

Compartiendo datos en Ecología: cómo añadir más valor a los datos

A. J. Pérez-Luque^{1,2,*}, A. Ros-Candeira^{1,2}

(1) Laboratorio de Ecología, Instituto Interuniversitario de Investigación del Sistema Tierra (IIISTA-CEAMA), Universidad de Granada, Avda. del Mediterráneo s/n, Granada 18006, España.

(2) Grupo de Ecología Terrestre, Departamento de Ecología, Universidad de Granada, Facultad de Ciencias, Campus Fuentenueva s/n, 18071, Granada, España.

* Autor de correspondencia: A. J. Pérez-Luque [ajperez@ugr.es]

> Recibido el 23 de agosto de 2019 - Aceptado el 25 de octubre de 2019

Pérez-Luque, A.J., Ros-Candeira, A. 2019. Compartiendo datos en Ecología: cómo añadir más valor a los datos. *Ecosistemas* 28(3):150-159. Doi.: 10.7818/ECOS.1838

Los datos constituyen uno de los productos valiosos de la ciencia. Teniendo en cuenta la importancia de las series de datos para comprender patrones ecológicos complejos y/o resolver problemas ambientales emergentes, su preservación, accesibilidad y reutilización en ecología resulta crucial. Aunque la mayoría de los investigadores reconocen la importancia de compartir datos, existen diferentes barreras sociales y tecnológicas que limitan la puesta en valor de los mismos, ralentizando posibles avances en la investigación en ecología. En este trabajo revisamos las distintas prácticas de publicación de datos en ecología y presentamos una serie de sugerencias para añadir más valor a los datos a través de su documentación y su publicación. Asimismo, presentamos un caso de estudio en el que mostramos diferentes aproximaciones a la hora de compartir datos. A lo largo de la revisión incluimos también diferentes recursos y herramientas que pueden ayudar a maximizar el valor de los datos que generamos y a facilitar el acceso a potenciales usuarios.

Palabras clave: publicación de datos; repositorios de datos; data papers; metadatos; principios FAIR

Pérez-Luque, A.J., Ros-Candeira, A. 2019. Sharing data in Ecology: how to add more value to data. *Ecosistemas* 28(3):150-159. Doi.: 10.7818/ECOS.1838

Data is one of the valuable products of science. Their preservation, accessibility and reuse in ecology is crucial, given the importance of data sets for understanding complex ecological questions and/or solving emerging environmental problems. Although most researchers recognize the importance of sharing data, there are different social and technological barriers that limit the value of data, slowing down possible advances in ecological research. In this paper we review the different practices of data publication in ecology and present a series of ideas to add "more" value to the data through their documentation and publication. We also present a case study showing different approaches to data sharing. Throughout the review we included different resources and tools to help us maximize the value of the data we generate and facilitate access to other users.

Key words: data publishing; data repository; data papers; metadata; FAIR principles

Introducción

La ciencia puede concebirse como una empresa social cuya finalidad es la acumulación de conocimiento como bien público (Munafò et al. 2017). Uno de los productos de esta gran empresa es, sin duda, los datos. Podría incluso afirmarse que se trata de uno de sus productos más valiosos, teniendo en cuenta que constituyen la base para la generación de información y conocimiento científico (Costello et al. 2013). Las series de datos a largo plazo permiten probar y desarrollar teorías, así como comprender patrones ecológicos y evolutivos complejos que a menudo surgen en amplias escalas temporales y espaciales (Kuebbing et al. 2018). En ecología, la importancia del acceso libre a los datos (*open data*) se ha reconocido más en algunos campos como la genómica o la biogeografía, mientras que muchos otros aún no han conseguido incorporar plenamente la filosofía de hacer disponibles sus datos (Culina et al. 2018).

En la transición hacia una ciencia abierta, la disponibilidad de los datos se perfila como un aspecto inherente a la reproducibilidad (Rodríguez-Sánchez et al. 2016), pues contribuye directa y positivamente a que un estudio científico sea más reproducible. Los beneficios de compartir datos están bien identificados y concretados (e.j. Whitlock 2011; Torres-Salinas et al. 2012; Hampton et al. 2015; Soranno 2019); entre otros figuran:

- Mejor aprovechamiento de los recursos invertidos en ciencia evitando, por ejemplo, la duplicación de proyectos para obtener datos similares.
- Optimización de metaanálisis y resolución de nuevas cuestiones.
- Reducción del riesgo de pérdida de los datos.
- Aumento de la transparencia en ciencia, permitiendo la reproducibilidad de análisis y verificación de resultados.
- Incremento en el crédito de los proveedores de datos.

Sin embargo, se necesita una mejora en la eficiencia del uso de los datos y pensar no solo en el uso concreto para el cual se han obtenido (Hampton et al. 2015; Díaz-Delgado 2016). En este sentido, en 2016 se llevó a cabo un esfuerzo por parte de la comunidad científica para establecer unos principios que permitan una gestión eficaz de los datos. Este marco de trabajo, conocido como FAIR (*Findable, Accessible, Interoperable, and Reusable*) (Wilkinson et al. 2016), ofrece una serie de recomendaciones para asegurar que los datos sean encontrables, accesibles, interoperables y reutilizables. Encontrables (*Findable*) se refiere a que los datos y los metadatos puedan ser descubiertos por la comunidad científica. Para ello se requiere que los datos y/o metadatos sean identificados de forma única y persistente, así como que estén indexados y descritos en detalle. Para que los datos sean accesibles (*Accessible*) una vez encontrados es necesario que se publiquen utilizando protocolos estandarizados. La utilización de estándares de intercambio permite así que los datos sean interoperables (*Interoperable*). Por último, los datos deben de ser reutilizables (*Reusable*), asegurando que se conserva su procedencia así como las condiciones de reutilización. En definitiva, los principios FAIR promueven que los datos, además de estar accesibles, estén bien documentados para poder reutilizarse, lo cual incrementará su valor original. Sin embargo, en la actualidad todavía muchos investigadores se encuentran con una baja accesibilidad a datos existentes, lo que ralentiza posibles avances en investigación en ecología (Wolkovich et al. 2012). En este contexto, la puesta en valor de los datos cobra especial relevancia.

¿Qué medidas concretas podemos adoptar para maximizar el valor de los datos que generamos y facilitar el acceso a los demás usuarios? En este trabajo revisamos algunas de las prácticas de publicación de datos en ecología y presentamos algunas sugerencias para añadir más valor a los datos a través de su documentación (metadatos) (Fegraus et al. 2005), su publicación (Costello et al. 2013) y la elaboración de artículos de datos (*data papers*) (Chavan y Penev 2011).

La publicación de datos: más allá del material suplementario

En los trabajos científicos, los conjuntos de datos suponen el punto de partida para la realización de análisis y la obtención de resultados. Sin embargo, aún prevalece la idea equivocada de que un proyecto de investigación finaliza cuando se escribe y publica un artículo científico (Goodman 2014), dejando en segundo plano el *qué hacer* con los datos originales y su relevancia para la reproducibilidad del estudio. Lo cierto es que la gran mayoría de los estudios científicos no son reproducibles y, en ocasiones, incluso la trazabilidad de los resultados es un misterio para los propios autores (Rodríguez-Sánchez et al. 2016). Una deficiente documentación y disponibilidad de los datos en las publicaciones científicas obstaculiza la evaluación crítica de sus resultados (Goodman 2014). Ante esta situación, incluir los datos como material suplementario de un artículo resulta insuficiente por diversas razones. A modo general, porque la información queda huérfana al no integrarse con otros conjuntos de datos que sustentan el conocimiento científico sobre un tema, dificultando su descubrimiento (Michener 2015); y a efectos prácticos, porque no puede garantizarse la disponibilidad de los datos a largo plazo, quedando expuestos al devenir de la editorial o revista que los aloja. La baja accesibilidad que afecta a los datos incluidos como material suplementario se debe a diversas causas: la variedad de formatos utilizada por los autores para enviar los datos; la falta de estándares que armonicen los conjuntos de datos; la inexistencia de herramientas de búsqueda que posibiliten su localización, etc. (Santos et al. 2005). Resulta evidente que incluir los datos de este modo no sustituye, de ninguna manera, su depósito en un repositorio público. En el caso concreto de la ecología, un incremento en la accesibilidad a los

