

# Costes asociados a la generación de un recurso de datos

David Galicia [dgalicia@unav.es](mailto:dgalicia@unav.es)  
Universidad de Navarra



Taller online GBIF.ES: Estandarización y publicación de datos de biodiversidad de regiones polares. Mayo 2021.



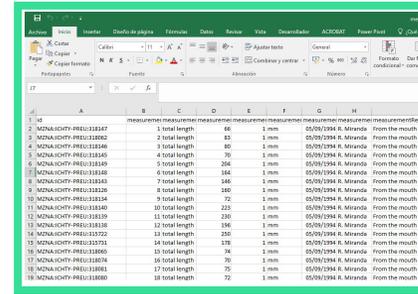
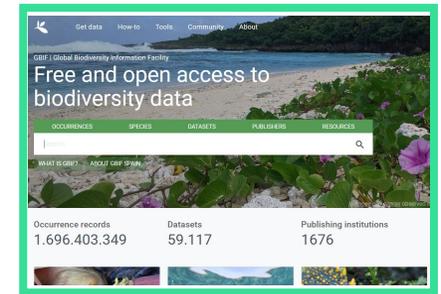
# Contenidos

- Tipos de fuentes de metadatos más habituales.
- Tipologías de datos en relación a su coste de digitalización.
- El flujo de la información en la investigación.



# El camino hasta generar un recurso de información

Generar un recurso de datos (un conjunto de información que sea útil para ser utilizado por terceras personas) es un camino largo y costoso.

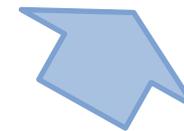
A screenshot of a Microsoft Excel spreadsheet showing a table of specimen data. The table has columns for ID, total length, sex, date, and location. The data is organized in a grid format with multiple rows and columns.

1. Recopilación de metadatos

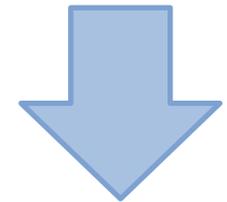
2. Digitalización y conservación

3. Depuración de la información

4. Accesibilidad universal



Gestión de Colecciones de Investigación  
(Museos y Centros de Investigación)



**RECURSO**

# El camino hasta generar un recurso de información

Generar un recurso de datos (un conjunto de información que sea útil para ser utilizado por terceras personas) es un camino largo y costoso.

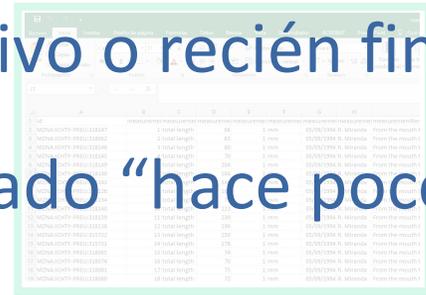


1. Recopilación de metadatos



- Proyecto en activo o recién finalizado.
- Proyecto finalizado “hace poco”.

2. Digitalización y conservación



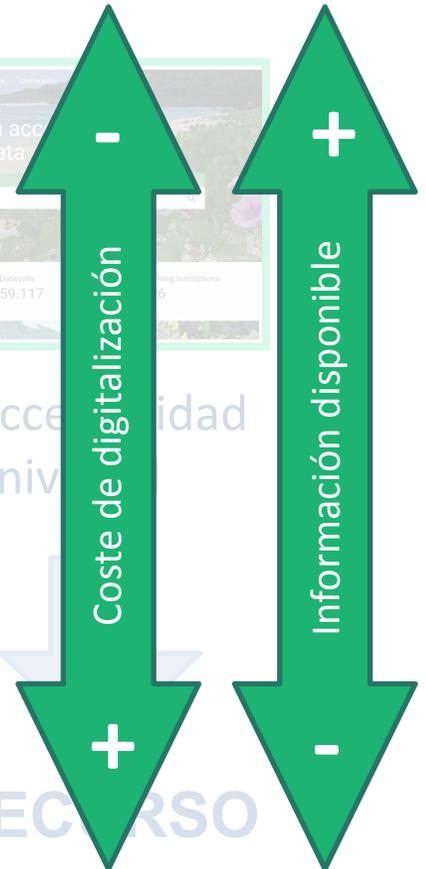
3. Depuración de la información

- Fuente única: donaciones, proyectos tesis...
- Fuente múltiple: cajón de sastre.



