

Taller GBIF.ES: Estandarización, documentación y publicación de datos de biodiversidad en colecciones

16 – 17 de septiembre de 2026, Institut Botànic, Barcelona

Introducción teórica a los data papers

Francisco Pando
GBIF.ES



**institut
botànic**

Centre mixt



¿Qué es un data paper?

1

Un documento científico que describe un dataset

Es un documento revisado por pares (peer-reviewed), publicado como artículo en una revista científica, cuyo objeto es describir de forma estructurada un conjunto de datos concreto.

2

No contrasta una hipótesis: documenta datos

A diferencia de un artículo científico convencional, su finalidad no es responder una pregunta de investigación, sino describir los datos y las circunstancias de su recogida, tratamiento y alcance.

3

Va enlazado a un dataset publicado (p. ej., en GBIF)

El data paper es el «artículo»; el dataset es «los datos». Ambos se publican por separado, cada uno con su propio DOI, y se citan y enlazan mutuamente.

¿Por qué publicar un data paper?

1

Reconocimiento para quien cura los datos

Da crédito académico formal al trabajo —a menudo invisible— de compilar, digitalizar, documentar y estandarizar colecciones de datos de biodiversidad.

2

Visibilidad y control de calidad

Alerta a la comunidad científica de que el dataset existe y está disponible, y añade una capa de revisión por pares que respalda la calidad y la documentación de los datos.

3

Citación y seguimiento del uso

El data paper se indexa y se cita como cualquier otra publicación (Web of Science, Google Scholar...), lo que permite medir el impacto real del dataset a lo largo del tiempo.

¿Cómo se elabora y se publica?

1

Punto de partida: los metadatos del dataset

El proceso arranca con los metadatos (formato Darwin Core / EML) ya cargados en el IPT (Integrated Publishing Toolkit) de GBIF.

2

Generación semiautomática del manuscrito

GBIF, junto con Pensoft, desarrolló una herramienta que genera un borrador del artículo (RTF) directamente a partir de esos metadatos; alternatively, el ARPHA Writing Tool permite escribir, revisar y publicar todo el data paper en una única plataforma web.

3

Resultado: dos publicaciones enlazadas

El dataset queda publicado en GBIF con su DOI, y el data paper se publica —tras revisión por pares— en la revista, con su propio DOI, remitiendo el uno al otro.

¿Dónde se publican los data papers?

GBIF mantiene una lista de revistas que aceptan data papers y admiten enlazar sus datos a GBIF:

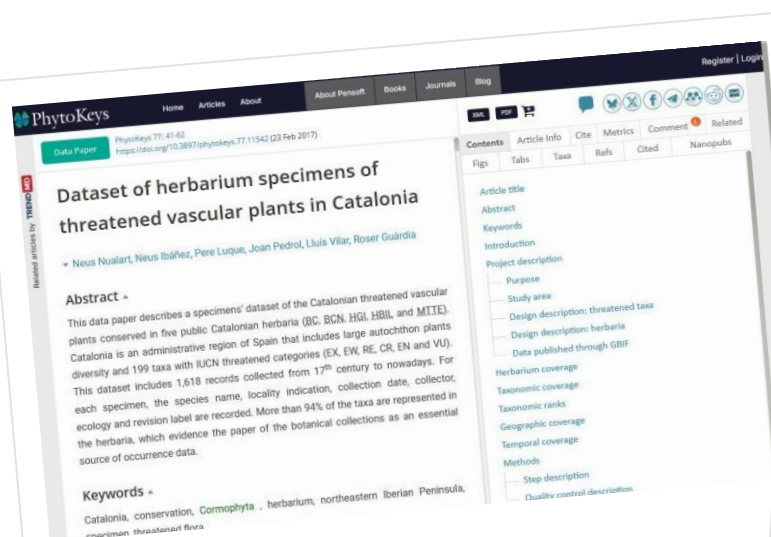
Grupo Pensoft / ARPHA

- PhytoKeys
<https://phytokeys.pensoft.net/>
- ZooKeys
<https://zookeys.pensoft.net/>
- MycoKeys
<https://mycokeys.pensoft.net/>
- Biodiversity Data Journal
<https://bdj.pensoft.net/>
- NeoBiota
<https://neobiota.pensoft.net/>
- Nature Conservation
<https://natureconservation.pensoft.net/>
- BioRisk
<https://biorisk.pensoft.net/>
- Metabarcoding and Metagenomics
<https://mbmg.pensoft.net/>

