



Gbif.es



Transferencia de capacidades técnicas asociadas a la publicación de datos de biodiversidad a través de GBIF



CESP: Regional Capacity Enhancement to Latinamerica by establishing Chile's node
23 al 25 abril 2018
Santiago - Chile



Leonardo Buitrago
Administrador Nodo GBIF Colombia,
Representante (S) Regional de Nodos GBIF (Latinoamérica y Caribe)
Líder Administración de contenidos SiB Colombia
albuitrago@humboldt.org.co

Taller Práctico

Estructuración DwC



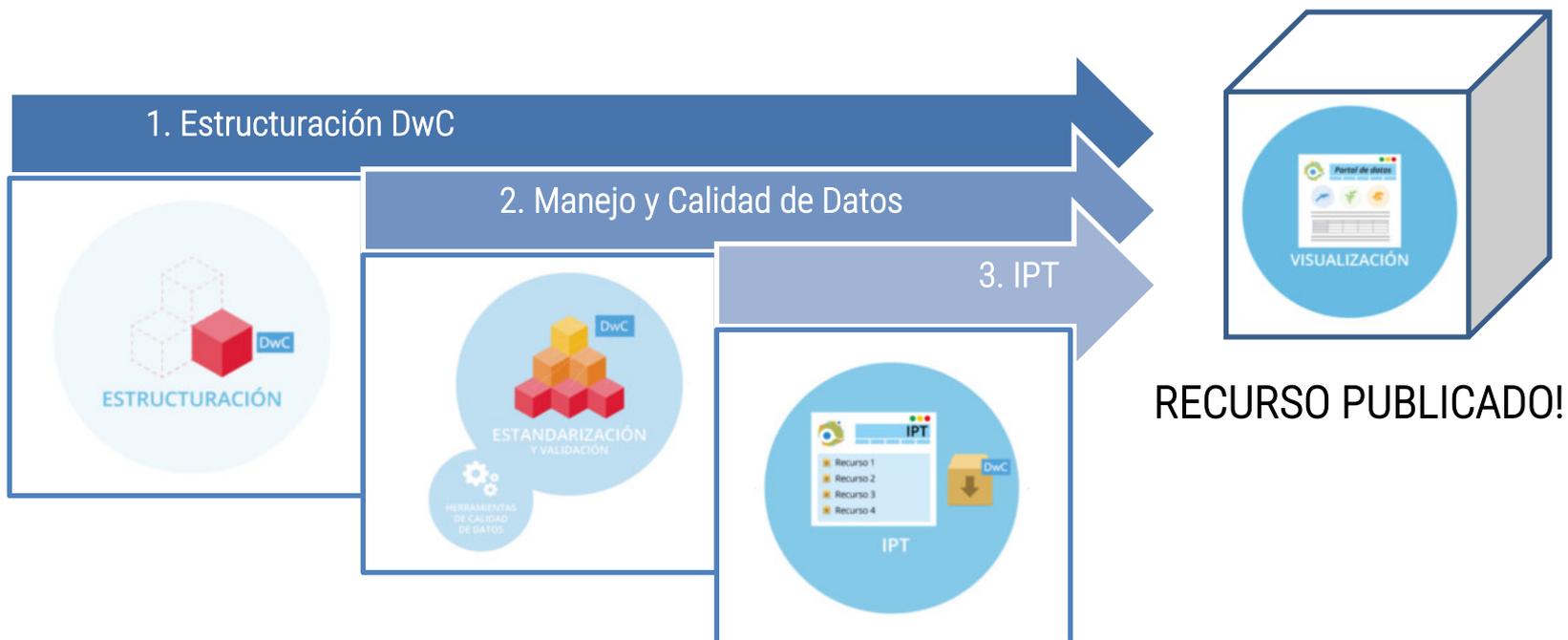
Gbif.es



Partiendo de un Caso de estudio

1. **Estructuración DwC** (MS Excel)
2. **Herramientas para el manejo y calidad de datos**
 - Open Refine
 - Validación taxonómica
 - Species Matching (GBIF)
 - Name Parser (GBIF España)
 - Validación geográfica
 - Conversión de coordenadas (GBIF Canadá)
 - * infoXY (CRIA - Brasil)
 - Validación de fechas (GBIF Canadá)
 - Validación conjunto de datos DwC (GBIF)
3. **Herramienta de publicación IPT**
 - Carga de datos
 - Elaboración de Metadatos

Camino hacia la publicación...



A estructurar datos en Dwc!