datos requiere la superación de dos retos: el tecnológico y el social (Reichman et al. 2011). El primer aspecto guarda relación con la complejidad inherente a la información ecológica, mientras que el segundo refleja la necesidad de superar ciertas reticencias sociales todavía presentes a la hora de compartir datos. Esto último es especialmente paradójico cuando se trata de investigaciones financiadas con fondos públicos, ya que la difusión de los datos se ve restringida por una malentendida propiedad sobre los mismos por parte de investigadores o grupos de investigación (Caldera-Serrano 2018). Autores como Couture et al. (2018) señalan el ambiente de competitividad para conseguir financiación y el hecho de reservarse futuras oportunidades de publicación como algunas de las causas sociales para no compartir datos. Michener (2015), en un análisis menos personalista, señala también como factores causantes la falta de apoyo institucional, la ausencia de recursos y experiencia en la gestión de datos o la complejidad de los derechos de propiedad intelectual.

En este sentido, Michener (2015) identifica a las revistas y editoriales, a los investigadores y a las entidades financiadoras como los tres impulsores clave de este cambio social. Los financiadores tienen más poder que las revistas para exigir el intercambio de datos a los investigadores, puesto que comparten, en cierto modo, "propiedad" sobre ellos (Couture et al. 2018). Sin embargo, no podemos obviar el gran protagonismo que las revistas tienen en la difusión (y evaluación) de los resultados científicos en el modelo de Ciencia actual (Caldera-Serrano 2018). En el caso de éstas últimas, pueden apoyar el cambio adoptando (y haciendo cumplir) políticas de datos que exijan a los autores la publicación de los datos en abierto e, incluso, fomentando la creación de revistas de datos (Tabla 1) (Michener 2015).

Una iniciativa importante tuvo lugar hace 9 años, cuando varias revistas clave en el campo de la evolución y la ecología adoptaron una nueva política de datos llamada *Joint Data Archiving Policy* (Whitlock 2011), la cual introducía como requisito indispensable para la publicación del artículo la accesibilidad a los datos que sustentaban la investigación. El panorama actual es variado, pero los cambios en las políticas de datos de las revistas muestran una tendencia general hacia una mayor disponibilidad de datos (Stodden 2013). Para hacernos una idea, en una revisión de las políticas de datos de las revistas de mayor impacto en el campo de la ecología (Tabla 2), hemos observado la existencia de un apoyo explícito (ej.: *Ecology Letters*) e implícito (ej.: *Methods in Ecology and Evolution* y *Ecological Monographs*) de algunas revistas a la reproducibilidad, poniendo énfasis en la necesidad de hacer públicos los datos para su reutilización, y exigiendo su depósito en un repositorio público como condición de publicación. Algunas, de hecho, son críticas con el uso del material suplementario como alternativa. Sin embargo, llama la atención el diferente grado de expectativa a la hora de compartir datos. Por ejemplo, en las ciencias ómicas (ej. genómica, proteómica, metabolómica, etc) se espera que los investigadores depositen ciertos datos, como las secuencias de ADN, mientras que es opcional para otro tipo de datos (ej. *The ISME Journal*). Este "trato diferente" evidencia que algunas disciplinas ya han superado los retos tecnológicos y sociales que en ecología aún son tarea pendiente. Un ejemplo vivo de este cambio de paradigma se manifiesta en el nacimiento de repositorios como GenBank, que almacena y comparte un gran volumen de secuencias genéticas (Reichman et al. 2011).

La proliferación de repositorios ha jugado un papel crucial en la disponibilidad de datos. Esta ingente cantidad de datos, que no se manejan fácilmente con las herramientas y prácticas habituales, ofrece oportunidades sin precedentes para el avance científico. En este sentido, la aplicación de técnicas a grandes volúmenes de datos (e.j. *Big Data*) puede servir para generar nuevo conocimiento útil para la toma de decisiones frente a problemas ambientales complejos (Hampton et al. 2013; LaDeau et al. 2017).

Tabla 1. Ejemplo de revistas que aceptan artículos de datos como modalidad de publicación. En negrita se indican las revistas de datos. Se indica si el acceso a la revista es Open Acces (OA) o híbrido Open Choice (H). Factor de impacto referido al 2017 (JCR, Clarivate Analytics 2018). Para el campo Repositorios los números indican: 1: listado amplio de repositorios generalistas y específicos; 2: escasos repositorios recomendados; 3: ningún repositorio recomendado; 4: recomienda consultar un directorio de repositorios; 5: aporta una lista de repositorios ya utilizados en los artículos de la revista como ejemplo.

Table 1. Example of journals that accept data papers as publication modality. Data journals are indicated in bold. The table indicates whether the access to the journal is Open Acces (OA) or hybrid Open Choice (H). The impact factor is referred to the year 2017 (JCR, Clarivate Analytics 2018). Numbers in the "Repositorios" column indicate: 1: a long list of generalist and specific repositories; 2: few recommended repositories; 3: no recommended repository; 4: it is recommended to consult a repository directory; 5: a list of repositories already used in the articles of the journal is provided as an example.

Revista	Acceso	Categorías JCR	Factor de Impacto	Tipo de Artículo	Cuartil	Nº total de artículos de datos publicados	Primer año de publicación de artículos de datos	Ejemplo	Instrucciones para autores y secciones del artículo	Repositorios	URL's útiles
Scientific Data	OA	Multidisciplinary sciences	5.311	Data Descriptor	Q1	680	2014	Escribano et al. (2018)	Estructura del manuscrito; Plantilla	1	Repositorios recomendados
Earth System Science Data	OA	Geosciences, Multifisciplinary Meteorology & Atmospheric sciences	8.792	Data Paper	Q1		2009	Serrano-Notivoli et al. (2017)	Estructura del manuscrito; Plantilla	5	Criterios de repositorio
Geoscience Data Journal	OA	Geosciences, Multifisciplinary Meteorology & Atmospheric sciences	1.867	Data Paper	Q3		2014	Batalla et al. (2018)	Estructura del manuscrito; Plantilla	1	Centros de datos o repositorios
Biodiversity Data Journal	OA			Data Paper		87	2013	Vízoso and Quesada (2015)	Penev et al. 2017	1	Normas para la publicación de datos
Data	OA			Data Descriptor		70	2016	Vicente-Serrano et al. (2017)	Estructura del manuscrito; Plantilla	4	Depósito de datos
Data in Brief	OA			Data Article		3950	2015	Torca et al. (2019)	Estructura del manuscrito; Plantilla	1	Repositorios de datos aprobados
Annals of Forest Science	H	Forestry	2.357	Data Paper	Q1		2016	Rebetez et al. (2018)	Estructura del manuscrito; Plantilla	2	Depósito de datos
ZooKeys	OA	Zoology	1.079	Data Paper	Q3	76	2011	Pérez-Luque et al. (2016a)	Penev et al. 2017	1	Normas para la publicación de datos
BMC Ecology	OA	Ecology	2.315	Database Article	Q2	3	2007	Tonk et al. (2013)	Estructura del manuscrito	2	Repositorio de datos
Biorisk	OA			Data Paper		0			Penev et al. 2017	1	Normas para la publicación de datos
Ecology	H	Ecology	4.617	Data Paper *solo el abstract se publica en Ecology	Q1			Bovendorp et al. (2017)	Estructura del manuscrito	2	
Nature Conservation	OA	Biodiversity conservation	1.367	Data Paper	Q3	5	2013	Grimm et al. (2014)	Penev et al. 2017	1	Normas para la publicación de datos
PLoS One	OA	Multidisciplinary sciences	2.766	Submissions describing databases	Q1			Marshall et al. (2018)	Estructura del manuscrito	1	Repositorios recomendados
BiolInvasions Records	OA	Biodiversity conservation	1.189	Data Paper	Q3			Semenchenko et al. (2016)	Estructura del manuscrito	2	
BMC Plant Biology	OA	Plant sciences	3.930	Database Article	Q1		2005	Wang et al. (2009)	Estructura del manuscrito	3	
PhytoKeys	OA	Plant sciences	1.393	Data Paper	Q3	18	2012	Pérez-Luque et al. (2015)	Penev et al. 2017	1	Normas para la publicación de datos
Botanical Studies	OA	Plant sciences	1.411	Database Article	Q3	0			Estructura del manuscrito	1	Repositorios recomendados
Global Ecology and Biogeography	H	Ecology Physical geography	5.958	Data Paper	Q1			Storchová and Hořák (2018)	Estructura del manuscrito	3	
GigaScience	OA	Multidisciplinary sciences	7.267	Data Note	Q1	162	2012	Yang et al. (2018)	Estructura del manuscrito	1	Repositorio de datos

