



Facultad de  
**Información y  
Comunicación**



UNIVERSIDAD  
DE LA REPÚBLICA  
URUGUAY

## **Maestría en Información y Comunicación**

Tesis para defender el título de la Maestría en  
Información y Comunicación

### **Comportamiento informativo colaborativo y prácticas al compartir datos sobre biodiversidad marina**

Autora: Andrea Cristiani Sienna

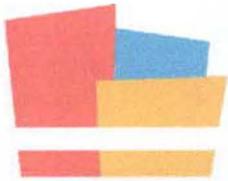
Directora de tesis: Dra. Martha Sabelli

Montevideo

Agosto de 2018



Esta obra se encuentra sujeta a la licencia Creative Commons  
Atribución – No Comercial – Sin Derivadas (CC BY-NC-ND 4.0)



Facultad de  
Información y  
Comunicación



UNIVERSIDAD  
DE LA REPÚBLICA  
URUGUAY

El Tribunal docente integrado por los abajo firmantes, aprueba la Tesis:

“Comportamiento informativo colaborativo y prácticas al compartir datos sobre biodiversidad marina”

Tesista: Andrea Cristiani Sienna

Maestría en Información y Comunicación

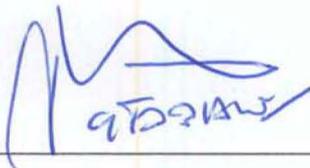
Fallo:

APROBADO CON MENCIÓN

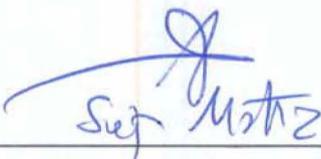


Tribunal:

Profesor/a:

  
\_\_\_\_\_

Profesor/a:

  
\_\_\_\_\_

Profesor/a:

  
\_\_\_\_\_ MARIO BARRIÉ

## **Agradecimientos**

A Martha por la dirección académica y también por la dirección de esta tesis de maestría.

A todos los investigadores que colaboraron con esta investigación: los que completaron la encuesta, los que accedieron a la entrevista y a todos los que además enviaron correos electrónicos con información adicional. Sin esos aportes no hubiera sido posible esta investigación.

Muchas gracias a los responsables de nodo OBIS que apartaron unos momentos de sus apretadas agendas para la entrevista. Las experiencias y opiniones compartidas en esas entrevistas han sido muy importantes para profundizar en la temática abordada.

Un agradecimiento muy especial a los investigadores que luego de completar la encuesta se comunicaron a mi correo electrónico para manifestar sus deseos de éxito y su interés por conocer los resultados de la investigación.

## Tabla de contenido

<b>Lista de acrónimos y siglas.....</b>	<b>7</b>
<b>Lista de gráficos .....</b>	<b>7</b>
<b>Lista de tablas .....</b>	<b>9</b>
<b>Resumen.....</b>	<b>10</b>
<b>Abstract .....</b>	<b>11</b>
<b>Introducción .....</b>	<b>12</b>
<b>Capítulo 1. Fundamentación del objeto de estudio, antecedentes y objetivos de la investigación .....</b>	<b>14</b>
1.1. Objeto de estudio .....	14
1.2. Antecedentes de estudios sobre comportamiento informativo en comunidades científicas .....	15
1.3. Antecedentes de estudios sobre compartir datos en comunidades científicas..	16
1.4. Preguntas de investigación .....	18
1.5. Objetivos.....	19
1.5.1. Objetivo general.....	19
1.5.2. Objetivos específicos .....	20
<b>Capítulo 2. Marco teórico .....</b>	<b>21</b>
2.1. Información y datos .....	22
2.2. Compartir información y comportamiento al compartir información .....	22
2.2.1. Prácticas de información.....	25
2.2.2. Prácticas al compartir datos ( <i>Data sharing practices</i> ) .....	27
En síntesis .....	28
<b>Capítulo 3. Datos e información sobre biodiversidad marina .....</b>	<b>30</b>
3.1. Los datos en biodiversidad marina .....	30
3.3. Compartir datos como práctica de información colaborativa.....	31

3.4. OBIS: el proyecto colaborativo que da marco a este estudio .....	32
<b>Capítulo 4. Metodología .....</b>	<b>36</b>
4.1. Encuesta a investigadores en biodiversidad marina .....	36
4.2. Análisis de contenido de los documentos de las reuniones del Grupo Directivo .....	38
4.3. Entrevistas a coordinadores de los nodos del sistema OBIS .....	44
<b>Capítulo 5. Resultados.....</b>	<b>45</b>
5.1. Encuesta.....	45
5.1.1. Análisis de los resultados de la encuesta .....	46
5.1.2. Factores que más inciden y prácticas más habituales al compartir datos en función de las variables edad y posición en el equipo de investigación.....	63
5.2. Análisis de contenido de los documentos de las reuniones del Grupo Directivo de OBIS .....	75
5.2.1. Factores que desalientan las actividades de compartir datos e información .....	78
5.2.2. Factores propicios a las actividades de compartir datos e información.....	79
5.3. Entrevistas a coordinadores de nodos OBIS.....	80
5.3.1. Principales problemas al compartir datos .....	80
5.3.2. Retención de datos .....	82
5.3.3. Prácticas para obtener datos.....	83
5.4. Discusión de los resultados de las tres técnicas aplicadas .....	84
5.4.1. Motivos y factores favorables a compartir datos .....	85
5.4.2. Motivos y factores que desalientan compartir datos.....	86
5.4.3. Prácticas para obtener datos.....	87
<b>Capítulo 6. Conclusiones y proyecciones .....</b>	<b>89</b>
6.1. Conclusiones.....	89

6.1.1. Motivaciones y factores que propician compartir datos sobre biodiversidad marina .....	90
6.1.2. Motivos y factores que desalientan compartir datos sobre biodiversidad marina .....	92
6.1.3. Principales prácticas para compartir datos sobre biodiversidad marina ....	94
6.1.4. Edad y posición en el equipo de investigación y comportamiento al compartir datos .....	96
6.2. Proyecciones .....	97
<b>Referencias Bibliográficas.....</b>	<b>98</b>
<b>Anexos.....</b>	<b>1</b>
Anexo 1. Formulario de la encuesta .....	2
Anexo 2. Lista de Nodos OBIS .....	7
Anexo 3. Mapa de instituciones proveedoras de datos a OBIS .....	9
Anexo 4. Datos recabados en la encuesta .....	10
Anexo 5. Ficha de registro - Análisis de contenido.....	22
Anexo 6. Textos correspondientes a pasajes de las entrevistas mencionados en esta tesis .....	25

## Lista de acrónimos y siglas

<b>ABCD</b>	<i>Access to Biological Collections Data</i>
<b>AC</b>	Análisis de Contenido
<b>CHI</b>	Comportamiento Humano Informativo
<b>DwC</b>	<i>Darwin Core</i>
<b>IODE</b>	<i>International Oceanographic Data and Information Exchange</i>
<b>OBIS</b>	<i>Ocean Biogeographic Information System</i>
<b>QA/QC</b>	<i>Quality Assessment / Quality Control</i>
<b>SG</b>	<i>Steering Group</i>
<b>TIC</b>	Tecnologías de la Información y la Comunicación
<b>TT</b>	<i>Task Team</i>
<b>UNESCO</b>	Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura
<b>WoRMS</b>	<i>World Register of Marine Species</i>

## Lista de gráficos

Gráfico 1. Distribución de los encuestados por género .....	46
Gráfico 2. Distribución etaria de los encuestados .....	47
Gráfico 3. Distribución de respuestas por región geográfica .....	47
Gráfico 4. Posición en el equipo investigación .....	48
Gráfico 5. Motivos para compartir datos .....	50
Gráfico 6. Factores que influyen positivamente al compartir datos sobre biodiversidad marina .....	52
Gráfico 7. Compartir datos: con y sin condiciones.....	53
Gráfico 8. Condiciones para compartir datos .....	54
Gráfico 9. Prácticas para obtener datos .....	55
Gráfico 10. Motivos para retener datos .....	57

Gráfico 11. Factores desfavorables a compartir datos.....	59
Gráfico12. Importancia de contar con un repositorio de datos según grupos de edades .....	63
Gráfico13. Importancia de contar con un repositorio de datos según la posición en el equipo de investigación .....	64
Gráfico14. Importancia de la existencia de una política de intercambio de datos según grupos de edades.....	65
Gráfico15. Importancia de la existencia de una política de intercambio de datos según posición en el equipo de investigación .....	65
Gráfico 16. Importancia de la utilización de estándares según grupo de edades .....	66
Gráfico 17. Importancia de la utilización de estándares según posición en el equipo de investigación .....	67
Gráfico 18. Frecuencia de búsqueda de datos sobre biodiversidad marina en línea según grupos de edad .....	68
Gráfico 19. Frecuencia de búsqueda de datos sobre biodiversidad marina en línea según posición en el equipo de investigación .....	68
Gráfico 20. Solicitudes personales de datos sobre biodiversidad marina según grupos de edades.....	69
Gráfico 21. Solicitudes personales de datos sobre biodiversidad marina según posición en el equipo de investigación.....	70
Gráfico 22. Frecuencia con que buscan en un repositorio central o centro de datos según grupo de edades .....	70
Gráfico 23. Frecuencia con que buscan en un repositorio central o centro de datos según posición en el equipo de investigación .....	71
Gráfico 24. Preocupación por el uso inadecuado de los datos según grupos de edad .	72
Gráfico 25. Preocupación por el uso inadecuado de los datos según posición en el equipo de investigación .....	72
Gráfico 26. Importancia del trabajo adicional requerido para compartir los datos según grupos de edad .....	73
Gráfico 27. Importancia del trabajo adicional requerido para compartir los datos según posición en el equipo de investigación .....	74
Gráfico 28. Importancia conferida a perder oportunidades de publicación por compartir datos según grupo de edades.....	74

Gráfico 29. Importancia conferida a perder oportunidades de publicación por compartir datos según posición en el equipo de investigación ..... 75

### **Lista de tablas**

Tabla 1. Región - Opción Otros..... 48

Tabla 2. Opción Otros - Pregunta 11 ..... 49

## Resumen

Se trata de un estudio descriptivo sobre las prácticas al compartir datos de los investigadores en biodiversidad marina, enmarcado en un sistema global de información que reúne datos biogeográficos de especies marinas. El ámbito de trabajo colaborativo de este sistema, donde los participantes comparten datos a través de plataformas que proveen acceso abierto, es muy propicio para llevar a cabo estudios sobre comportamiento informativo colaborativo. Se identifican y estudian las prácticas al compartir datos, los factores favorables que alientan estas prácticas y los factores desfavorables que desalientan el compartir datos, a través de las opiniones y experiencias que aportan los propios investigadores. Las técnicas utilizadas son encuesta, entrevista y análisis de contenido. Se utilizan enfoques específicos de la teoría del comportamiento informativo colaborativo y de las prácticas de información colaborativas para interpretar los resultados. Las conclusiones de la investigación pretenden aportar insumos para contribuir a solucionar las complejidades actuales al compartir datos sobre biodiversidad marina y que ello redunde en el desarrollo de mejores productos y servicios de información basados en datos para esta área del conocimiento.

**Palabras clave:** comportamiento informativo, investigadores, biodiversidad marina, prácticas al compartir datos, colaboración científica

## **Abstract**

This is a descriptive study on data sharing practices of marine biodiversity researchers, framed in a global information system that gathers biogeographical data of marine species. The collaborative work environment of this system, where participants share data via open access platforms, is very suitable to carry on collaborative information behavior studies. Data sharing practices are identified and studied, regarding to both, favorable factors that encourage those practices and unfavorable factors that discourage the sharing of data, through the opinions and experiences of the researchers. The techniques applied are survey, interview and content analysis. Specific approaches to collaborative information behavior and information sharing practices are used to interpret the results. The conclusions of the research are intended to provide inputs to help solving the current complexities of sharing data on marine biodiversity, which will benefit the development of better products and information services based on data for this area of knowledge.

**Keywords:** information behavior, marine biodiversity, researchers, data sharing practices, scientific collaboration

## Introducción

La producción de conocimientos científicos basada en la colaboración entre grupos de diferentes instituciones y países y, actualmente, las basadas en la constitución de redes de investigación no son ajenas a ninguna disciplina. La tendencia al predominio de la colaboración se puede apreciar en las formas de organización de proyectos y redes. Se manifiesta de variadas formas y en distintas etapas de la producción de conocimiento, desde la recolección y generación de los datos hasta la publicación de los resultados.

Cuando los datos sobre biodiversidad marina se comparten de forma abierta, se logran productos y servicios de información muy valiosos para la comunidad científica, los tomadores de decisión y el público en general. Sin embargo, el compartir datos primarios de investigación sobre biodiversidad marina en cualquiera de sus modalidades (ya sea en abierto de forma total o parcial, restringida o luego de un período de embargo) tiene sus complejidades y problemáticas propias.

Las tecnologías de información, con sus herramientas para gestionar metadatos y repositorios, propagan y facilitan compartir datos e información. Sin embargo, aún existen asuntos complejos por resolver (interoperabilidad de los sistemas, compatibilidad de los formatos, propiedad de los datos, derechos de uso, dispersión y duplicación, por mencionar algunos), más cuando los datos y la información se comparten de forma global.

La principal justificación del estudio es ampliar el conocimiento sobre los comportamientos informativos de colaboración y el uso de la información por parte de los científicos que participan en proyectos internacionales en biodiversidad marina.

Se ha escogido como marco del estudio un proyecto internacional de carácter global en el que se comparten datos sobre biodiversidad marina en un contexto de trabajo científico colaborativo: el sistema de información *Ocean Biogeographic Information System* (OBIS)<sup>1</sup>, cuyos objetivos y características aparecen detalladas en el Capítulo 3, apartado 3.4.

---

<sup>1</sup> *Ocean Biogeographic Information System* (OBIS). Disponible en Internet: <http://www.iobis.org/>

En esta investigación se pretende una aproximación a las dificultades y complejidades de compartir datos desde la mirada del comportamiento informativo. La estrategia para abordar la problemática estudiada se basa en la identificación y el estudio de prácticas de datos concretas, así como también de los múltiples factores que inciden en esas prácticas. Se utilizan tanto métodos cuantitativos como cualitativos para recabar puntos de vista, preferencias, opiniones y motivos de los investigadores respecto de compartir datos e información sobre biodiversidad en la era de la información electrónica. El identificar y estudiar esos aspectos permite profundizar en comportamientos informativos que son de relevancia para el desarrollo de nuevos servicios y sistemas de información enfocados a académicos trabajando en ambientes colaborativos, y también pretende aportar insumos que contribuyan con mejores soluciones respecto a servicios de información y comunicación ya existentes para esta área del conocimiento.

Los lectores potenciales de esta investigación son principalmente los especialistas en información y en gestión de datos, especialmente quienes planifican sistemas de información para el área biología marina, además de quienes investigan o estudian comportamiento informativo y prácticas de información. Los investigadores en biodiversidad marina y aquellos que planifican y dirigen proyectos que implican compartir datos en esa área del conocimiento también son potenciales lectores de este trabajo.