4. Acceso universal

Gestión de Colecciones de Investigación  
(Museos y Centros de Investigación)



RECURSO

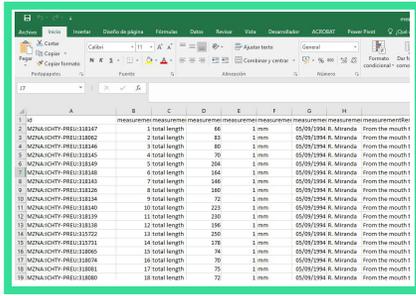
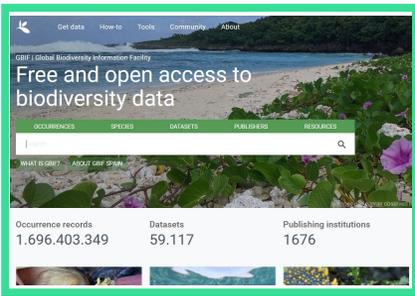
# El camino hasta generar un recurso de información

Generar un recurso de datos (un conjunto de información que sea útil para ser utilizado por terceras personas) es un camino largo y costoso.



# El camino hasta generar un recurso de información

Generar un recurso de datos (un conjunto de información que sea útil para ser utilizado por terceras personas) es un camino largo y costoso.

- La depuración, limpieza y validación de los datos está relacionada con su *Calidad*.  
  

  - La publicación de la información implica una *Estandarización* e incorporación a un *Repositorio*.  

1. *metadatos*      2. *conservación*
3. Depuración de la información
4. Accesibilidad universal
- Gestión de Colecciones de Investigación  
(Museos y Centros de Investigación)
- RECURSO

# El camino hasta generar un recurso de información

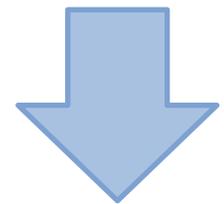
Generar un recurso de datos (un conjunto de información que sea útil para ser utilizado por terceras personas) es un camino largo y costoso.



- Las posibilidades de que el recurso sea utilizado están muy relacionadas con la calidad de los propios **datos** de biodiversidad que estamos sirviendo y de los **metadatos** que los acompañan

- Unos metadatos de calidad implican una adecuada *organización* en la generación de recursos.