Tabla 2. Requerimientos en materia de datos de las revistas con mayor impacto en ecología. Búsqueda de revistas en Web of Science [consultado 2019/01/23] y revisión de política de datos para aquellas cuya publicación principal es el artículo de investigación. Obligatoriedad de la política de datos: verde: se obliga tanto a depositar los datos en un repositorio como a incluir información sobre su accesibilidad; rosa: se obliga a incluir información sobre la accesibilidad de los datos y se alienta a los autores a depositarlos en un repositorio público; amarillo: se alienta a los autores a depositar los datos en un repositorio público y a incluir información sobre su accesibilidad.

Table 2. Data sharing requirements of the journals with the greatest impact on Ecology. A search of journals by rank using Web of Science [accessed 2019/01/23] and data policy review for those whose main publication is the research article. The data policy requirements: green: it is required to deposit the data in a repository and include a data accessibility statement; pink: it is required to include a data accessibility statement and the authors are encouraged to deposit the data in a public repository; yellow: it is encouraged to deposit the data in a public repository and to include information about its accessibility.

Ranking	Revista	Factor de Impacto	Tipo de Artículo Principal	URL's útiles. Criterios de repositorio	URL's útiles. Política de datos
1	Trends in Ecology and Evolution	15.938	Review		
2	Annual Review of Ecology Evolution and Systematics	10.160	Review		
3	ISME Journal	9.520	Research	Repositorios recomendados por Springer Nature	Política de datos de investigación
4	Ecology Letters	9.137	Research	Normas para los autores-Archivo de datos	Política de intercambio de datos de Wiley
5	Global Change Biology	8.997	Research		Normas para los autores-Intercambio de datos y accesibilidad
6	Frontiers in Ecology and the Environment	8.302	Communication		
7	Ecological Monographs	7.828	Research	Normas para los autores-Política de datos	Política de disponibilidad de datos
8	Molecular Ecology Resources	7.059	Resource		
9	Methods in Ecology and Evolution	6.363	Research	Política de datos y repositorios comunes	
10	Molecular Ecology	6.131	Research	Normas para los autores-Accesibilidad, almacenamiento y documentación de datos	

La variedad de repositorios a los que tenemos acceso los investigadores es muy alta (Tabla 3), desde repositorios más generalistas que almacenan datos de diversas ramas científicas, como por ejemplo Dryad; a otros más específicos, como GBIF (Global Biodiversity Information Facility) para datos de biodiversidad o el ITRDB (International Tree-Ring Data Bank) para datos dendrocronológicos. Ante esta gran cantidad de repositorios, existen herramientas que nos ayudan a identificar el repositorio más apropiado, por ejemplo OpenDOAR o RepositoryFinder, que realiza búsquedas en el registro internacional de repositorios de datos de investigación re3data. Organizaciones como la British Ecological Society proporcionan en las políticas de datos de sus revistas un listado de repositorios recomendados. Asimismo, existen listados de repositorios categorizados por disciplinas y por tipo de datos.

El impulso de las políticas que promueven el archivo público de datos junto con la existencia de repositorios institucionales de datos, están consiguiendo un aumento significativo de la accesibilidad a los datos. Sin embargo, el depósito de los datos *per se* no siempre asegura su reutilización, tal como han puesto de manifiesto Roche et al. (2015) quienes, tras evaluar la accesibilidad de los datos asociados a 100 publicaciones en Ecología y Evolución, encontraron que más de la mitad de los conjuntos de datos estaban incompletos o no estaban bien documentados.

Los metadatos: hableme de “tus” datos y...

Para promover la reutilización de los datos y la reproducibilidad, es necesario dotar a los datos de una mínima información adicional (metadatos) (Alonso y Valladares 2006). Este ejercicio de documentar los datos, además de constituir una buena defensa contra el olvido y el tiempo que pasa desde su colecta, ayuda a una interpretación correcta de los mismos y hace viable su reutilización por otros usuarios (Fegraus et al. 2005). Una documentación completa

de los datos implica un cambio en las prácticas de los investigadores, poniendo énfasis en su posible reutilización, concibiendo los datos como un elemento clave en la producción de conocimiento y no limitándose a la publicación de un artículo científico. Para ello, las autoras Baker y Millerand (2010) apelan a implementar un cambio en el tratamiento de datos, donde tradicionalmente los investigadores se limitan a su análisis, para poner mayor énfasis en su procesado y en la producción de metadatos, así como en su organización y preparación para publicación en un repositorio.

Para facilitar el intercambio, la integración y síntesis de los datos se han desarrollado diversos estándares de metadatos, en función del propósito y la disciplina científica (Tabla 4). En una revisión para el campo de la ecología, Alves et al. (2018) encontraron que los más utilizados son: Ecological Metadata Language (EML) (Fegraus et al. 2005), Darwin Core (Wieczorek et al. 2012); y la Directiva Europea INSPIRE para datos espaciales. Asimismo, para facilitar la tarea a los investigadores, algunas iniciativas como la Environmental Data Initiative, ofrecen plantillas para generar los metadatos de un conjunto de datos de acuerdo al estándar EML.

Para la creación de metadatos existe un amplio abanico de herramientas (Tabla 4). La red europea de seguimiento a largo plazo LTER (Long-Term Ecological Research) ha creado una herramienta llamada DEIMS-SDR (Dynamic Ecological Information Management System - Site and dataset registry) (Wohner et al. 2019), que permite documentar conjuntos de datos ecológicos y proporcionarlos en diferentes estándares (EML, Biological Data Profile, INSPIRE). En R existen varios paquetes que nos permiten trabajar con el estándar EML, entre ellos destacamos: EML (Boettiger et al. 2019) y emld (Boettiger 2019) para la creación de metadatos, y el paquete EMLAssemblyLine (Smith 2019) que incorpora además un flujo de trabajo. Asimismo, existen otros que facilitan la visualización de los metadatos creando páginas web sencillas para ello, como el paquete emldown.

Tabla 3. Listado de repositorios.**Table 3.** List of repositories.

Nombre	Enlace
Repositorios	
Dryad	https://datadryad.org
GBIF	https://www.gbif.org
ITRDB	https://www.ncdc.noaa.gov/data-access/paleoclimatology-data/datasets/tree-ring
Directorios de repositorios	
OpenDOAR	http://v2.sherpa.ac.uk/opendoar/
RepositoryFinder	https://repositoryfinder.datacite.org/
re3data	https://www.re3data.org/
Repositorios categorizados por disciplinas	
Repositorios recomendados por la revista Scientific Data	https://www.nature.com/sdata/policies/repositories
Repositorios categorizados por tipo de datos	
Thessen y Patterson 2011	https://doi.org/10.3897/zookeys.150.1766

Tabla 4. Herramientas de gestión de metadatos en Ecología.**Table 4.** Management tools for metadata in Ecology.