La tesis se organiza en seis capítulos. El capítulo 1 presenta el objeto de estudio de la investigación, la fundamentación, las preguntas de investigación, los objetivos, y los antecedentes. El capítulo 2 desarrolla el marco teórico y define los conceptos fundamentales del enfoque que guía posteriormente el análisis de los resultados de la investigación. El capítulo 3 se ocupa del ámbito en el que se realiza la investigación y se exponen allí las características y objetivos del sistema de información que da marco a esta investigación. En el capítulo 4 se explican la metodología y las técnicas aplicadas en el estudio, con un apartado destinado a cada una de las tres técnicas que luego permitirán triangular los resultados. El capítulo 5 está dedicado al análisis de los resultados de los datos recabados con las herramientas presentadas en el capítulo 4. Al final del capítulo 5 se incluye un análisis que integra los resultados de las tres técnicas empleadas. A modo de cierre, en el capítulo 6 se presentan las conclusiones y las proyecciones del estudio.

## **Capítulo 1. Fundamentación del objeto de estudio, antecedentes y objetivos de la investigación**

En este capítulo se presenta el objeto de estudio de la investigación, fundamentando su relevancia. Luego se exponen los antecedentes de estudios previos sobre comportamiento informativo en comunidades científicas, y a continuación se incluyen las preguntas de investigación y los objetivos propuestos para la misma.

### **1.1. Objeto de estudio**

El objeto de estudio son las prácticas de información relacionadas con compartir datos de los investigadores que se dedican a estudiar la biodiversidad marina.

En primer lugar, se incluyen las definiciones de los términos datos y biodiversidad, para luego pasar a explicar qué se entiende por prácticas de información.

Los datos son representaciones precisas y bien definidas de observaciones, descripciones o mediciones de un referente (objeto, fenómeno o evento) registradas de alguna manera estándar y bien especificada (AnthroDPA Metadata Working Group, 2009) (mi traducción).

La biodiversidad o diversidad biológica se define en el Convenio de las Naciones Unidas sobre la Diversidad Biológica de la siguiente forma:

Es la variabilidad entre los organismos vivos de todos los orígenes, incluidos entre otros, terrestres, marinos y otros ecosistemas acuáticos y los complejos ecológicos de los que forman parte: esto incluye la diversidad dentro de las especies, entre las especies y de los ecosistemas (United Nations Environment Programme, 1992, p.3) (mi traducción).

"Prácticas de información" es un concepto que se discute a menudo en Ciencia de la Información y cuyo significado se abordará más detenidamente en el capítulo 2. Si se compara con el término "comportamiento informativo", el foco de las "prácticas de información" se encuentra en las dimensiones sociales y culturales de las actividades relacionadas con la información y en la naturaleza de la relación de estas prácticas con su contexto (Byström & Lloyd, 2012, p.4) (Savolainen, 2007, p.126).

Cuando se alude al comportamiento informativo, se hace referencia al comportamiento con respecto a la información: a la información que se necesita, a la información que se

genera, a la información que se comparte y a la que se difunde en el contexto del trabajo científico y de investigación.

## **1.2. Antecedentes de estudios sobre comportamiento informativo en comunidades científicas**

Las formas que tienen las personas de interactuar con la información, pasando por la necesidad, la búsqueda, la gestión y el uso de información, han sido objeto de estudio de muchas investigaciones a nivel nacional e internacional desde las teorías sobre Comportamiento Humano Informativo (CHI), un área de estudio bastante desarrollada actualmente dentro de la Ciencia de la Información.

No se encontraron antecedentes en la literatura sobre la aplicación del enfoque de las prácticas de información al estudio de actividades colaborativas en las que se comparten datos sobre biodiversidad marina. Si bien no se han encontrado antecedentes específicos, si los hay relacionados con estudios sobre comportamiento informativo colaborativo en comunidades académicas en otras áreas de la actividad científica.

Los antecedentes de investigación correspondientes a trabajos nacionales sobre comportamiento informativo en comunidades académicas corresponden a una investigación de Pérez y Sabelli (2003) titulada “Uso de información electrónica por los académicos de la Universidad de la República”, que estudió el comportamiento informativo de académicos de todas las Facultades de la Universidad respecto al uso de información en formato electrónico.

De la revisión de bibliografía internacional, centrada específicamente en casos en los que se ha estudiado el comportamiento informativo colaborativo en comunidades académicas, surgen interesantes enfoques y conceptos, algunos de los cuales se retoman en detalle en el capítulo 2, Marco Teórico.

Talja (2002) desarrolla un marco conceptual para describir varios tipos y niveles al compartir información en comunidades académicas; Bao y Bouthillier (2007) en su artículo *Information sharing as a type of information behavior* abordan la naturaleza de este comportamiento informativo y se plantean el desafío de establecer un indicador cuantitativo para medirlo; Poteri (2007) en su tesis de grado sobre comportamiento

informativo colaborativo estudia el comportamiento en la búsqueda de información de dos grupos de investigación en ciencias sociales y humanas en Finlandia; Reddy y Jansen (2008) desarrollan un modelo de comportamiento informativo colaborativo para equipos de salud; Hyldegård (2009) estudia los procesos colaborativos de búsqueda de información en grupos de estudiantes de Ciencia de la Información; Soulier, Tamine y Bahsoun (2014) estudian la recuperación de información en ambientes colaborativos en base a un modelo propio que tiene en cuenta la incidencia del nivel de “*expertise*” del usuario en un determinado dominio.

### **1.3. Antecedentes de estudios sobre compartir datos en comunidades científicas**

Si bien la investigación científica en el siglo XXI requiere más datos y es más colaborativa que en el pasado, siendo el compartir datos una parte valiosa del método científico pues permite la verificación de los resultados y el desarrollo de nuevas investigaciones, hay una evidente preocupación por la existencia de barreras para compartir datos.

Las barreras para compartir y preservar datos de forma efectiva están profundamente arraigados en las prácticas y la cultura del proceso de investigación, así como en los propios investigadores (Tenopir et al., 2011, p. e0021101) (mi traducción).

Blumenthal et al. (2006) advierten que el proceso de intercambio de información en la ciencia es complejo y variado, y muchas formas de retención de datos pueden afectar la velocidad y la dirección del progreso científico.

Mientras mejor comprendamos el fenómeno de la retención de datos, es más probable que los esfuerzos para alentar el intercambio de esos datos sean exitosos (Blumenthal et al., 2006) (mi traducción).

Kim & Adler (2015) investigaron el comportamiento al compartir datos en científicos sociales, centrando su estudio en el rol que las motivaciones individuales, las presiones institucionales y los repositorios de datos tienen en dicho comportamiento. El mismo Kim, esta vez en coautoría con Burns, estudia el papel de los metadatos, los repositorios, y los requerimientos de las revistas y los patrocinadores de investigación en ciencias biológicas y concluyen que esos factores inciden en las normas para compartir datos, que a su vez influyen el comportamiento al compartir datos (Kim & Burns, 2016, p.230).

Tenopir y sus colegas (2015) realizaron un estudio de alcance global para determinar los cambios y las percepciones en el intercambio de datos dentro de las comunidades científicas, llegando a interesantes resultados y conclusiones sobre cómo quienes están fuera de una comunidad científica específica usan y dan sentido a los datos compartidos por otra comunidad, y como se sienten los científicos respecto al uso de datos recopilados por otros. Encontraron algunas diferencias entre los investigadores de distintos grupos etarios, áreas geográficas y disciplinas. Por ejemplo, los investigadores más jóvenes expresaron una mayor disposición a compartir datos y a utilizar datos de otros. Sin embargo, los comportamientos reales en cuanto a compartir datos contradicen esa percepción (Tenopir et al., 2015, p.10).

Los resultados del estudio de Tenopir y sus colaboradores apuntan a “una mayor aceptación y voluntad de participar en compartir datos” (mi traducción). Sin embargo, también se detectó la existencia de un mayor riesgo percibido asociado a compartir datos, y la persistencia de “barreras” específicas para el intercambio de datos. Este estudio muestra que tanto las restricciones como los factores facilitadores del uso compartido y la reutilización de datos se manifiestan de manera diferente en todas las disciplinas.

Dentro de las causas mencionadas por varios autores como problemáticas al compartir datos se encuentran la falta de incentivos (Piwowar & Chapman, 2008, p.12) (Borgman, 2012, p. 1072), así como también la preocupación de los investigadores generadores de los datos respecto a que otros puedan realizar antes que ellos los análisis proyectados sobre esos datos (Piwowar & Vision, 2013, p. e175).

En un estudio sobre las expectativas de publicación de datos en investigadores de diversas disciplinas desde ciencias naturales a ciencias sociales, Kratz & Strasser (2015) encontraron que éstas se centran en la disponibilidad de los datos, generalmente a través de una base de datos abierta o repositorio. Los investigadores encuestados reconocieron la importancia de los metadatos y de las citas formales en la lista de referencias como forma correcta de dar crédito a los creadores de los conjuntos de datos. El recuento de citas aparece en los resultados de este estudio como la medida de impacto más útil, y el recuento de descargas se consideró casi tan valioso como el número de citas.

Una encuesta realizada a 307 investigadores del área telemetría de peces para comprender sus perspectivas y experiencias sobre compartir datos arrojó interesantes resultados respecto a la relación entre la tendencia positiva a compartir datos y algunas características y actitudes personales de estos investigadores. Aplicando una regresión logística binaria se concluyó que el compartir datos se relaciona positivamente a investigadores con tendencias colaborativas, que pertenecen a una red de telemetría, que publican mucho y que expresan motivos altruistas para su investigación (Nguyen et al., 2017, p. 1264).

En una investigación reciente realizada por Zhao, Yan y Li (2018) sobre citas y menciones a conjuntos de datos en una muestra de 600 *papers* publicados en PLoS ONE, en distintas áreas de investigación de diferentes disciplinas entre enero del 2014 y diciembre de 2015, llama la atención uno de los descubrimientos realizados: la reutilización de datos tuvo lugar en menos del 30% de las publicaciones que utilizaron datos. Estos resultados sugieren que los investigadores todavía están más dispuestos a crear y usar sus propios conjuntos de datos, que a reutilizar datos generados por otros.

Otro estudio del año 2013, realizado por Wallis, Rolando y Borgman, pone en duda si los investigadores en ciencia y tecnología reutilizarían datos no generados por ellos.

[...] se sabe poco sobre cómo administran los investigadores sus datos, o sobre cuándo o cómo compartirán sus datos. Se sabe aún menos sobre cuándo, cómo y si utilizarían datos que no han sido recopilados por ellos mismos. (Wallis, Rolando, & Borgman, 2013, p. e67332) (mi traducción).

#### **1.4. Preguntas de investigación**

¿Cuáles son las principales dificultades que se presentan al compartir datos sobre biodiversidad marina?

¿Qué fomenta el compartir datos e información sobre los datos y qué lo limita?

¿Qué datos se comparten habitualmente y cuáles no?

¿De qué formas se localizan datos recogidos por otras investigaciones?

¿Cuáles son las principales prácticas para compartir datos en biodiversidad marina?  
¿Varían esas prácticas con la edad del investigador o la posición en el equipo de investigación?

¿Qué es más habitual como práctica?: ¿compartir datos sin restricciones o establecer condiciones para compartirlos?

## **1.5. Objetivos**

### **1.5.1. Objetivo general**

Se busca aportar información actual sobre comportamiento informativo colaborativo al compartir datos en biodiversidad marina, a través del estudio de las prácticas de información que tienen lugar en un ámbito donde las actividades de colaboración entre organizaciones, instituciones y personas son parte primordial del sistema de información del que forman parte.

Esta investigación indaga en los factores que intervienen al compartir datos e información sobre esos datos: factores que favorecen y factores que dificultan compartir datos de investigación, pues éstos influyen en los comportamientos informativos.

A continuación, algunas definiciones para aclarar el sentido de los términos utilizados al expresar el objetivo general:

Compartir datos: poner los datos recabados en una investigación (publicados o no publicados) a disposición de otros, utilizando fundamentalmente los medios proporcionados por las Tecnologías de la Información y la Comunicación (TIC).

Actividades de colaboración: refiere a las actividades específicas que se llevan a cabo cuando se comparten datos e información sobre los datos por parte de los integrantes de OBIS. Comprende actividades de intercambio de datos con otros integrantes del sistema, así como también actividades relacionadas con productos y servicios que se desarrollan partiendo de los datos y la información compartida.

Factor (RAE): 2. m. Elemento o causa que actúan junto con otros.

### **1.5.2. Objetivos específicos**

- 1) Identificar las motivaciones y los factores que propician compartir datos sobre biodiversidad marina y cuáles de ellos son los más habituales relacionados con comportamientos favorables a compartir datos.
  
- 2) Identificar los motivos y los factores que desalientan compartir datos sobre biodiversidad marina y cuáles de ellos son los más habituales incidiendo en comportamientos no favorables a compartir datos.
  
- 3) Explorar cuáles son las prácticas más utilizadas para obtener datos sobre biodiversidad marina y las condiciones que ponen los investigadores para compartir datos.
  
- 4) Descubrir posibles relaciones entre los factores más habituales incidiendo en el comportamiento respecto a compartir datos y las variables edad y posición en el equipo de investigación.

## Capítulo 2. Marco teórico

El punto de vista del investigador respecto de un objeto de investigación específico, así como la forma de conceptualizarlo y de pensar en él, provienen de las ideas y concepciones filosóficas detrás de las teorías. Las investigaciones en Ciencia de la Información no son la excepción: las metateorías posibilitan determinar qué tipo de entidades son, por ejemplo, información, conocimiento, usuarios y sistemas de información. Este capítulo contiene los enfoques, conceptos y teorías que luego se utilizan para el análisis de los resultados en el capítulo 5.

*Metatheory can be seen as the philosophy behind the theory, the fundamental set of ideas about how phenomena of interest in a particular field should be thought about and researched (Bates, 2005, p. 2).*

La teoría y los modelos que guían este estudio, la formulación de las preguntas de investigación y las direcciones en la interpretación de los resultados están expresamente relacionadas con la metateoría construccionista.

La perspectiva dialógica que aporta el construccionismo postula que nuestras interpretaciones son construcciones lingüísticas basadas en diálogos dentro de comunidades (organizaciones, etcétera) o tradiciones (por ejemplo, paradigmas). Según Bates (2005, p. 9), el enfoque construccionista o discursivo-analítico –con raíces en las ciencias humanas y las ciencias sociales– asume que el discurso de una sociedad condiciona las respuestas de los individuos dentro de esa sociedad, incluyendo la comprensión social de la información.

*Constructionism stresses the dialogic and contextual nature of knowledge production and the dialogic and contextual nature of users, information needs and relevance criteria. The information user makes the same pieces of knowledge or document mean different things depending on what kind of social action he or she is performing with the help of language in a specific interactional and conversational context (Tuominen, Talja & Savolainen, 2002, p. 277).*

De acuerdo con Talja, Tuominen, & Savolainen (2005, p. 89), desde el construccionismo los humanos producimos y organizamos la realidad social juntos, utilizando el lenguaje. Para el construccionismo, el lenguaje es visto como un elemento fundamental para la formación de significados. Pero a la vez proporciona un punto de

vista crítico a la visión de que el lenguaje describe los procesos internos y la realidad externa de una manera referencial y representacionista (Tuominen et al., 2002, p. 281). Este enfoque permite, como bien lo expresa Sandoval Moya (2010, p.33) “desarrollar una perspectiva alternativa al enfoque individual del conocimiento, que permite analizar el rol que juega el saber compartido por una comunidad en la mantención y reproducción de la realidad”.