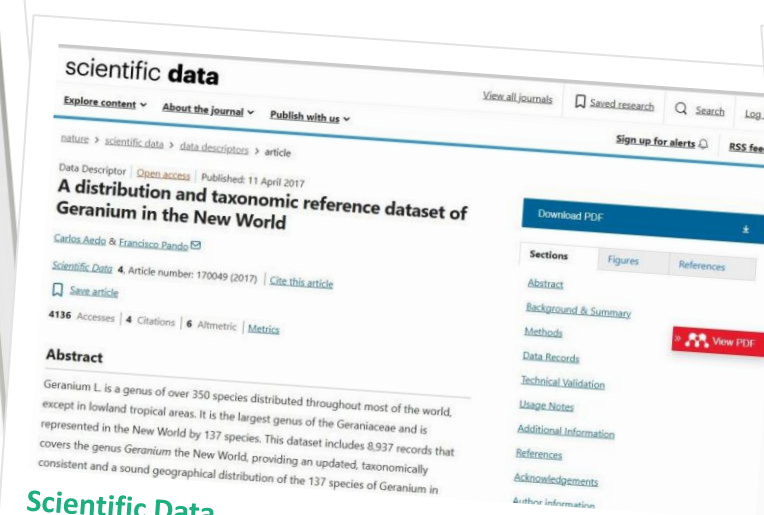
Otras editoriales

- Scientific Data (Nature Publishing Group)
<https://www.nature.com/sdata/>
- Data (MDPI)
<https://www.mdpi.com/journal/data>
- Data in Brief (Elsevier)
<https://www.sciencedirect.com/journal/data-in-brief>
- GigaScience (Oxford Univ. Press)
<https://academic.oup.com/gigascience>
- GigaByte (GigaScience Press)
<https://gigabytejournal.com/>
- Earth System Science Data (Copernicus)
<https://www.earth-system-science-data.net/>
- Biota Colombiana (Instituto Humboldt)
<https://revistas.humboldt.org.co/index.php/biota>
- PLOS ONE
<https://journals.plos.org/plosone/>

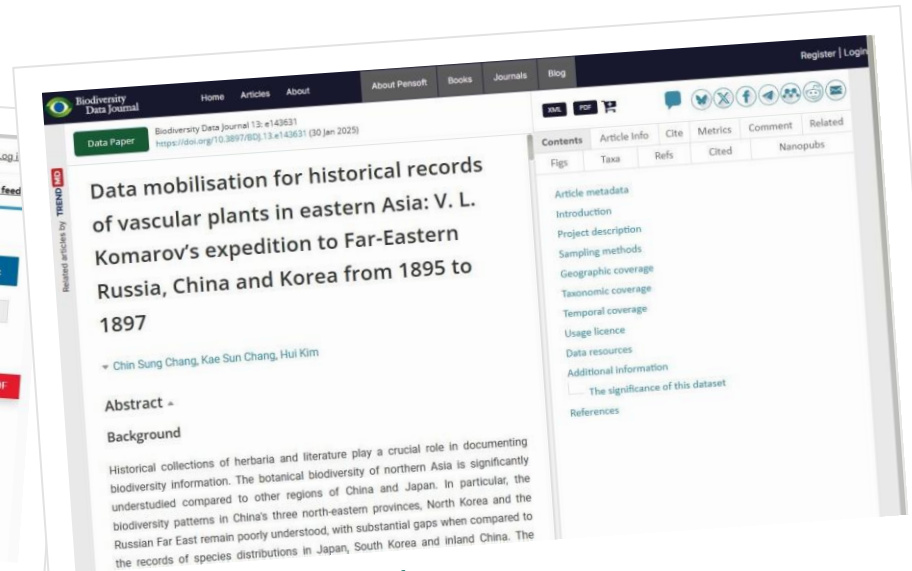
Ejemplos: data papers botánicos en distintas revistas