Estándares de Metadatos	Enlaces	Referencias
Ecological Metadata Language (EML)	http://knb.ecoinformatics.org/software/eml/	Fegraus et al. 2005
Darwin Core	https://dwc.tdwg.org/	Wieczorek et al. 2012
Directica INSPIRE	https://inspire.ec.europa.eu/	
Creación y Visualización de Metadatos		
DEIMS-SDR (Dynamic Ecological Information Management System - Site and dataset registry)	https://deims.org/	Wohner et al. 2019
paquete de R EML	https://ropensci.github.io/EML	Boettiger et al. 2019
emld	https://docs.ropensci.org/emld	Boettiger 2019
EMLAssemblyLine	https://github.com/EDlorg/EMLassemblyline	Smith 2019
emldown	https://github.com/ropenscilabs/emldown	
Plantilla de metadatos en EML	https://environmentaldatainitiative.org/welcome/submit-data/	Environmental Data Initiative

Algunos investigadores ya están incorporando esta aproximación de documentación de datos y depósito de los mismos en repositorios internacionales siguiendo el estándar EML. Por ejemplo, [Parmenter et al. \(2018\)](#), en un trabajo sobre factores ambientales y endógenos que afectan a la vejería en bosques de Nuevo México, realizaron una documentación de los datos mediante el estándar EML y su posterior integración dentro de repositorio de la Red LTER.

El *data paper*: conjugando datos y metadatos en una publicación científica

Ir un paso más allá de publicar los datos junto con los metadatos en un repositorio puede ser escribir un artículo de datos (*data paper*) sobre ese conjunto de datos. El *data paper* es una publicación científica que documenta detalladamente uno o varios conjuntos de datos accesibles, describiendo el contexto en el que fueron generados y su contenido. Respecto a su estructura, no sigue el esquema típico de un artículo científico, pues su propósito no es exponer una investigación, sino describir conjuntos de datos de forma entendible y estructurada. El *data paper* es, por tanto, un producto que pone en valor el recurso de datos que documenta, dándole visibilidad a su contenido y reconocimiento académico a sus

proveedores ([Chavan y Penev 2011](#)). [Costello et al. \(2013\)](#) resaltan esta modalidad de publicación como una garantía de accesibilidad y calidad de los datos y metadatos, ya que requiere una revisión por pares del producto. Esto implica que, en algunos casos, los artículos de datos cubren una de las carencias que frecuentemente presenta el proceso de revisión por pares de un artículo científico: la falta de revisión de los datos que subyacen a la investigación que se publica. A pesar de haber avanzado en sus políticas de datos, las revistas todavía no cuentan con directrices claras que asignen responsabilidades sobre quién debe revisar y asegurar la calidad de los datos ([Sholler et al. 2019](#)). En los últimos años, ha aumentado el número de revistas dedicadas exclusivamente a la publicación de *data papers* (*data journals*) y, con mayor frecuencia, revistas de larga trayectoria incorporan el *data paper* como modalidad de publicación ([Tabla 1](#)).

Como hemos visto a lo largo de estas tres secciones, existen distintos pasos ([Fig. 1](#)) que podemos llevar a cabo para mejorar la documentación y accesibilidad a los datos: desde otorgarles protagonismo publicándolos como un recurso propio y citable, hasta estandarizar sus metadatos. Ante la problemática de accesibilidad que enfrentamos, [Soranno \(2019\)](#) resume cómo los autores pueden incorporar seis sencillos pasos en su flujo de trabajo para compartir los datos asociados a artículos: decidir la autoría de los datos

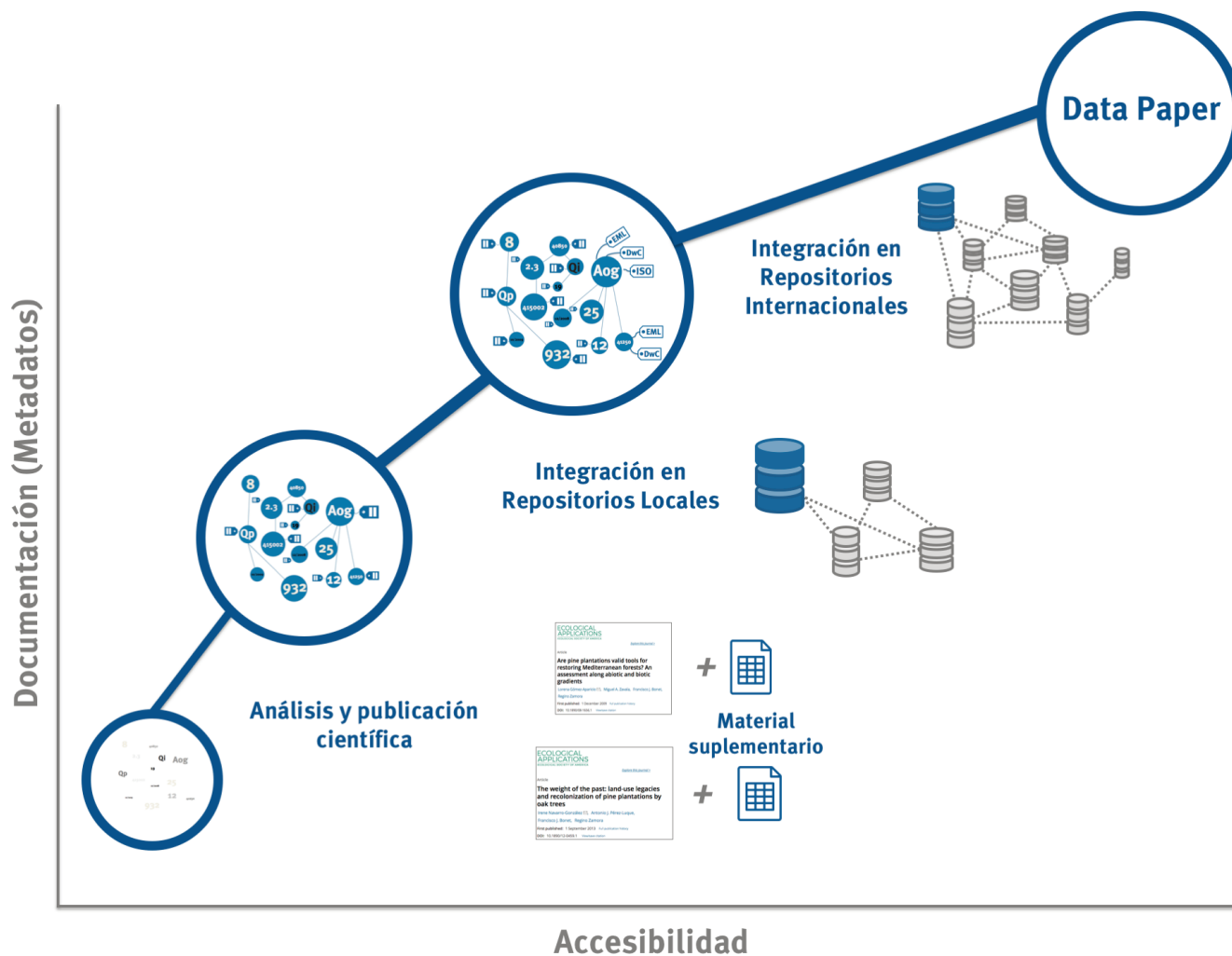


Figura 1. Acciones que incrementan la documentación y accesibilidad de los datos.

Figure 1. Actions that increase the documentation and accessibility of the data.

por un lado y la del artículo de investigación por otro; utilizar formatos simples para organizar los datos; escribir los metadatos; depositar los datos y los metadatos en un repositorio; escribir una declaración de disponibilidad de datos (*data availability statement*) en el artículo; y citar los datos tanto en la sección de Métodos como en la Bibliografía. Asimismo, algunas iniciativas como la Environmental Data Initiative, proporciona una guía con diferentes fases en la publicación los datos (<https://environmentaldatainitiative.org/resources/five-phases-of-data-publishing/>).