## **2.1. Información y datos**

Información, el concepto central en Ciencia de la Información, visto desde la perspectiva del construccionismo es una construcción social de naturaleza contextual: se crea en comunicación con otros, en contextos interactivos o argumentativos, y sus versiones se asocian los correspondientes propósitos prácticos que se quieren lograr.

*[...] Information is a communicative construct which is produced in a social context [...] The contextual nature of information means that the way in which a version of information is constructed always depends on the interactive or argumentative context of talk, as well as on the pragmatic social purposes this version is designed to accomplish (Tuominen & Savolainen, 1997, p. 89).*

Para los objetivos de nuestro estudio y en el contexto de esta investigación, estamos refiriéndonos a datos cuando hablamos de información que no está procesada ni analizada. En opinión de Borgman, Wallis & Mayernik (2012, p. 485), los datos son el "pegamento" de una colaboración, por lo tanto estudiar cómo se producen y usan los datos es un lente a través del cual estudiar la efectividad de tales colaboraciones. En esta investigación, es precisamente estudiar las prácticas respecto a compartir datos lo que permite identificar y describir comportamientos informativos característicos de la comunidad OBIS.

## **2.2. Compartir información y comportamiento al compartir información**

Las tecnologías de la información y la comunicación son empleadas para facilitar el intercambio de información. No es de extrañar entonces que las actividades que implican que los humanos compartan información se hayan abordado principalmente desde un punto de vista tecnológico. Sin embargo, el compartir información y sus prácticas asociadas, pueden entenderse además como prácticas contextualizadas y

colectivas que deben tener en cuenta el punto de vista social. Si bien la naturaleza social del trabajo académico se ve potenciada por muchas herramientas, interfaces y sistemas tecnológicos que permiten el trabajo en colaboración, desde correo electrónico, las intranets e Internet, hasta las plataformas virtuales, es el enfoque que aporta el construccionismo social el que proporciona a esta investigación la perspectiva de considerar a la tecnología como un producto de y para la sociedad que cubre determinadas necesidades sociales.

En el trabajo académico ocurre frecuentemente un tipo específico de comportamiento informativo, el comportamiento colaborativo al compartir información, cuya denominación breve es comportamiento informativo colaborativo (*Collaborative Information Behavior*, CIB por sus siglas en inglés), el cual se ha definido de varias maneras. Para aclarar esas diferencias, se define a continuación que se entiende por comportamiento de colaboración y por compartir información.

Un comportamiento de colaboración se caracteriza por: 1) diferentes actores trabajando juntos en una actividad; 2) trabajando en un objeto común; 3) pudiendo (o no) compartir un objetivo común en esa actividad (Bao & Bouthillier, 2007, p.3) (mi traducción).

El término "compartir información" (*information sharing* en inglés) está estrechamente vinculado con otras actividades: las relacionadas con obtener información (en el sentido *information seeking* del inglés, que incluye tanto la búsqueda activa como pasiva) y con el uso de la información. Según Talja & Hansen (2006, p. 114) el compartir información incluye intercambios de información activos y explícitos, así como también intercambios implícitos, y otros intercambios menos centrados en metas.

Para Talja & Hansen (2006, p. 123) la colaboración relacionada con procurar y recuperar información puede incluir: compartir la misma necesidad de información, compartir las estrategias de búsqueda, compartir los resultados, compartir los objetos de información encontrados, continuar procesando juntos los objetos de información encontrados (interpretación, filtro, síntesis) y archivar información potencialmente útil en el repositorio del grupo.

Talja (2002, p. 145) define compartir información como un concepto paraguas que comprende una variedad de comportamientos de colaboración que van desde compartir información encontrada accidentalmente hasta colaborar en la formulación de una

estrategia de consulta y recuperación de la información. Además, un comportamiento colaborativo puede darse en el acto de compartir información de forma *ad hoc* entre los miembros del equipo, pero también cuando existe una división planificada del trabajo en la que algunos miembros trabajan juntos en encontrar información para una tarea específica, mientras que otros se ocupan de otros temas (Talja y Hansen, 2006, p. 114).

Para Talja (2002, p. 148) el compartir información es un proceso bidireccional en el que los académicos más experimentados y los que se inician se benefician mutuamente de los resultados de las búsquedas de los demás. Desde esa perspectiva, el comportamiento al compartir información se puede definir como la colaboración entre dos grupos de actores con el fin de intercambiar información con el propósito de cumplir con sus intereses individuales o bien con sus intereses comunes.

Talja & Hansen (2006, p. 122) avanzan en una clasificación para analizar el comportamiento informativo colaborativo. Ellos plantean, luego de realizar una revisión de literatura respecto a artículos que tratan la temática compartir información, siete “dimensiones”: actividades síncronas y asíncronas, colaboraciones co-localizadas y remotas, actividades estrechamente vinculadas o ligeramente vinculadas, colaboración planeada o no planeada, colaboración intra-grupo o inter-grupos, colaboración directa o indirecta, actividades coordinadas o diferenciadas.

Para Bao y Bouthillier (2007, p. 12), compartir información (*information-sharing behavior*) es un tipo de comportamiento colaborativo basado en la comunidad, impulsado por beneficios mutuos. Uno de los resultados del estudio realizado en 2007 por estos autores apunta a que el comportamiento al compartir información puede ser un tipo de comportamiento auto-reforzado (comportamiento que se refuerza cuando produce buenos resultados y satisfacción), no un tipo de comportamiento basado exclusivamente en el contexto. Por lo tanto, desde su perspectiva, en el estudio del comportamiento humano al compartir información es muy útil identificar y explorar las “fuerzas impulsoras” más importantes, con la finalidad de promover mejoras en las actividades en las que se comparte información.

Según Pilerot y Limberg (2011, p. 314) lo que se intercambia en el acto de compartir información es información adquirida, y también incluyen en esa concepción el acto de

proporcionar a alguien instrucciones para obtener la información u obtener información sobre la información.

Widén & Hansen (2012) realizan un interesante aporte cuando exploran cómo los procesos colaborativos de manejo de información están relacionados con las prácticas en el ambiente de trabajo. Ellos parten de la base de que la colaboración y la interacción están cada vez más presentes en la vida laboral actual, por lo cual la gestión de los distintos aspectos de la colaboración y el intercambio interactivo de información en los diferentes procesos es importante para apoyar la toma de decisiones y el desarrollo de sistemas de información tanto técnicos como no técnicos que faciliten el trabajo en las organizaciones. Si bien lo que estos autores desarrollan en su artículo es una discusión teórica basada en el punto de vista estratégico de la información en las organizaciones, resulta muy significativo el planteo que realizan acerca de la importancia de integrar las perspectivas de la cultura de la información en la organización con el comportamiento informativo colaborativo. Las políticas, normas y demás aspectos relacionados con la gestión de los datos y la información que hacen a la cultura de la información son elementos que aportan mucho al análisis de los resultados en esta investigación.

### **2.2.1. Prácticas de información**

El concepto de prácticas de información (también llamadas prácticas informacionales) se puede encontrar en trabajos precursores como los de Sanna Talja (1997, 2005), Talja y Hansen (2006), y posteriormente en los trabajos de Reijo Savolainen (2007, 2008, 2012) y de Andrew M. Cox (2012).

Para Savolainen, el concepto de práctica de información (*information practice* en inglés), se define como:

Conjunto de formas social y culturalmente establecidos de identificar, buscar, utilizar y compartir la información disponible en diversas fuentes, tales como la televisión, los periódicos, y el Internet. Estas prácticas suelen ser habituales y pueden ser identificadas tanto en contextos laborales como no laborales (Savolainen, 2008, p.2-3)(mi traducción).

Todas las prácticas sociales implican uso, creación y búsqueda de información, pero esto no las convierte en prácticas de información, porque solo unas pocas prácticas están

específicamente orientadas a la información (Cox, 2012, p.185). Desde la perspectiva neopragmatista asociada al concepto de prácticas de información, todo proceso de búsqueda de información es una práctica social (Savolainen, 2007, p. 125).

A pesar de esta diversidad, parece que la mayoría de los teóricos entienden las prácticas como algo concreto, mediado por la acción humana (actividades), y organizado centralmente en torno al entendimiento compartido [...] Por lo tanto, es fundamental para las teorías de la práctica que la acción o actividades y hábitos humanos, y no los “comportamientos” sean los principales constituyentes de las prácticas (Savolainen, 2008, p. 24) (mi traducción).

El paradigma construccionista pone énfasis en las prácticas sociales, o sea las actividades concretas y situadas de las personas que interactúan. Esto es determinante en el enfoque de esta investigación, pues la decisión de centrarnos en el estudio de las prácticas significa que el énfasis del análisis pasa de lo cognitivo a lo social, y el foco se sitúa en las personas interactuando con la información y los datos en su contexto social.

Las actividades relacionadas con información deben estudiarse en los contextos de trabajo y de la vida cotidiana, que son los que proporcionan los motivos y los objetivos de las actividades de búsqueda y uso de información. Para Talja & Nyce (2015) “La situación, la tarea y la práctica son algunas de las formas de conceptualizar el contexto de las actividades de información” (mi traducción).

Tanto para el trabajo como para la vida cotidiana, la colaboración para obtener información es tan común como las situaciones en que se procura información de forma individual. Es así como las tareas de adquirir y filtrar información son distribuidas entre varios individuos, y son planeadas como esfuerzos colectivos y colaborativos. La colaboración distribuida (a veces entre individuos separados por grandes distancias geográficas) crece como práctica de trabajo, favorecida por las TIC. Para muchos equipos de investigación, colaborar es un componente esencial de su forma de trabajar. Desde el abordaje propuesto, se hace fundamental situar las prácticas de información en el contexto histórico-cultural en que éstas ocurren e inseparables de la estructura social.

Las acciones de informarse, producir o diseminar información están sometidas a esa dialéctica entre el *habitus* y la situación vivida que lleva al sujeto a buscar, producir o

diseminar información. Esto sucede en un contexto social donde el sujeto ocupa determinada posición y utiliza de su capital específico, el capital informacional, para llevar a cabo sus prácticas de información. Así, las acciones de producción, búsqueda, recepción y apropiación de las informaciones deben ser comprendidas a partir de las posiciones ocupadas por los sujetos en la estructura social, pues éstas determinan el *quantum* de capital informacional de que esos sujetos disponen para sus acciones cotidianas (Pinto & Araújo, 2012, p. 225) (mi traducción).

Para Talja (2002, p. 143-144) una explicación plausible de por qué los sistemas de recuperación de información no son utilizados en todos los casos en que los académicos requieren información, es porque no captan adecuadamente sus prácticas sociales y colectivas actuales en la búsqueda, recuperación y uso de información.

Se recurre además para esta investigación a los aportes de la Teoría de la Actividad Histórico-Cultural, pues situar a las personas en su contexto histórico, social y cultural es muy importante para los objetivos de este estudio sobre prácticas de información.

Esencialmente, la Teoría de la Actividad ubica a los individuos en su contexto histórico, social y cultural, que demarca e influye en la manera que estos desarrollan la suma de acciones y operaciones que componen su actividad, y en el que ejercen influencia sobre un objeto determinado a través de herramientas de mediación dentro de un conjunto de reglas que rigen la comunidad en la que los individuos actúan según los roles y/o jerarquías que poseen (Rodríguez, 2017, p.73).

Para la Teoría de la Actividad, las actividades son emprendidas por los sujetos, motivadas por un propósito. Son los sujetos mediante sus acciones los que transforman el objeto. Además, postula que el objeto puede ser compartido por una comunidad de actores que trabajan juntos para lograr el resultado deseado, existiendo herramientas, reglas y diferentes roles que median la relación entre sujetos, comunidad y objeto.

### **2.2.2. Prácticas al compartir datos (*Data sharing practices*)**

El "compartir datos" ocurre cuando los científicos intencionalmente ponen sus propios datos a disposición de otras personas para su uso en investigaciones u otras actividades científicas relacionadas (Tenopir et al., 2015, p.3). Compartir datos implica proporcionar acceso a datos anteriormente almacenados de forma privada. Los productores de datos tienen una serie de opciones para compartir datos, desde hacer que los datos sean de acceso totalmente abierto (es decir, público) hasta limitar su

distribución a investigadores individuales. Se pueden compartir datos de varias formas: incluyéndolos en artículos publicados, publicándolos en sitios web institucionales o personales, depositándolos en repositorios o enviándolos en respuesta a solicitudes personales de otros investigadores.

Aunque los científicos frecuentemente comparten datos, el intercambio a menudo se limita a redes de colaboradores o colegas cercanos en pequeña escala y no a la comunidad en general (Cragin et al., 2010, p. 4031) (mi traducción).

En general, existen cuatro razones para compartir datos: 1) verificar y (o) reproducir la investigación; 2) poner a disposición del público los resultados de investigaciones financiadas con fondos públicos; 3) permitir que otros investigadores hagan preguntas nuevas y diferentes usando los datos; 4) avanzar en el estado de la investigación y la innovación, proporcionando nuevos conocimientos y comprensión (Borgman, 2012; Poisot, Mounce & Gravel, 2013; Tenopir et al., 2011).

El compartir datos también se asocia a beneficios personales para quienes han compartido datos, relacionados con una mayor visibilidad y relevancia de los resultados de la investigación, más oportunidades de publicación a través de colaboraciones, e incremento de las citas de los conjuntos de datos en publicaciones primarias (Piwowar, Day & Fridsma, 2007; Poisot et al., 2013).

### **En síntesis**

Desde la postura del construccionismo social se ha cuestionado la perspectiva individualista que se le ha dado al papel del usuario de información en los enfoques tradicionales. Proponemos para esta investigación un enfoque que destaca la dimensión social del usuario y las relaciones del usuario con su contexto histórico-cultural al estudiar sus prácticas de información colaborativas.

Resulta muy importante, en este caso para el estudio de las prácticas de información colaborativas, tener en cuenta los significados atribuidos por los sujetos durante las acciones de buscar, usar y compartir informaciones y datos en el contexto en el que se dan tales prácticas.

Recurrir a los postulados de la Teoría de la Actividad, que incluye la perspectiva histórico-cultural, permite visualizar a las actividades relacionadas con compartir datos e información siempre ligadas a la estructura social de la comunidad.

Las comprobaciones de estudios anteriores sobre el comportamiento al compartir información en diferentes comunidades científicas constituyen un valioso punto de partida para esta investigación, muy útiles para pensar el contenido de los instrumentos de recolección de datos para esta investigación (encuestas, entrevistas) a desarrollar en el capítulo 4: Metodología.