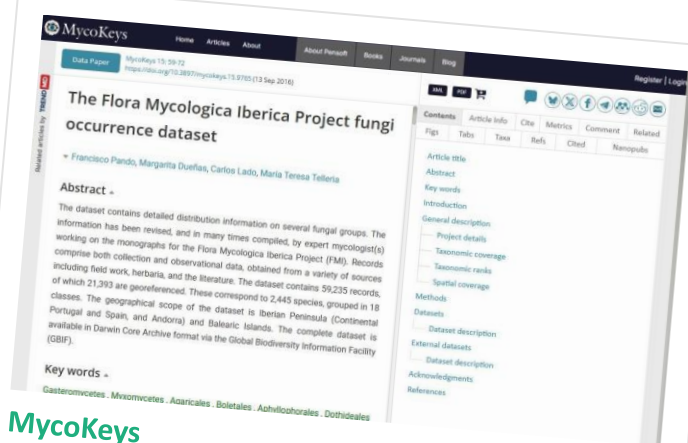
PhytoKeys
doi.org/10.3897/phytokeys.77.11542



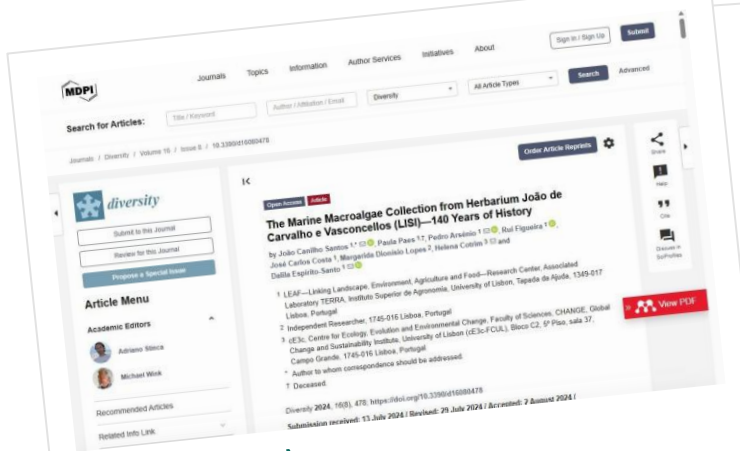
Scientific Data
doi.org/10.1038/sdata.2017.49



Biodiversity Data Journal
doi.org/10.3897/BDJ.13.e143631



MycoKeys
doi.org/10.3897/mycokeys.15.9765



Diversity (MDPI)
doi.org/10.3390/d16080478

<https://ipt.gbif.es/resource?r=threatenedCAT>

Inicio / Recurso

Dataset of herbarium specimens of threatened vascular plants in Catalonia

Registro biológico

Última versión publicado por Herbarium specimens of threatened vascular plants in Catalonia el nov 30, 2018

Editar

Descargue la última versión de los datos como un Archivo Darwin Core (DwC-A) o los metadatos como EML o RTF:

Datos como un archivo DwC-A [descargar](#) 1.618 registros en Inglés (109 KB) - Frecuencia de actualización: desconocido

Metadatos como un archivo EML [descargar](#) en Inglés (13 KB)

Metadatos como un archivo RTF [descargar](#) en Inglés (12 KB)

Descripción

Registros

Versiones

Cómo referenciar?

Descripción

This dataset includes the specimens of the Catalanian threatened vascular plants conserved in five public Catalanian herbaria (BC, BCN, HGI, HBIL and MTTE). Catalonia is an administrative region of Spain that includes large autochthon plants diversity and 199 taxa with IUCN threatened categories (EX, EW, RE, CR, EN and VU). This dataset includes 1,634 records collected from 17th century to nowadays. For each specimen, the species name, locality indication, collection date, collector, ecology and label are recorded. More than 94% of the taxa are represented in the herbaria, which evidence the paper of the botanical collections as an essential source of occurrence data.

