ID in WFO	
		With IPNI ID in WCVP		No IPNI ID in WCVP	With IPNI ID in WCVP	No IPNI ID in WCVP
		With the Same IPNI IDs	With Different IPNI IDs			
In WFO not in WCVP		N/A	N/A	Include name from WFO with WFO IPNI ID	N/A	Include name from WFO (no IPNI ID)
In WFO and WCVP	Exact Match or Family not Matched or Status not Matched	Keep name from WCVP with WCVP IPNI ID	Keep name from WCVP with WCVP IPNI ID	Keep name from WCVP with WFO IPNI ID	Keep name from WCVP with its IPNI ID	Keep name from WCVP (no IPNI ID)
	Authors or Publication don't Match	Add to Review List (to confirm IPNI ID)	Add to Review List (to choose IPNI ID)	Add to Review List (to assign WFO IPNI ID)	Add to Review List (to assign WCVP IPNI ID)	Keep name from WCVP (no IPNI ID)
In WCVP not in WFO		N/A	N/A	N/A	Include name from WCVP with its IPNI ID	Include name from WCVP (no IPNI ID)

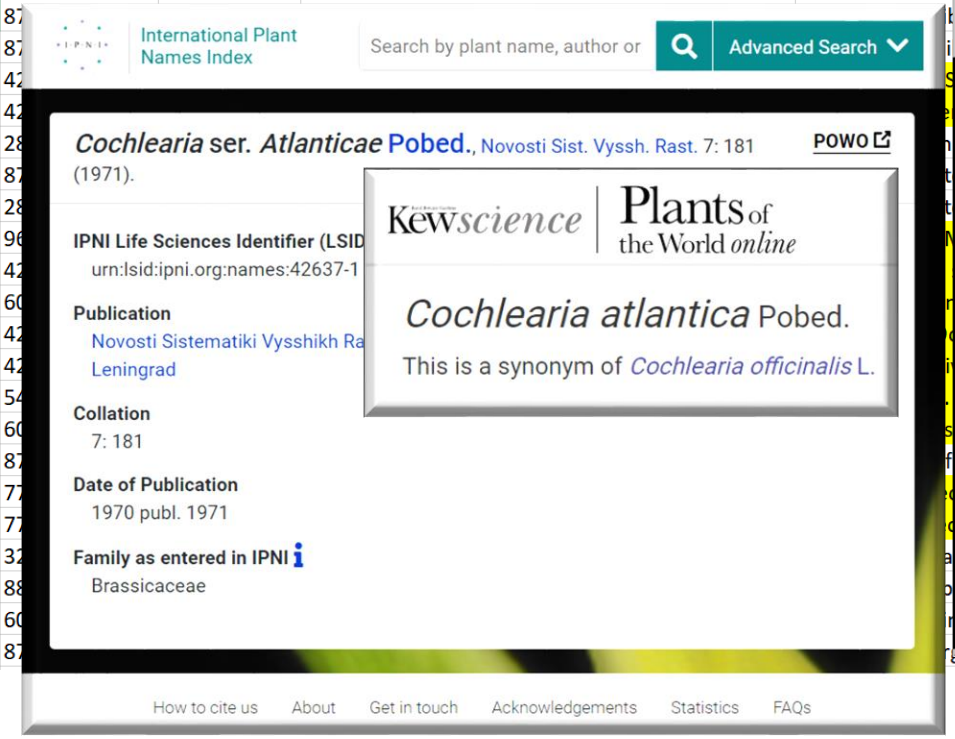
Updates in the Pre-Harvest DB IPNI for the TENs: Arecaceae, Musaceae, Styracaceae, Euphorbiaceae

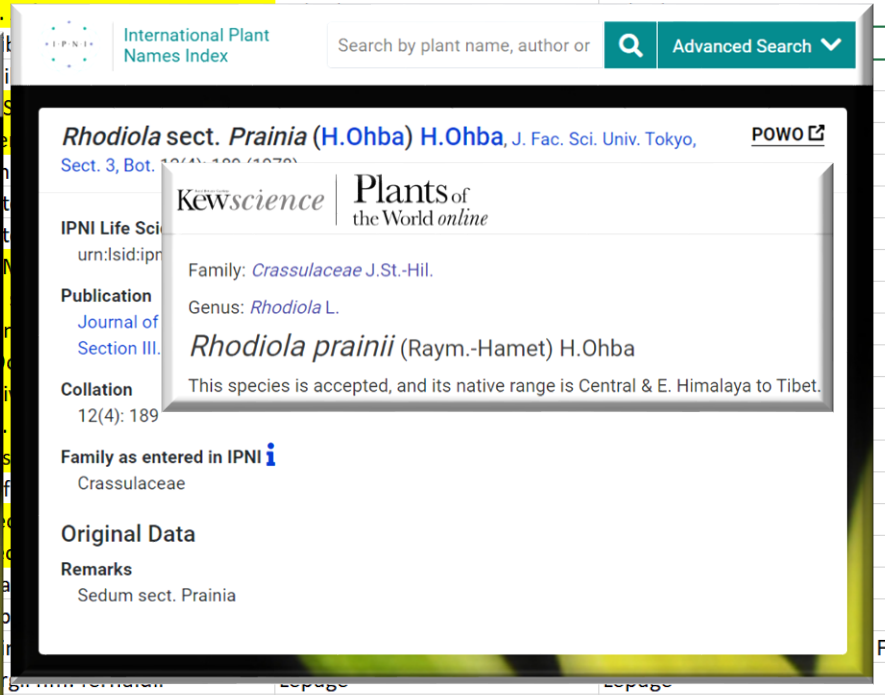
Name		With IPNI ID in WFO			No IPNI ID in WFO	
		With IPNI ID in WCVP		No IPNI ID in WCVP	With IPNI ID in WCVP	No IPNI ID in WCVP
		With the Same IPNI IDs	With Different IPNI IDs			
In WFO not in WCVP		N/A	N/A	Nothing to Do (leave WFO IPNI ID)	N/A	Nothing to Do (leave IPNI ID blank)
In WFO and WCVP	Exact Match or Family not Matched or Status not Matched	Nothing to Do (leave WFO IPNI ID)	Update to WCVP IPNI ID(*)		Update to WCVP IPNI ID	
	Authors or Publication don't Match	Add to Review List (to confirm IPNI ID)	Add to Review List (to choose IPNI ID)	Add to Review List (to assign WFO IPNI ID)	Add to Review List (to assign WCVP IPNI ID)	
In WCVP not in WFO		N/A	N/A	N/A	Update to WCVP IPNI ID	

IPNI & WCVP Matching

- WCVP names in WFO through IPNI

WCVP_IPNIID	WFO_IPNIID	WCVP_scientificName	WFO_scientificName	WCVP_Author	WFO_Author
42585-1	42585-1	Rhodiola prainii	Rhodiola sect. Prainia	(Raym.-Hamet) H.Ohba	(H.Ohba) H.Ohba
42637-1	42637-1	Cochlearia atlantica	Cochlearia ser.		