Para identificar factores que dificultan el compartir datos de investigación sobre biodiversidad marina desde el enfoque de las prácticas de información se pretende indagar sobre las dificultades específicas al compartir esos datos. Relacionándolo con las “barreras” que mencionan Tenopir et al. (2015), interesa recabar información sobre características y factores que condicionan el comportamiento informativo de los investigadores de forma desfavorable a compartir datos.

De la misma forma que es importante estudiar las barreras, también lo es para los objetivos de esta investigación estudiar los aspectos que favorecen el compartir datos, con qué otras variables se relacionan, y qué motivaciones tienen los investigadores para compartir datos. Si la tendencia positiva a compartir datos se relaciona con algunas características personales como plantean Nguyen y sus colegas (2017), resulta interesante poder comprobar si esto se confirma para los investigadores de la biodiversidad marina, además de otras características que puedan surgir durante la investigación y sean propias de esta comunidad en particular.

Agrupar los aspectos que favorecen compartir datos por un lado y los aspectos que no son favorables por otro se constituye en una buena alternativa a utilizar más adelante, en el capítulo 5, para organizar el análisis de los resultados de la investigación.

## Capítulo 3. Datos e información sobre biodiversidad marina

Este capítulo contiene las definiciones de conceptos esenciales para esta investigación, en relación con el sentido y los usos que le asigna la comunidad estudiada. Para finalizar, se realiza una síntesis de las principales características del sistema internacional de información que da marco al estudio.

### 3.1. Los datos en biodiversidad marina

Los datos son mediciones, observaciones o descripciones creadas o recopiladas por un investigador. Los tipos de elementos descritos (referentes) varían ampliamente según la disciplina (AnthroDPA Metadata Working Group, 2009, p.3). En las ciencias físicas y biológicas, los investigadores recopilan y producen la mayoría de los datos mediante observaciones, experimentos o modelos (Borgman, 2012, p. 488-489).

Los datos sobre el referente (objeto, fenómeno o evento) pueden tener distintos formatos, como texto (descripciones), mapas, imágenes o grabaciones de audio, señales, símbolos o números. Los datos son representaciones precisas y bien definidas, registrados de alguna manera estándar especificada. Se organizan en “conjuntos de datos” (*datasets*), los cuales tienen al menos “cuatro aspectos comunes: agrupación, contenido, relación y propósito” (Renear, Sacchi, & Wickett, 2010, p.1) (mi traducción) y a su vez, esos aspectos presentan múltiples categorías.

Otra cuestión importante a tener en cuenta cuando hablamos de compartir datos son los metadatos. Los metadatos están compuestos por la descripción esencial para documentar el proceso de creación, recopilación, gestión y preservación de los datos.

Los metadatos proporcionan información sobre el referente original, los procesos de recopilación, las reglas de recopilación, así como las descripciones de los procesos de gestión de datos y las disposiciones para el acceso y uso de los datos (o sea, la licencia para especificar los usos permitidos) (AnthroDPA Metadata Working Group, 2009, p.3) (mi traducción).

Existen varios estándares para *datasets* sobre biodiversidad, los más utilizados son *Darwin Core*<sup>2</sup> (*DwC*) y *Access to Biological Collections Data (ABCD) Schema*<sup>3</sup>.

### 3.3. Compartir datos como práctica de información colaborativa

Los investigadores siempre han creado y recopilado datos, pero la creciente disponibilidad de datos digitales en las redes informáticas globales, además del aumento en el tamaño y complejidad de los conjuntos de datos han creado nuevos desafíos para el uso, la reutilización, el almacenamiento, el acceso futuro y la recuperación de esos datos. Además, en biología marina como en otras disciplinas, el uso creciente de instrumentos y sensores para captura de datos ha multiplicado el volumen de los datos de investigación.

Compartir datos e información resultantes de la investigación científica es una práctica que ha evolucionado lenta y desigualmente. Los patrocinadores de investigación, los editores y las sociedades científicas tienen mucho que ver en este cambio sociocultural.

Hay una creciente presión de la comunidad científica, incluidos los organismos de financiación, las revistas y los pares, para que los autores publiquen los datos de biodiversidad utilizados en artículos y otra literatura científica publicada (Costello & Wieczorek, 2014, p. 68) (mi traducción).

Por supuesto, siempre puede haber casos excepcionales. Por ejemplo, no se pueden liberar todos los datos sobre una especie amenazada porque expone a la población de la especie a la recolección ilegal o la caza. En opinión de Chapman & Grafton (2008) aún en esos casos en que la información sobre la ubicación de la especie deba ser generalizada o retenida, otros datos si pueden ser compartidos.

El entorno institucional opera como determinante importante en las normas, actitudes y comportamientos de los científicos al compartir datos. Ese entorno comprende los recursos institucionales, las presiones de los organismos que financian la investigación y las editoriales.

---

<sup>2</sup> *Darwin Core (Version 2014-11-08)*. *Biodiversity Information Standards (TDWG)*, disponible en: <http://doi.org/10.5281/zenodo.12694>

<sup>3</sup> *Access to Biological Collection Data (ABCD), Version 2.06*. *Biodiversity Information Standards (TDWG)*, disponible en: <http://www.tdwg.org/standards/115>

Es de esperar que hacer disponibles los datos en línea aumente el reconocimiento de los científicos, pues eso favorece el aumento en la cantidad de citas y su productividad será más visible. Sin embargo, citar conjuntos de datos aún es considerado como una práctica emergente.

La cita de datos es una práctica emergente, pero tiene el potencial de convertirse en una norma de atribución académica, acompañada de una mejor gestión de datos, acceso, descubrimiento y reutilización. A nivel institucional, el apoyo para la citación de datos incluye el desarrollo de infraestructura y el suministro de orientación a los investigadores, como parte de una estrategia más amplia de mejora en la gestión y el depósito de los datos (Simons, Visser, & Searle, 2013, p. 8) (mi traducción).

El desarrollo de una teoría y práctica para la cita de datos es fundamental para considerar los datos como objetos de investigación de primera clase con la misma relevancia y centralidad de los productos científicos tradicionales (Silvello, 2018, p.6) (mi traducción).

Para citar es necesario que los conjuntos de datos estén disponibles en un lugar accesible y permanente, pues eso posibilita enlazar la referencia y garantizar el acceso. Las plataformas de acceso que integran conjuntos de datos, ya sean repositorios institucionales abiertos en línea, repositorios centrales o centros de datos, constituyen una muy buena alternativa para publicar datos en línea, más si los conjuntos de datos a publicar cuentan con suficientes metadatos. Los centros de datos proporcionan un valor agregado por los controles de calidad que realizan a los datos.

### **3.4. OBIS: el proyecto colaborativo que da marco a este estudio**

El cometido del sistema de información OBIS es ser una plataforma global para la colaboración internacional en biodiversidad marina. Hay una intención implícita y un compromiso de compartir información y datos en las instituciones, organizaciones y programas científicos que se incorporan a este sistema. Esto genera una comunidad de académicos pertenecientes a cientos de instituciones de varios países (ver Anexo 3: Mapa de instituciones proveedoras de datos a OBIS) que comparten actividades, información y datos.

Desde el enfoque de la Teoría de la Actividad, comunidad es “un colectivo de individuos y grupos cuya actividad está orientada hacia el objeto compartido”

(Engeström, 2001) (mi traducción). El proyecto *Ocean Biogeographic Information System* (OBIS)<sup>4</sup> de la Comisión Oceanográfica Intergubernamental (COI) de UNESCO está organizado como una red global de nodos<sup>5</sup> que publican datos estandarizados sobre biodiversidad marina, localizados geográficamente, a través de un único portal.

Existen muchas bases de datos académicas sobre biodiversidad en acceso abierto. La mayoría proporciona información sobre especies como FishBase<sup>6</sup>, AlgaeBase<sup>7</sup> y la base de datos de especies invasoras globales (GISD)<sup>8</sup>, e imágenes, como Ceta-Base<sup>9</sup> para identificar ballenas a partir de sus fotografías. Solo algunas proporcionan datos de distribución estandarizados, como el Sistema OBIS (Costello et al., 2013, p.456) (mi traducción).

Un beneficio importante de la publicación a través de centros de datos (por ejemplo, PANGAEA, OBIS) es que su personal verifica el formato de los conjuntos de datos antes de ponerlos a disposición en línea. Otra ventaja de la mayoría de los centros de datos es que asignan un identificador único que permite la localización y el seguimiento del conjunto de datos (Costello & Wiczorek, 2014, p.72) (mi traducción).

Dentro de los objetivos de OBIS se encuentra proporcionar datos, información y herramientas para apoyar la identificación de hábitats marinos y costeros biológicamente importantes para el desarrollo de planes de manejo de áreas marinas. También el proporcionar información y orientación sobre el uso de datos de biodiversidad para la educación y la investigación. Es de esperar que los participantes de este sistema estén comprometidos con estos objetivos comunes respecto a los datos y la información. Son objetivos compartidos centrales para este sistema colaborativo que deben estar muy presentes al estudiar las prácticas al compartir datos por parte de los miembros de esta comunidad.

El sistema OBIS cuenta con más de 30 nodos en todo el mundo (ver Anexo 2).

Los nodos OBIS son proyectos, programas, institutos, organizaciones o centros de datos oceánicos a nivel nacional, o bien proyectos, programas e instituciones que operan a nivel regional e internacional llevando a cabo funciones de gestión de datos oceánicos. Los nodos OBIS son responsables de representar todos los aspectos de OBIS en una

---

<sup>4</sup> *Ocean Biogeographic Information System* (OBIS). Disponible en Internet: <http://www.iobis.org/>

<sup>5</sup> Nodo: componente de la red que tiene la capacidad de contribuir a sus objetivos. Los nodos obtienen y aportan datos, infraestructura técnica y experiencia científica.

<sup>6</sup> FishBase, disponible en Internet: <http://www.fishbase.org/>

<sup>7</sup> AlgaBase, disponible en Internet: <http://www.algaebase.org/>

<sup>8</sup> *Global Invasive Species Database* (GISD), disponible en Internet: <http://www.iucngisd.org/gisd/>

<sup>9</sup> Ceta-Base, disponible en Internet: <http://www.cetabase.org/>

región particular o dentro de un dominio taxonómico. Es el nodo el que establece relaciones con los proveedores de datos clave dentro de su área geográfica (o taxonómica) de responsabilidad y se encarga de brindar los datos y metadatos correspondientes a la base de datos global para ser compartidos con la comunidad OBIS (UNESCO. IODE, 2016) (mi traducción).

El principal producto de información del sistema OBIS, elaborado en base a los conjuntos de datos que allí se comparten, es un sitio web o portal que sirve de punto de acceso a información sobre biodiversidad, distribución y abundancia de especies marinas. Es una base de datos que funciona como repositorio central, y además cosecha metadatos de otros repositorios.

Los conjuntos de datos disponibles en OBIS contienen información estandarizada sobre dónde y cuándo se han registrado especies marinas. Es posible realizar búsquedas en todos los conjuntos de forma simultánea utilizando nombres de especies, nivel taxonómico, área geográfica, profundidad del mar y fecha. También es posible acceder a las variables ambientales asociadas a los lugares geográficos (Fuente: Sitio web de OBIS).

Centrarnos en los aspectos colectivos del comportamiento con la información al estudiar las prácticas de datos en un contexto de trabajo científico colaborativo, significa tener muy presentes las normas establecidas para compartir y utilizar los datos en ese contexto. Las normas y directivas de OBIS, publicadas en línea en el documento *Guidelines on the sharing and use of data in OBIS*, adoptadas en 2015, contienen el acuerdo para compartir datos y el acuerdo para el uso de los datos (*Data sharing agreement* y *Data use agreement*<sup>10</sup>, respectivamente).

En el acuerdo para compartir datos entre el repositorio central y los proveedores de datos se establece que los proveedores de datos conservan todos los derechos y responsabilidades asociadas a los datos que ponen a disposición de OBIS a través de los nodos OBIS. Los nodos OBIS a su vez, deben llevar a cabo los acuerdos necesarios con los proveedores de datos originales para poder ponerlos a disposición bajo las siguientes

---

<sup>10</sup> Ambos documentos están incluidos en *Guidelines on the sharing and use of data in OBIS* disponible en Internet: <http://www.iobis.org/data/policy/>

licencias *Creative Commons*: CC-0, CC-BY y CC-BY-NC (Fuente: Sitio web de OBIS<sup>11</sup>).

En ese acuerdo, los proveedores de datos son responsables de la integridad de los datos y de los perfiles de metadatos. Mediante el acuerdo, se otorga permiso a OBIS para: distribuir los datos a través de su portal de datos e información, construir una base de datos integrada, utilizar los datos para fines de control de calidad, complementar los datos con otros datos (ejemplo: datos del clima), y crear productos y servicios de información de valor agregado para la ciencia y la toma de decisiones.

En cuanto al acuerdo de uso de datos, los datos en OBIS están disponibles gratuitamente para todos, siguiendo los principios de acceso equitativo y distribución de beneficios, apoyando el desarrollo de capacidades, y la participación de todos los estados miembros de la Comisión Oceanográfica Intergubernamental de UNESCO en programas globales.

Está establecido en el acuerdo de uso que los usuarios de los datos deben proporcionar la atribución correspondiente a los proveedores de datos (reconocer la autoría) y que se debe dar un uso legítimo a los datos. Esto último refiere a que el usuario debe citar como fuente al portal OBIS, no estando permitida la redistribución en línea de la base de datos OBIS.

---

<sup>11</sup> *Ocean Biogeographic Information System* (OBIS). Disponible en Internet: <http://www.iobis.org/>

## **Capítulo 4. Metodología**

En este capítulo explicamos la estrategia de investigación, las herramientas y los métodos a utilizar. La teoría del construccionismo social asume que los humanos construimos versiones de la realidad y que el conocimiento es algo que las personas crean juntas, no una posesión individual. Esta visión tiene implicaciones directas en la forma de estudiar la información y en la metodología aplicada a esta investigación.

Para lograr los objetivos propuestos en el capítulo 1, se pretende relevar las características, actitudes y opiniones de investigadores de la biodiversidad marina respecto a las prácticas al compartir datos, utilizando tres herramientas diferentes. La metodología de investigación incluye una encuesta a investigadores participantes del sistema OBIS, con el objetivo de recabar sus opiniones, valoraciones y experiencias respecto a compartir datos. Además, se aplican entrevistas abiertas semi-estructuradas a referentes calificados: los coordinadores responsables de los nodos que participan en el proyecto. Las entrevistas tienen el propósito de profundizar en la problemática estudiada desde las opiniones, experiencias y comentarios de los referentes (responsables regionales o temáticos) de OBIS. La tercera herramienta es el análisis de contenido de los documentos de las reuniones del Grupo Directivo de OBIS. El propósito de aplicar esta técnica es estudiar los factores que dificultan y los factores que favorecen compartir datos e información en relación con las actividades de colaboración mencionadas en esos documentos para el período 2016 - 2017.

En síntesis, estamos proponiendo una estrategia que incluye tres componentes:

- 1) Encuesta a miembros de equipos de investigación (aproximadamente 90 - 100 investigadores vinculados a OBIS de diferentes instituciones de todo el mundo).
- 2) Entrevista a responsables de la coordinación de nodo OBIS (aproximadamente 10 coordinadores de un total de 33 nodos).
- 3) Análisis de documentos de las reuniones anuales de OBIS (2016 - 2017).