<https://www.gbif.org/dataset/dbc709b9-e36e-4dd7-ab5b-c3cb08c2779d>

<https://phytokeys.pensoft.net/article/11542/>

Get data ▾ How-to ▾ Tools ▾ Community ▾ About ▾

DOI 10.15470/sqanal

Occurrence dataset | Registered 13 March 2017

Dataset of herbarium specimens of threatened vascular plants in Catalonia

Published by Herbarium specimens of threatened vascular plants in Catalonia

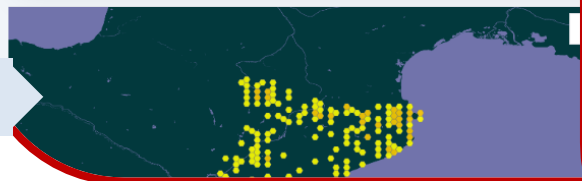
Nualart N [Licence](#) CC BY 4.0

About Metrics Download

Description

Publication date: 30 November 2018

This dataset includes the specimens of the Catalanian threatened vascular plants conserved in five public Catalanian herbaria (BC, BCN, HGI, HBIL and MTTE). Catalonia is an administrative region of Spain that includes large autochthon plants diversity and 199 taxa with IUCN threatened categories (EX, EW, RE, CR, EN and VU). This dataset includes 1,634 records collected from 17th century to nowadays. For each specimen, the species name, locality indication, collection date, collector, ecology and label are recorded. More than 94% of the taxa are represented in the herbaria, which evidence the paper of the botanical collections as an essential source of occurrence data.



PhytoKeys

Regist

Data Paper [PhytoKeys 77: 41-62](#)
<https://doi.org/10.3897/phytokeys.77.11542> (23 Feb 2017)

Dataset of herbarium specimens of threatened vascular plants in Catalonia

Neus Nualart, Neus Ibáñez, Pere Luque, Joan Pedrol, Lluís Vilar, Roser Guàrdia

Abstract

This data paper describes a specimens' dataset of the Catalanian threatened vascular plants conserved in five public Catalanian herbaria (BC, BCN, HGI, HBIL and MTTE). Catalonia is an administrative region of Spain that includes large autochthon plants diversity and 199 taxa with IUCN threatened categories (EX, EW, RE, CR, EN and VU). This dataset includes 1,618 records collected from 17th century to nowadays. For each specimen, the species name, locality indication, collection date, collector, ecology and revision label are recorded. More than 94% of the taxa are represented in the herbaria, which evidence the paper of the botanical collections as an essential source of occurrence data.

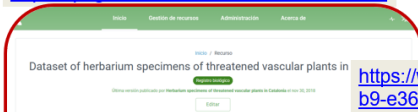
Keywords

Catalonia, conservation, **Cormophyta**, herbarium, northeastern Iberian Peninsula, specimen, threatened flora

<https://doi.org/10.15470/sqanal>

<https://doi.org/10.3897/phytokeys.77.11542>

<https://ipt.gbif.es/resource?r=threatenedCAT>



<https://www.gbif.org/dataset/dbc709b9-e36e-4dd7-ab5b-c3cb08c2779d>

<https://phytokeys.pensoft.net/article/11542/>

<https://doi.org/10.15470/sqanal>

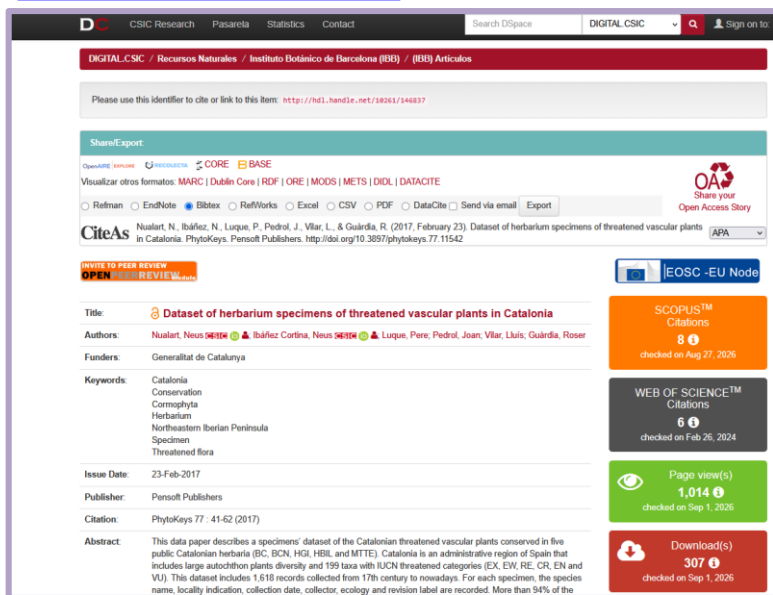
<https://doi.org/10.3897/phytokeys.77.11542>

