

---

## DATA-SHARING EN URUGUAY, LA VISIÓN DE LOS COLECTORES Y USUARIOS DE DATOS

Florencia Grattarola<sup>1,2\*</sup> & Daniel Pincheira-Donoso<sup>2</sup>

<sup>1</sup>School of Life Sciences, University of Lincoln, Brayford Campus, Lincoln, Lincolnshire LN6 7DL, United Kingdom.

<sup>2</sup>MacroBiodiversity Lab, School of Biological Sciences, Queen's University Belfast, 19 Chlorine Gardens, Belfast, BT9 5DL, United Kingdom.

\*Autor para correspondencia: fGrattarola@lincoln.ac.uk

### RESUMEN

El intercambio de datos (*data-sharing*) se ha convertido en un protocolo clave en la ciencia moderna, con numerosas ventajas para colectores y usuarios. Sin embargo, en Uruguay permanece una práctica poco común, dada la falta de fuentes de información sobre biodiversidad primaria abiertas y disponibles públicamente. Las causas son aún desconocidas. Este estudio explora, mediante una encuesta online, la visión de los colectores y usuarios de datos analizando tanto las razones de la reticencia a compartirlos, como las motivaciones para hacerlo. Los investigadores están dispuestos a compartir sus datos si reciben crédito adecuado por su esfuerzo y conocimiento. Para poder traducir esta voluntad en una práctica efectiva de ciencia abierta, se debe trabajar en las estructuras de incentivo y motivaciones, comunicar la importancia y los beneficios de la apertura e intercambio, y brindar herramientas técnicas y capacitación en todos los aspectos de la gestión y el intercambio de datos. La práctica de *data-sharing* debe resultar en un fortalecimiento de la cultura de colaboración científica que no solo beneficia a los investigadores individualmente, sino al progreso del conocimiento en sí mismo. Este trabajo ofrece una exploración inicial al tema, que evidencia su importancia y la necesidad de que sea mayormente discutido.

**Palabras clave:** *data-sharing*; biodiversidad; base de datos; encuesta online.

### ABSTRACT

**Data-sharing in Uruguay, the vision of collectors and data users.** Data-sharing has become a key issue in modern science, with numerous advantages for data collectors and user. However, the practice of sharing data in Uruguay is still uncommon given there are no primary biodiversity datasets open or publicly available. The reasons are still unknown. This study explores through an online survey the vision of data collectors and users analysing the reasons for both the reluctance to share data and the motivations for doing so. Over-all, scientists are willing to share their research data if they receive adequate credit for their effort and knowledge. In order to translate the will into an effective practice of open science, we must work on the incentives and motivation structures behind, communicate the importance and benefits of

data openness and exchange, and provide technical tools and training on all aspects of data management and sharing. Importantly, data-sharing practice must result in the reinforce of a scientific collaboration culture that benefits not only researchers at the individual level, but the progress of knowledge itself. This work represents a valuable initial approach to the subject that proofs the importance that data availability has for researchers in Uruguay and the need for it to be mostly discussed.

**Key words:** data-sharing; biodiversity; database; online survey.

### INTRODUCCIÓN

La era digital ha revolucionado la producción de conocimiento científico (Hey *et al.*, 2009), fundamentalmente permitiendo la generación de capacidades de obtención, acceso y análisis de enormes volúmenes de datos en magnitudes sin precedentes. El intercambio de datos (*data-sharing*) es una práctica relevante en todas las áreas científicas que ha aumentado considerablemente durante la última década (Fecher *et al.*, 2015). Compartir datos consiste en hacer posible la disponibilidad pública de información para el acceso de otros, y depende mayormente de depositarlos y asegurar su preservación en el tiempo. Sin embargo, puede también abarcar actos tan variados como informar sobre la existencia de datos, publicarlos en un sitio web, o mantenerlos en repositorios minuciosamente curados (Borgman, 2012). Es ampliamente aceptado que compartir datos implica múltiples ventajas (Fienberg *et al.*, 1985), tales como la posibilidad de verificar y reproducir resultados de investigación, aumento de la transparencia científica, optimización de recursos por reutilización y minimización de la recolección, atribución de créditos cuando los datos son utilizados por otros, acceso mejorado a datos de investigación y nuevas oportunidades para la colaboración, entre otros. Pese a estos dominantes beneficios, el intercambio de datos también plantea desafíos a nivel técnico y sociocultural (Guralnick *et al.*, 2007; Borgman, 2012; Michener, 2015). Algunos impedimentos incluyen el tiempo y labor necesarios para compartir datos,

falta de experiencia en gestión de datos de investigación y capacitación insuficiente, falta de conocimiento de las normas, falta de apoyo institucional y recursos de financiación para la gestión de datos, limitaciones legales y complejidad de los derechos de propiedad intelectual de los datos. Sin embargo, las barreras para el intercambio efectivo de datos están profundamente arraigadas con las prácticas y la cultura del proceso de investigación, así como con los propios investigadores (Linek *et al.*, 2017). En definitiva, el grado y el éxito del intercambio de datos en biodiversidad depende de las actitudes y prácticas de las personas que recolectan y curan especímenes, identifican especies y publican sus resultados y datos, dado que de ellos depende esta decisión. Colectivamente, la práctica de compartir datos ha permitido la producción de ciencia y de respuestas a preguntas de primaria importancia para el mundo, en una magnitud que ha acelerado los avances sustancialmente.

En Uruguay la apertura de datos de biodiversidad es pobremente practicada. Que un dato sea abierto implica que se encuentre disponible en un formato estándar y libre de restricciones, de manera de poder ser reutilizado y redistribuido sin mediar un complejo protocolo formal de solicitud (Wilkinson *et al.*, 2016). En la actualidad, las fuentes públicas (gubernamentales y académicas) de información sobre biodiversidad no son *open-access* ni están disponibles públicamente. Asimismo, Uruguay tiene uno de los niveles más bajos de disponibilidad de datos en América Latina en GBIF (Global Biodiversity Information Facility - [www.gbif.org/](http://www.gbif.org/)), la plataforma digital más grande en número de registros de especies en el mundo, con alrededor de 185,000 registros localizados en el país, pertenecientes a más de 570 conjuntos de datos extranjeros. Las razones detrás de esta baja disponibilidad de datos permanecen desconocidas. Para comprender la posición de Uruguay en este tema, es importante comenzar por identificar cuáles son las percepciones de los investigadores (colectores y usuarios). El objetivo de este estudio es revisar la opinión de diferentes investigadores de Uruguay sobre la recopilación y uso de datos de biodiversidad, particularmente su disposición a compartir datos, aspectos técnicos sobre el intercambio de datos y las motivaciones e inquietudes u objeciones principales para hacerlo.

## MATERIAL Y MÉTODOS

Se generó un cuestionario online anónimo utilizando Microsoft Forms™, basado en Enke *et al.* (2012), que cubrió una población de 93 expertos. El cuestionario, el cual estuvo activo durante 33 días, representa una valiosa aproximación inicial al tema

siendo la primera encuesta hecha en Uruguay (y en la región) en este sentido. El objetivo fue examinar la opinión de diferentes expertos nacionales (definidos como funcionarios activos de instituciones localizadas en el país, independiente de su nacionalidad de origen), en relación a la colecta y uso de datos de biodiversidad. El cuestionario consistió en 27 preguntas de respuesta múltiple y una última pregunta abierta a realizar comentarios (ver en Apéndice 1). En la mayor parte de las preguntas los encuestados podían elegir más de una respuesta y, por tanto, los resultados de estas preguntas no suman 100%. En otras, solo sí o no fueron respuestas permitidas (e.g.: disposición a depositar sus datos de manera pública). La encuesta fue enviada a través del correo electrónico intentando cubrir un amplio rango de investigadores uruguayos que utilizan datos de biodiversidad.