### **4.1. Encuesta a investigadores en biodiversidad marina**

La forma de distribución de la encuesta, auto administrada vía formulario en línea, es la opción adecuada por ser OBIS una comunidad global. Las fuentes para obtener la

información de contacto para distribuir la encuesta a los investigadores son los sitios web de los proveedores de datos OBIS y el directorio *OceanExpert: The Global Directory of Marine and Freshwater Professionals*<sup>12</sup>. *OceanExpert* es un directorio que posee mecanismos de actualización periódica y mantiene un registro actualizado de los investigadores y los proyectos en que éstos participan.

El diseño de la encuesta está pensado para relevar los datos correspondientes a variables y categorías previamente definidas, de forma que no le lleve mucho tiempo al encuestado completarlo (la menor cantidad de preguntas posible, cortas y utilizando un lenguaje preciso), pero que a su vez permitan recabar suficiente información para realizar las necesarias vinculaciones entre variables.

El formulario de la encuesta<sup>13</sup> presenta siete afirmaciones redactadas en primera persona (preguntas 1 a 5, 7 y 8) en las cuales se pide que indiquen las opciones que correspondan. Las preguntas 6 y 9 están orientadas a recabar otra información no contemplada por las opciones ofrecidas. Se están recabando de esta forma las experiencias relacionadas con prácticas sobre compartir datos o bien las opiniones del encuestado respecto a factores que favorecen o que dificultan esas prácticas. Además, en las preguntas 2, 5 y 8 se está pidiendo una valoración para las opciones que se señalen. Para ello se emplea la escala de Likert, utilizada habitualmente para medir actitudes en base a un conjunto de enunciados sobre los cuales el encuestado debe mostrar su nivel de acuerdo o desacuerdo.

## Variables

- 1) Demográficas:** edad, género, región, posición en el equipo de investigación.
- 2) Actitudes personales** de los investigadores respecto a compartir datos. Esta variable tiene dos categorías: las relacionadas con actitudes al compartir datos propios (datos generados en la investigación o investigaciones que lleva o ha llevado a cabo) y las relacionadas con actitudes y prácticas para acceder a datos que otros investigadores compartan.

---

<sup>12</sup> OceanExpert, disponible en internet: <https://www.oceanexpert.net/>

<sup>13</sup> Ver Anexo 1, Formulario de la encuesta

Respecto a los datos propios, se establecen dos categorías para la variable:

1. Los comparte: tanto si los comparte total o parcialmente, interesa saber si pone condiciones para compartirlos y cuáles son.
2. No los comparte: interesa conocer qué fue lo que llevó a no compartirlos (retenerlos).

Respecto al acceso a datos generados por otros investigadores, interesa saber en caso de necesitarlos cómo los consigue y con qué frecuencia utiliza cada forma de conseguirlos. Escala utilizada: 0 = *Never*, 1 = *Rarely*, 2 = *Sometimes*, 3 = *Very often*, 4 = *Always*

### **3) Factores que inciden al compartir datos**

Esta variable tiene dos categorías: factores que favorecen y factores que dificultan compartir datos. En la primera categoría se pide la opinión de los encuestados sobre cuáles son para ellos los factores favorables para compartir datos (a través de diferentes opciones que pueden marcar directamente en el formulario: pertenecer a una red, contar con apoyo institucional, etcétera, y se deja un espacio para que coloquen otros) y una valoración sobre los factores que ellos consideren favorables. Escala utilizada: 0 = *Not important at all*, 1 = *Of little importance*, 2 = *Of average importance*, 3 = *Very important*, 4 = *Absolutely essential*. En la segunda categoría se recaban opiniones sobre los factores que desalientan al entrevistado a compartir datos (también aquí se mencionan diferentes opciones y se brinda un espacio para que escriban otras opciones) utilizando la misma escala de 0 a 4 para valorar la importancia que le asignan a cada factor.

## **4.2. Análisis de contenido de los documentos de las reuniones del Grupo Directivo**

### **4.2.1. Documentos a analizar y objeto de referencia**

Para aplicar análisis de contenido (en adelante AC) tiene que haber una narración a propósito de algo (objeto de referencia).

El análisis de contenido se utiliza para estudiar cualquier tipo de documento en el que esté transcrito algún relato, relativo a cualquier objeto de referencia (Bernete, 2013, p. 222).

*Content analysis is a research technique for making replicable and valid inferences from texts (or other meaningful matter) to the contexts of their use* (Krippendorff, 2008, p.18).

El material textual contenido en los documentos de las reuniones anuales (*reports*) que nos proponemos analizar es originado en las deliberaciones de los coordinadores de nodos (*node managers*) pertenecientes a instituciones y organizaciones que integran el sistema de información OBIS. Estos documentos son editados y publicados por UNESCO, estando disponibles en acceso abierto en su sitio web.

Las actividades en que los integrantes del sistema comparten datos e información aparecen mencionadas en estos documentos. Con la aplicación de la técnica de AC pretendemos descubrir ciertas pautas (orden subyacente) a la problemática estudiada, a través de la descomposición del objeto de estudio en variables y atributos (opciones de la variable).

Interesa todo lo relacionado con cómo los nodos y demás participantes de la red comparten datos e información en el desarrollo de sus actividades. Con las deducciones que se logren a partir de los datos que están explícitos y las justificaciones que resulten del conocimiento que podamos obtener de las menciones al contexto de esas actividades, es esperable inferir algunas de las prácticas de información en ese ámbito y acercarnos a conocer las situaciones en que se comparten o no se comparten los datos y la información.

#### **4.2.2. Objetivo del análisis de contenido**

La posibilidad de compartir datos e información se ha visto potenciada por las TIC y los científicos marinos utilizan a diario estas herramientas en sus actividades. El cambio tecnológico y las complejidades del intercambio hacen que tanto los factores que favorecen como los que dificultan el compartir datos e información sean variados, cambiantes y obedezcan a diferentes causas.

El objetivo del AC apunta a identificar esos factores y su enunciado es el siguiente: investigar las prácticas al compartir información identificando los factores que dificultan y los factores que favorecen compartir datos e información en relación con las actividades de colaboración en el período (2016 - 2017).<sup>14</sup>

#### 4.2.3. Corpus o universo de estudio

Los productos comunicativos que se han seleccionado para el análisis contienen narraciones en un estilo indirecto en la mayoría de los casos, o sea que no reproducen exactamente las palabras de quienes intervienen, sino que allí se cuenta lo que han expresado.

Los documentos a analizar son los informes de las reuniones del Grupo Directivo de OBIS de 2016 y 2017, los cuales aparecen como *IODE Steering Group for OBIS (SG-OBIS) (2016)* y *IODE Steering Group for OBIS (SG-OBIS) (2017)* respectivamente, en las referencias bibliográficas de esta tesis.

La función de estos productos es dejar testimonio escrito de los asuntos tratados en cada reunión, pero también comunicar esos asuntos a toda la comunidad y al público interesado. Los documentos contienen datos de referencia que describen logros, actividades, opiniones y evaluaciones relacionadas con el acontecer de OBIS en el período entre sesiones. Las representaciones de la realidad contenidas allí son muy útiles para los objetivos de este AC.

En cuanto al criterio de selección de los documentos, la homogeneidad de los *reports* se puede establecer por la similar estructura de su presentación. Esto último se explica por la existencia de un modelo o plantilla que aplica a todos los *reports* de las reuniones de los programas y proyectos UNESCO.

La pertinencia de estos documentos se debe a que son “[...] adecuados como fuente de información con arreglo al objetivo que suscita el análisis” (Bardin, 1986, p. 73). La correspondencia de los objetivos propuestos y los datos que podamos recoger de los

---

<sup>14</sup> Las definiciones de compartir datos, actividades de colaboración y factor son las aportadas anteriormente en la p.19 para el objetivo general de esta investigación.

relatos incluidos en los documentos seleccionados es muy importante para la exitosa aplicación de la técnica de AC.

#### 4.2.4. Ficha de registro

En cada ficha se registran todos los datos pertinentes, de acuerdo al objetivo planteado para el AC. La ficha utilizada para este análisis se adjunta en el Anexo 4.

Los datos por recabar han sido agrupados y clasificados en categorías a partir de identificar los temas o dimensiones que aparecen en los documentos que son de utilidad para nuestro objetivo. A partir de la ficha de registro se transforma el contenido de cada unidad de registro en información codificada. Se llena una ficha por unidad de análisis. Cada ficha tiene un número de referencia y se completa en su totalidad. En caso de que no aparezcan datos en el documento analizado para esa unidad de registro, se codificará con un cero.

Las características contempladas para las categorías propuestas son: que sean homogéneas, mutuamente excluyentes (con significados claramente diferenciados) y exhaustivas (que cubran todas las posibilidades de una variable).

#### 4.2.5. Libro de códigos

En base a categorías definidas de antemano, se construye el libro de códigos que se utiliza para estandarizar el proceso de codificación. En él se han listado los elementos a recoger y sus agrupaciones, asignándole un número de código consecutivo a cada uno. Otras agrupaciones se podrán considerar con posterioridad, si presentan ventajas para el análisis. También se incluyen en este libro las definiciones explícitas de las categorías, con sus reglas de codificación.

**0. Datos básicos de identificación.** Códigos (1) al (3). Comprende el número dado a cada unidad, el título del *report*, y el año en que tuvo lugar la reunión.

**1. Actividades de colaboración.** Códigos (4) al (8). Incluye todas las actividades que aparecen mencionadas en los documentos en las cuales los integrantes de OBIS comparten datos e información. Consignar en cada categoría la/s expresión/es

utilizada/s en el texto y anotar el número de veces que se menciona cada una en el documento analizado. Ejemplo: Cód. (4), F. (Frecuencia) 2, Expresiones: “*nodes data exchange*”, “*to share data for product development*”.

1.1. Colaboración con OBIS: proveer datos y colaborar en el desarrollo de productos y servicios de información OBIS. Ejemplo de producto: el sitio web de OBIS.

1.2. Colaboración con otros editores y proveedores de datos (*data contributor = data provider*): colaboración con proyectos, programas u organizaciones nacionales o internacionales no pertenecientes a OBIS. Ejemplo: actividades de colaboración con otras organizaciones como *World Register of Marine Species (WoRMS)*.

1.2. Actividades orientadas al mantenimiento y crecimiento de OBIS: colaborar en restablecer nodos inactivos, propiciar el establecimiento de nuevos nodos.

1.3. Actividades para brindar capacitación. Ejemplos: proveer instrucción (*training activities = training courses = training workshops*) en evaluación y gestión para compartir datos e información, asistir a otros nodos con la utilización de estándares.

1.5. Promover la colaboración: tanto interna como externa a OBIS. Incluye participación en actividades promocionales que buscan ampliar los horizontes para compartir datos e información y asistencia a eventos (reuniones, conferencias) con la misma finalidad expresada anteriormente.

**2. Factores que desalientan las actividades relativas a compartir datos e información.** Códigos (9) al (14). Consignar en cada categoría la/s expresión/es utilizada/s en el texto y anotar el número de veces que se menciona cada una en el documento analizado. Ejemplo: Cód. (9), F. 3, Expresiones: “*fitness for use*”, “*limitations of use*”, “*ownership disputes*”.

2.1. Consideraciones legales - Incluir aquí menciones sobre si los datos están listos para ser usados, si existen limitaciones de uso o de propiedad de los datos. Trazabilidad de los datos y procedencia. Ejemplo: no se comparten para evitar disputas de propiedad en caso de datos pertenecientes a un equipo de investigación.

2.2. Baja participación de los nodos - Incluir aquí menciones sobre desvinculación de nodos del sistema.

2.3. Dificultades técnicas - Incluye menciones a problemas y limitaciones técnicas de los nodos y de sus proveedores de datos.

2.4. Falta de consistencia de los datos - menciones específicas a inconsistencias y también referencias a la falta de estándares establecidos o la carencia de una estructura de datos aceptada por todos.

2.5. Baja calidad de los datos - Carencia de documentación del procedimiento de recogida, baja calidad de los datos, los metadatos y menciones a controles de calidad.

2.6. Otros factores que dificultan.

**3. Factores propicios a las actividades de compartir datos e información.** Códigos (15) al (22). Consignar en cada categoría la/s expresión/es utilizada/s en el texto y anotar el número de veces que se menciona cada una en el documento analizado. Ejemplo: Cód. (20), F. 2, Expresiones: “*established standards*”, “*accepted by all standard*”.

3.1. Variedad de herramientas, productos y servicios - los actuales y los que estén en desarrollo.

3.2. Interoperabilidad de los sistemas.

3.3. Compatibilidad de los metadatos.

3.4. Fortalecimiento de redes regionales.

3.5. Políticas de datos.

3.6. Uso de estándares. Ejemplos: *Darwin Core (DwC)*, *Dublin Core (DC)*. Incluye los estándares y las modificaciones que sean compatibles con ellos. Incluir aquí menciones y propuestas tendientes a estandarizar políticas y procedimientos para los datos de investigación.

3.7. Actualización - Incluye capacitación específica en tecnologías para compartir datos e información y actividades organizadas con la finalidad de brindar capacitación.

3.8. Otros factores que favorecen.

### 4.3. Entrevistas a coordinadores de los nodos del sistema OBIS

La pauta de entrevista consiste en tres preguntas abiertas sobre compartir datos:

- 1) ¿Cuáles son actualmente las principales dificultades cuando se comparten datos sobre biodiversidad marina?
- 2) En su experiencia, ¿qué tipo de datos es más probable que sean retenidos, incluso si no existen impedimentos o restricciones para compartir esos conjuntos de datos?
- 3) Al buscar conjuntos de datos o fuentes de datos generadas por otros, ¿cuáles son las estrategias y/o herramientas más usuales que utiliza para ubicar datos?

Al solicitar la entrevista, se pone en conocimiento de los entrevistados la completa confidencialidad y se establece específicamente que sus nombres no serán asociados a las respuestas. Las respuestas son analizadas en conjunto y no queda registrada la identidad del entrevistado en la tesis, artículos u otros trabajos académicos que se publiquen como resultado de esta investigación.

Las entrevistas a los coordinadores de nodo OBIS que aceptan participar del estudio se realizan vía video conferencia o correo electrónico debido a las distancias geográficas.

## Capítulo 5. Resultados

Este capítulo se centra en el análisis de los datos recabados con la aplicación de las herramientas presentadas en el capítulo 4: encuestas a investigadores, entrevistas a responsables de nodos OBIS y análisis de contenido de los documentos de las reuniones del Grupo Directivo de OBIS. Al final del capítulo presentamos un análisis que integra los resultados de las técnicas empleadas, poniendo énfasis en resultados en los que existen coincidencias para las tres técnicas en algunos o en todos los aspectos analizados.