<https://www.recercat.cat/handle/2072/359353>

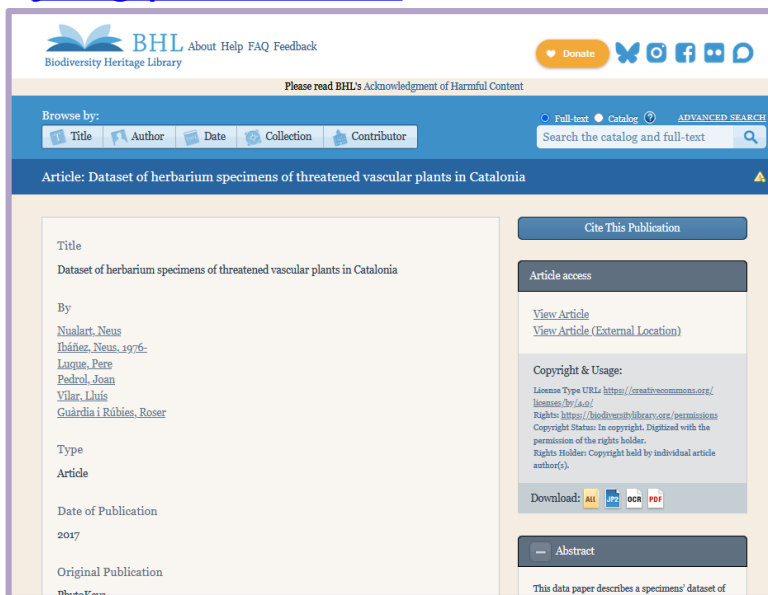
<https://hdl.handle.net/2072/359353>



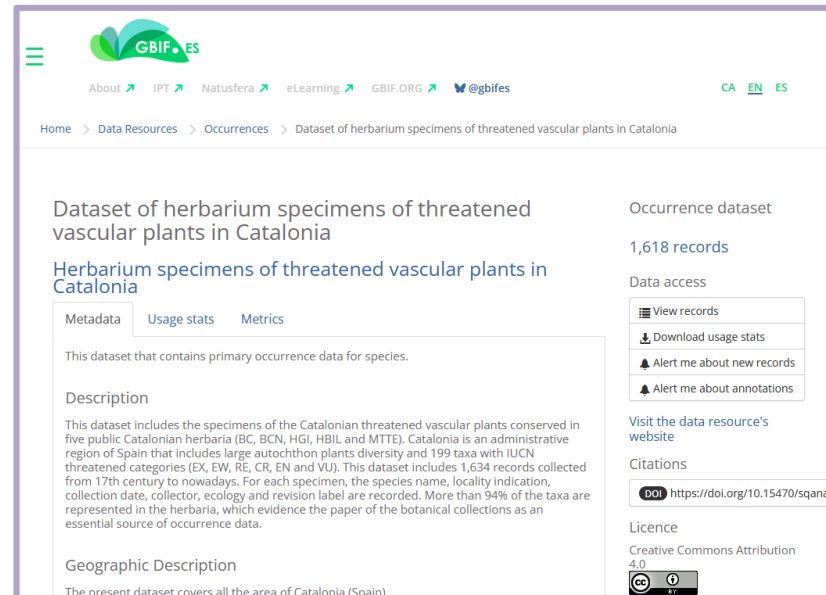
<https://digital.csic.es/handle/10261/146837>



<https://www.biodiversitylibrary.org/part/219452>

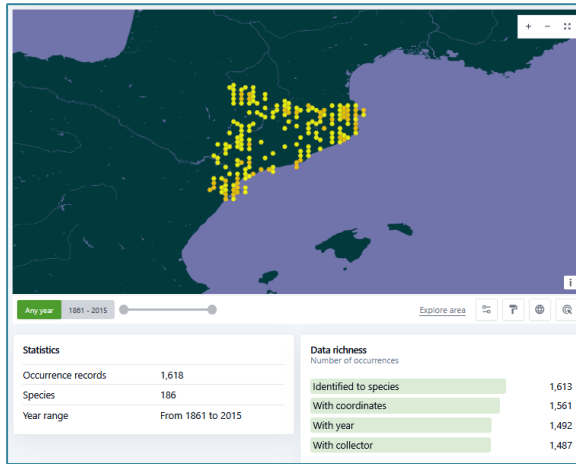


<https://colecciones.gbif.es/public/showDataResource/dr299>



<http://hdl.handle.net/10261/146837>

Metadatos del juego de datos vs. Artículo de datos



IUCN Red List Category
Species that are categorized as Near Threatened (NT) or at higher risk according to the IUCN Red List of Threatened Species.