IPNI & WCVP Matching

IPNI Names matched with WFO

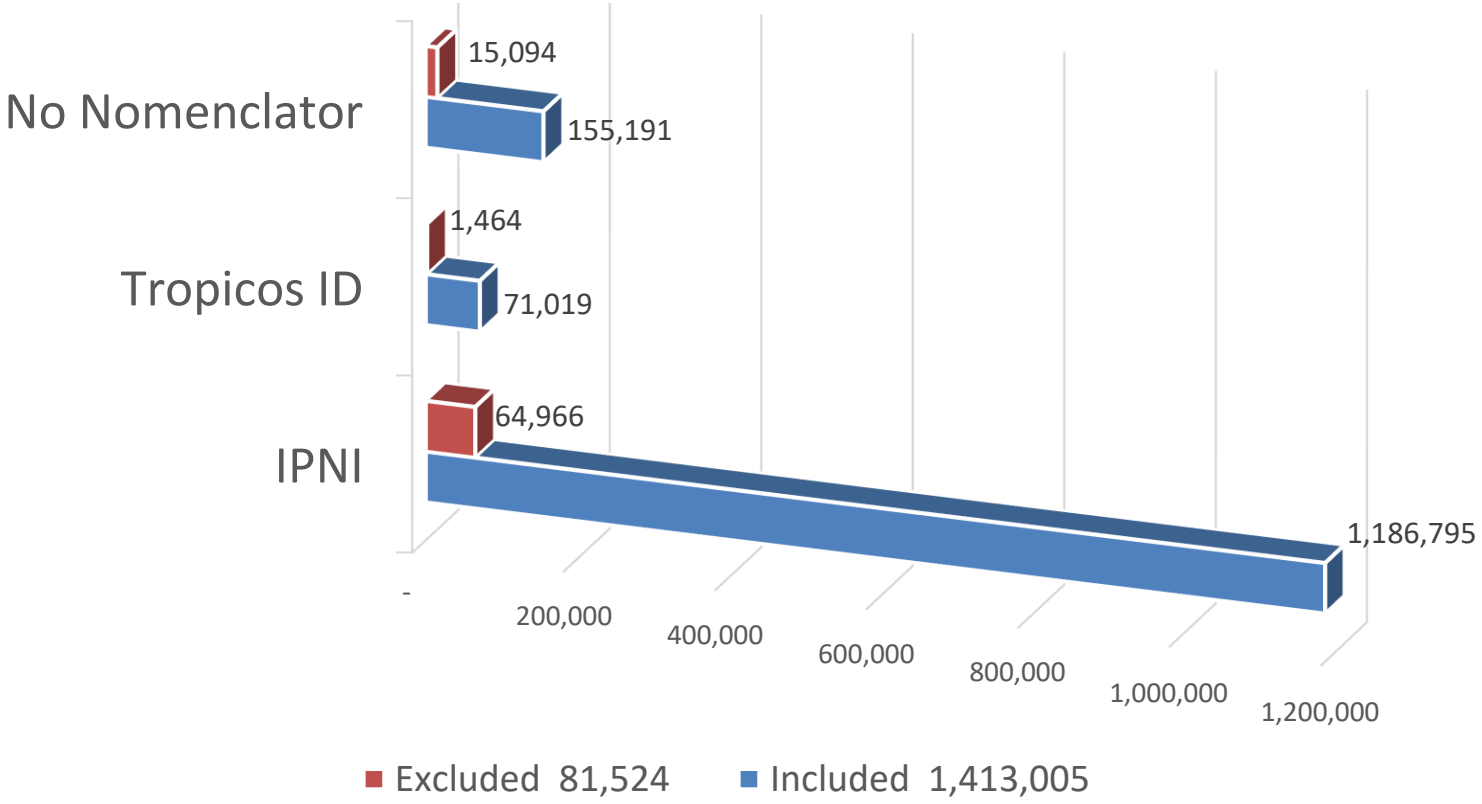
Total	Records Matched: with WFOIDs Assigned		Records Not Matched: (No WFOID)	Match %
	Included in WFO	Excluded from WFO		
1,608,732	1,311,631	7,426	289,675	**82%**

WCVP Names matched with WFO

Total	Records Matched: with WFOIDs Assigned		Records Not Matched: (No WFO-ID)	Match %
	Included in WFO	Excluded from WFO		
1,356,707	1,139,107	11,666	205,931	**84%**

IPNI and Tropicos Nomenclators

	WFO	IPNI	Tropicos ID	No Nomenclator
Included	1,413,005	1,186,795	71,019	155,191
Excluded	81,524	64,966	1,464	15,094
	1,494,529	**1,251,761**	**72,483**	**170,285**