### 5.1. Encuesta

Se creó una base de datos con 735 contactos, correspondientes a científicos pertenecientes a equipos de investigación de instituciones que figuran como proveedoras de datos (*data providers*) en el sitio web de OBIS (Ver Anexo 3. Mapa de instituciones proveedoras de datos a OBIS). La encuesta se distribuyó exitosamente a 656 recipientes de correo electrónico, mediante un correo personalizado, refiriendo a cada destinatario por su nombre, cargo e institución. De esos 656 recipientes, 498 abrieron el mensaje y de éstos, 123 completaron y enviaron el formulario de encuesta. La encuesta estuvo disponible en línea, abierta a la recepción de respuestas desde el 12 de febrero al 13 de marzo de 2018.

El cuestionario está compuesto de 13 preguntas (Ver Anexo 1. Formulario de la encuesta). Ninguna de las preguntas requiere una respuesta obligatoria para proceder al envío del formulario. Las preguntas que recaban opinión son las nueve primeras, las últimas cuatro (números 10 al 13 del cuestionario) están orientadas a conocer región geográfica, posición en el equipo de investigación, género y edad de los encuestados.

En las preguntas con múltiples opciones, las opciones han sido codificadas con el número de la pregunta y un dígito decimal consecutivo para cada opción. Ejemplos: pregunta 1, opción 1 se codifica como 1.1; pregunta 2 opción 3, se codifica como 2.3. En las preguntas para las cuales se requiere una valoración para cada opción, las valoraciones se codificaron de 0 a 4.

Las preguntas 6 y 9 son abiertas. Todas las respuestas a estas preguntas están disponibles en el Anexo 4. Parte 3. Respuestas a las preguntas abiertas de la encuesta (preguntas 6 y 9). En las preguntas 1, 4 y 7, la última opción para cada una de ellas es abierta y permite escribir otras respuestas mediante una casilla de texto denominada “Otros”. Están codificadas como 1.7, 4.4 y 7.8 respectivamente, y sus respuestas, por motivos de extensión también se incluyen en el Anexo 4.

### 5.1.1. Análisis de los resultados de la encuesta

El total de investigadores que respondió la encuesta es de 123. Como se puede apreciar en el Gráfico 1, 47 son mujeres (codificado como 12.1), 75 hombres (codificado como 12.2) y uno no aporta datos de género.

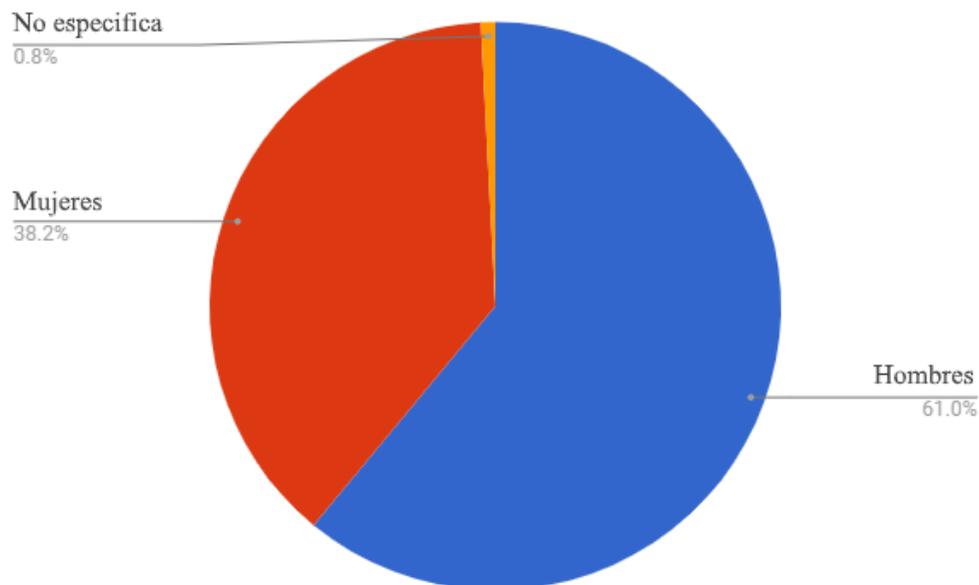


Gráfico 1. Distribución de los encuestados por género

El promedio de edades de los encuestados es de 47 años y medio. Como lo muestra el Gráfico 2, la mayor cantidad de respuestas recibidas (80) se da en las edades entre 36 y 55 años, que corresponde a científicos en pleno desarrollo de su carrera, interesados por la problemática que plantea este estudio.

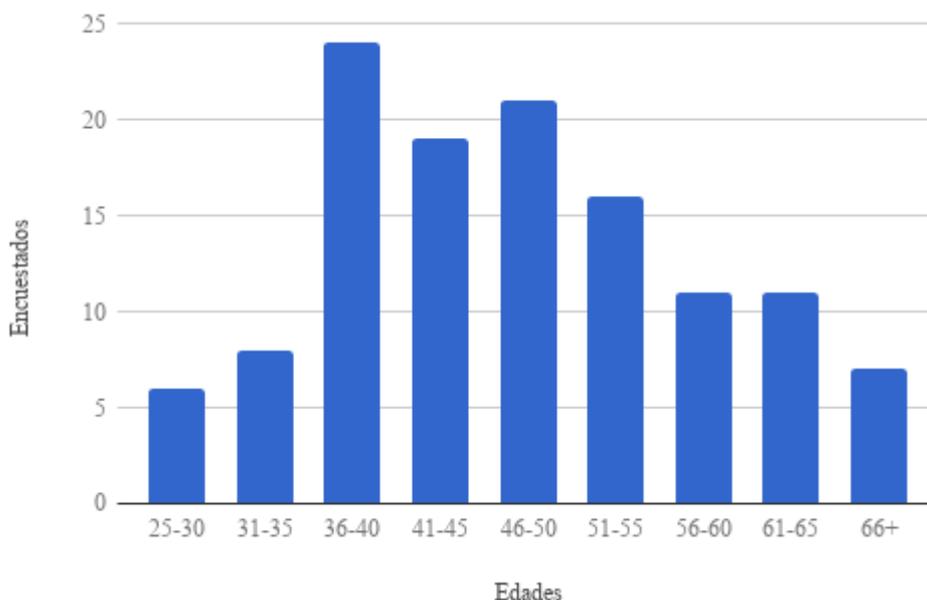


Gráfico 2. Distribución etaria de los encuestados

En cuanto al alcance geográfico, las respuestas llegaron de todas las regiones previstas en el diseño de la investigación, por lo cual se logró el cometido de que la encuesta sea global. Como se ve en el Gráfico 3, la cantidad de respuestas varía según el área geográfica, contándose mayor cantidad de formularios enviados para las regiones de Europa y Sudamérica. Esto se justifica por el mayor número de instituciones cooperantes con OBIS (*Data providers*) en Europa, y por la buena acogida que tuvo la encuesta en Sudamérica.

10.1	África.....	13
10.2	Asia.....	10
10.3	Centroamérica y Caribe..	6
10.4	Europa.....	33
10.5	Norteamérica.....	12
10.6	Oceanía.....	8
10.7	Sudamérica.....	36
10.8	Otros .....	5

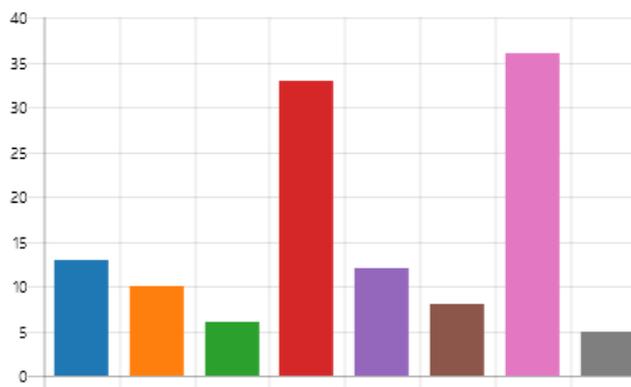


Gráfico 3. Distribución de respuestas por región geográfica

La mayoría de los investigadores que respondieron la encuesta se identificaron con las regiones previstas en el cuestionario.

Sin embargo, la opción “Otros”, como se puede apreciar en la Tabla 1, también fue utilizada para indicar otras regiones donde los investigadores están desarrollando estudios sobre biodiversidad, y en el caso de uno de los investigadores, para indicar que actualmente no realiza actividades de investigación.

Tabla 1. Región - Opción Otros

R002	Seychelles
R003	None - stopped active research
R026	Global
R030	Arctic
R062	Antarctica

Respecto al puesto en el equipo de investigación de los encuestados, los más interesados en la temática de la encuesta, como lo muestra el Gráfico 4, han sido los niveles de dirección y conducción: directores de proyecto, coordinadores de investigación y los líderes de equipos de investigación.

<b>11.1</b> Integrante de equipo de investigación .....	<b>16</b>
<b>11.2</b> Líder de equipo de investigación .....	<b>29</b>
<b>11.3</b> Investigador asistente .....	<b>9</b>
<b>11.4</b> Investigador doctoral.....	<b>12</b>
<b>11.5</b> Investigador Postdoctoral.....	<b>15</b>
<b>11.6</b> Coordinador de investigación / Director de proyectos.....	<b>31</b>
<b>11.7</b> Otros .....	<b>8</b>

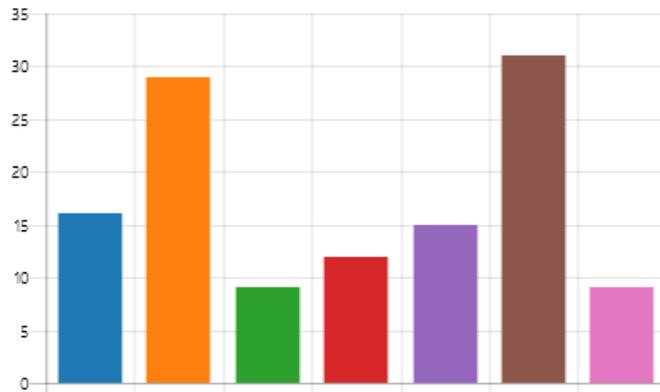


Gráfico 4. Posición en el equipo investigación

Era previsible que la categoría “Otros” para una encuesta que se distribuye de forma global podía llegar a contener información muy variada sobre puestos de trabajo, entre otras razones porque la denominación de los puestos de trabajo difiere de un país a otro. Alguna de las respuestas que aparecen en la Tabla 2 bien podrían caer en las categorías previstas, como R114 que es un líder de equipo. En el caso de los demás encuestados

que aparecen en la Tabla 2, sus puestos efectivamente no fueron contemplados en el formulario.

Son los casos de R009: Oficial de datos; R016 y R022: Investigador asociado; R025: Investigador principal; R041: Jubilado; R002: Administrador de proyectos; R063: Investigador senior. En algunos casos tienen un doble rol y utilizan la opción otros para especificar ambos roles (R022 y R025).

Tabla 2. Opción Otros - Pregunta 11

R002	Project Manager.
R009	Biodiversity Data officer at GBIF national node.
R016	Research Associate.
R022	Associate researcher and data manager.
R025	Lecturer/PI and co-PI.
R041	Retired.
R063	Senior Researcher.
R114	Government research scientist in charge of a team and a program.

## Opiniones de los encuestados (P1 a P9: preguntas 1 a 9 del cuestionario)

### P1. Razones por la cuales están dispuestos a compartir datos (con o sin restricciones)

Los motivos para compartir datos que más opiniones a favor obtuvieron fueron “Ayuda a evitar la duplicación de esfuerzos en la recogida de datos” y “Mejora la eficiencia en el uso de los fondos”, seguidos por razones como la utilidad que representa compartir datos para entrenar nuevos investigadores y la ayuda que proporciona en identificar errores en los datos.

<b>1.1</b> Ayuda a evitar la duplicación de esfuerzos en la recogida de datos .....	<b>83</b>
<b>1.2</b> Desalienta el plagio.....	<b>25</b>
<b>1.3</b> Mejora la eficiencia en el uso de los fondos.....	<b>78</b>
<b>1.4</b> Es útil para entrenar nuevos investigadores.....	<b>54</b>
<b>1.5</b> Implica beneficios para mi carrera.....	<b>30</b>
<b>1.6</b> Ayuda a identificar errores en los datos.....	<b>50</b>
<b>1.7</b> Otros .....	<b>18</b>

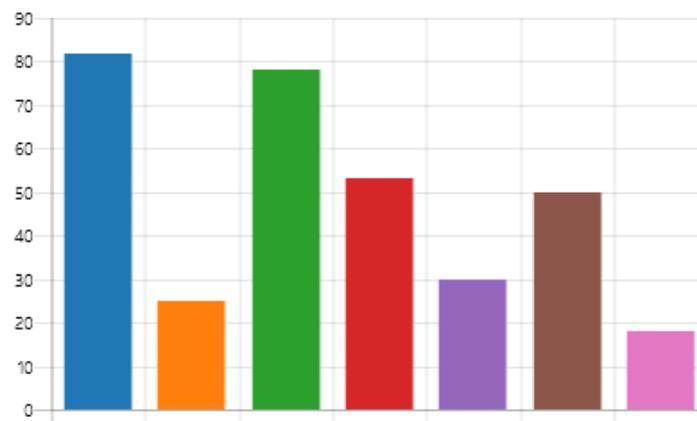


Gráfico 5. Motivos para compartir datos

Las respuestas completas a la opción “Otros” de la pregunta 1 se encuentran en el Anexo 4. Los investigadores aportan en esta opción varios motivos que los llevan a compartir datos no mencionados dentro de las opciones del formulario. Agrupamos los motivos más mencionados en tres categorías que facilitan el análisis:

**Compartir datos beneficia a otros investigadores** (9 menciones): beneficia a otros investigadores (*R020, R022, R041, R066, R0108*), especificando los beneficios de la reutilización, pues los datos son útiles para otros análisis u aplicaciones hasta ese momento no contempladas (*R022, R066, R091, R102*).

**Compartir datos es un beneficio para la comunidad científica** (4 menciones): en la misma línea de pensamiento de la categoría anterior, pero expresado de forma más general aparecen motivos como maximizar los beneficios de los datos para la comunidad científica (*R101*), contribuir al avance del conocimiento y de la investigación (*R031, R089*) y que los datos se utilicen lo más ampliamente posible (*R001*).

**Compartir datos alienta la colaboración** (3 menciones): *R019, R091, R053*.

Otros motivos que tienen una mención cada uno son: por ser una exigencia para la publicación de *papers* (*R033*), porque beneficia la conservación (*R016*), porque ayuda a la visibilidad de los datos (*R063*), por alentar el trabajo en equipo (*R028*) y motivos

que podríamos denominar como de reciprocidad -"*I would also like to have access to other's data*"- (R088).

## **P2. Factores que influyen positivamente al compartir datos sobre biodiversidad marina**

Se obtuvieron 122 respuestas para esta pregunta. De acuerdo con lo expresado por los investigadores encuestados, el factor que más influye al momento de compartir datos de investigación sobre biodiversidad marina es la disponibilidad de un repositorio: 81.8% lo considera muy importante o absolutamente esencial (47,1% muy importante y 34.7% absolutamente esencial).