7 results

VU *Salix tarraconensis* 47

Domain Eukaryota > Kingdom Plantae > Subkingdom Pteridobiotina > Phylum Tracheophyta > Class Magnoliopsida > Order Malpighiales > Family Salicaceae > Tribe Saliceae > Genus Salix > Species Salix tarraconensis

EN *Silene sennenii* 21

Domain Eukaryota > Kingdom Plantae > Subkingdom Pteridobiotina > Phylum Tracheophyta > Class Magnoliopsida > Order Caryophyllales > Family Caryophyllaceae > Tribe Sileneae > Genus Silene > Species Silene sennenii

NT *Juncus heterophyllus* 10

Domain Eukaryota > Kingdom Plantae > Subkingdom Pteridobiotina > Phylum Tracheophyta > Class Liliopsida > Order Poales > Family Juncaceae > Genus Juncus > Species Juncus heterophyllus

LC *Prunus lusitanica* 10

Domain Eukaryota > Kingdom Plantae > Subkingdom Pteridobiotina > Phylum Tracheophyta > Class Magnoliopsida > Order Rosales > Family Rosaceae > Subfamily Amygdaloideae > Tribe Amygdaleae > Genus Prunus > Species Prunus lusitanica

NT *Elatine alsinastrum* 5

Domain Eukaryota > Kingdom Plantae > Subkingdom Pteridobiotina > Phylum Tracheophyta > Class Magnoliopsida > Order Malpighiales > Family Elatinaceae > Genus Elatine > Species Elatine alsinastrum

NT *Elatine bronchonii* 5

Domain Eukaryota > Kingdom Plantae > Subkingdom Pteridobiotina > Phylum Tracheophyta > Class Magnoliopsida > Order Malpighiales > Family Elatinaceae > Genus Elatine > Species Elatine bronchonii

NT *Gagea pratensis* 2

Domain Eukaryota > Kingdom Plantae > Subkingdom Pteridobiotina > Phylum Tracheophyta > Class Liliopsida > Order Liliales > Family Liliaceae > Genus Gagea > Species Gagea pratensis

Taxonomic distribution of occurrences

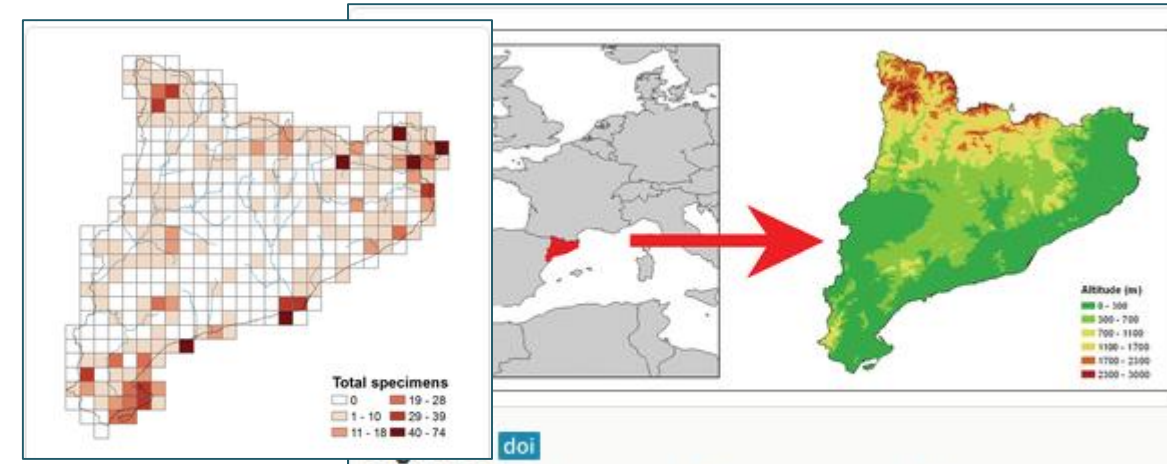
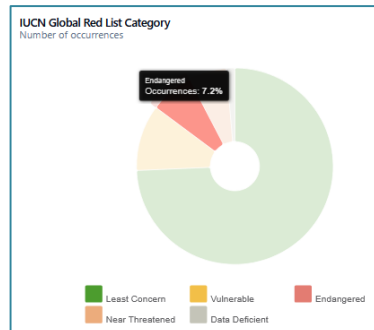
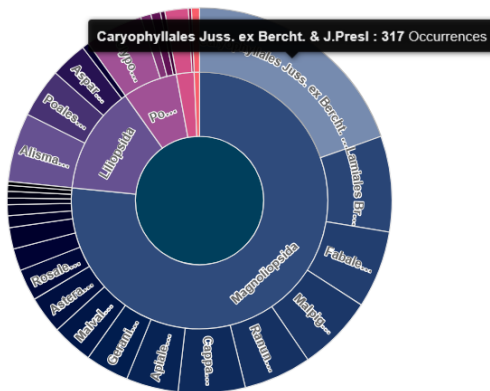
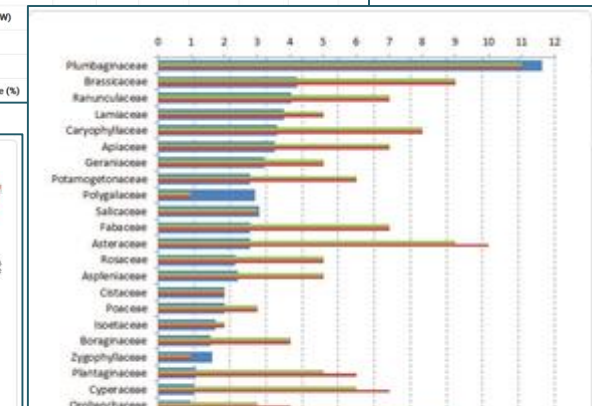
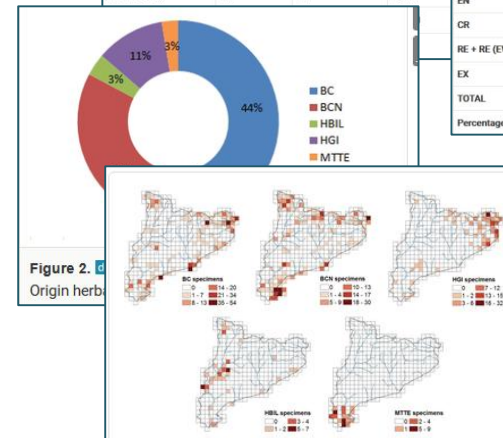