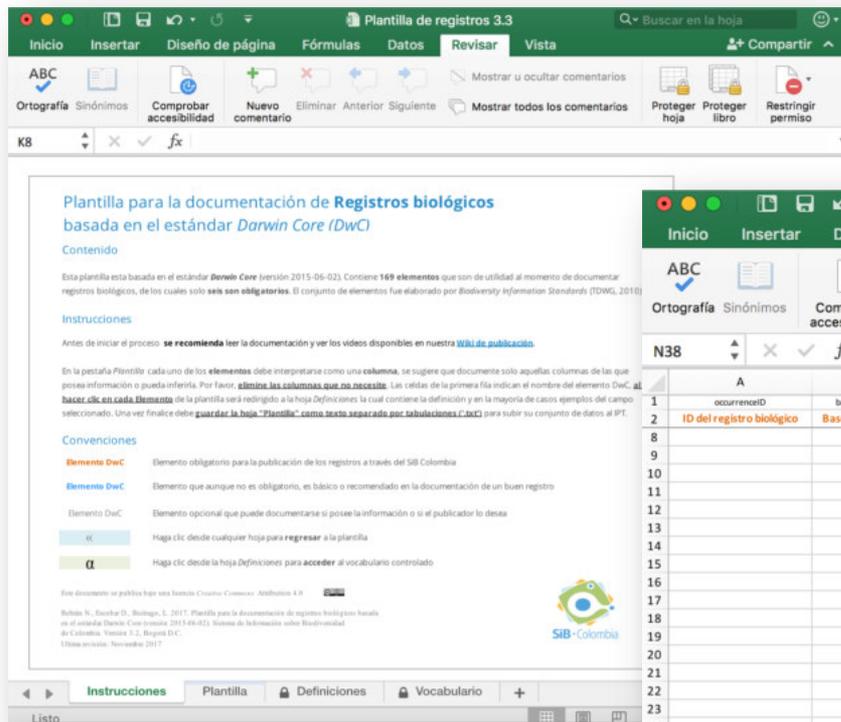


A estructurar datos en DwC!



DwC, cómo abordarlo?

1. Uso de Plantillas estándar (todos los elementos)



Plantilla para la documentación de Registros biológicos basada en el estándar Darwin Core (DwC)

Contenido

Esta plantilla está basada en el estándar **Darwin Core** (versión 2015-06-02). Contiene **169 elementos** que son de utilidad al momento de documentar registros biológicos, de los cuales solo **seis son obligatorios**. El conjunto de elementos fue elaborado por **Biodiversity Information Standards (DwC, 2010)**.

Instrucciones

Antes de iniciar el proceso, se recomienda leer la documentación y ver los videos disponibles en nuestra [Wiki de publicación](#).

En la pestaña **Plantilla** cada uno de los **elementos** debe interpretarse como una **columna**, se sugiere que documente solo aquellas columnas de las que posea información o pueda inferirla. Por favor, **elimine las columnas que no necesite**. Las celdas de la primera fila indican el nombre del elemento DwC, **al hacer clic en cada elemento** de la plantilla será redirigido a la hoja **Definiciones** la cual contiene la definición y en la mayoría de casos ejemplos del campo seleccionado. Una vez finalice debe **guardar la hoja "Plantilla" como texto separado por tabulaciones (.txt)** para subir su conjunto de datos al IPT.

Convenciones

- Elemento DwC** Elemento obligatorio para la publicación de los registros a través del SiB Colombia
- Elemento DwC** Elemento que aunque no es obligatorio, es básico o recomendado en la documentación de un buen registro
- Elemento DwC** Elemento opcional que puede documentarse si posee la información o si el publicador lo desea