Los resultados de la encuesta fueron evaluados clasificando las respuestas según (i) tipo de institución en la que trabajan (Universidad, Gobierno, Instituto de Investigación, ONG, etc.), (ii) estatus profesional (Investigador, Docente, Estudiante, Voluntario, etc.), (iii) campo de estudio (Ecología, Genética, Taxonomía, Conservación, etc.) y (iv) tipo de dato con el que trabajan (morfológicos, moleculares, comportamentales, listas de especies, etc.). Esta categorización se utilizó para analizar diferencias en uso y disposición a compartir datos entre los diferentes investigadores. Los datos fueron analizados de manera fundamentalmente descriptiva, con el objetivo de presentar una síntesis que muestre la distribución de frecuencia de las múltiples respuestas y los patrones generales emergentes de este análisis. Todos los análisis fueron realizados utilizando software R (R Core Team, 2018).

## RESULTADOS

Del total de 93 investigadores que respondieron a la encuesta online, la mayoría trabajaba en Universidades (62.4%), en el gobierno (30.1%) o en organizaciones sin fines de lucro (16.1%). Una cantidad menor (Tabla 1), trabajaba en museos, empresas privadas, institutos de investigación, eran trabajadores independientes, o trabajaban en Educación no universitaria. La mayoría de los participantes fueron investigadores o docentes (58.1%) y el resto estudiantes de doctorado, maestría o PostDoc (41.9%) (Tabla 2). Los perfiles científicos de los encuestados fueron amplios (por ejemplo, Ciencias Agrícolas (8.6%) y Geografía (5.4%)), mientras la mayoría provenía de las ciencias biológicas y trabajaba en campos como Ecología (53.8%), Zoología (41.9%), Biología de la Conservación (35.5%), Biogeografía (18.3%) y Sistemática (15.1%) (Tabla 3).

**Tabla 1.** Tipo de institución en la que trabajan los entrevistados (en %).

Institución en que trabajan	%
Universidad	59.1
Gobierno	28.0
Organización sin fines de lucro	16.1
Museo	12.9
Empresa privada	7.5
Instituto de Investigación	4.3
Independiente	3.2
Educación (no Universidad)	2.2

**Tabla 2.** Estatus profesional de los entrevistados (en %).

Estatus profesional	%
Investigador/a	54.8
Docente	37.6
Científico/a trabajando en el gobierno	29.0
Estudiante de Maestría	20.4
Estudiante de Doctorado	20.4
Voluntario/a en una Organización sin fines de lucro	12.9
Científico/a trabajando en el sector privado	9.7
Científico/a Ciudadano/a (citizen scientist)	9.7
PostDoc	1.1

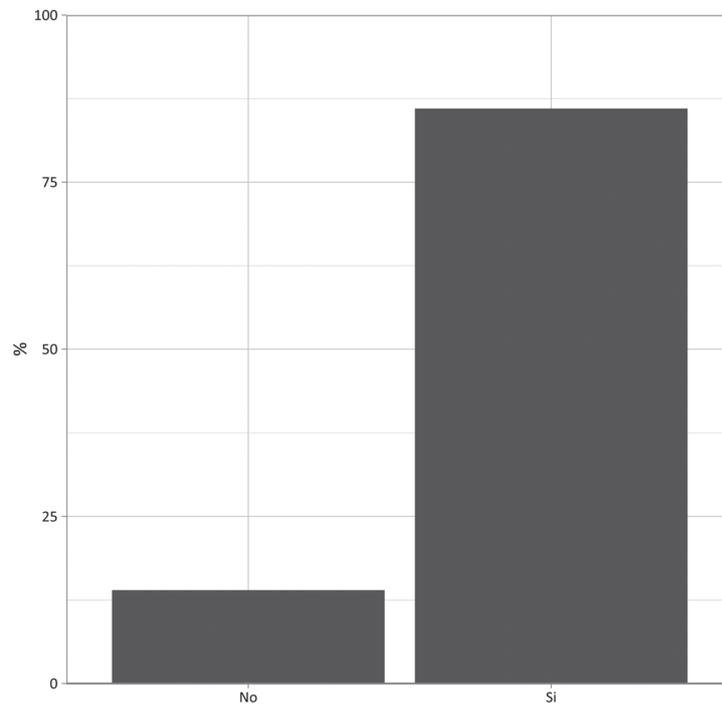
**Tabla 3.** Campo de investigación de los encuestados (en %).

Campo de investigación	%
Ecología	53.8
Zoología	41.9
Biología de la Conservación	35.5
Biogeografía	18.3
Sistemática	15.1
Biología del Comportamiento	14.0
Evolución	14.0
Educación Ambiental	12.9
Taxonomía	11.8
Oceanografía	11.8
Genética	10.8
Botánica	8.6
Ciencias Agrícolas	8.6
Geografía	5.4
Paleontología	3.2
Bioinformática	2.2
Biología Molecular	2.2
Bioquímica	2.2
Entomología	2.2
Microbiología	1.1
Geología	1.1
Virología	1.1
Biología del Desarrollo	1.1
Fisiología	1.1
Neurobiología	1.1

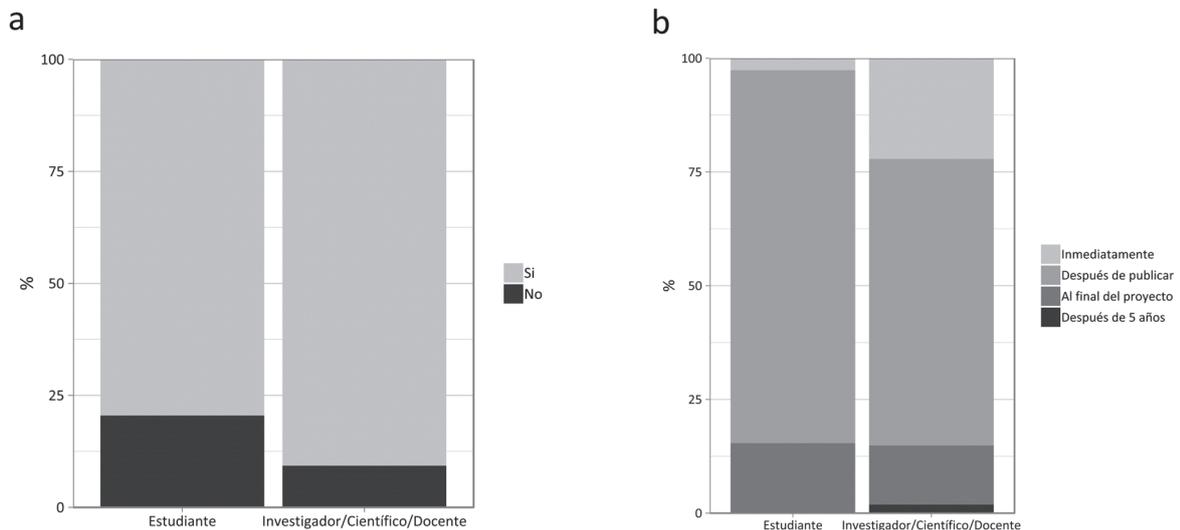
### Voluntad de compartir datos

Frente a la pregunta sobre la disposición de los encuestados a depositar sus datos en bases de datos disponibles públicamente, 86% de los participantes estuvo de acuerdo (Fig. 1). El porcentaje de personas que se mostró reticente a compartir datos se hizo mayor en personas que trabajaban en empresas privadas (50%) u organizaciones sin fines de lucro (20%) y menor en aquellas que trabajaban en el gobierno (7.7%), universidades (12.7%) y museos (9.1%). Respecto del estatus profesional, observamos una mayor negativa de estudiantes que de científicos, investigadores o docentes no estudiantes (Fig. 2a).