A continuación, presentamos un caso que ejemplifica cómo puede abordarse el manejo de los datos desde distintas aproximaciones, una de ellas más cercana a la reproducibilidad, la cual recomendamos ampliamente.

Manejo de datos sobre diversidad en los bosques de Sierra Nevada, transición de lo tradicional a lo reproducible

Sierra Nevada es una región montañosa situada en el sur de Europa, considerada un hotspot de diversidad vegetal en la cuenca Mediterránea. Alberga 2353 taxones de flora vascular, lo que constituye el 33.2 % de la flora de España (incluyendo las islas) (Lorite 2016). En esta región, que cuenta con una larga tradición de seguimiento de los procesos ecológicos a largo plazo (Pérez-Luque et al. 2016b), se realizó en el año 2004 un exhaustivo inventario forestal en 600 parcelas distribuidas a lo largo de gradientes ecológicos y altitudinales. Este inventario, conocido como Sinfonevada

(Pérez-Luque et al. 2014), ha sido utilizado por diversos investigadores para avanzar en el campo de la ecología de las repoblaciones (Apéndice 1), identificando los factores ecológicos clave que afectan a la regeneración natural bajo repoblaciones de coníferas, como por ejemplo la proximidad a la fuente semillera, las características intrínsecas de la plantación o la importancia del uso del suelo previo a la repoblación (Gómez-Aparicio et al. 2009; González-Moreno et al. 2011; Navarro-González et al. 2013).

Estos estudios han seguido lo que hemos denominado un manejo “tradicional” de los datos, compartiendo el siguiente hilo común: se llevó a cabo un análisis de los datos del inventario forestal (parte de ellos o en su totalidad) para testar algunas hipótesis, se publicaron los resultados en una revista científica y se añadieron los datos como material suplementario del artículo. Si bien se mejoró el acceso potencial a dichos datos a pesar de no estar depositados en ningún repositorio (179 citas de los tres trabajos que han utilizado directamente parte de este conjunto de datos) (Fig. 2a), su utilización estaba limitada por la escasa documentación sobre dicho inventario, que aparecía a lo sumo dispersa y no estructurada.

Con el fin de mejorar el acceso a ese conjunto de datos y aumentar la posibilidad de reutilización del mismo se llevó a cabo una normalización del inventario forestal y su posterior integración en repositorios de datos tanto locales como internacionales (Fig. 2b). Para ello, se realizó una documentación del conjunto de datos que permite conocer todos sus detalles (desde su génesis hasta su procesado) y, posteriormente, se integró en el Observatorio de Cambio

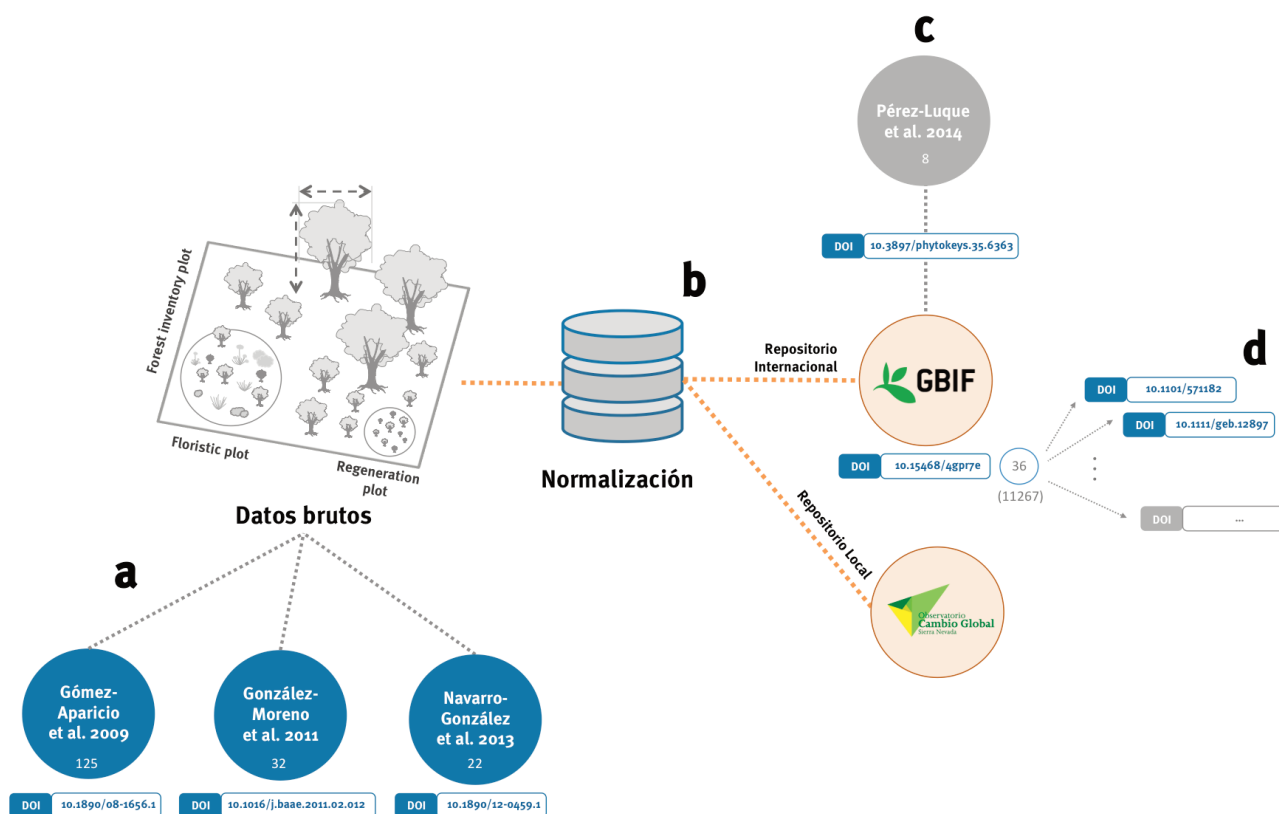


Figura 2. Tránsito del manejo de datos tradicional hacia el manejo de datos reproducible. Ejemplo de los datos del inventario forestal de Sierra Nevada (Pérez-Luque et al. 2014). **a:** publicación de artículos científicos con datos como material suplementario. **b:** normalización del conjunto de datos original e integración en repositorios locales e internacionales. **c:** Publicación de los metadatos del conjunto de datos original en forma de Data Paper. **d:** (re)utilización del conjunto de datos original (total o parcialmente) por otros trabajos científicos con su correspondiente DOI.

Figure 2. The transition from traditional to reproducible data management. Example of Sierra Nevada forest inventory data (Pérez-Luque et al. 2014). **a:** publication of scientific articles with data as supplementary material. **b:** standardisation of the original data set and integration in local and international repositories. **c:** metadata publication of the data set as a Data Paper. **d:** (re)use of the data set (totally or partially) by other scientific works with their corresponding DOI.

Global de Sierra Nevada (repositorio local) (<http://obsnev.es/>). Además, para permitir su integración en iniciativas internacionales como GBIF y que otros investigadores conociesen el conjunto de datos, se utilizaron estándares de documentación y de datos (EML y Darwin Core). Así, el conjunto de datos inicial, una vez normalizado y documentado, se incluyó en distintos repositorios, se documentó utilizando estándares y se permitió su reutilización para fines diferentes. En nuestro ejemplo, el conjunto de datos Sinfonevada integrado en GBIF, ha sido citado 36 veces [Datos obtenidos en Mayo de 2019], y lo más interesante es que se ha descargado (total o parcialmente) más de 11267 veces (Fig. 2d), aumentando en varios órdenes de magnitud el uso inicial de dicho conjunto de datos. Además de la publicación de los datos, se consideró la publicación de una descripción detallada de Sinfonevada en formato *data paper* que, tras un proceso de revisión por pares, se publicó en una revista indexada (Pérez-Luque et al. 2014) (Fig. 2c).