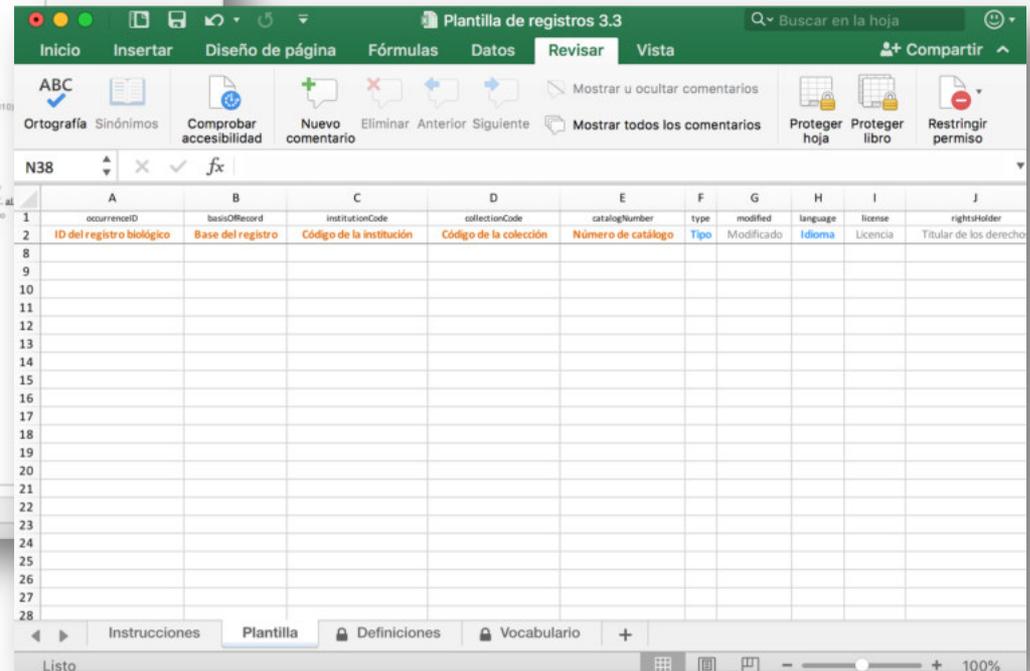
 Haga clic desde cualquier hoja para **regresar** a la plantilla
 Haga clic desde la hoja **Definiciones** para **acceder** al vocabulario controlado

Este documento se publica bajo una licencia Creative Commons Atribución 4.0 

Roberto N. Escobar D., Biólogo, L. 2017. Plantilla para la documentación de registros biológicos basada en el estándar Darwin Core (versión 2015-06-02). Sistema de Información sobre Biodiversidad de Colombia. Versión 3.2, Bogotá D.C.
 Última revisión: Noviembre 2017



Instrucciones | **Plantilla** | Definiciones | Vocabulario | +



	A	B	C	D	E	F	G	H	I	J
1	occurrenceID	basisOfRecord	institutionCode	collectionCode	catalogNumber	type	modified	language	license	rightsHolder
2	ID del registro biológico	Base del registro	Código de la institución	Código de la colección	Número de catálogo	Tipo	Modificado	Idioma	Licencia	Titular de los derechos
3										
4										
5										
6										
7										
8										
9										
10										
11										
12										
13										
14										
15										
16										
17										
18										
19										
20										
21										
22										
23										
24										
25										
26										
27										
28										

Instrucciones | **Plantilla** | Definiciones | Vocabulario | +

169 elementos DwC

DwC, cómo abordarlo?

2. Generar plantillas DwC (definido por el publicador)

INVENTARIO FORESTAL NACIONAL
COLOMBIA

F4.1 Colección botánica ID Conglomerado: 142922

Diligenciado por: Antonio Buzza Fecha: 2017-08-03

Tamaño de individuo	Nº Individuo ID	Nombre común	Determinación en campo o Igualación	Observaciones del individuo	Iniciales y No. Colector	Colectada
F	41	Platanote		abundante latex blanco	CCF 1045	SI
F	42	N.N.19	CF POLYGONACEAE		CCF 1046	SI
F	43	laurel		-Corteza externa café interna amarilla intensa olor Benzopiraud	CCF 1047	SI
F	44	Palma	= 39			
F	45	Anime	= 32			
6F	46	N.N.20				

Sección Formato IDEAM	Elementos IDEAM	Tipo de Dato IDEAM	Core DwC	Extensión DwC	Elementos DwC
4.1.1	ID Conglomerado	—	—	—	fieldNumber
4.1.2	Diligenciado por	Alfanumérico	—	—	—
4.1.3	Fecha	AAAAMMDD	—	—	—
4.1.4	Tamaño del Individuo	Numérico	Occurrence	measurementOrFacts	measurementType
	—				measurementValue
4.1.5	Número de Individuo ID	Alfanumérico	Occurrence	—	organismID
4.1.6	Nombre común	Alfanumérico	Occurrence	—	vernacularName

DwC, cómo abordarlo?