Frente a la pregunta de cuándo estarían dispuestos a compartir sus datos, el 71% de los encuestados respondió que estaban dispuestos a hacer disponibles sus datos de investigación después de publicarlos y solo 6,5% inmediatamente luego de obtenidos. Al analizar esto según el estatus profesional de los encuestados (Fig. 2b), la mayor parte de los estudiantes estuvo dispuesta a hacerlo una vez publicados los datos (82.1%) y solo una pequeña proporción estuvo dispuesta a compartirlos inmediatamente luego de obtenidos (2.6%). Entre docentes, científicos e investigadores esa relación se hizo menor, 63.0% los compartiría luego de tener una publicación y 22.2% inmediatamente luego de obtenidos los datos.



**Fig. 1.** Disposición de los encuestados a depositar sus datos en bases de datos disponibles públicamente. En el gráfico se muestra que el 86% de los participantes estuvo de acuerdo, mientras el 14% se encontró en desacuerdo. Solo sí y no fueron posibles respuestas.



**Fig. 2.** Análisis según el estatus profesional de los encuestado de: (a) la disposición de los encuestados a depositar sus datos en bases de datos disponibles públicamente, mostrando que el 79.5% de los estudiantes estuvieron dispuestos, mientras 20.5% no lo estuvieron, y entre investigadores, científicos y docentes el 90.7% estuvieron de acuerdo mientras un 9.3% no lo estuvo; y (b) el período de tiempo en el cual estarían dispuestos a compartir sus datos, mostrando que 2.6% de los estudiantes estuvo dispuesto a liberar sus datos inmediatamente, 82% después de publicar y 15.4% al finalizar el proyecto, mientras que 22.1% de los científicos, investigadores y docentes estuvo dispuesto a hacerlo inmediatamente, 63% después de publicar, 13% al final del proyecto y 1.9% después de 5 años.

### Inquietudes y objeciones

Las principales objeciones de los encuestados con respecto a compartir sus datos en bases de datos online fueron la falta de reconocimiento por su esfuerzo (44.1%) y el tiempo y esfuerzo que deben invertir para depositar sus datos (36.6%) (Fig. 3). Similarmente, 34.4% de los encuestados manifestó que la 'pérdida de control' de sus datos representaba un impedimento para hacerlos disponibles, mientras la falta de estándares de datos también representaba un obstáculo para el 32.3% de los investigadores.

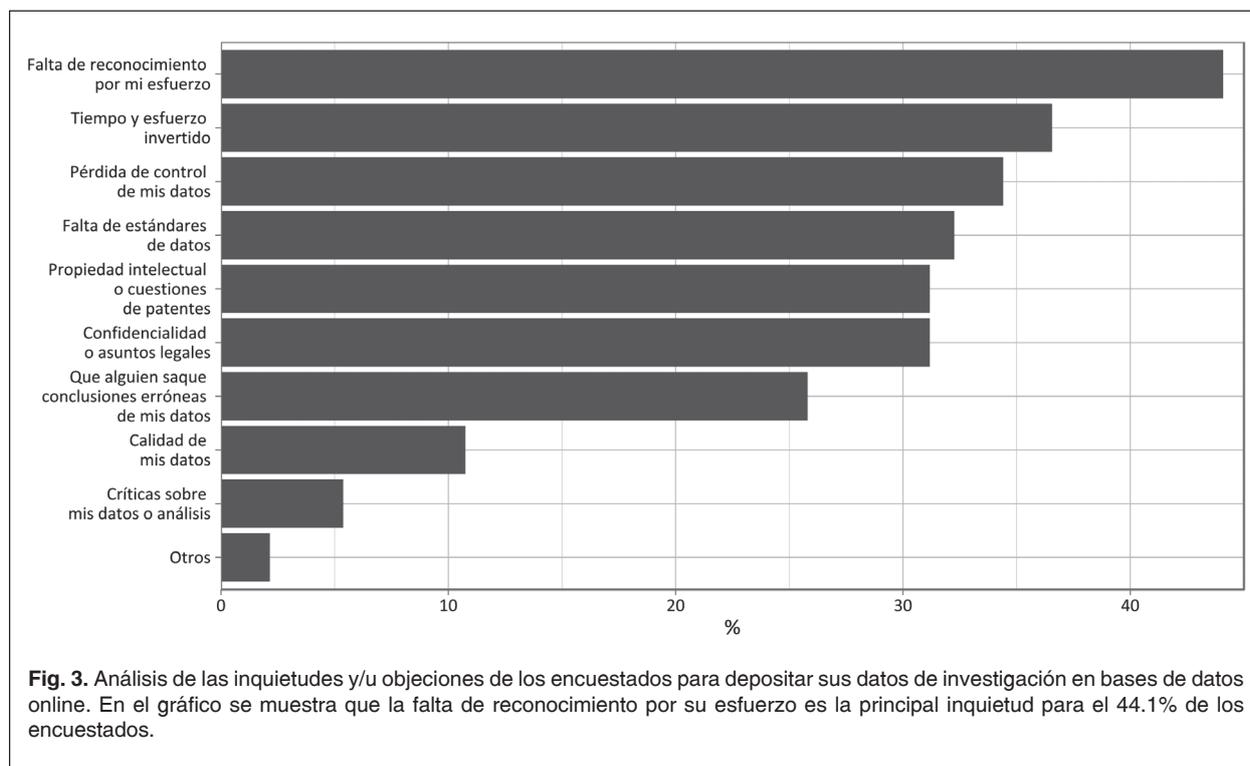
El 63.4 % de los investigadores dijo conocer bases de datos donde depositar sus datos, mientras que el restante 36.6% dijo no conocer ningún repositorio. Entre las bases de datos más mencionadas estuvieron GenBank, eBird, GBIF y diversas colecciones nacionales. También fueron excepcionalmente mencionadas otras bases como Orthoptera Species File Online, SALVIAS, Morphbank, Dryad, TreeBASE, World Spider Catalogue y BOLD Systems. El 63.8% de los investigadores dispuestos a compartir datos afirmaron que existía una base de datos en la que podían depositar sus datos. Similarmente, 61.5% de los que no estaba dispuestos a hacerlo sí conocían una base de datos en la que podían depositar sus datos. El porcentaje de investigadores generando datos moleculares que conocían bases de datos en las que podían depositar sus datos fue elevado (95.5%), mientras solo el 51% de los investigadores que generan datos taxonómicos o morfológicos conocían una base de datos para depositar sus datos.