Metaxya rostrata

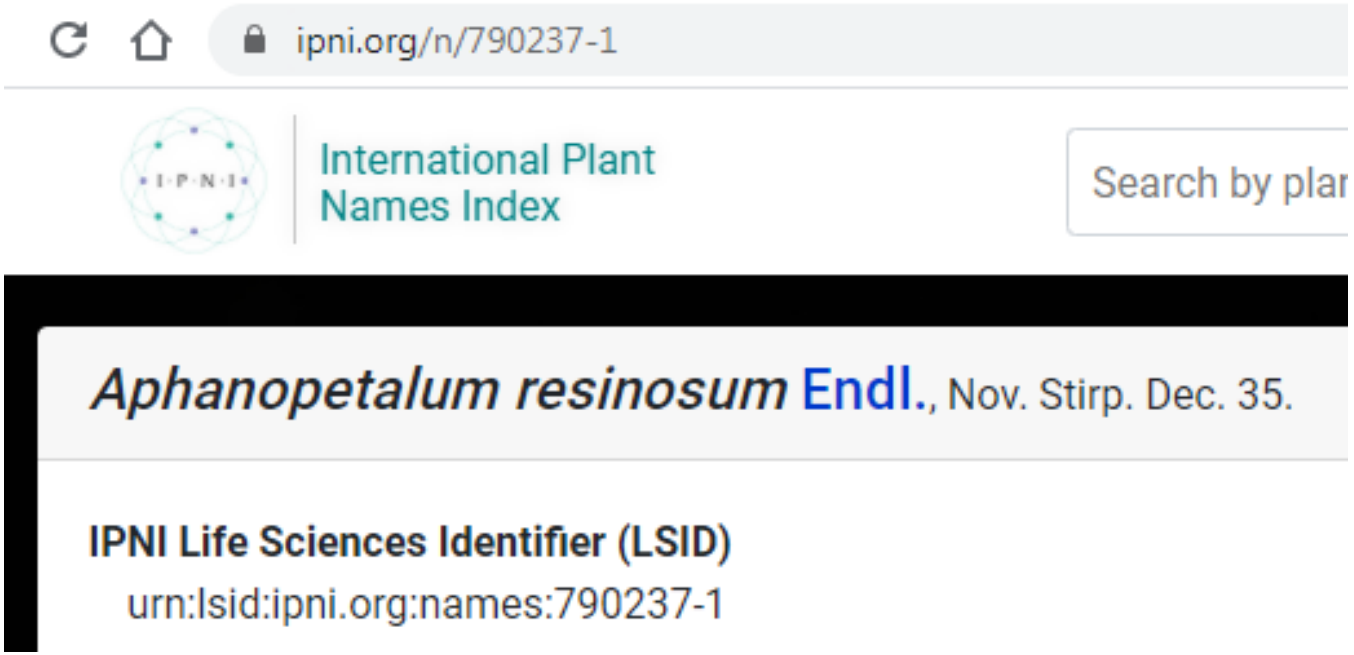
/ipni/17147720-1

/tpl/tro-26606568



Aphanopetalum resinosum

/ipni/790237-1 /tpl/kew-2643409



Nuevos Nombres agregados a WFO

- Nuevos Nombres contribuídos por Proveedores de Datos, tanto Taxonómicos, como de Contenido Descriptivo.

- Nuevos Nombres de Endémicas
  (**Madagascar**, **Brasil**, **Flore d'Afrique Centrale**, *South Africa, Colombia)*

Nuevos Nombres de Endémicas agregadas

por Rango Taxonómico y Fuente a Junio 2020

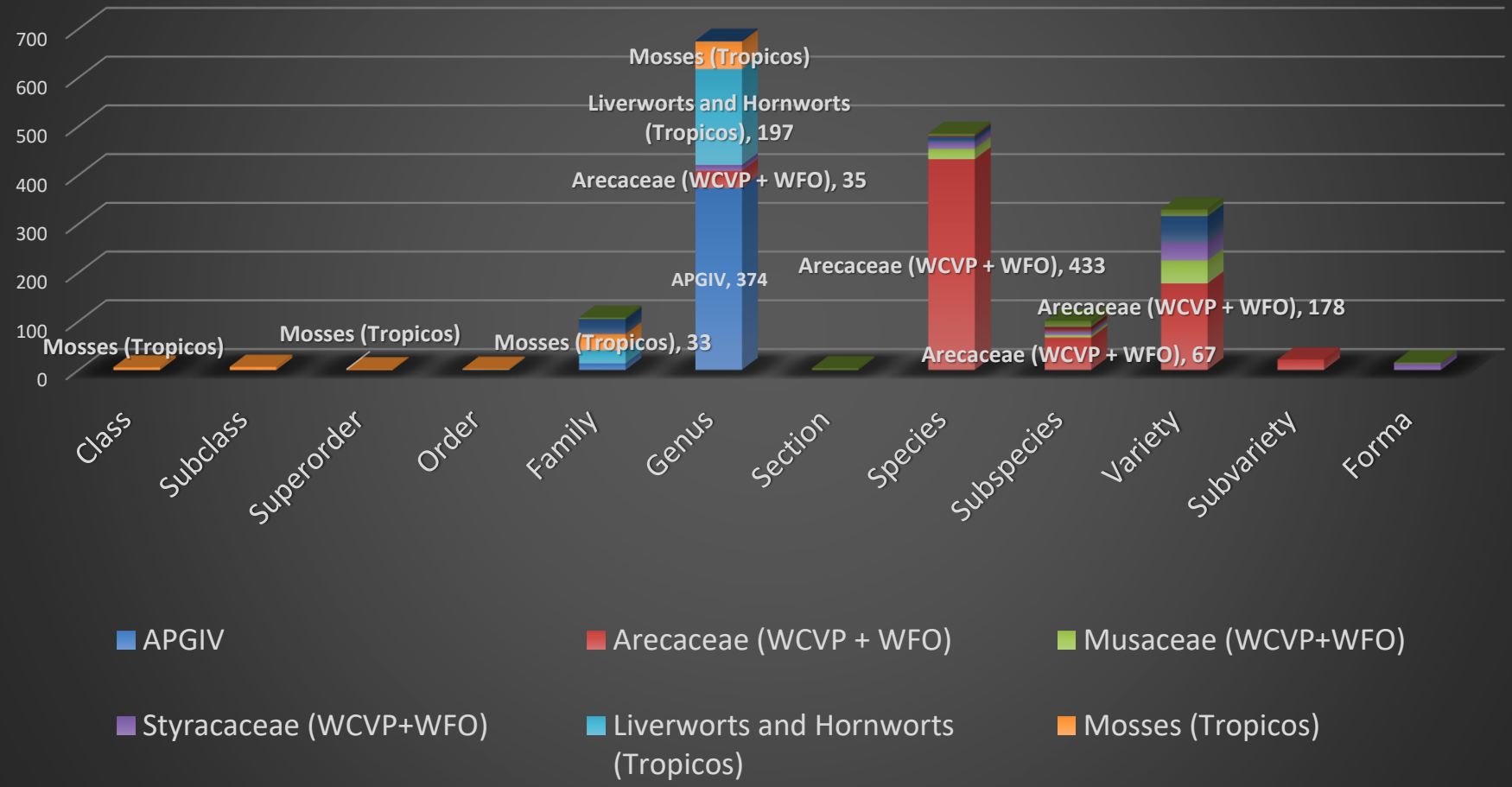
	APGIV	Arecaceae (WCVP + WFO)	Musaceae (WCVP+WFO)	Styracaceae (WCVP+WFO)	Liverworts and Hornworts (Tropicos)	Mosses (Tropicos)	Flore d'Afrique Centrale	Flora Helvetica	Flora of Nepal	TOTAL
Class						6				6
Subclass						7				7
Superorder						2				2
Order					1	2				3
Family	14				27	33	31		3	108
Genus	374	35		12	197	56	2			676
Section									4	4
Species		433	21	13			11	3	4	485
Subspecies		67	5	7				10	12	101
Variety		178	47	37			54		14	330
Subvariety		22								22
Forma				13					3	16
TOTAL	388	735	73	82	225	106	98	13	40	1760