Table 1.
Representation of the Catalanian threatened taxa in the dataset.

Taxa	Red List	Herbarium	%
VU	91	90	98.90
EN	54	53	98.15
CR	37	33	89.19
RE + RE (EW)	16	11	68.75

Table 2.
Number of taxa according to the specimens' range and the IUCN category.

	Number of specimens						
	1-5	6-10	11-15	16-20	20-25	25-30	>30
VU	43	19	12	5	2	2	7
EN	27	11	6	7	1	-	1
CR	20	8	3	2	-	-	-
RE + RE (EW)	-	-	-	-	-	-	-
EX	-	-	-	-	-	-	-
TOTAL	-	-	-	-	-	-	-
Percentage (%)	-	-	-	-	-	-	-



<https://doi.org/10.15470/sqanal>

<https://doi.org/10.3897/phytokeys.77.11542>

Metadatos del juego de datos vs. artículo de datos

459 citations

1,618 occurrences

On this page

Description

Geographic scope

Temporal scope

Taxonomic scope

Methodology

Contacts

GBIF registration

Citation

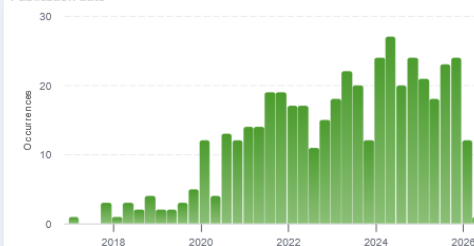
459 citations

1,618 occurrences

About Metrics Download

Citations

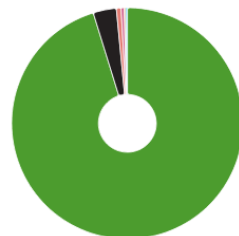
Publication date



100% of all records have a value

Relevance

Number of occurrences



Country or area of researcher			
Number of occurrences			
93 results			
United States of America	171		
Germany	102		
United Kingdom	98		
Switzerland	68		
China	66		
Australia	64		
Spain	64		
France	57		
Canada	54		
Netherlands (Kingdom of the)	51		

Topic			
Number of occurrences			
18 results			
Ecology	144		
Climate change	86		
Evolution	66		
Biodiversity science	64		
Conservation	54		
Biogeography	44		
Invasives	43		
Phylogenetics	40		
Species distributions	25		
Taxonomy	14		