En nuestra opinión, esta forma de compartir datos (ecológicos o de otras disciplinas), conlleva una serie de beneficios, entre los que destacamos:

- Aumento de la longevidad de los datos. Al documentar el conjunto de datos e integrarlos en repositorios institucionales evitamos el peligro de que se pierdan en el cajón de los investigadores o por causa de una rotura de discos de almacenamiento.
- Durante la documentación de los datos se realizan varios procesos de control de calidad. De esta forma, aumenta la fiabilidad y la seguridad sobre los datos, puesto que han sido validados.

- Aumento de la accesibilidad a los datos, tanto a nivel local como internacional.
- Incremento de la reutilización de los datos, incluso con propósitos diferentes a los originales.

Consideraciones finales

Al igual que las especies, existen datos que se encuentran “en peligro de extinción” y es importante dedicar esfuerzos a su preservación, mediante su documentación y depósito en repositorios oficiales (Specht et al. 2018). ¿Cómo hacemos más sencillo este camino? El reto no sólo está en manos de las revistas y las editoriales, también los investigadores podemos contribuir a ello documentando los datos, publicándolos y apostando por productos como los *data papers*. Los esfuerzos por hacer de los datos y sus metadatos productos publicables y citables están provocando una puesta en valor de la información que repercute en una serie de beneficios no sólo para el investigador, sino para la comunidad científica en general. En los últimos años han surgido diferentes recomendaciones y buenas prácticas que aportan consejos sobre la gestión de los datos y cómo compartirlos, muchos de ellos centrados en el campo de la ecología (Kervin et al. 2013; Goodman 2014; Michener 2015). Estos hábitos, saludables en términos de reproducibilidad, posibilitan la reutilización de los datos dando lugar a nuevos estudios capaces de lograr tanto impacto como las investigaciones originales o incluso reutilizándolos con un propósito insospechado en el momento de su colecta, nutriendo así la ciencia del futuro (Editorial Nature-Communications 2018; Specht et al. 2018).

Como vemos, existen muchos argumentos a favor de archivar y documentar los datos, sin embargo, no podemos olvidarnos de los costes, tanto en tiempo como en dinero, asociados a la puesta en valor de datos históricos y actuales (Costello et al. 2013; Godhill 2014; Michener 2015; Specht et al. 2018). Un cambio fundamental que reside en los investigadores es pasar de “Yo soy el propietario de los datos” a “Recopilo y comparto los datos en nombre de la comunidad científica y la sociedad” (Hampton et al. 2015). Pero este cambio necesariamente tiene que estar acompañado de un mayor reconocimiento en el mundo académico e institucional, lo cual contribuiría a reducir las barreras que existen para compartir datos (Michener 2015). Para un reconocimiento efectivo es fundamental que la tríada científicos, datos e investigaciones se encuentre conectada con el fin de provocar un impacto positivo de las reutilizaciones sobre aquellos que generaron los datos (Pierce et al. 2019).

Otro aspecto imprescindible es la formación. Muchos investigadores están dispuestos y animados a compartir datos, pero encuentran un freno en la falta de experiencia e incluso apoyo institucional para hacerlo (Michener 2015). Se necesitan programas formativos relativos al manejo y puesta en valor de los datos destinados a los investigadores, independientemente de la etapa investigadora en la que se encuentren. Este aspecto resulta crucial, sobre todo teniendo en cuenta que muchas instituciones incluyen como requisito en sus convocatorias de financiación de la investigación la existencia de un Plan de Gestión de Datos (*Data Management Plan*, Miksa et al. 2019) que asegure, entre otros, la viabilidad de los datos. Algunas sociedades científicas ecológicas, como la *British Ecological Society* (<https://www.britishecologicalsociety.org/publications/guides-to/>), han desarrollado una guía que ayuda a los investigadores a realizar una gestión de datos.

La transición hacia un mayor intercambio de datos en ecología es reflejo de un cambio en la forma en que se lleva a cabo la investigación y en el papel que desempeñan los datos en la producción de conocimiento científico (Sholler et al. 2019). Este cambio requiere un compromiso por parte de los diferentes actores implicados: investigadores, instituciones y revistas. En este sentido, consideramos que sería muy positivo que la revista *Ecosistemas* participase activamente en este cambio, apoyando de forma explícita la documentación y puesta en valor de los datos, así como su inclusión en repositorios oficiales.

Agradecimientos

Agradecemos la financiación recibida por los proyectos ECO-POTENTIAL (EU Horizon 2020 grant agreement No 641762): *Improving future ecosystem benefits through earth observations*; y LIFE-ADAPTAMED (LIFE14 CCA/ES/000612): *Protección de servicios ecosistémicos clave amenazados por el cambio climático mediante gestión adaptativa de socioecosistemas mediterráneos*. Este trabajo ha sido desarrollado bajo el *Convenio de colaboración entre la Consejería de Medio Ambiente y Ordenación del Territorio de la Junta de Andalucía y la Universidad de Granada para el desarrollo de actividades vinculadas al Observatorio de Cambio Global de Sierra Nevada*. Quisiéramos agradecer los comentarios de J. Durán y de un revisor anónimo que han mejorado sustancialmente el manuscrito. Igualmente a I. Bartomeus, S. Echeverría, C. Puerta-Piñero y F. Rodríguez-Sánchez que aportaron comentarios en una versión preliminar del manuscrito.