Taxonomía de Rangos Superiores (719)

WCVP + WFO (890)

Nuevos Nombres Agregados de Floras (151)

Number of New Names added by Taxonomic Rank per Data Source by Jun.2020

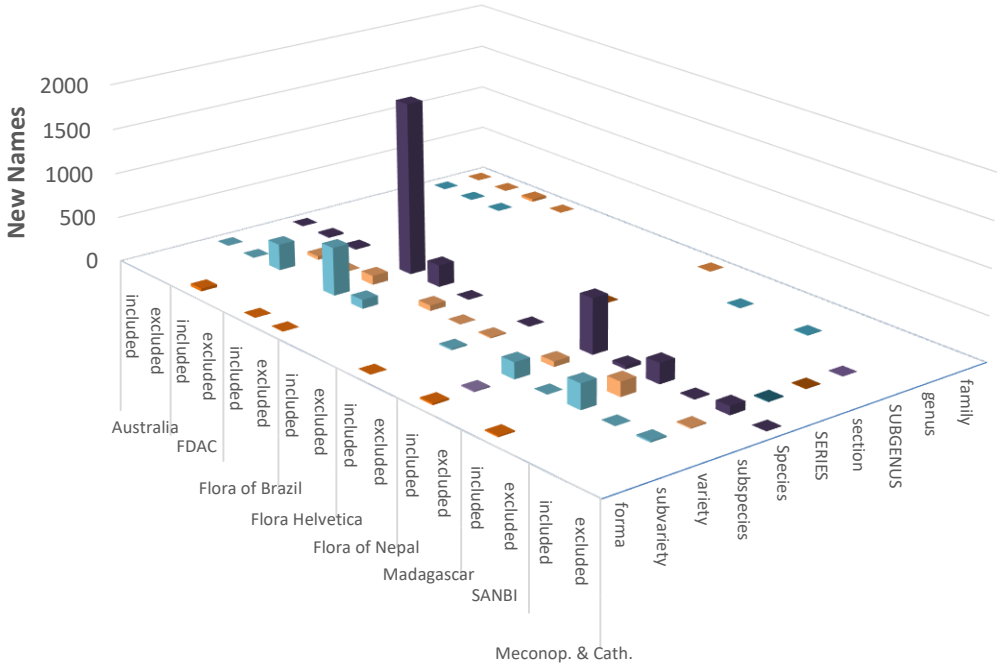


Nuevos Nombres de Endémicas agregadas

por Rango Taxonómico y Fuente a Junio 2020

	Flora do Brasil 2020	Flore d'Afrique Centrale	Catalogue of Madagascar	TOTAL
Genus			4	4
Species	2150		634	2784
Subspecies	100		60	160
Variety	633	194	184	1011
Subvariety			13	13
Forma	7		19	26
TOTAL	2890	194	914	**3998**

New Names in the WFO Taxonomic Backbone from Regional Floras, July 2021



included	4791
excluded	429
Total	**5220**

	Australia		FDAC		Flora of Brazil		Flora Helvetica		Flora of Nepal		Madagascar		SANBI		Meconop. & Cath.	
	included	excluded	included	excluded	included	excluded	included	excluded	included	excluded	included	excluded	included	excluded	included	excluded
family	1	4	29	2					3							
genus	1	4	2								5		7			
SUBGENUS															2	
section									4						7	
SERIES															12	
Species	1	18	17		1912	242	4		4		607	37	226	10	96	1
subspecies			40	1	100		61	3	12		60		163		11	
variety	2	3	291		534	99			14		183	1	280	1	16	
subvariety											13					
forma			32		4	3			3		19		13			

Trabajo con los duplicados en IPNI

Name	IPNI ID	WFO ID	TPL
Lacinaria boykinii	134349-2	wfo-0000004153	gcc-104001
Lacinaria × boykinii	134349-2	wfo-0000052925	gcc-151499
Scorzonera musili	243307-1	wfo-0000004324	gcc-104164
Scorzonera musilii	243307-1	wfo-0000048917	gcc-147425

*There are only **219 names*** included in the pre-harvest DB from IPNI that already matched with a name in WFO (scientificName, Author and Family) but had a different IPNI ID already, so a new WFO record was created.

*There are **1,832 names*** in WFO (species group) that were inserted from IPNI and whose IPNI ID already existed in WFO pre-harvest DB.