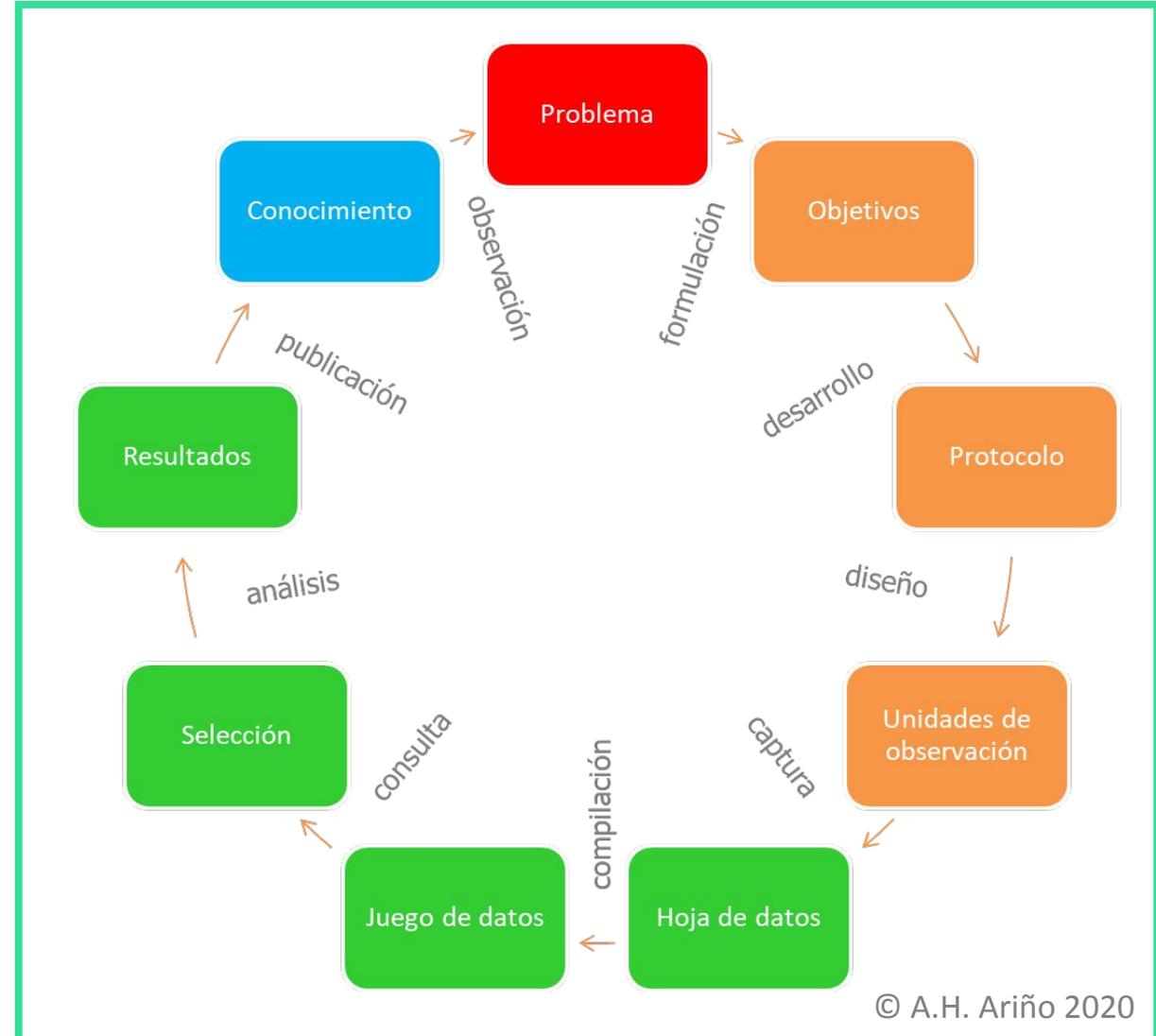
Gestión de Colecciones de Investigación  
(Museos y Centros de Investigación)



**RECURSO**

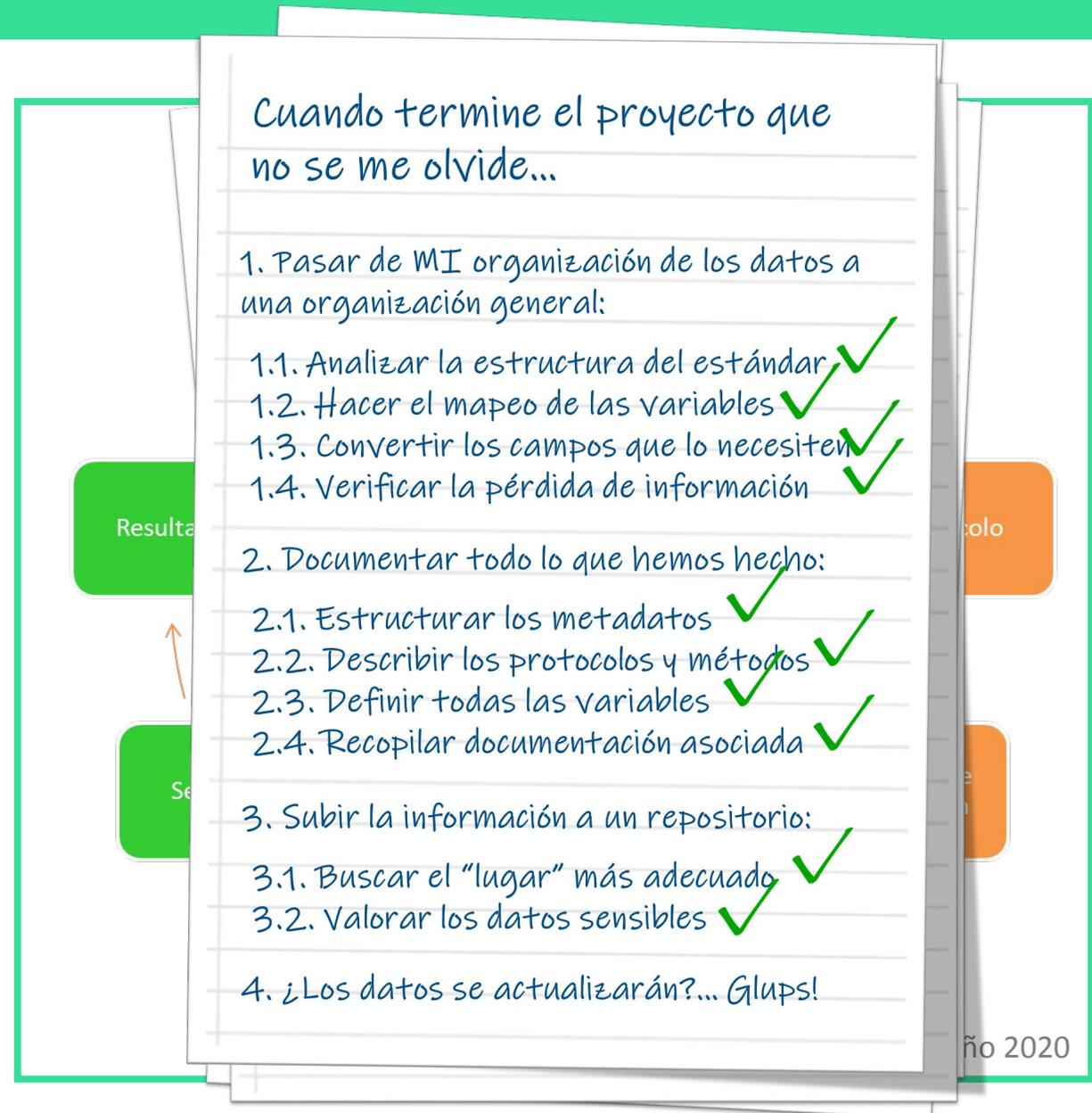
# El ciclo de la vida de un proyecto de investigación