El segundo factor en importancia de acuerdo con su influencia positiva en compartir datos es la existencia de una política de intercambio de datos: 77.3% lo considera muy importante o absolutamente esencial (53.8% muy importante y 23.5% absolutamente esencial).

La utilización de estándares de biodiversidad para el intercambio de datos aparece como un factor positivo importante, estando en concordancia con el porcentaje de importancia obtenido por los dos factores mencionados con anterioridad (disponibilidad de un repositorio y existencia de políticas de intercambio de datos). El 63.8% considera la utilización de estándares muy importante o absolutamente esencial (42% muy importante y % 21,8 absolutamente esencial).

En cuarto lugar, siendo uno de los factores considerados como muy positivos para el intercambio de datos en la literatura, se encuentra el pertenecer a una red o bien a un proyecto sobre biodiversidad marina. El 61.5% lo considera muy importante o absolutamente esencial (52.5% muy importante y 9% absolutamente esencial).

Como se puede apreciar en el Gráfico 6, los otros dos factores referidos a la presión para publicar datos (por parte de agencias de financiación y por parte de las revistas científicas) han sido los que obtuvieron menos porcentaje en cuanto a importancia otorgada por los encuestados.

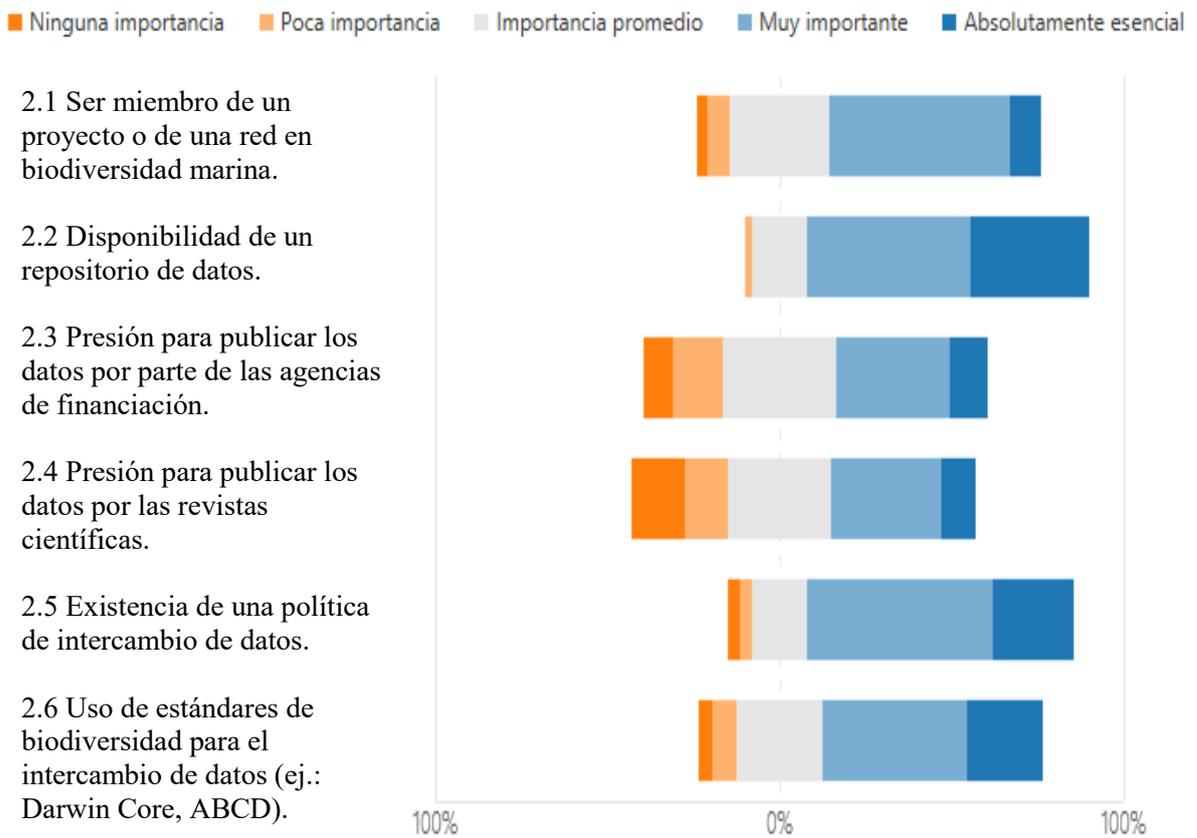


Gráfico 6. Factores que influyen positivamente al compartir datos sobre biodiversidad marina

El resultado más sorprendente del análisis de esta pregunta es que la influencia positiva que atribuyen a integrar una red o proyecto sobre biodiversidad sobre compartir datos si bien es alta, no es tanta como podría esperarse y hay otros tres factores que aparecen como más importantes.

Otro resultado que destaca es que no se atribuye tanta importancia en cuanto a factor positivo sobre compartir datos a las opciones relacionadas con presión para publicar datos.

### P3 y P4. Compartir conjuntos de datos y condiciones de uso o distribución

El total de respuestas para la pregunta 3 es de 122. El 39 % (32 encuestados) manifiesta no poner ninguna condición para compartir datos, frente a un 61% (90 encuestados) que pone condiciones para el uso y la distribución de los datos. Pero en realidad estos porcentajes no son tales pues cuando se dan opciones acerca de algunas condiciones

específicas para compartir datos (Pregunta 4), se puede apreciar que 28 de los 32 encuestados que se decantaron por la opción “No pongo ninguna condición” para la pregunta 3, luego señalan una (11 investigadores) o más de una de las opciones (10 investigadores) e incluso utilizan la opción Otros (7 investigadores) para responder a la pregunta 4. Solo 4 investigadores (R004, R025, R044 y R059) dejan sin contestar o indican “no requirements” o “none of the above” en la opción otros de la pregunta 4. En consecuencia, los porcentajes reales para esta pregunta son los que muestra el Gráfico 7.

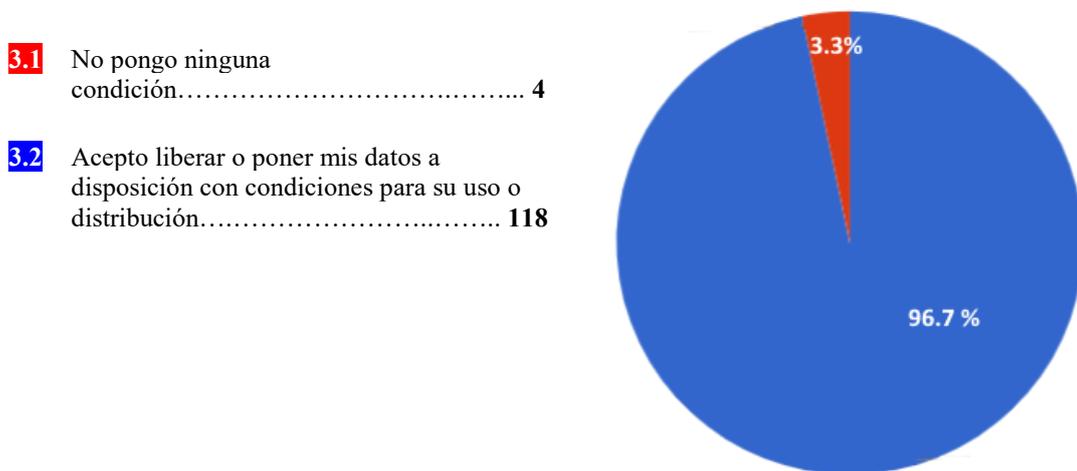


Gráfico 7. Compartir datos: con y sin condiciones

La condición que destaca en primer lugar dentro de las opiniones de los encuestados referentes a las condiciones para liberar o poner datos a disposición de otros es el permiso y/o el reconocimiento de los autores (44,7%), en segundo lugar, que los datos no sean utilizados con fines comerciales (31,2%) y en tercer lugar requerir la coautoría en publicaciones que surjan de los datos (17,2%). La opción Otros (codificada como 4.4) fue utilizada por el 7% de los investigadores (Gráfico 8).

4.1	Los datos no se utilicen sin permiso y/o sin reconocimiento del autor/es .....	96
4.2	Los datos no se deben usar con fines comerciales.....	67
4.3	Cualquier uso incluya mi coautoría en publicaciones que surgen de los datos.....	37
4.4	Otros .....	15

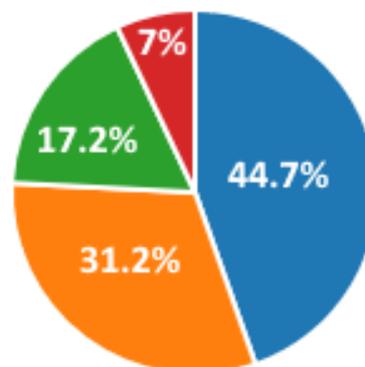


Gráfico 8. Condiciones para compartir datos

Las respuestas completas a Otros se encuentran en el Anexo 4. Los investigadores aportan en esta opción varias condiciones para compartir datos no mencionadas dentro de las opciones del formulario. Agrupamos las condiciones más mencionadas en tres categorías que facilitan el análisis:

**Reconocimiento** (5 menciones): de los autores (*R066, R088, R101, R102*), y de la institución y de los financiadores de la investigación (*R109*).

**Hacer los datos disponibles solo luego de utilizarlos para una publicación** (3 menciones): “*Make it available only after publishing*” (*R032*), “Ya hayan sido usados en una publicación de mi autoría” (*R117*), “Su utilización luego de la publicación de los resultados por nuestro grupo” (*R033*).

**Proponer o preguntar por coautoría o colaboración conjunta** en una publicación que incluya los datos (2 menciones): “*May inquire about co-authorship but wouldn't require it*” (*R016*), “*Any use includes proposing joint collaboration towards publications that include the data*” (*R063*).

Otras condiciones o requerimientos que tienen una mención cada uno son: que los datos sean usados apropiadamente para las preguntas a responder (*R022*), autorización de otras agencias si los datos fueron generados bajo contrato con otra agencia (*R020*), respetar los acuerdos contractuales y de propiedad intelectual (*R020*) y que se incluya un apartado con agradecimientos para el proyecto que financió la recogida de los datos (*R072*).

## P5. Prácticas para obtener datos sobre biodiversidad marina generados por otros científicos

Se obtuvieron 121 respuestas a esta pregunta, los resultados pueden observarse en el Gráfico 9. Buscar en línea resultó ser la práctica más utilizada para conseguir datos: 66,7% de los investigadores utiliza esta estrategia siempre o frecuentemente (29,2% siempre y 37,5% lo hace frecuentemente). La segunda práctica más utilizada para conseguir datos es la solicitud personal a los colegas: 58,6% de los investigadores lo hace siempre o frecuentemente (19,8% siempre y 38,8% lo hace frecuentemente). La tercera práctica en importancia para obtener datos es buscar en un repositorio central o centro de datos: 43,2% de los investigadores lo hace siempre o frecuentemente (12,7% siempre y 30,5% lo hace frecuentemente).

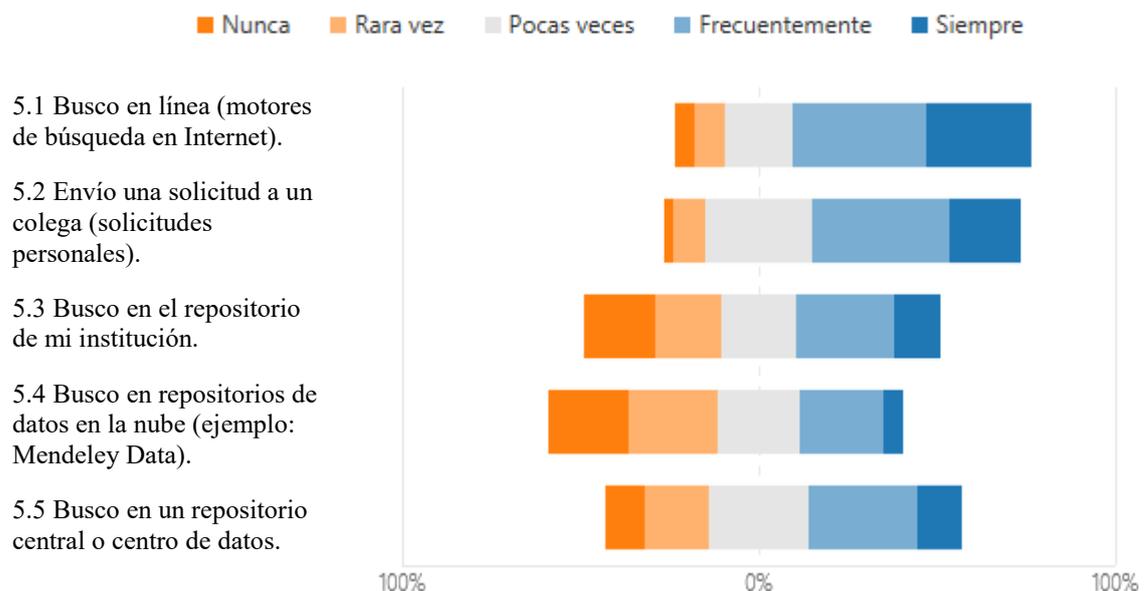


Gráfico 9. Prácticas para obtener datos

Buscar en el repositorio de la propia institución o en repositorios en la nube resultaron ser las opciones menos utilizadas por los investigadores. Buscar en repositorios de la institución lo hace el 40,3% siempre o frecuentemente, mientras que el restante 60,7% pocas veces, nunca o rara vez lo hace. Buscar en repositorios de datos en la nube es la estrategia menos utilizada: 70,6% nunca, rara vez o pocas veces la utiliza (22,7% nunca, 22,7% pocas veces y 25,2% rara vez lo hace).

## **P6. Otros recursos o estrategias para conseguir datos utilizados (prácticas no contempladas en las opciones de la pregunta 5).**

Se obtuvieron 24 respuestas a esta pregunta, de las cuales dos no contienen respuesta, solo un guión (*R012, R104*) y otra en la cual el encuestado considera que no hay opciones no contempladas en la pregunta 5 (*R086*). Los textos completos de las respuestas, por motivos de extensión, se han incluido en el Anexo 4, Parte 3: Respuestas a las preguntas abiertas.

Para facilitar el análisis de las estrategias y recursos más mencionados, éstos se presentan agrupados en cuatro categorías según la coincidencia en las respuestas obtenidas:

**Contactar colegas** (5 menciones): *R041, R046, R073, R100, R121*.

**Contactar organizaciones e instituciones** (4 menciones): que patrocinan investigación (*R001*), que recogen datos para otros propósitos que no son la investigación (*R058*), con las que se tiene convenios de cooperación (*R073, R098*).

**Buscar en bases de datos o bancos de datos** (4 menciones): *R033, R047, R069, R090*.

**Leer publicaciones** (3 menciones): literatura actual relacionada a recolección de datos (*R022*), "Publicaciones" (*R050*), Informes técnicos de acceso público, publicaciones seriadas (*R111*).

Otras prácticas para conseguir datos que tienen una mención cada una son: asistir a conferencias o discusiones sobre tópicos específicos (*R010*), utilizar los sistemas de información institucionales (*R034*), utilizar Google (*R084*), muestreo directo e interacción con contactos locales (*R094*), "*Collaborative projects specifically designed to bring together data (or metadata) for analysis*" (*R112*), "Datos en paqueterías de R." (*R117*).