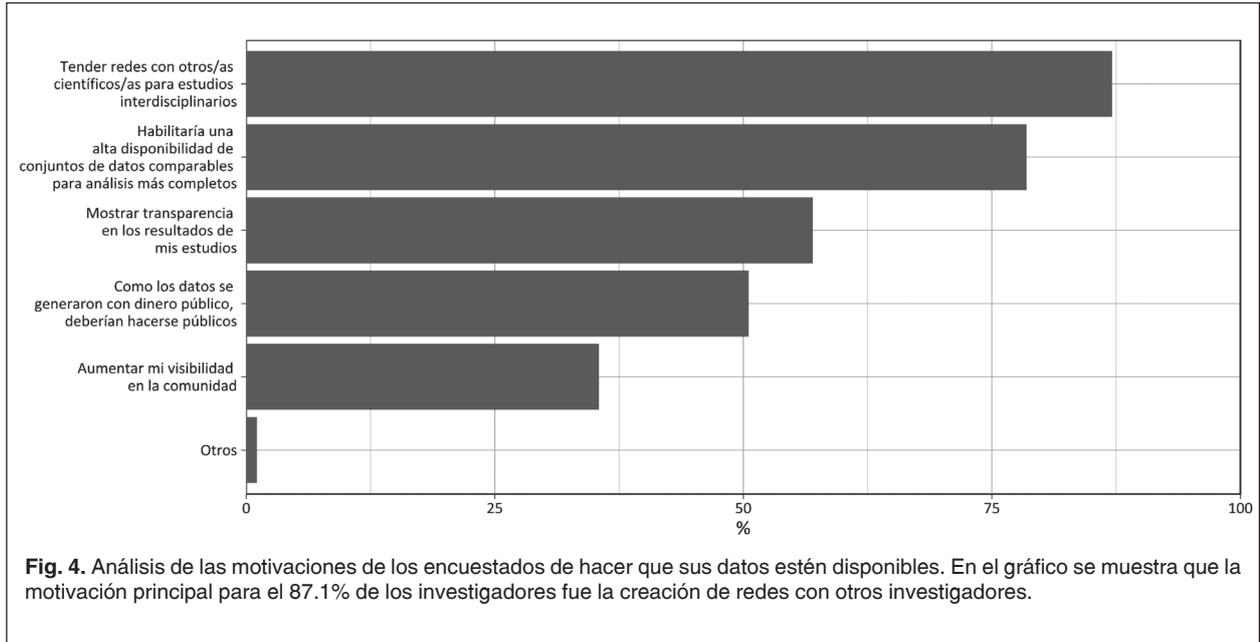
El 78.5% de los entrevistados declaró no tener un plan de gestión de datos, mientras 10.8% dijo no saberlo y 10.8% no tenerlo. En la mayoría de las disciplinas, la disposición de planes de gestión de datos fue similar; no obstante, solo el 6.3% de los investigadores que generaban datos fisiológicos tenían un plan de gestión de datos que fue más común en quienes generaban datos moleculares (23.1%). Aquellos que sí tenían un plan de gestión de datos, manifestaron que éstos eran importantes (30%), muy importantes (30%) y extremadamente importantes (40%). Por otro lado, los que declararon no tenerlo dijeron que estos eran importantes (45.2%), muy importantes (31.5%), extremadamente importantes (11%), algo importantes (11%) y no importaban en absoluto (1.4%).

La mayoría de los participantes dijeron que los datos que producen le pertenecían a su institución (38.7%) o que ellos eran los propietarios (23.7%). Por otro lado, uno de cada diez encuestados indicó que los datos no le pertenecían a nadie y por tanto eran públicos (9.7%).

### Motivaciones

La motivación principal para que los investigadores compartieran datos fue la creación de redes con otros investigadores (87.1%), mientras que la disponibilidad de conjuntos de datos comparables para análisis completos fue también importante (78.5%) (Fig. 4). Otras razones para

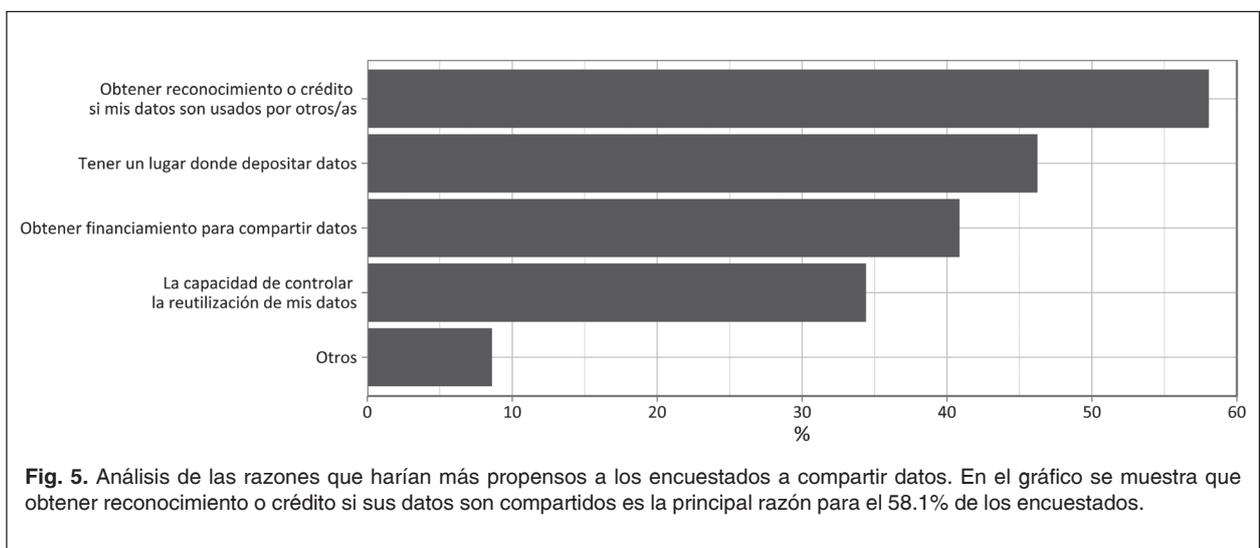




compartir datos fueron la transparencia de los resultados (57%), que la generación de datos de investigación a menudo se financiaba con dinero público (50.5%) y que el intercambio de datos aumentaba la visibilidad en la comunidad (35.5%). También, la posibilidad de citar conjuntos de datos como publicaciones (57%), los requisitos de las revistas (38.7%) y las agencias de financiación (34.4%) representaron importantes motivadores para los investigadores.

Ser reconocido o citado por compartir datos con otros motivaría más a los investigadores a hacerlo

(58.1%) (Fig. 5). Además, un lugar (46.2%) y fondos (40.9%) para depositar datos alentarían el intercambio de datos entre los investigadores. Otro criterio importante fue la capacidad de controlar la reutilización de datos (34.4%). Cuando se les preguntó cómo les gustaría ser reconocidos por compartir sus datos, la mayoría de los participantes respondió en la bibliografía como una publicación (59.1%), al 17.2% le gustaría aparecer como coautor de cualquier publicación que usara sus conjuntos de datos y para un 12.9% simplemente ser mencionado en los agradecimientos sería suficiente.