2. Generar plantillas DwC (definido por el publicador)

B	C	D	E	F	G	H	I
basisOfRecord	institutionCode	collectionCode	catalogNumber	recordedBy	eventDate	samplingProtocol	individualCount
En Colección	UCN	MH-ORNIT	46-2300MI2008AV0021	Velásquez C; Correa E; Ramírez J	2001/03/19	Red de Niebla	1
Avistamiento en campo	UCN	MH-CAMPO	46-2300MI2008MA0003	Vargas I	19/03/2001		1
Avistamiento en campo	UCN	MH-CAMPO	46-2300MI2008MA0030	Vester H; Castro B	Feb 10, 2001		1
Avistamiento en campo	UCN	MH-CAMPO	46-2300MI2008MA0093	Vester H; Castro Bernardo	2001		1
Avistamiento en campo	UCN	MH-CAMPO	46-2300MI2008MA0352	Arias J; Ramírez M	19/03/2001		1
En Colección	UCN	MH-ORNIT	46-2300MI2008AV0046	Velásquez C; Correa E; Ramírez J	20/03/2001	Red de Niebla	1
En Colección	UCN	MH-ORNIT	46-2300MI2008AV0066	Velásquez C; Correa E; Ramírez J	Mar 2001	Red de Niebla	1
En Colección	UCN	MH-ORNIT	46-2300MI2008AV0118	Marulanda O; Montoya G; Toro G; Giraldo A	2001/03/19	Red de Niebla	1
En Colección	UCN	MH-ORNIT	46-2300MI2008AV0130	Marulanda O; Montoya G; Toro G; Giraldo A	Feb 2001	Red de Niebla	1
En Colección	UCN	MH-ORNIT	46-2300MI2008AV0145	Vargas I	May 2001	Red de Niebla	1
En Colección	UCN	MH-ORNIT	46-2300MI2008AV0197	Benavides A	Mar 19, 2001	Red de Niebla	1
En Colección	UCN	MH-ORNIT	46-2300MI2008AV0223	Benavides J; Trujillo P	2001	Red de Niebla	1
En Colección	UCN	MH-ORNIT	46-2300MI2008AV0255	Vargas I	19/03/2001	Red de Niebla	1
En Colección	UCN	MH-ORNIT	46-2300MI2008AV0287	Roldán F; Marulanda O; Flórez M; Osorio L	19/03/2001	Red de Niebla	1
En Colección	UCN	MH-ORNIT	46-2300MI2008AV0322	Velásquez C; Correa E; Ramírez J	Feb 10, 2001	Red de Niebla	1
En Colección	UCN	MH-ORNIT	46-2300MI2008AV0343	Galeano M; Martínez X; Matapí A; Anzola F; Pavón M	2001	Red de Niebla	1
En Colección	UCN	MH-ORNIT	46-2300MI2008AV0390	Fonnegra R	19/03/2001	Red de Niebla	1
En Colección	UCN	MH-ORNIT	46-2300MI2008AV0429	D Giraldo	Mar 2001	Red de Niebla	1
En Colección	UCN	MH-ORNIT	46-2300MI2008AV0455	Beltran N	Mar 19, 2001	Red de Niebla	1
En Colección	UCN	MH-ORNIT	46-2300MI2008AV0549	Forero E; Jaramillo R	Mar 19, 2001	Red de Niebla	1

45 elementos DwC

DwC, cómo abordarlo?

2. Generar plantillas DwC (definido por el publicador)

DwC Excel Template Generator

Español Set language

Check the boxes next to the terms you want to include in your template and click the **Create template** button.

[Darwin Core Terms: A quick reference guide](#)

RECORDLEVEL

- modified
- institutionID
- collectionID
- datasetID
- institutionCode
- collectionCode
- datasetName
- ownerInstitutionCode
- basisOfRecord
- informationWithheld

collectionID [ID de la colección]

Un identificador de la colección o conjunto de datos del que procede el registro. Para especímenes físicos, la práctica recomendada es utilizar el identificador de un índice de colección tal como el Índice de Colecciones de Biodiversidad (<http://www.biodiversitycollectionsindex.org/>).

Ejemplo: "urn:lsid:biocol.org:col:34818"

<https://data.gbif.no/dwcexcel>

Generador de plantillas DwC – GBIF Noruega

Dwc, cómo abordarlo?

3. Exploración y conocimiento profundo del Dwc



Biodiversity Information Standards
T C W B

Darwin Core Terms: A quick reference guide

Title: Darwin Core Terms: A quick reference guide
Date Issued: 2009-02-12
Date Modified: 2015-06-02

Abstract: This document is a quick reference for all recommended Darwin Core terms. For complete historical term information, including version changes and pre-standard terms, see [HISTORY]. For a comparative table of elements from pre-standard versions of Darwin Core to the current terms in the standard, see [VERSIONS].

Contributors: John Wicczorek (MVZ), Markus Döring (GBIF), Renato De Giovanni (CRJA), Tim Robertson (GBIF), Dave Vieglais (KUNHM)

Legal: This document is governed by the standard legal, copyright, licensing provisions and disclaimers issued by the Taxonomic Databases Working Group.