Referencias

- Alonso, B., Valladares, F. 2006. Bases de datos y metadatos en ecología: compartir para investigar en cambio global. *Ecosistemas* 15(2): 83-88.
- Alves, C., Castro, J., Ribeiro, C., Honrado, J., Lomba, Â. 2018. Research data management in the field of Ecology: an overview. En: *Proceedings of the DCMI International Conference on Dublin Core and Metadata Applications, Porto, Portugal, 10-13 September 2018*, pp. 87-94. Dublin Core Metadata Initiative.
- Baker, K.S., Millerand, F. 2010. Infrastructuring Ecology: Challenges in Achieving Data Sharing. En: Parker, J.N., Vermeulen, N., Penders, B. (eds.), *Collaboration in the New Life Sciences*, pp. 111-138. Taylor and Francis Group. Londres, Reino Unido.
- Batalla, M., Ninyerola, M., Catalan, J. 2018. Digital long-term topoclimate surfaces of the Pyrenees mountain range for the period 1950–2012. *Geoscience Data Journal* 5: 50-62.
- Boettiger, C. 2019. Ecological Metadata as Linked Data. *Journal of Open Source Software* 4: 1276.
- Boettiger, C., Jones, M.B., Maier, M., Mecum, B., Salmon, M., Clark, J. 2019. *EML: Read and Write Ecological Metadata Language Files* [on-line]. Disponible en: <https://cloud.r-project.org/web/packages/EML/index.html>
- Bovendorp, R.S., Villar, N., de Abreu-Junior, E.F., Bello, C., Regolin, A.L., Percequillo, A.R., Galetti, M. 2017. Atlantic small-mammal: a dataset of communities of rodents and marsupials of the Atlantic forests of South America. *Ecology* 98: 2226-2226
- Caldera-Serrano, J. 2018. Repositorios Públicos Frente a La Mercantilización de La Ciencia: Apostando Por La Ciencia Abierta y La Evaluación Cualitativa. *Métodos de información* 9: 74-101.
- Chavan, V., Penev, L. 2011. The data paper: a mechanism to incentivize data publishing in biodiversity science. *BMC Bioinformatics* 12 Suppl 15: S2.
- Costello, M.J., Michener, W.K., Gahegan, M., Zhang, Z.-Q., Bourne, P.E. 2013. Biodiversity data should be published, cited, and peer reviewed. *Trends in Ecology and Evolution* 28: 454-461.
- Couture, J.L., Blake, R.E., McDonald, G., Ward, C.L. 2018. A Funder-Imposed Data Publication Requirement Seldom Inspired Data Sharing. *PLOS ONE* 13: e0199789.
- Culina, A., Baglioni, M., Crowther, T.W., Visser, M.E., Woutersen-Windhouver, S., Manghi, P. 2018. Navigating the Unfolding Open Data Landscape in Ecology and Evolution. *Nature Ecology and Evolution* 2: 420-426.
- Díaz-Delgado, R. 2016. La investigación y seguimiento ecológico a largo plazo (LTER). *Ecosistemas* 25(1): 01-03. Doi.: 10.7818/ECOS.2016.25-1.01.
- Editorial Nature-Communications. 2018. Data sharing and the future of science. *Nature Communications* 9: 2817.
- Escribano, N., Oscoz, J., Galicia, D., Cancellario, T., Durán, C., Navarro, P., Ariño, A.H. 2018. Freshwater macroinvertebrate samples from a water quality monitoring network in the Iberian Peninsula. *Scientific Data* 5: 180108.
- Fegraus, E.H., Andelman, S., Jones, M.B., Schildhauer, M. 2005. Maximizing the Value of Ecological Data with Structured Metadata: An Introduction to Ecological Metadata Language (EML) and Principles for Metadata Creation. *The Bulletin of the Ecological Society of America* 86: 158-168.
- González-Moreno, P., Quero, J., Poorter, L., Bonet, F., Zamora, R. 2011. Is spatial structure the key to promote plant diversity in Mediterranean forest plantations? *Basic and Applied Ecology* 12: 251-259.
- Goodhill, G.J. 2014. Practical costs of data sharing. *Nature* 509: 33-33.
- Goodman, A., Pepe, A., Blocker, A.W., Borgman, C.L., Cranmer, K., Crosas, M., et al. 2014. Ten Simple Rules for the Care and Feeding of Scientific Data. *PLOS Computational Biology* 10: 1-5.
- Gómez-Aparicio, L., Zavala, M.A., Bonet, F.J., Zamora, R. 2009. Are pine plantations valid tools for restoring Mediterranean forests? An assessment along abiotic and biotic gradients. *Ecological Applications* 19: 2124-2141.
- Grimm, A., Ramírez, A.M.P., Moulherat, S., Reynaud, J., Henle, K. 2014. Life-history trait database of European reptile species. *Nature Conservation* 9: 45-67.
- Hampton, S.E., Anderson, S.S., Bagby, S.C., Gries, C., Han, X., Hart, E.M., Jones, M.B., et al. 2015. The Tao of open science for ecology. *Ecosphere* 6: art120.
- Hampton, S.E., Strasser, C.A., Tewksbury, J.J., Gram, W.K., Budden, A.E., Batcheller, A.L., Duke, C.S., Porter, J.H. 2013. Big Data and the Future of Ecology. *Frontiers in Ecology and the Environment* 11: 156-162.
- Kervin, K., Michener, W., Cook, R. 2013. Common Errors in Ecological Data Sharing. *Journal of eScience Librarianship* 2: e1024.
- Kuebbing, S.E., Reimer, A.P., Rosenthal, S.A., Feinberg, G., Leiserowitz, A., Lau, J.A., Bradford, M.A. 2018. Long-Term Research in Ecology and Evolution: A Survey of Challenges and Opportunities. *Ecological Monographs* 88: 245-258.

- LaDeau, S.L., Han, B.A., Rosi-Marshall, E.J., Weathers, K.C. 2017. The Next Decade of Big Data in Ecosystem Science. *Ecosystems* 20: 274-283.
- Lorite, J. 2016. An updated checklist of the vascular flora of Sierra Nevada (SE Spain). *Phytotaxa* 261: 1-57.
- Marshall, H.H., Griffiths, D.J., Mwanguhya, F., Businge, R., Griffiths, A.G.F., Kyabulima, S., Mwesige, K., et al. 2018. Data collection and storage in long-term ecological and evolutionary studies: The Mongoose 2000 system. *PLOS ONE* 13: 1-15.
- Michener, W.K. 2015. Ecological data sharing. *Ecological Informatics* 29: 33-44.
- Miksa, T., Simms, S., Mietchen, D., Jones, S. 2019. Ten Principles for Machine-Actionable Data Management Plans. *PLOS Computational Biology* 15: e1006750.
- Munafó, M.R., Nosek, B.A., Bishop, D.V.M., Button, K.S., Chambers, C.D., Percie du Sert, N., Simonsohn, U., et al. 2017. A manifesto for reproducible science. *Nature Human Behaviour* 1: 0021.
- Navarro-González, I., Pérez-Luque, A.J., Bonet, F.J., Zamora, R. 2013. The weight of the past: land-use legacies and recolonization of pine plantations by oak trees. *Ecological Applications* 23: 1267-1276.
- Parmenter, R.R., Zlotin, R.I., Moore, D.I., Myers, O.B. 2018. Environmental and endogenous drivers of tree mast production and synchrony in piñon-juniper-oak woodlands of New Mexico. *Ecosphere* 9: e02360.
- Pérez-Luque, A.J., Bonet, F.J., Pérez-Pérez, R., Aspizua, R., Lorite, J., Zamora, R. 2014. Sinfonevada: Dataset of floristic diversity in Sierra Nevada forests (SE Spain). *PhytoKeys* 35: 1-15.
- Pérez-Luque, A.J., Sánchez-Rojas, C.P., Zamora, R., Pérez-Pérez, R., Bonet, F.J. 2015. Dataset of phenology of Mediterranean high-mountain meadows flora (Sierra Nevada, Spain). *PhytoKeys* 46: 89-107.
- Pérez-Luque, A.J., Barea-Azcón, J.M., Álvarez-Ruiz, L., Bonet-García, F.J., Zamora, R. 2016a. Dataset of Passerine bird communities in a Mediterranean high mountain (Sierra Nevada, Spain). *ZooKeys* 552: 137-154.
- Pérez-Luque, A.J., Bonet, F.J., Zamora, R., Barea-Azcón, J.M., Aspizua, R., Sánchez-Gutiérrez, F.J. 2016b. Señales del cambio global en el sitio LTER-Sierra Nevada. *Ecosistemas* 25(1): 65-71. Doi.: 10.7818/ECOS.2016.25-1.08.
- Pierce, H.H., Dev, A., Statham, E., Bierer, B.E. 2019. Credit data generators for data reuse. *Nature* 570: 30-32.
- Rebetez, M., von Arx, G., Gessler, A., Pannatier, E.G., Innes, J.L., Jakob, P., Jetel, M, et al. 2018. Meteorological data series from Swiss long-term forest ecosystem research plots since 1997. *Annals of Forest Science* 75: 41.
- Reichman, O.J., Jones, M.B., Schildhauer, M.P. 2011. Challenges and Opportunities of Open Data in Ecology. *Science* 331: 703-705.
- Roche, D.G., Kruuk, L.E.B., Lanfear, R., Binning, S.A. 2015. Public Data Archiving in Ecology and Evolution: How Well Are We Doing? *PLOS Biology* 13: e1002295.
- Rodríguez-Sánchez, F., Pérez-Luque, A.J., Bartomeus, I., Varela, S. 2016. Ciencia reproducible: qué, por qué, cómo? *Ecosistemas* 25(2): 83-92. Doi.: 10.7818/ECOS.2016.25-2.11
- Santos, C., Blake, J., States, D.J. 2005. Supplementary data need to be kept in public repositories. *Nature* 438: 738.
- Semenchenko, V., Son, M.O., Novitski, R., Kvach, Y., Panov, V.E. 2016. Checklist of non-native benthic macroinvertebrates and fish in the Dnieper River basin. *BiolInvasions Records* 5: 185-187.
- Serrano-Notivol, R., Beguería, S., Saz, M.A., Longares, L.A., de Luis, M. 2017. SPREAD: a high-resolution daily gridded precipitation dataset for Spain – an extreme events frequency and intensity overview. *Earth System Science Data* 9: 721-738.
- Sholler, D., Ram, K., Boettiger, C., Katz, D.S. 2019. Enforcing Public Data Archiving Policies in Academic Publishing: A Study of Ecology Journals. *Big Data and Society* 6: 1-18.
- Smith, C. 2019. *EMLassemblyline: A workflow for creating EML*. [on-line]. Disponible en: <https://github.com/EDLorg/EMLassemblyline/blob/master/DESCRIPTION>
- Soranno, P.A. 2019. Six Simple Steps to Share Your Data When Publishing Research Articles. *Limnology and Oceanography Bulletin* 28: 41-44.
- Specht, A., Bolton, M.P., Kingsford, B., Specht, R.L., Belbin, L. 2018. A story of data won, data lost and data re-found: the realities of ecological data preservation. *Biodiversity Data Journal* 6: e28073.
- Stodden, V., Guo, P., Ma, Z. 2013. Toward Reproducible Computational Research: An Empirical Analysis of Data and Code Policy Adoption by Journals. *PLOS ONE* 8: 1-8.
- Storchová, L., Hořák, D. 2018. Life-history characteristics of European birds. *Global Ecology and Biogeography* 27: 400-406.
- Thessen, A., Patterson, D. 2011. Data issues in the life sciences. *ZooKeys* 150: 15-51.
- Tonk, L., Bongaerts, P., Sampayo, E.M., Hoegh-Guldberg, O. 2013. SymbioGBR: a web-based database of Symbiodinium associated with cnidarian hosts on the Great Barrier Reef. *BMC Ecology* 13: 7.
- Torca, M., Campos, J.A., Herrera, M. 2019. Species composition and plant traits of south Atlantic European coastal dunes and other comparative data. *Data in Brief* 22: 207-213.
- Torres-Salinas, D., Robinson-García, N., Cabezas-Clavijo, Á. 2012. Compartir los datos de investigación en ciencia: introducción al *data sharing*. *El Profesional de la Información* 21: 173-184.
- Vicente-Serrano, S.M., Tomas-Burguera, M., Beguería, S., Reig, F., Latorre, B., Peña-Gallardo, M., Luna, M.Y., et al. 2017. A high resolution dataset of drought indices for Spain. *Data* 2 (3): 22.
- Vizoso, M.T., Quesada, C. 2015. Catalogue of type specimens of fungi and lichens deposited in the Herbarium of the University of Granada (Spain). *Biodiversity Data Journal* 3: e5204.
- Wang, N., Sherwood, A.R., Kurihara, A., Conklin, K.Y., Sauvage, T., Presting, G.G. 2009. The Hawaiian Algal Database: a laboratory LIMS and online resource for biodiversity data. *BMC Plant Biology* 9: 117.
- Whitlock, M.C. 2011. Data archiving in ecology and evolution: best practices. *Trends in Ecology and Evolution* 26: 61-65.
- Wieczorek, J., Bloom, D., Guralnick, R., Blum, S., Döring, M., Giovanni, R., Robertson, T., Vieglais, D. 2012. Darwin core: An evolving community-developed biodiversity data standard. *PLoS ONE* 7: e29715.
- Wilkinson, M.D., Dumontier, M., Aalbersberg, I.J., Appleton, G., Axton, M., Baak, A., Blomberg, N. et al., 2016. The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data* 3: 160018.
- Wohner, C., Peterseil, J., Poursanidis, D., Kliment, T., Wilson, M., Mirtl, M., Chrysoulakis, N. 2019. DEIMS-SDR – A web portal to document research sites and their associated data. *Ecological Informatics* 51: 15-24.
- Wolkovich, E.M., Regetz, J., O'Connor, M.I. 2012. Advances in Global Change Research Require Open Science by Individual Researchers. *Global Change Biology* 18: 2102-2110.
- Yang, H., Wong, G.K.-S., Cheng, S., Sun, W., Liu, X., Fu, Y., Xu, X. et al. 2018. 10KP: A phylodiverse genome sequencing plan. *GigaScience* 7 (3): gij013.