Proveedores de Contenido Descriptivo

	Provider	Flora	Initial	Test	Production	Status
1	Solanaceae	Solanaceae Source	1,956	1,956	**1,956**	Included new names from more up-to-date dataset
2	SANBI	Flora of South Africa	79,130	79,130	**79,130**	
3	Kew POWOP	Flora of West Tropical Africa	15,982	15,982	**15,982**	
4	Kew POWOP	Flora of Tropical East Africa	78,165	78,164	**78,164**	
5	Kew POWOP	Flora Zambesiaca	70,085	70,084	**70,084**	
6	MBG	Flora Mesoamericana	8,510	8,510	**8,510**	
7	MBG	Flora de Nicaragua	9,016	9,016	**9,016**	
8	MBG	Flora of Pakistan	4,666	4,662	**4,662**	
9	MBG	Flora de Panama	17,290	17,290	**17,290**	
10	eFloras	Flora of North America	16,438	16,438	**16,438**	
11	eFloras	Flora of China	37,029	37,029	**37,029**	
12	MBG	Manual de Plantas de Costa Rica	5,933	5,933	**5,933**	
13	MBG	Moss Flora of China	1,722	1,721	**1,721**	
14	MBG	Moss Flora of Central America	704	704	**704**	
15	MBG	Novon	277	275	**277**	
16	Turkey	Illustrated Flora of Turkey	282	282	**282**	
17	FloraColombia	Flora de Colombia	847	847	**847**	
18	RBGE	Rhododendron Resource Center	857	855	**855**	
19	NYBG	Flora of NorthEastern US	5,804	5,575	***5,085***	
20	NYBG	Flora Neotropica	18,661	18,541		
21	NYBG	Brittonia	11,272	11,200		
22	NYBG	Memoirs	22,133	21,562		
23	Geneva	Flora Helvetica (German)	6,069	5,938	5,938	Descriptions truncated. New complete dataset received.
24	Geneva	Flora Helvetica (French)	6,069			Complete dataset received.
25	Geneva	African Plants Database	*15,831*			Descriptions truncated. New complete dataset expected.
26	Geneva	eFloramaghreb.org				
27	Meise	Flora of Central Africa	17,507	16,784	**16,784**	Harvested in Production Portal. Name Match required
28	FloraBrazil	Flora do Brasil 2020	105,119	76,908	**76,908**	Habit and distribution in. New descriptions available.
29	Bahamas	Leon Levy Nature Preserve Flora	1,396	1,104	**1,396**	647 taxa
30	IrishFlora	Flora of Ireland	1,388			Reviewed Name Match again.
31	Australia	Flora of Australia	418	406		First set of data received and harvested in Test (418).
32	FNA	New FNA data	65,794			Name Matching process redo. Pending harvesting.
33	eMonocot	Araceae	8,373			
34	eMonocot	PalmWeb.org	5,915			
35	RBGE	Flora of Nepal				
36	DNP	Flora of Thailand				Draft descriptions file within a few weeks - Jun.14th
37	KRNatlArboretum	Flora of Korea				Would upload Korean information at the end of the year
38	MNHN	Floras MNHN				In progress...
			640,638	**506,896**	**454,991**	



African Plants Names Database

First batch of descriptions provided

	Included	*Excluded*	Total
With a WFO ID	174,693	2,339	177,032
Without a WFO ID	26,920	131	27,051
Total	201,613	2,470	204,083

Algunos Proveedores de Contenido

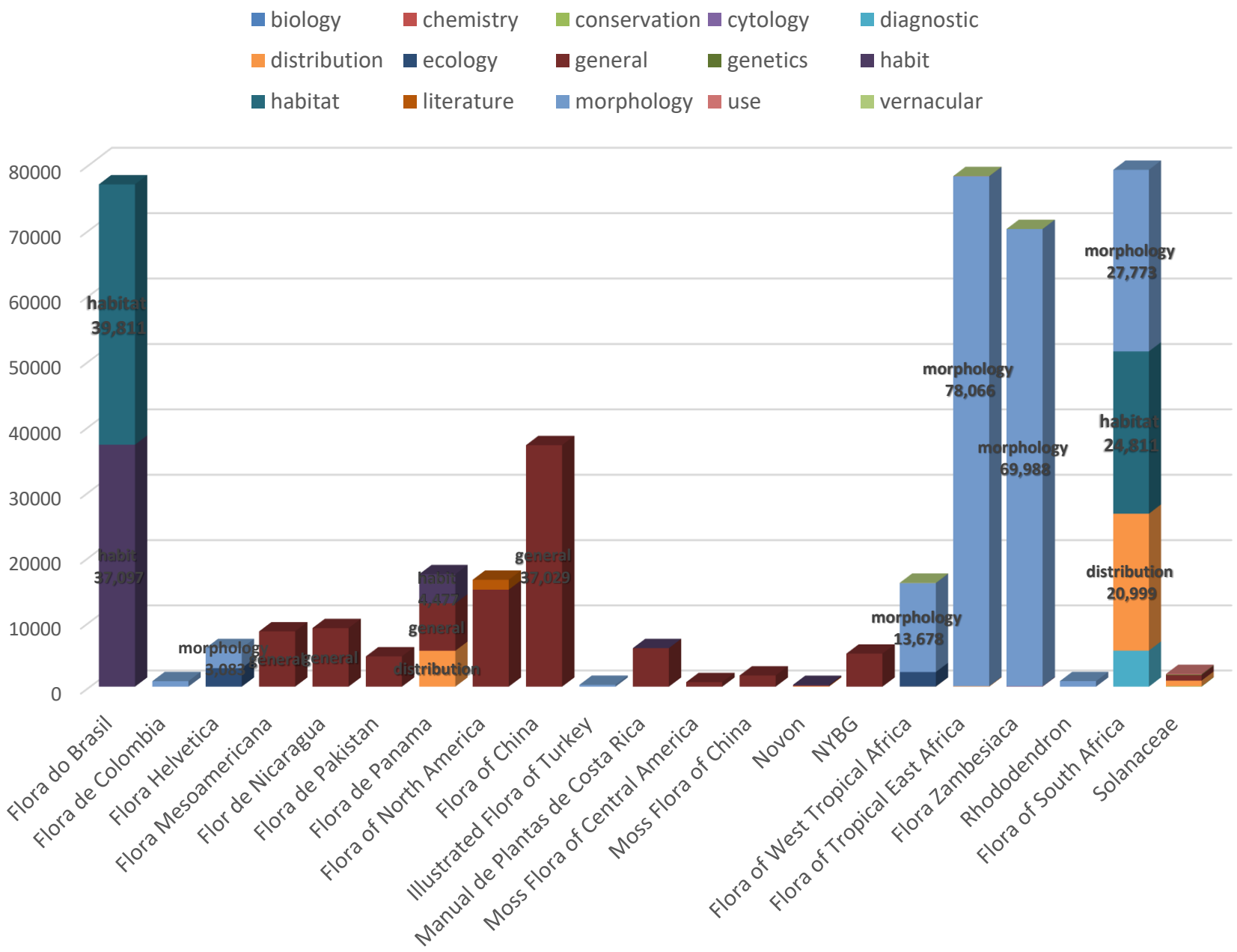
- **MBG Floras:**
  - Flora of Pakistan — 4,662 descriptions ✓
  - Moss Flora of China — 1,721 descriptions ✓
  - Moss Flora of Central America — 704 descriptions ✓
  - Rhododendron — 855 descriptions ✓
- **Flora Helvetica (German):**
  - Ecololgy — 2,855 descriptions ✓
  - Morphology — 3,083 descriptions ✓
  - Vernacular Names — 10,419 vernacular names ✓
- **NYBG** *(5,085 already available in the WFO Portal in Production)*
  - Brittonia *(available in WFO Portal in Test)* — 11,272 descriptions ✓
  - *Flora Neotropica – Pending* — *18,671 descriptions* ✓
  - *Memoirs NYBG – Pending* — *26,182 descriptions* ✓
  - *NE Flora – Pending* — *5,833 descriptions* ✓