## Estructuras técnicas

Los datos producidos por los investigadores se dividieron en varias categorías: datos morfológicos, fisiológicos, moleculares, taxonómicos y anatómicos, así como de comunidades y ecosistemas, y la mayoría de los investigadores generaban una combinación de estos. Los 3 tipos de datos más comúnmente colectados fueron listas de especies (presencia/ausencia) (73.1%), datos de ecosistemas (55.9%) y taxonómicos (50.5%). Los metadatos que acompañan a los datos generalmente incluyeron localidad (92.4%), coordenadas geográficas (89.2%), fecha de colecta (88.2%) y nombre del organismo (84.9%). Solamente 1 de cada 5 investigadores registró la configuración del equipo de colecta (22.5%) y la precisión de las medidas que toma (21.5%).

El formato estándar utilizado para trabajar los datos fue de tablas (95.7%). Además, también fueron muy utilizados los documentos de texto (74.1%) y las fotos (62.4%). Asimismo, la gran mayoría de los investigadores almacenaba sus datos en tablas (93.5%) y una gran parte de los encuestados lo hizo escritos a mano en papel (41.9%). Los datos fueron guardados mayormente en las computadoras personales de los investigadores (88.2%) o en discos externos (63.4%), y en menor medida en servidores institucionales (17.2%) o en la nube (11.8%). El respaldo de los datos de investigación fue hecho poco frecuentemente por casi la mitad de los investigadores entrevistados (47.3%), mientras 29% dijo hacerlo todos los meses y 6.4% nunca.

Las condiciones técnicas necesarias para que los investigadores depositaran sus datos en repositorios públicos fueron mayormente ser contactados si alguien quiere usar sus datos (65.6%) y poder editar los datos después de subirlos (64.5%). También fue importante la seguridad en la calidad de los datos y que el historial de uso y usuarios de los datos fuera visible (61.3%).

## Reutilización de datos

El 63.4% de los investigadores reutilizaban datos generados en otros laboratorios, 30% no estaban seguros de hacerlo y 8,6% no lo hizo. El 38,5% de los que no estaban dispuestos a disponibilizar sus datos reutilizó datos de otros investigadores, mientras 61,5% no lo hizo o no sabe si lo hizo. La calidad de los datos que los entrevistados dijeron requerir para reutilizar datos fue cómo se colectaron los datos (89.2%), dónde (73.1%), quién los recolectó (67.7%), cuándo se colectaron (66.7%) y, en menor medida, para qué fueron colectados (53.8%).

## DISCUSIÓN

Nuestro artículo presenta el primer estudio disponible investigando el estado de la cultura de los investigadores de instituciones uruguayas a compartir abiertamente

sus datos científicos. Nuestros análisis revelan que los investigadores están considerablemente dispuestos a compartir sus datos públicamente pero que no lo hacen debido a un obstáculo fundamental, que el tiempo y esfuerzo invertido para la generación de los datos no se traduzca en reconocimiento, en tanto crédito académico. Por lo tanto, este estudio ofrece el primer paso hacia la necesidad de promover una cultura de *data-sharing* con el objetivo primario de reforzar redes de colaboración y optimizar los beneficios individuales vía el desarrollo de sistemas de trabajo comunitario. Esta estrategia permitirá en última instancia que la información que se produce fundamentalmente a partir de actividades financiadas por fondos estatales pueda hacerse disponible para su uso en áreas que van más allá de la ciencia en sí misma, tales como la educación y el desarrollo de políticas públicas y planes de acción. Nuestro estudio además revela las legítimas preocupaciones que impulsan a numerosos investigadores a evitar la disponibilización de sus datos. Por ello, sugerimos que cada comunidad o grupo de trabajo establezca claros protocolos de control, embargo y retribución de quienes toman parte de la publicación de sus datos de forma abierta.

## Voluntad de compartir datos

El alto nivel de disposición de los investigadores a compartir datos, similar al encontrado entre pares europeos (Enke *et al.*, 2012), es contrastante con la escasa cantidad de planes de gestión de datos que se implementan (~20% de los encuestados). Considerando que la capacidad de compartir datos está limitada a las particularidades con la que los datos son generados, la falta de planes de gestión y la alta cantidad de datos que son registrados en papel pueden explicar en parte por qué existen números tan bajos de datos disponibles. El trabajo que implica para los investigadores curar los datos para cumplir estándares que les permitan cargarlos en repositorios globales, se vuelve excesivo cuando los datos no son creados en primera instancia con un foco en datos abiertos. Por otro lado, a nivel mundial y también en Uruguay, muchos datos de biodiversidad son generados por investigadores individuales que no forman parte de una posible estrategia institucional de curación de datos. Sumado a esto, las mediciones generalmente se alojan en computadoras y entornos de software que rápidamente se vuelven obsoletos y, por tanto, son proclives a no recuperarse para su reutilización (Griffin, 2015). En este contexto se torna clave el desarrollo e implementación de planes de gestión de datos que sean acordados por las distintas comunidades científicas (investigadores individuales e instituciones) y que progresivamente permitan mejorar el panorama de apertura de datos de biodiversidad nacional.

A primera vista resulta llamativo observar que la voluntad de compartir datos es menor entre estudiantes que entre científicos, investigadores o docentes. Esto se relaciona con el tiempo que expresaron sería necesario para hacer sus datos disponibles, que entre estudiantes es principalmente luego de publicar sus resultados y entre profesionales se observan más personas dispuestas a hacerlos disponibles inmediatamente. Una posible explicación es que ambos se encuentran en fases académicas diferentes y para aquellos en procesos formativos aun en desarrollo, consolidar su trayectoria a través de publicaciones podría ser más fundamental. Schmidt *et al.* (2016) vieron que el deseo de publicar resultados antes de publicar datos fue algo más frecuente en las primeras etapas de una carrera de investigación, resaltando esta necesidad como una barrera importante a la liberación de datos. Asimismo, similares resultados fueron encontrados por Tenopir *et al.* (2011), quienes observaron que los investigadores más jóvenes son menos proclives a compartir sus datos, aunque es más probable que acepten compartirlos mientras existan controles sobre su acceso. Por esto, vemos que la apertura de datos (actual y futura) está fuertemente condicionada a la posibilidad de publicación y por tanto de retribución académica y una consecuencia de esto es que todos aquellos datos generados a través de actividades o proyectos que no son incluidos en publicaciones académicas se perderían de (re)utilizar.

Por todo lo anterior, los datos de biodiversidad del Uruguay podrían caracterizarse como datos en riesgo (*"Data at Risk"*) (Griffin, 2015). Son denominados datos en riesgo aquellos que no están disponibles por falta de un diseño acorde en su colecta o por negligencia, o aquellos en riesgo de ser eliminados por degradación física, destrucción por desastres naturales, o debido a una supuesta falta de valor científico (Data Rescue Interest Group - Research Data Alliance - <https://rd-alliance.org/groups/data-rescue.html>). Para contrarrestar esta potencial pérdida sería necesario realizar el rescate de datos (*"Data Rescue"*) (Griffin, 2018), para así recuperar la información y colocarla en un entorno protegido donde los datos puedan ser curados y compartidos. Una vez más, se vuelve urgente trabajar en el desarrollo de estrategias colaborativas.

