ESTRUCTURACIÓN A DARWIN CORE

Taller GBIF.ES: Publicación de datos de biodiversidad en GBIF y en revistas científicas



INTRODUCCIÓN

Al compartir datos sobre biodiversidad, es importante utilizar ciertas normas comunes para facilitar el entendimiento y uso de los datos. Una de las normas más utilizadas es es estándar Darwin Core. En este ejercicio realizaremos un mapeo de campos con nombres originales a los términos correspondientes en el estándar.

Nos pondremos como objetivos del ejercicio los siguientes puntos:

* Familiarizarse con los términos del estándar Darwin Core y sus definiciones.
* Lograr un correcto mapeo de los campos contenidos en las planillas de datos a términos en el estándar Darwin Core.
* Discutir potenciales diferencias entre mapeos y conocer las mejores prácticas llevadas a cabo en la comunidad.
* Discutir las opciones de mapeo de otros campos presentes en distintas bases de datos de biodiversidad.

ANTES DE EMPEZAR

El [estándar Darwin Core](http://rs.tdwg.org/dwc/terms/index.htm) es desarrollado y mantenido por la comunidad a través del [Biodiversity Information Standards](http://www.tdwg.org/) (conocido también como TDWG). El estándar consiste en una lista de términos y sus definiciones para compartir datos sobre biodiversidad. Es importante familiarizarse con el estándar y utilizarlo, siguiendo sus definiciones y en lo posible también las recomendaciones que provee sobre el uso de determinados valores para ciertos campos (i.e., vocabularios controlados).

NECESITARÁS

El juego de datos de prueba[*Caso 01\_EstructuraciónDarwinCore\_PecesMexico.csv*](https://drive.google.com/open?id=1aaQSBnahu4c-2iNTburGIHzM82TUj2nMrF66TvPlj0I)*.*

*La planilla tiene formato .csv. Descárguela e impórtela en MS Excel. (No trabaje desde el archivo en Google Drive para no superponerse con el trabajo de otro).*

URL del estándar [Darwin Core](http://rs.tdwg.org/dwc/terms/index.htm); <http://rs.tdwg.org/dwc/terms/index.htm>

INSTRUCCIONES

1. Inspeccione la planilla de datos provista. Identifique los nombres de los campos y su contenido general.
2. Mapee cada campo a un término en Darwin Core (puede hacerlo agregando una fila por encima del encabezado o reemplazando los nombres de los campos).
3. En los casos en que no se pueda realizar un mapeo directo, indique qué acciones tomaría sobre los campos para poder realizar un mapeo correcto (e.g., dividir un campo en dos, concatenar campos, etc.).
4. Haga una lista de otros posibles campos que deberían estar presentes para poder compartir los datos (puede incluir algunos derivados del punto anterior).

PARA SABER MÁS

Para familiarizarte más con el estándar Darwin Core, puedes visitar los siguientes sitios:

- **Página del estándar:** <http://rs.tdwg.org/dwc/terms/index.htm>

- **Darwin Core Questions & Answers Site:** <https://github.com/tdwg/dwc-qa>

- **Wieczorek et al. 2012. Darwin Core**: An Evolving Community-Developed Biodiversity Data Standard. Plos One, [**https://doi.org/10.1371/journal.pone.0029715**](https://doi.org/10.1371/journal.pone.0029715)

- **Videos:**

- Estándares: generalidades: <https://www.youtube.com/watch?v=9Y2KgUDag2M&t=8s>

- Estándares: Darwin Core: <https://www.youtube.com/watch?v=a01GktP0M_Q>