Apéndice 1. *Publicaciones que han utilizado el Inventario Forestal Sinfonevada antes de su integración en un repositorio.***Appendix 1.** *Publications that have used the Sinfonevada Forest Inventory before its integration into a repository.*

- González-Moreno, P. 2009. *Natural regeneration and biodiversity in pine plantations of Southern Spain: a landscape approach*. Tesis de maestría. Wageningen University, Wageningen, Países Bajos.
- González-Moreno, P., Quero, J., Bonet-García, F.J., Poorter, L., Zamora, R. 2009. El papel de la estructura espacial de la vegetación en la naturalización de pinares de repoblación en Sierra Nevada (Granada-Almería). En: *Actas del V Congreso Forestal Español, 21-25 Septiembre 2009, Ávila, España*. Sociedad Española de Ciencias Forestales y Junta de Castilla y León. Valladolid, España.
- González-Moreno, P., Quero, J., Bonet-García, F.J., Poorter, L., Zamora, R. 2010. The influence of spatial structure on natural regeneration and biodiversity in Mediterranean pine plantations: a nested landscape approach. En: *Proceedings of the IUFRO Landscape Ecology Working Group International Conference: Forest landscapes and global change new frontiers in management, conservation and restoration, 21-27 September 2010, Bragança, Portugal*, pp. 52-57. Instituto Politécnico de Bragança, Braganza, Portugal.
- González-Moreno, P., Quero, J., Poorter, L., Bonet, F., Zamora, R. 2011. Is spatial structure the key to promote plant diversity in Mediterranean forest plantations? *Basic and Applied Ecology* 12: 251-259.
- Gómez-Aparicio, L., Zavala, M.A., Bonet, F.J., Zamora, R. 2009. Are pine plantations valid tools for restoring Mediterranean forests? An assessment along abiotic and biotic gradients. *Ecological Applications* 19: 2124-2141.
- Gómez-Aparicio, L., Zavala, M.A., Bonet-García, F.J., Zamora, R. 2009. Regeneración y diversidad en pinares de repoblación: un análisis a través de gradientes ambientales. En: *Actas del V Congreso Forestal Español, 21-25 Septiembre 2009, Ávila, España*. Sociedad Española de Ciencias Forestales y Junta de Castilla y León. Valladolid, España.
- Navarro-González, I. 2009. *Efectos del uso pasado del suelo en la regeneración actual de la vegetación natural bajo las plantaciones de pinares de Sierra Nevada. Implicaciones en la gestión forestal*. Tesis de maestría Universidad Autónoma de Madrid. Madrid, España.
- Navarro-González, I., Bonet-García, F.J., Zamora, R. 2011. Current Mediterranean forest regeneration depends on land use in the recent past. En: *Proceedings of the 12th European Ecological Federation Congress: Responding to Rapid Environmental Change, 25-29 September 2011, Ávila, Spain*, pp. 420. AEET, Madrid, España.
- Navarro-González, I., Pérez-Luque, A.J., Bonet, F.J., Zamora, R. 2013. The weight of the past: land-use legacies and recolonization of pine plantations by oak trees. *Ecological Applications* 23: 1267-1276.
- Pérez-Luque, A.J. 2011. *Análisis multivariante ambiental de los melojares de Quercus pyrenaica Willd. de Sierra Nevada*. Tesis de maestría. Universidad de Granada, Granada, España.
- Pérez-Luque, A.J., Bonet-García, F.J., Benito, B.M., Zamora, R. 2013. Caracterización ambiental de los robledales de Quercus pyrenaica Willd. de Sierra Nevada. En: *Libro de resúmenes de comunicaciones del XI Congreso Nacional de la AEET: Invitación a la Ecología. Estrechando lazos con la Sociedad, 6-10 mayo de 2013, Pamplona/Iruña*, pp. 92. AEET, Madrid, España.