### Inquietudes y objeciones

La principal objeción de los investigadores a la hora de disponer sus datos públicamente es la falta de reconocimiento por el esfuerzo que implica, siendo la falta de estructuras de incentivos la

principal dificultad que enfrentan. En otros estudios, también se encontró que el esfuerzo en tiempo fue uno de los mayores obstáculos a la hora de compartir datos (Tenopir *et al.*, 2011; Enke *et al.*, 2012), lo que muestra que esto representa un obstáculo también a nivel internacional. Las cuestiones de propiedad intelectual y los asuntos legales también representan obstáculos mayores. La incertidumbre sobre la ley y el vacío existente a nivel de reglamentaciones institucionales aporta a complejizar la disponibilidad de información y extiende, en definitiva, los impedimentos de compartir datos. Por otra parte, es interesante observar que en Uruguay los aspectos financieros no se perciben como dificultades mayores, algo que de antemano podría advertirse como un obstáculo sustancial dado que, por ejemplo, este tipo de actividades no son en general objeto específico de financiación.

Es interesante señalar la visión de algunos entrevistados que marcaron como un obstáculo sentirse potencialmente perjudicados por contar con menos oportunidades y herramientas que en definitiva postergan la posibilidad de generar publicaciones a partir de los datos colectados. El temor pasa por haber dedicado un gran esfuerzo a coleccionar datos, compartirlos, pero no tener el tiempo suficiente para analizarlos y por tanto perder cierto crédito sobre los mismos. Si bien la apertura de datos implica en teoría acelerar el ritmo de la ciencia en un marco de oportunidades de colaboración, en la práctica existe una brecha entre las posibilidades de producción de conocimiento en instituciones académicas entre países que no puede ser descuidada. En este sentido, Meyer *et al.* (2015) resaltan la urgente necesidad de integrar fuentes de datos no occidentales y de intensificar la cooperación para abordar de manera más efectiva las necesidades de información de la sociedad sobre la biodiversidad. Actualmente, ciertas regiones del mundo están muy sobre representadas a nivel de datos en relación a otras (McRae *et al.*, 2017). Además, las regiones pobres en datos tienden a ser las áreas más ricas en biodiversidad que experimentan las amenazas más profundas (Schmeller *et al.*, 2017). En tiempos de disminución persistente de la biodiversidad, transformación de la tierra y cambio climático, es esencial que los países generen mecanismos para medir cuantitativamente la biodiversidad y monitorear sus cambios. La ausencia de datos puede llevar a la sobre o subrepresentación de la riqueza de especies y de su distribución, lo que finalmente puede traer errores en la priorización de ciertas áreas para la conservación (Boakes *et al.*, 2016). Por esto, si bien en Uruguay las urgencias obligaron a tener que definir políticas ambientales sin conocer debidamente la biodiversidad del país, esta falta de

información debe ser seriamente considerada a la hora de planificar y definir futuras iniciativas vinculadas a la conservación de la diversidad biológica. En este sentido, la colecta masiva de datos y a escalas cada vez más finas de resolución, tanto por parte de especialistas pero especialmente por parte de la ciudadanía (i.e., ciencia ciudadana), se plantea como una práctica transformadora ya que conducirá a una elaboración de políticas más informada y basada en pruebas (Macey, 2013).

### Motivaciones

Tender redes con otros científicos para estudios interdisciplinarios y la posibilidad de tener una alta disponibilidad de conjuntos de datos comparables para análisis más completos, resultaron los principales incentivos de los científicos para compartir sus datos. Esto posiciona a las estructuras colaborativas como herramientas sustanciales para permitir prácticas de generación de conocimiento en un marco de ciencia abierta.

A nivel mundial, el cambio sociocultural y el movimiento hacia una ciencia más abierta han evolucionado más rápidamente en las últimas dos décadas en respuesta a nuevos requisitos establecidos por organizaciones gubernamentales, editoriales y sociedades científicas. En algunos países, depositar datos de investigación generados a partir de fondos públicos es un requisito, por ejemplo, en Alemania a través de la German Science Foundation (DFG) y la National Science Foundation (NSF) en Estados Unidos. También, las revistas internacionales en el área de biodiversidad cada vez más requieren que los datos utilizados en el marco de sus publicaciones sean liberados en repositorios accesibles. En este sentido varias iniciativas internacionales de gran importancia han sido desarrolladas para permitir que los datos sean reconocibles, libremente reutilizables y citables, como son: figshare (<https://figshare.com/>), Dryad (<http://datadryad.org>), DataONE (<https://dataone.org>), Scientific Data (<https://nature.com/sdata/>) y FAIRsharing (<http://fairsharing.org>), entre otras. En Uruguay no existen iniciativas relacionadas con reconocer, recompensar o retribuir a quienes comparten datos científicos. En particular, las agencias de financiación pública nacionales, como la Agencia Nacional de Investigación e Innovación (ANII) o la UdelaR a través de su Comisión Sectorial de Investigación Científica (CSIC), no poseen exigencias que impongan presiones a los investigadores respecto a la apertura de sus datos. Asimismo, si bien la ley 18.381 del 2008 establece el derecho de acceso a la información pública en Uruguay y a nivel gubernamental existe AGESIC (Agencia para el Desarrollo del Gobierno de Gestión Electrónica y la Sociedad de la Información y del

Conocimiento), actualmente no existen políticas institucionales transversales en relación al acceso a datos e información científica financiada con fondos públicos.

### CONCLUSIONES

Los investigadores en Uruguay parecen dispuestos a compartir sus datos de investigación, sin embargo, en la práctica esto no termina de concretarse debido a dos grandes impedimentos: la falta de planes de gestión de datos y el desconocimiento de estructuras de retribución por los datos generados. A esto se le suma el escaso estímulo a nivel gubernamental y académico por la apertura de datos científicos, que termina por configurar un panorama crítico que nos posiciona entre los países con menor apertura de datos en América.

En este escenario, se torna imprescindible mejorar la infraestructura que apoya la reutilización de datos derivados de la investigación científica en el país. Es vital entender que la comunidad de colectores y usuarios de datos requiere incentivos efectivos para asegurar su participación, que ayude a superar las barreras sociales y culturales existentes y promueva un marco de producción de conocimiento más colaborativo. Por otro lado, es necesario avanzar hacia un cambio de cultura ya que muchos investigadores en Uruguay no comparten sus datos debido a que “por defecto” es un hábito que no se practica, incluso a pesar de que reutilizan datos de otros investigadores. Se necesita difundir la importancia y los beneficios de la apertura e intercambio de datos para mejorar el conocimiento sobre cómo compartir y así poder finalmente traducir la voluntad en una práctica efectiva de ciencia abierta. Asimismo, se torna imperativo para el país buscar una solución a la potencial pérdida de datos, garantizando que los datos de biodiversidad sean rescatados, archivados y puestos a disposición de la comunidad científica y ciudadana. En consecuencia, el desarrollo de políticas institucionales de acceso a datos científicos financiados con fondos públicos permitiría optimizar el uso de los recursos nacionales involucrados en la generación de conocimiento y, también, asegurar el amplio acceso y la preservación en el tiempo de los datos.