Número de Descripciones en WFO

por Fuente y Tipo a Junio 2020

DATA PROVIDER	biology	chemistry	conservation	cytology	diagnostic	distribution	ecology	general	genetics	habit	habitat	literature	morphology	use	vernacular	TOTAL
Flora do Brasil										37,097	39,811					76,908
Flora de Colombia													847			847
Flora Helvetica							2,855						3,083			5,938
Flora Mesoamericana								8,510								8,510
Flor de Nicaragua								9,016								9,016
Flora de Pakistan								4,662								4,662
Flora de Panama						5,532		7,281		4,477						17,290
Flora of North America								14,931				1,507				16,438
Flora of China								37,029								37,029
Illustrated Flora of Turkey													282			282
Manual de Plantas de Costa Rica								5,931		2						5,933
Moss Flora of Central America								704								704
Moss Flora of China								1,721								1,721
Novon						84		131		62						277
NYBG								5,085								5,085
Flora of West Tropical Africa				2		11	2,248						13,678	9	34	15,982
Flora of Tropical East Africa		1		2		68	3						78,066	22	2	78,164
Flora Zambesiaca				67		19	7	1					69,988	1	1	70,084
Rhododendron													855			855
Flora of South Africa					5,547	20,999					24,811		27,773			79,130
Solanaceae	1		73			855		848	84					95		1,956
TOTAL	1	1	73	71	5,547	27,568	5,113	95,850	84	41,638	64,622	1,507	194,572	127	37	436,811

Number of Descriptions per Source and Type of Description in the World Flora Online as per June 2020



World Flora Online
www.worldfloraonline.org

Cantidad de Descripciones

por Rango Taxonómico

Taxon Rank	Descriptions
FAMILY	5,587
GENUS	41,692
SPECIES	449,487
SUBSPECIES	5,777
VARIETY	14,968
FORM	175

TOTAL	**517,686**

Cantidad de Nombres con Descripción

por Rango Taxonómico

Taxon Rank	Taxa
FAMILY	406
GENUS	7,214
SPECIES	135,229
SUBSPECIES	3,135
VARIETY	9,449
FORM	92

TOTAL	**155,525**

Cantidad de Nombres con Descripción

por Rango y Estado Taxonómico

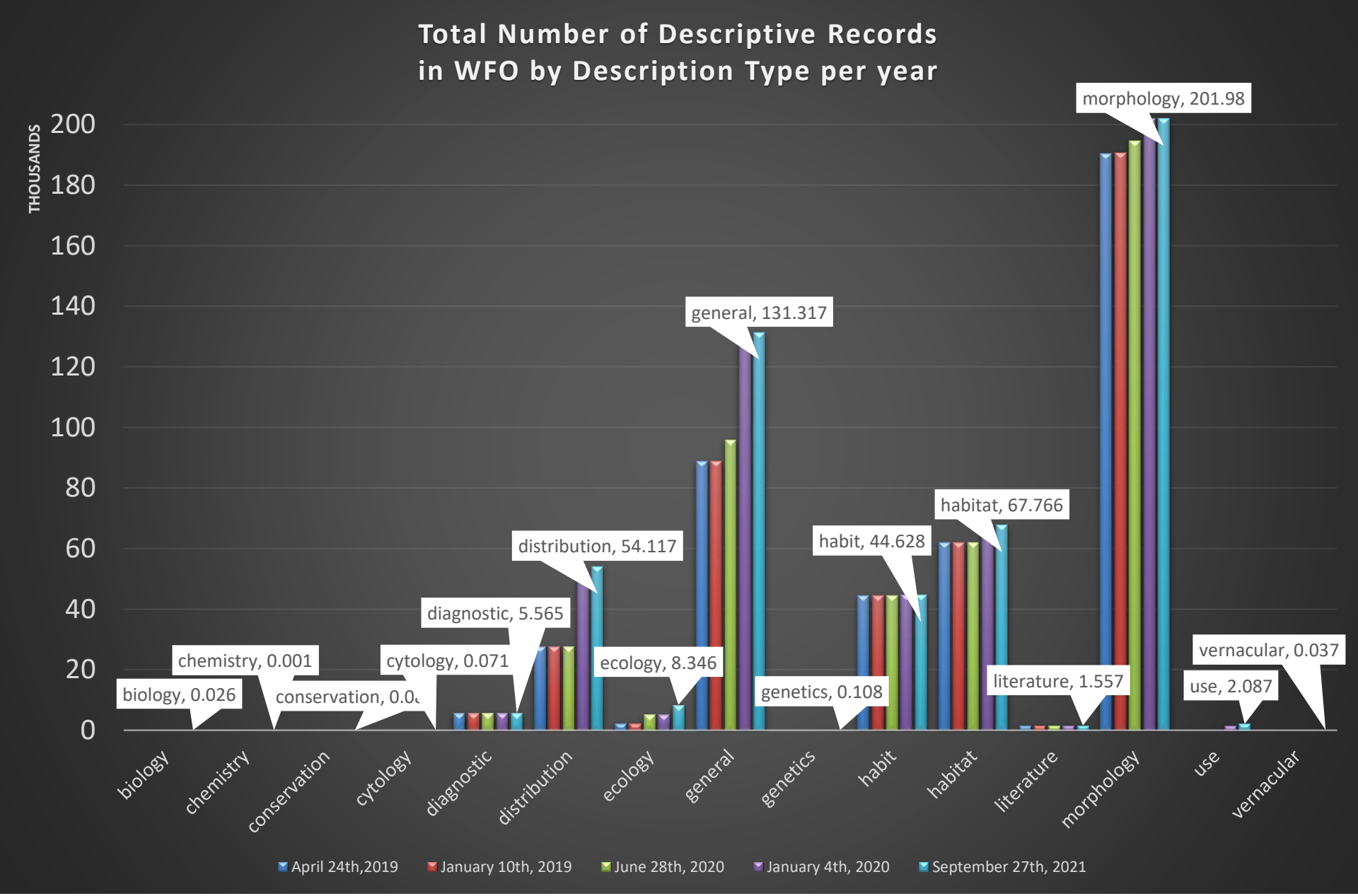
Taxon Rank	Taxonomic Status	Taxa	Total
FAMILY	Accepted	405	406
	Synonym	1	
GENUS	Accepted	6,922	7,214
	Synonym	209	
	Heterotypic Synonym	4	
	Doubtful	79	
SPECIES	Accepted	114,600	135,229
	Synonym	14,641	
	Heterotypic Synonym	63	
	Homotypic Synonym	16	
	Doubtful	5,909	