Part of TDWG: <http://www.tdwg.org/standards/450/>
Standard:
Creator: Darwin Core Task Group
Identifier: <http://rs.tdwg.org/dwc/2015-03-19/terms/>
Latest Version: <http://rs.tdwg.org/dwc/terms/>
Replaces: <http://rs.tdwg.org/dwc/2014-11-08/terms/>

Document Status: Current Standard

1. Term Index

Audience: This document is targeted toward those who need a reference (lists and definitions) to the essential information about the current recommended Darwin Core terms.

This quick guide provides a list of all current terms of the Darwin core. The terms are organized by categories (in bold) in the index. The categories correspond to Darwin Core terms that are classes (terms that have other terms to describe them). The terms that describe a given class (the class properties) appear in the list immediately below the name of the category in the index. The index provides links to the term descriptions in the table below the index.

Names that begin with dcterms: are terms managed by the Dublin Core Metadata Initiative [DCMI], which are reproduced here and which may include a description and comments adapted for the biodiversity community. The original definition of such terms can be found by following the link given in the Details section for the term.

Terms under the Record-level Terms section apply to the whole record regardless of the record type. For example, a record for a camera trap bird sighting would fall into the Dublin Core dcterms:type category StillImage and at the same time into the Darwin Core basisOfRecord category MachineObservation - a particular type of Occurrence, so the value of dcterms:type term for the record would be "StillImage" and the value of the basisOfRecord term would be "MachineObservation" to show which kind of Occurrence record it is. All of the rest of the record-level terms would apply to the Occurrence record.

The terms found under the Auxiliary Terms section can only be meaningfully implemented in an application that supports relational structures. For example, none of the auxiliary terms will be found in the specification of the Simple Darwin Core [SIMPLEDWC], which is "flat" in structure.

Record-level Terms

dcterms:type | dcterms:modified | dcterms:language | dcterms:license | dcterms:rightsHolder | dcterms:accessRights | dcterms:bibliographicCitation | dcterms:references | institutionID | collectionID | datasetID | institutionCode | collectionCode | datasetName | ownerInstitutionCode | basisOfRecord | informationWithheld | dataGeneralizations | dynamicProperties

Occurrence

occurrenceID | catalogNumber | recordNumber | recordedBy | individualCount | organismQuantity | organismQuantityType | sex | lifeStage | reproductiveCondition | behavior | establishmentMeans | occurrenceStatus | observations | disposition | associatedMedia | associatedReferences | associatedSequences | associatedTaxa | otherCatalogNumbers | occurrenceRemarks

Organism

organismID | organismName | organismScope | associatedOccurrences | associatedOrganisms | previousIdentifications | organismRemarks

MaterialSample | LivingSpecimen | PreservedSpecimen | FossilSpecimen

materialSampleID

Event | HumanObservation | MachineObservation

<http://rs.tdwg.org/dwc/terms/index.htm>

DwC, cómo abordarlo?

3. Exploración y conocimiento profundo del DwC

Term Name	reproductiveCondition
Identifier:	http://rs.tdwg.org/dwc/terms/reproductiveCondition
Class:	http://rs.tdwg.org/dwc/terms/Occurrence
Definition:	The reproductive condition of the biological individual(s) represented in the Occurrence. Recommended best practice is to use a controlled vocabulary.
Comment:	Examples "non-reproductive", "pregnant", "in bloom", "fruit-bearing". For discussion see http://terms.tdwg.org/wiki/dwc:reproductiveCondition
Details:	reproductiveCondition
Term Name	behavior
Identifier:	http://rs.tdwg.org/dwc/terms/behavior
Class:	http://rs.tdwg.org/dwc/terms/Occurrence
Definition:	A description of the behavior shown by the subject at the time the Occurrence was recorded. Recommended best practice is to use a controlled vocabulary.
Comment:	Examples: "roosting", "foraging", "running". For discussion see http://terms.tdwg.org/wiki/dwc:behavior
Details:	behavior

DwC, cómo abordarlo?

3. Exploración y conocimiento profundo del DwC

I'M ONE WITH
THE FORCE
AND THE FORCE
IS WITH ME

Vámos con la opción 2

Generar plantilla DwC (definido por el publicador)

Plantilla DwC Simple

41 elementos DwC





Gbif.es



CESP: Regional Capacity Enhancement to Latinamerica by establishing Chile's node
23 al 25 abril 2018
Santiago - Chile

Muchas gracias por su atención

Preguntas?



Leonardo Buitrago
Administrador Nodo GBIF Colombia,
Representante (S) Regional de Nodos GBIF (Latinoamérica y Caribe)
Líder Administración de contenidos SiB Colombia
albuitrago@humboldt.org.co

Transferencia de capacidades técnicas asociadas a la publicación de datos de biodiversidad a través de GBIF