Por último, resulta importante mencionar reglamentaciones como el reciente Decreto N° 115/018 y su modificación 269/018 sobre la Ley 19.175 de Recursos Hidrobiológicos, ya que van en contra de la disponibilidad de datos científicos (Sommaruga *et al.*, 2018). Concretamente, el Decreto establece que «Los resultados y datos obtenidos en lo que respecta a las competencias de Dirección Nacional de Recursos Acuáticos [DINARA] deberán ser puestos a disposición de este organismo antes de su utilización o divulgación

en cualquier medio». De acuerdo con esta reglamentación, no se podrán publicar datos derivados de investigaciones relacionadas con recursos hidrobiológicos, pesca, acuicultura y ambiente, sin autorización expresa de la DINARA. Este tipo de limitaciones en la publicación y el acceso a la información ambiental van diametralmente en contra de nuevos avances políticos a nivel regional como el Acuerdo de Escazú sobre el Acceso a la Información, la Participación Pública y el Acceso a la Justicia en Asuntos Ambientales en América Latina y el Caribe (CEPAL, 2018), y la Declaración de Panamá sobre Ciencia Abierta en el marco del Foro Abierto de Ciencias de América Latina y el Caribe. En este sentido, resulta imprescindible demandar en Uruguay el diseño de políticas, leyes, normas y prácticas que no impongan restricciones y limitaciones al proceso de generación de conocimiento y la apropiación asimétrica de los resultados, sino que, por el contrario, procuren avanzar hacia una investigación abierta, reproducible y replicable para todos y todas.

## AGRADECIMIENTOS

La autora quiere agradecer a todos los participantes de la encuesta electrónica por su tiempo, interés y valiosos comentarios. También, agradecemos a los revisores por sus sugerencias y aportes para la discusión política de las implicancias en la apertura de datos en Uruguay. La investigación que da origen a los resultados presentados en la presente publicación recibió fondos de la Agencia Nacional de Investigación e Innovación bajo el Código POS\_EXT\_2016\_1\_136663.

## BIBLIOGRAFÍA

- Boakes E.H., Fuller R.A., McGowan P.J.K. & G.M. Mace. 2016. Uncertainty in identifying local extinctions: the distribution of missing data and its effects on biodiversity measures. *Biology Letters*, 12.
- Borgman C.L. 2012. The conundrum of sharing research data. *Journal of the American Society for Information Science and Technology*, 63:1059-1078.
- CEPAL Comisión Económica para América Latina y el Caribe. 2018. Acuerdo Regional sobre el Acceso a la Información, la Participación Pública y el Acceso a la Justicia en Asuntos Ambientales en América Latina y el Caribe.
- Enke N., Thessen A., Bach K., Bendix J., Seeger B. & B. Gemeinholzer. 2012. The user's view on biodiversity data sharing — Investigating facts of acceptance and requirements to realize a sustainable use of research data —. *Ecological Informatics*, 11:25-33.
- Fecher B., Friesike S. & M. Hebing. 2015. What Drives Academic Data Sharing? *PLoS One*, 10:e0118053.
- Fienberg S.E., Martin M.E. & M.L. Straf. 1985. Sharing research data. National Academy Press, Washington, DC, US.
- Griffin E. 2018. The Data that Get Forgotten. Pages 183-198 *Data Science Landscape*. Springer.
- Griffin R.E. 2015. When are Old Data New Data? *GeoResJ*, 6:92-97.
- Guralnick R.P., Hill A.W. & M. Lane. 2007. Towards a collaborative, global infrastructure for biodiversity assessment. *Ecology letters*, 10:663-672.
- Hey T., Tansley S. & K.M. Tolle. 2009. The fourth paradigm: data-intensive scientific discovery. Microsoft research Redmond, WA.
- Linek S.B., Fecher B., Friesike S. & M. Hebing. 2017. Data sharing as social dilemma: Influence of the researcher's personality. *PLoS One*, 12:e0183216.
- Macey G.P. 2013. The architecture of ignorance. *Utah Lab Review*:1627.
- McRae L., Deinet S. & R. Freeman. 2017. The diversity-weighted living planet index: controlling for taxonomic bias in a global biodiversity indicator. *PLoS One*, 12:e0169156.
- Meyer C., Kreft H., Guralnick R. & W. Jetz. 2015. Global priorities for an effective information basis of biodiversity distributions. *Nature Communications*, 6:8221.
- Michener W.K. 2015. Ecological data sharing. *Ecological Informatics*, 29:33-44.
- R Core Team. 2018. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.
- Schmeller D.S., Böhm M., Arvanitidis C., Barber-Meyer S., Brummitt N., Chandler M., Chatzinikolaou E., Costello M.J., Ding H., García-Moreno J., Gill M., Haase P., Jones M., Juillard R., Magnusson W.E., Martin C.S., McGeoch M., Mihoub J.-B., Pettorelli N., Proença V., Peng C., Regan E., Schmedel U., Simaika J.P., Weatherdon L., Waterman C., Xu H. & J. Belnap. 2017. Building capacity in biodiversity monitoring at the global scale. *Biodiversity and Conservation*, 26:2765-2790.
- Schmidt B., Gemeinholzer B. & A. Treloar. 2016. Open Data in Global Environmental Research: The Belmont Forum's Open Data Survey. *PLoS One*, 11:e0146695.
- Sommaruga R., Rodríguez-Graña L., Calliari D., Lercari D. & L. Aubriot. 2018. Legislation restricts research in Uruguay. *Science*, 362:1008-1008.
- Tenopir C., Allard S., Douglass K., Aydinoglu A.U., Wu L., Read E., Manoff M. & M. Frame. 2011. Data sharing by scientists: practices and perceptions. *PLoS One*, 6:e21101.
- Wilkinson M.D., Dumontier M., Aalbersberg I.J.,

Appleton G., Axton M., Baak A., Blomberg N., Boiten J.-W., da Silva Santos L.B., Bourne P.E., Bouwman J., Brookes A.J., Clark T., Crosas M., Dillo I., Dumon O., Edmunds S., Evelo C.T., Finkers R., Gonzalez-Beltran A., Gray A.J.G., Groth P., Goble C., Grethe J.S., Heringa J., 't Hoen P.A.C., Hooft R., Kuhn T., Kok R., Kok J., Lusher S.J., Martone M.E., Mons A., Packer A.L., Persson B., Rocca-Serra

P., Roos M., van Schaik R., Sansone S.-A., Schultes E., Sengstag T., Slater T., Strawn G., Swertz M.A., Thompson M., van der Lei J., van Mulligen E., Velterop J., Waagmeester A., Wittenburg P., Wolstencroft K., Zhao J. & B. Mons. 2016. The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*, 3:160018.