- Recopilamos información para enriquecer nuestro conocimiento del mundo que nos rodea.
- El Método Científico requiere tomar datos que permitan refutar las hipótesis planteadas.
- El ciclo de vida del proyecto finaliza con la *publicación de los resultados* que abren la posibilidad de plantear nuevas hipótesis.



# El ciclo de la vida de un proyecto de investigación

- Hasta ahora, el foco de atención de la publicación de información está en los *resultados*, no en los *datos*.
- El prestigio y reconocimiento viene de la publicación de artículos científicos (resultados) en revistas de investigación de “alto impacto”.
- La publicación de los datos que han dado lugar a esos resultados es considerado un esfuerzo *extra*.



# El ciclo de la vida de un proyecto de investigación

- Hasta ahora, el foco de atención de la publicación de información está en los resultados, no en los datos.
  - En proyectos que generan registros de Historia Natural esta práctica repercute en un crecimiento de la cantidad de registros pendientes de movilizar.
- El prestigio y reconocimiento viene de la publicación de artículos científicos (resultados) e impactos de investigación de alto impacto.
  - Hay que incorporar la generación de recursos de datos en el diseño y presupuesto de los proyectos y en los protocolos de gestión de las colecciones.
- La publicación de los datos que han dado lugar a esos resultados es considerado un esfuerzo *extra*.

Quando termine el proyecto que no se me olvide...

1. Pasar de MI organización de los datos a una organización general:

1.1. Analizar la estructura del estándar ✓

1.2. Establecer la estructura de las variables ✓

1.3. Convertir los campos que lo necesiten ✓

1.4. Crear un protocolo de información ✓

2. Documentar todo lo que hemos hecho:

2.1. Crear un protocolo de datos ✓

2.2. Describir los protocolos y métodos ✓

2.3. Crear un protocolo de datos ✓

2.4. Recopilar documentación asociada ✓

3. Subir la información a un repositorio:

3.1. Buscar el "lugar" más adecuado ✓

3.2. Valorar los datos sensibles ✓

4. ¿Los datos se actualizarán?... Glups!

# Taller online GBIF.ES: Estandarización y publicación de datos de biodiversidad de regiones polares

¡Gracias!

[dgalicia@unav.es](mailto:dgalicia@unav.es)