6

Total citations

Article title

Abstract

Keywords

Introduction

Project description

Purpose

Study area

Design description: threatened taxa

Design description: herbaria

Data published through GBIF

Herbarium coverage

Taxonomic coverage

Taxonomic ranks

Geographic coverage

Temporal coverage

Methods

Step description

Quality control description

Dataset description

Acnowledgements

References

Dimensions



6 Total citations
0 Recent citations
0.69 Field Citation Ratio
0.09 Relative Citation Ratio

Scite

scite_

Publications 8
Supporting 0
Mentioning 11
Contrasting 0

scite_



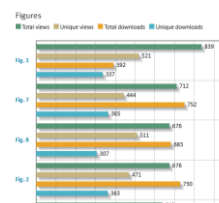
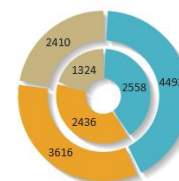
I 2
M 0
R 0
D 1
O 4

Sections
See more details

Article views

Unique views: 6318 | Total views: 10519

HTML PDF XML



[Inicio](#) / [Recurso](#)

A Distribution and Taxonomic Reference Dataset of Geranium (Geraniaceae) in the New World

[Registro biológico](#)
[Especímen](#)

Última versión publicado por Real Jardín Botánico (CSIC)

Inicio: <http://www.geranium.es/>
 GBIF UUID: 26d72d3b-4544-4645-aa56-27aa8a669c6f
 Fecha de publicación: 22 de diciembre de 2021
 Publicado por: Real Jardín Botánico (CSIC)
 Licencia: CC-BY 4.0

¿Cómo referenciar? [DOI](#) 10.15470/ibqzdo

Descargue la última versión de los datos como un Archivo Darwin Core (DWC-A), EML o RTF:

Datos como un archivo **DWC-A** [descargar](#) 8.937 registros en actualización: no planeado

Metadatos como un archivo **EML** [descargar](#) en Inglés (38 KB)

Metadatos como un archivo **RTF** [descargar](#) en Inglés (35 KB)

395 citations

8,937 occurrences

4144

Article Accesses

4

Citations

[Obtener datos](#)
[How-to](#)
[Herramientas](#)
[Comunidad](#)
[Acerca de](#)

[Inicio sesión](#)

[DOI](#) 10.15470/ibqzdo

[Conjunto de datos de registros biológicos](#) | Registrado

A Distribution and Taxonomic Reference Dataset of Geranium (Geraniaceae) in the New World

Publicado por Real Jardín Botánico (CSIC)

Aedo C • Pando F

[www.geranium.es](#)

Licencia [CC BY 4.0](#)

395 citas 8937

[Acerca de](#)
[Estadísticas](#)
[Descargar](#)

Descripción

Fecha de publicación: 22 de diciembre de 2021

The dataset includes 8937 records that covers ge sound geographical distribution of the 137 species monograph was based:

Aedo, Carlos (2012). Revision of Geranium (Geraniaceae) in the New World. Real Jardín Botánico (CSIC).

Figure 1: Map of georeferenced specimens.



scientific data

View all journals

Search

Log in

Explore content

About the journal

Publish with us

[nature](#) > [scientific data](#) > [data descriptors](#) > article

[Download PDF](#)

Data Descriptor | [Open access](#) | Published: 11 April 2017

A distribution and taxonomic reference dataset of Geranium in the New World

DOI: 10.1038/s41598-017-170049-2

7 Altmetric



<https://doi.org/10.15470/ibqzdo>

<https://doi.org/10.1038/s41598-017-170049-2>



Francisco Pando

Unidad de Coordinación de GBIF.ES

CSIC

Joaquín Costa 22, 28002 Madrid

pando@gbif.es



<https://creativecommons.org/licenses/by/4.0/deed.es>

GBIF-ES es el Nodo Nacional de Información en Biodiversidad patrocinado por el Ministerio de Ciencia, Innovación y Universidades, gestionado por el Consejo Superior de Investigaciones Científicas (CSIC). <https://www.gbif.es/> <https://datos.gbif.es>, <https://elearning.gbif.es>



MINISTERIO
DE CIENCIA, INNOVACIÓN
Y UNIVERSIDADES



CSIC
CONSEJO SUPERIOR DE INVESTIGACIONES CIENTÍFICAS