## APÉNDICE 1

### 1. ¿En qué tipo de institución trabajás?

*Marcá todas las opciones que correspondan*

- Museo  
 Universidad  
 Gobierno  
 Organización sin fines de lucro  
 Empresa privada

### 2. Estatus profesional

*Marcá todas las opciones que correspondan*

- Científico/a trabajando en el gobierno  
 Científico/a trabajando en el sector privado  
 Investigador/a  
 Científico/a Ciudadano/a (citizen scientist)  
 Estudiante de Maestría  
 Estudiante de Doctorado  
 PostDoc  
 Docente  
 Voluntario/a en una Organización sin fines de lucro  
 Otro

### 3. Seleccioná tu campo de investigación

*Marcá todas las opciones que correspondan*

- Biofísica  
 Biogeografía  
 Bioinformática  
 Biología Celular  
 Biología de la Conservación

- Biología del Comportamiento  
 Biología del Desarrollo  
 Biología Molecular  
 Bioquímica  
 Botánica  
 Ciencias Agrícolas  
 Ciencias Veterinarias  
 Ecología  
 Educación Ambiental  
 Evolución  
 Fisiología  
 Genética  
 Geografía  
 Geología  
 Micología  
 Microbiología  
 Neurobiología  
 Oceanografía  
 Paleontología  
 Sistemática  
 Taxonomía  
 Virología  
 Zoología  
 Otro

## 4. Seleccioná tu(s) grupo(s) de estudio

*Marcá todas las opciones que correspondan*

- Mamíferos  
 Aves  
 Reptiles  
 Anfibios  
 Peces  
 Insectos  
 Arácnidos  
 Crustáceos  
 Moluscos  
 Plantas vasculares  
 Hongos  
 Briofitas/Pterofitas  
 Otro

## 5. ¿Con qué tipo de datos trabajás regularmente?

*Marcá todas las opciones que correspondan*

- Listas de especies (presencia/ausencia)  
 Datos del ecosistema  
 Datos moleculares  
 Datos morfológicos  
 Datos fisiológicos  
 Datos taxonómicos  
 Otro

## 6. ¿Qué tipo de metadatos recopilás?

*Marcá todas las opciones que correspondan*

- Nombre del organismo  
 Localidad  
 Descripción de la localidad  
 Coordenadas geográficas  
 Fecha de colecta  
 Colector/a  
 Número de colección  
 Configuración del equipo de colecta  
 Precisión de las mediciones  
 Otro

## 7. ¿Con qué formatos de datos trabajás?

*Marcá todas las opciones que correspondan*

- Texto plano (.txt, .doc, .docx, .odt)  
 Tabla (.xls, .xlsx, .ods, .csv)  
 Foto (.jpg, .tif, .gif, .png)  
 Audio (.wav, .wma, .mp3, .m4a, .mpeg, .aac)  
 Video (.avi, .flv, .wmv, .mov, .mp4)  
 Secuencias (.fas, .fasta)  
 Otro

## 8. ¿Hay bases de datos en las que podés depositar tus datos?

- Si  
 No

## 9. Si existen bases de datos en la que depositar tus datos, especificá cuáles

## 10. ¿Cuánto tiempo pasás al año con la curación y el cargado de datos?

- Menos de 1 día  
 De 1 a 3 días  
 Una semana  
 Más de una semana  
 Otro

## 11. ¿A quién pertenecen tus datos?

- Me pertenecen a mi  
 A quien me financia  
 A mi institución  
 Al gobierno  
 A nadie  
 No lo sé  
 Otro

12. ¿En qué formato guardás tus datos?

*Marcá todas las opciones que correspondan*

- Escritos a mano (por ejemplo en un manual de campo)
- Tabla de Excel o Calc
- Documento de Word o Writer
- Base de datos Access o Base
- Otra base de datos
- Otro

13. ¿Dónde guardás tu copia local de los datos?

*Marcá todas las opciones que correspondan*

- Computadora propia
- Computadora institucional
- Servidor institucional
- Disco duro externo
- Memoria USB
- Otro

14. ¿Con qué frecuencia hacés una copia de seguridad de tus datos?

- Todos los días
- Todas las semanas
- Todos los meses
- Poco frecuentemente
- Nunca
- Otro

15. ¿Estarías dispuesto/a a depositar tus datos en bases de datos disponibles públicamente?

- Si
- No

16. ¿Cuál dirías que es la principal dificultad que enfretás para compartir datos?

- Problemas tecnológicos
- Cuestiones institucionales
- Cuestiones financieras y presupuestarias
- Asuntos legales y políticos
- Falta de estructuras de incentivo
- Otro

17. ¿Qué tipo de inquietudes y/u objeciones tenés para depositar tus datos de investigación en bases de datos online?

*Marcá todas las opciones que correspondan*

- Propiedad intelectual o cuestiones de patentes
- Confidencialidad o asuntos legales
- Pérdida de control de mis datos
- Tiempo y esfuerzo invertido
- Falta de reconocimiento por mi esfuerzo
- Calidad de mis datos
- Falta de estándares de datos
- Que alguien saque conclusiones erróneas de mis datos
- Críticas sobre mis datos o análisis
- Otro

18. ¿Cuándo estarías dispuesto/a a compartir tus datos?

- Inmediatamente
- Después de publicar
- Al final del proyecto
- Después de 5 años
- Después de 10 años
- Luego de mi retiro de la ciencia
- Otro

19. ¿Qué te haría más propenso/a a compartir datos?

*Marcá todas las opciones que correspondan*

- Tener un lugar donde depositar datos
- Obtener reconocimiento o crédito si mis datos son usados por otros/as
- Obtener financiamiento para compartir datos
- La capacidad de controlar la reutilización de mis datos
- Otro

20. Si tus datos se usaran, ¿cómo te gustaría que te citaran?

- En los agradecimientos
- En la Bibliografía
- Como co-autor
- Otro

21. ¿Qué condiciones requerirías para depositar tus datos?

*Marcá todas las opciones que correspondan*

- Poder editar los datos después de subirlos
- Poder eliminar datos
- Visualización del historial de uso (quién, cuándo, para qué)
- Ser contactado/a si alguien quiere usar mis datos
- Garantía de mantenimiento de mis datos a largo plazo
- Seguridad en la calidad de los datos
- Carga de datos fácil de usar
- Pautas para la reutilización de datos
- Otro

22. ¿Qué podría motivarte a hacer que tus datos estén disponibles?

*Marcá todas las opciones que correspondan*

- Mostrar transparencia en los resultados de mis estudios
- Habilitaría una alta disponibilidad de conjuntos de datos comparables para análisis más completos
- Tender redes con otros/as científicos/as para estudios interdisciplinarios
- Aumentar mi visibilidad en la comunidad
- Como los datos se generaron con dinero público, deberían hacerse públicos
- Otro

23. ¿Qué otras cosas te motivarían?

*Marcá todas las opciones que correspondan*

- Que sea un requisito de las revistas
- Que sea un requisito de quienes me financian
- Que los conjuntos de datos se clasifiquen como publicaciones
- Otro

24. ¿Reutilizás datos generados en otros laboratorios?

- Sí
- No
- No lo sé

25. ¿Qué tipo de información necesitás para evaluar la calidad de los datos de otras personas?

*Marcá todas las opciones que correspondan*

- ¿Quién recolectó los datos?
- ¿Cómo se recolectaron los datos?
- ¿Cuándo se recolectaron los datos?
- ¿Para qué se recolectaron los datos?
- ¿Dónde se recolectaron los datos?
- Otro

26. ¿Tenés un plan de gestión de datos?  
"Data management plan" en inglés

- Sí
- No
- No lo sé

27. ¿Qué tan importantes creés que son los planes de gestión de datos?

- No importan en absoluto
- Algo importantes
- Importantes
- Muy importantes
- Extremadamente importantes

28. ¿Quisieras agregar algún comentario final?

*Fecha de Recepción: 20 de agosto de 2018  
Fecha de Aceptación: 20 de diciembre de 2018*