Número de Nombres con Descripción

por Rango y Estado taxonómico

Taxon Rank	Taxonomic Status	Taxa	Total
SUBSPECIES	Accepted	1,606	3,135
	Synonym	1,466	
	Heterotypic Synonym	1	
	Homotypic Synonym	2	
	Ambiguous	60	
VARIETY	Accepted	3,349	9,449
	Synonym	5,818	
	Heterotypic Synonym	4	
	Doubtful	278	
FORM	Accepted	11	92
	Synonym	81	

TOTAL			**155,525**


Total Number of Descriptive Records in WFO by Description Type per year
THOUSANDS
200
180
160
140
120
100
80
60
40
20
0
biology, 0.026
chemistry, 0.001
conservation, 0.0
cytology, 0.071
diagnostic, 5.565
distribution, 54.117
ecology, 8.346
general, 131.317
genetics, 0.108
habit, 44.628
habitat, 67.766
literature, 1.557
morphology, 201.98
use, 2.087
vernacular, 0.037
biology
chemistry
conservation
cytology
diagnostic
distribution
ecology
general
genetics
habit
habitat
literature
morphology
use
vernacular
April 24th,2019
January 10th, 2019
June 28th, 2020
January 4th, 2020
September 27th, 2021

TAXON
The Journal of the International Association for Plant Taxonomy

METHODS AND TECHNIQUES | Open Access

World Flora Online: Placing taxonomists at the heart of a definitive and comprehensive global resource on the world's plants

Thomas Borsch, Walter Berendsohn, Eduardo Dalcin, Maïté Delmas, Sebsebe Demissew, Alan Elliott, Peter Fritsch, Anne Fuchs, Dmitry Geltman, Adil Güner, Thomas Haevermans, Sandra Knapp, M. Marianne le Roux, Pierre-André Loizeau, Chuck Miller, James Miller, Joseph T. Miller, Raoul Palese, Alan Paton, John Parnell, Colin Pendry, Hai-Ning Qin, Victoria Sosa, Marc Sosef, Eckhard von Raab-Straube, Fhatani Ranwashe, Lauren Raz, Rashad Salimov, Erik Smets, Barbara Thiers, Wayt Thomas, Melissa Tulig, William Ulate, Visotheary Ung, Mark Watson, Peter Wyse Jackson, Nelson Zamora, ... See fewer authors

First published: 26 November 2020 | https://doi.org/10.1002/tax.12373

Associate Editor: Dirk C. Albach

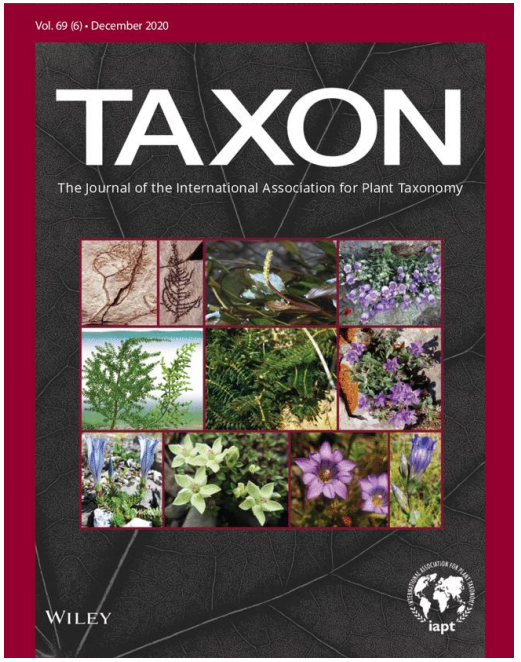
Abstract

It is time to synthesize the knowledge that has been generated through more than 260 years of botanical exploration, taxonomic and, more recently, phylogenetic research throughout the world. The adoption of an updated Global Strategy for Plant Conservation (GSPC) in 2011 provided the essential impetus for the development of the World Flora Online (WFO) project. The project represents an international, coordinated effort by the botanical community to achieve GSPC Target 1, an electronic Flora of all plants. It will be a first-ever unique and authoritative global source of information on the world's plant diversity, compiled, curated, moderated and updated by an expert and specialist-based community (Taxonomic Expert Networks – "TENs" – covering a taxonomic group such as family or order) and actively managed by those who have compiled and contributed the data it includes. Full credit and acknowledgement will be given to the original sources, allowing users to refer back to the primary data. A strength of the project is that it is led and endorsed by a global consortium of more than 40 leading botanical institutions worldwide. A first milestone for producing the World Flora Online is to be accomplished by the end of 2020, but the WFO Consortium is committed to continuing the WFO programme beyond 2020 when it will develop its full impact as the authoritative source of information on the world's plant biodiversity.

TAXON Journal

DOI: 10.1002/tax.12373

World Flora Online: *Placing taxonomists at the heart of a definitive and comprehensive global resource on the world's plants*



World Flora Online
www.worldfloraonline.org

Trabajo Futuro

Lista de Pendientes (ToDo list)…

Status	Activity	Last Action	Current Status	Next Step
W	Fl. Helvetica	Check DwCA	Sent corrections	Harvest Content in Test
…	***Fl. Central Africa***	***Name Matched***	***Waiting for Content***	***Harvest Content in Test***
0	Fl. Malesiana		No Input Yet	
0	Euro+Med		No Input Yet	
…	***Fl. Nepal***	***Name Matched***	***Waiting for Content***	***Harvest Content in Test***
W	**EMonocot (PalmWeb)**	**Content received**	**Pending Name Matching**	**Harvest Content in Test**
W	**EMonocot (Araceae)**	**Content received**	**Pending Name Matching**	**Harvest Content in Test**
0	EMonocot (GrassBase)		No Input Yet	
0	EMonocot (Aeridinae)		No Input Yet	
0	EMonocot (Cypripedioideae)		No Input Yet	
0	EMonocot (Eulophiinae)		No Input Yet	
0	EMonocot (Orchideae)		No Input Yet	
0	Endemic Fl. Chile		No Input Yet	
W	**(W.) Fl. Pakistan**	**Harvested in Test**	**Pending Review**	**Upload in Production**
W	**(W.) Moss Fl. China**	**Harvested in Test**	**Pending Review**	**Upload in Production**
W	**(W.) Mos Fl. Central America**	**Harvested in Test**	**Pending Review**	**Upload in Production**
0	African Plant Database		No Input Yet	
W	**(W?) Fl. Ireland - Unstructured Word docs received**	**Case Explained**	**Pending Research**	**Research**
W	**(W?) Frictionless Data**	**Case Explained**	**Pending Research**	**Research**
S	**(S) IUCN v.3**	**New Format Harvested**	**Pending Research**	**Research how to substitute previous version**
S	**(S?) Identification Keys**	**Previous Format Harvested**	**Pending Research**	**Research how to harvest new format**
W	**(W?) Google Analytics - Create report**	**Case Explained**	**Pending Research**	**Research how to join Analytics stats with Provider info**
?	**(?) Images Metadata**	**Case Explained**	**Pending Research**	**Research how to load more Image metadata**
W	**(W?) DOIs**	**Case Explained**	**Pending Research**	**Research how to assign DOIs to Result sets**
W	***NYBG***	***Name Matched***	***Waiting for Updated Content***	***Harvest DwCA in Test***
W	**(W?) GTA (Distribution)**	**Case Explained**	**Pending Research**	**Research how to harvest distribution and other info**
W	**(S) Fl. Brazil - New and Existing Endemics**	**Names included**	**Pending Harvesting**	**Get new content**
…	***(W.) Fl. Colombia***	***Names Matched***	***Waiting for Response***	***Impement recommendations***
W	**(W.) Fl. Australia**	**Names Received**	**Pending Endemics Names Matching**	**Send Report and Updated Guidelines**
…	***Zingiberaceae***	***Names Matched***	***Waiting for Response***	***Implement recommendations***
W	**(W?) Solanaceae - New input**	**Content Received**	**Pending Name Matching**	**Name Match Taxonomic Names**
W	***SANBI - New and Existing Descriptions***	***Harvested***	***Waiting for Endemics corrections***	***Implement recommendations***
W	**Irvingiaceae**	**DwCA generated**	**Pending Harvesting**	**Harvest DwCA in Test**
SX	Musaceae - WCVP + WFO done. Send for BOTALISTA.	WCVP + WFO done	Done	
SX	Styracaceae - WCVP + WFO done.	WCVP + WFO done	Done	
SX	Aracaceae - WCVP + WFO done. Name Match observa	WCVP + WFO done	Done	
S	Sapotacea	WCVP + WFO done	Preparing Report	Send Report
S	Gesneriaceae	WCVP + WFO done	Preparing Report	Send Report
…	***Picramniaceae - Report sent, waiting corrections***	***Names Matched***	***Waiting for Response***	***Implement recommendations***
…	***Cordiaceae - Report sent, waiting corrections***	***Names Matched***	***Waiting for Response***	***Implement recommendations***
S	**(W.) Bryophytes (Mosses, Liverworts, and Hornwor**	**Received comments from Alan**	**Pending Implementation**	**Implement recommendations**
X	GeoServer - Pending Fix	Map is working	Checking errors	Modify Puppet scripts
W	**WCVP Distribution**	**Content received**	**Pending Name Matching**	**Harvest distribution and other extensions available**
W	**(W.) Rhododendron**	**Harvested in Test**	**Pending Review**	**Upload in Production**

Potenciales Fuentes de Datos

- Flora Malesiana
- Euro+Med
- African Plant Database
- Illustrated Treatments for Korea
- Endemic Flora Chile

- **eMonocot**:
  - Grassbase (10,752)
  - Aeridinae(70)
  - Cypripedioideae(158)
  - Eulophiinae(140)
  - Orchideae(122)

- Digital Flora Phillipines
- eFlora of Maghreb
- Flora of Somalia
- Flora of Marquesas
- Flora of Hawaii
- Flora of Gabon
- Flora of the Guianas
- Flora Iberica
- Flora of Arabia

- Global Tree Assessment Distribution

BISS Journal
DOI: 10.3897/biss.5.75343

TDWG 2021

#TDWG2021

Challenges, Solutions, and Workflows Developed for the Taxonomic Backbone of the World Flora Online

William Ulate[1], Alan Elliott[2] and Sunitha Katabathuni[1]







[1]Missouri Botanical Garden, USA [2]Royal Botanical Gardens Edinburgh



William Ulate william.ulate@mobot.org

Fotografía: Playa Prieta, Costa Rica por W. Ulate, 2021