## **P7. Motivos para retener datos**

Se obtuvieron 120 respuestas para esta pregunta. Los principales motivos por los cuales los investigadores encuestados retienen datos son: no tener tiempo suficiente para

formatear y preparar los datos para compartirlos (53 encuestados), por motivos de derechos de autor (45 encuestados), porque los patrocinadores de la investigación no aceptan hacer disponibles los datos (37 encuestados), por no tener el apoyo de organización para gestionar los datos (36 encuestados), y porque compartirlos disminuye su propia ventaja competitiva (27 encuestados). Estos resultados pueden observarse en el Gráfico 10.

<b>7.1</b> Los datos tenían derechos de autor y no se pueden compartir libremente....	<b>45</b>
<b>7.2</b> No tuve el apoyo de mi organización para administrar mis datos.....	<b>36</b>
<b>7.3</b> No tuve tiempo suficiente para formatear y preparar mis datos para compartirlos.....	<b>53</b>
<b>7.4</b> Sentí que compartir los datos disminuiría mi propia ventaja competitiva.....	<b>27</b>
<b>7.5</b> Mis conclusiones basadas en los datos pueden ser objetadas si los datos estuvieran disponibles.....	<b>5</b>
<b>7.6</b> La desidentificación de datos sensibles resultaba compleja.....	<b>9</b>
<b>7.7</b> Los patrocinadores de mi investigación no aceptaron hacer disponibles los datos.....	<b>37</b>
<b>7.8</b> Otros.....	<b>20</b>

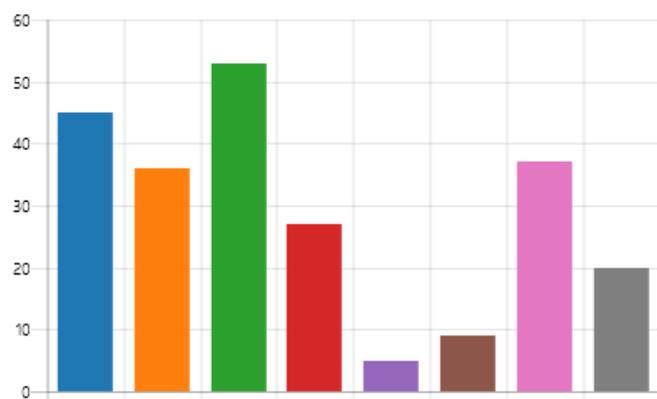


Gráfico 10. Motivos para retener datos

Motivos para retener datos como la desidentificación de datos sensibles o que las conclusiones basadas en los datos pueden ser objetadas, no resultaron ser motivos de retención muy habituales.

La opción Otros de esta pregunta (codificada como 7.8) obtuvo 20 respuestas. Cinco encuestados utilizaron la opción Otros para indicar que la pregunta no aplica en su caso (R016, R027, R115, R109), o que nunca retuvieron datos (R072). En un caso fue utilizada para hacer una reflexión sobre experiencias propias: “Sospecho que la mayoría de las razones anteriores han estado detrás de mis solicitudes fallidas de datos” (R112) (mi traducción), refiriéndose a las opciones 7.1 a 7.7 de esta pregunta. En suma, se analizaron 14 respuestas que sí aportan motivos para esta opción, agrupando los que tienen más de una mención en categorías según la coincidencia en las respuestas obtenidas:

**Por estar trabajando en publicaciones** (5 menciones): *R031, R041, R049, R053, R101*.

**Por causa de los permisos para compartir** (3 menciones): obtener permisos de colegas toma tiempo (*R022*), colegas que contribuyeron con datos se niegan luego a compartirlos (*R057*), trabajar con compilaciones de datos requiere los permisos de todos los proveedores originales (*R102*).

**Por desconfianza sobre el uso que otros puedan dar a los datos** (2 menciones): *R091, R114*.

Otros motivos para retener datos que tienen una mención cada uno son: para evitar usos comerciales de los datos (*R108*), por no ser consciente de la importancia de compartir datos (*R093*) y por el tipo de datos: "Para datos moleculares, no es posible retener los datos si la intención es una publicación científica. Para aplicaciones biotecnológicas retengo los datos" (*R033*).

## **P8. Factores desfavorables a compartir datos**

Las preocupaciones sobre el uso inadecuado de los datos son la principal causa que desalienta a compartir datos, motivo que luego profundizaremos pues vuelve a manifestarse en la siguiente pregunta del cuestionario (Pregunta 9, abierta). El 45.4 % lo considera muy importante o absolutamente esencial (15.1 % muy importante y 30.3 % absolutamente esencial) (Gráfico 11).

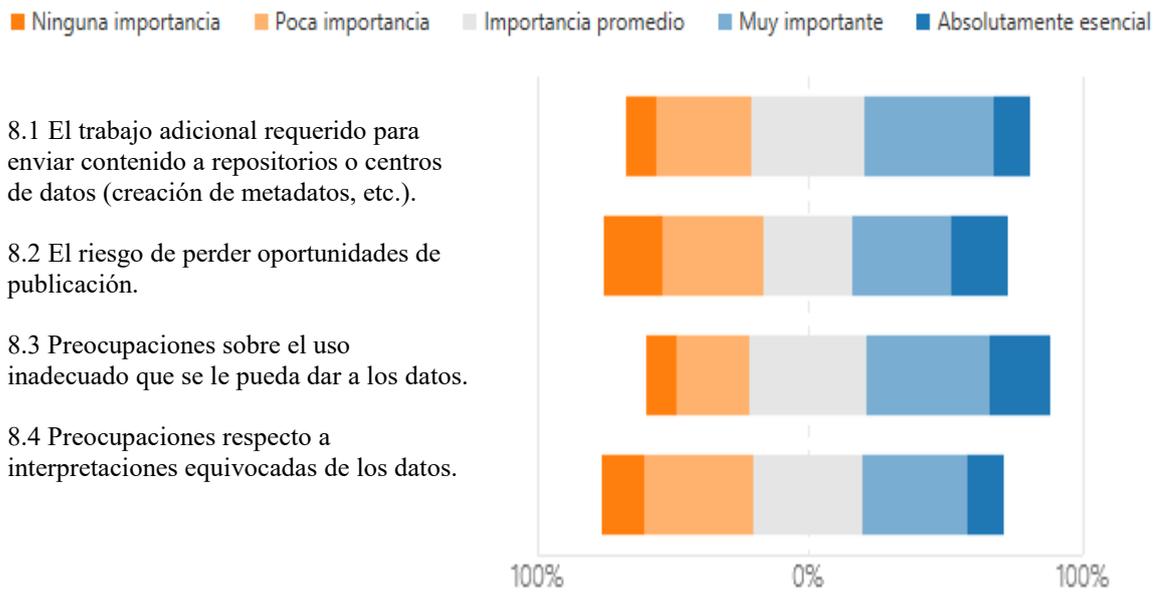


Gráfico 11. Factores desfavorables a compartir datos

El factor desfavorable al momento de compartir datos que aparece en segundo lugar según las opiniones recabadas en la encuesta es el trabajo adicional requerido para poner los datos a disposición en repositorios o centros de datos (31.7% lo considera un factor muy importante y 9.2 % absolutamente esencial). El riesgo de perder oportunidades de publicación también aparece como un factor importante al momento de compartir datos (24.2 % lo considera un factor muy importante y 12.4 % absolutamente esencial). Las preocupaciones por las interpretaciones equivocadas de los datos tienen una tendencia más equilibrada, pues si bien 26.1% lo considera un factor muy importante y 9.2 % absolutamente esencial, hay un 26.9% que le atribuye una importancia promedio, y un 37.8% que le atribuye poca o ninguna importancia a este factor (26.9% poca y 10.9% ninguna importancia).

### P9. Comentarios sobre experiencias personales en compartir datos u opiniones sobre otros factores que alientan o desalientan a compartir datos

Las respuestas a esta pregunta abierta complementan la información recabada en las preguntas 2 y 8, cuyas tendencias se ven confirmadas y en muchos casos los encuestados aportan detalles sobre aspectos específicos. Los textos completos de las respuestas a esta pregunta abierta, por motivos de extensión, aparecen en el Anexo 4. Respuestas a las preguntas abiertas.

Se obtuvieron 29 respuestas a esta pregunta, de las cuales se analizan a continuación 25, pues cuatro encuestados utilizaron esta opción de respuesta abierta para expresar que no tenían comentarios (R035, R094, R104, R086).

Son mayoría los comentarios sobre factores desfavorables al compartir datos, incluyendo en algunos casos detalles sobre malas experiencias al respecto (13 comentarios): R001, R016, R047, R053, R057, R058, R069, R087, R098, R111, R112, R114, R120). La confianza es el factor más mencionado (R053, R058, R069, R111, R114), muy vinculada a experiencias previas de apropiación indebida de los datos.

*Data appropriation without consent (stealing data) is a common issue in the Brazil, and as is a problem in sharing valuable information with colleagues (R058).*

*[...] unscrupulous researchers used the data as their own without any form of acknowledgement of the data producer or project/institution. Their argument for doing so was that the data was "freely available". This highlights the fact that clear conditions for the use of the data need to be in place (R069).*

*La confianza en que el investigador usará los datos apropiadamente y compartirá autoría en publicaciones es fundamental (R111).*

*I was much more open to freely sharing data in the past but am now extremely cautious, due to three bad experiences [...] These experiences have totally discouraged me from sharing data outside of a close circle of trusted colleagues. I will not put data on any publicly accessible site where I have no control over its use (R114).*

Otras causas que se mencionan repetidas veces en detrimento de compartir datos son: actitudes de resistencia a compartir datos (R016, R087, R112) y no querer compartirlos por tratar los datos de investigación con recelo, gestionándolos como propiedad personal cuando en muchos casos no lo son.

*Many of the potential contributors are hesitant or unwilling to share data that they've been sitting on for ~15-20 years that is still unpublished. Reasons range from concern about missing publishing opportunities to having previous bad experiences with data sharing (R016).*

*In most cases ownership of data belongs to organizations (or funding agencies) not individual scientists, but this does not seem to stop some scientists treating data as their own (R112).*

Aparecen también las alusiones a la mala calidad de los datos como factor desfavorable al compartir. Esto sucede por distintas razones, entre ellas el tiempo que toma hacer una

evaluación de la calidad / control de calidad de los datos, o porque muchas veces lo que se ofrece son paquetes de datos incompletos que dificultan un trabajo serio a partir de ellos (R001, R057).

*Mostly the issue is time [...] broad collaboration requires a good amount of QA/QC. That is the biggest barrier to more widely distributed data (R001).*

*I have previously found errors in biodiversity data (omission of zero observations, duplication of records, insufficient metadata describing the sampling methods, taxonomic errors, etc.) that make it hard to use the data for serious work (R057).*

También son de destacar las menciones específicas a la falta de una cultura de compartir datos como factor desfavorable (R047, R120), con especial mención a la reciprocidad como valor esencial que debe formar parte de esa cultura. La necesidad de compartir datos de una forma más permanente, sin importar si se forma parte de una red de especialistas o no, es una reflexión que de alguna manera también apunta a denotar esa carencia de una cultura de compartir datos (R047).

*Reciprocity is essential, clear data publishing rules by funding agencies is needed and, when existent must be enforced. At least in Mexico we have a huge lack of data sharing rules and culture (R088).*

*It lacks support of my institution for the preparation of the data and also lacks the culture of sharing data in a constant flow (R120).*

En segundo lugar, teniendo en cuenta el número de respuestas (6 comentarios), están los encuestados que utilizan esta opción de respuesta abierta para realizar aportes y reflexiones sobre compartir datos.

La mayor cantidad de aportes en esta línea son los que tienen que ver con prioridades al compartir datos (R031, R034, R046), ya sea diferenciándolos por tipos de datos a compartir, o por el origen de los fondos que apoyaron la recolección de esos datos.

*Hay datos de base (por ejemplo, registros de temperaturas en el mar) que he compartido inmediatamente y sin restricciones. Otros datos biológicos más específicos (ej., resultados de censos de abundancia animal) se pueden hacer públicos una vez que se han publicado (R031).*

*Entregar datos brutos o muestrales, pueden ser mal interpretados, por lo tanto, se sugiere depurar y realizar el análisis exploratorio previo antes de entregar o publicar (R034).*

En algunos casos, debería priorizarse el uso compartido de datos. Por ejemplo, institutos de investigación nacional, financiados por el estado, tienen más obligaciones prioritarias con el público en general (R046).

Las reflexiones se concentran en la utilidad de compartir datos (R073) y en plantear cambios que se consideran necesarios (R016, R087).

*I think for our attitudes towards data sharing to change we need to think critically about making published datasets carry just as much "academic weight" as a publication. And how we can create a better system to reward data collection (R016).*

*I believe that misuse of data, extra work providing it to a repository, or the possibility of misuse are all real. However, in order to bring about change, all of these issues are critical in order to advance the field of marine biodiversity (R087).*

En lo personal creo que compartir y requerir datos es esencial para fortalecer líneas de investigación y generar redes de apoyo entre distintos investigadores con el fin de mejorar el conocimiento actual que se puede tener sobre un tema (R073).

En tercer lugar, teniendo en cuenta el número de respuestas (4 comentarios), están los encuestados que utilizan esta opción de respuesta abierta para reforzar su posición de estar dispuestos y no tener ningún problema en compartir datos y lo fundamentan (R084, R100, R105, R112).

*I share my data freely as it gives my work exposure, offers opportunities to develop publications, and increasingly because it is required by funding organizations anyway (R112).*

Solo dos encuestados se centran en comentar los factores que en su opinión favorecen el compartir datos (R033, R053).

Es esencial disponer de sistemas de almacenamiento de datos fáciles de utilizar, preferentemente nacionales [...] (R033).

*Encourage: easy to use platforms for sharing, easy to use metadata forms, help from collaborators, errors to be fed back on by collaborators, users to send any issues/requests back to the owner, so the data can be updated and cleaned accordingly (R053).*

### 5.1.2. Factores que más inciden y prácticas más habituales al compartir datos en función de las variables edad y posición en el equipo de investigación

A continuación, se analizan los factores que favorecen compartir datos sobre biodiversidad marina que resultaron en primer, segundo y tercer lugar en los resultados de la encuesta (sub-sección A, apartados A.1, A.2 y A.3), las tres prácticas más utilizadas para obtener datos (sub-sección B, apartados B.1, B.2 y B.3) y los tres factores principales que desalientan compartir datos (sub-sección C, apartados C.1, C.2 y C.3) respecto de las variables edad y posición en el equipo de investigación de los encuestados.

#### A. Principales factores que favorecen compartir datos sobre biodiversidad marina

##### A.1. Disponibilidad de un repositorio de datos

La disponibilidad de un repositorio donde alojar los datos es muy apreciada por investigadores de todas las edades, resultando los menores de 35 años los que más relevancia le dan a este factor, y destacándose que un alto porcentaje del grupo etario de 55 a 64 lo considera absolutamente esencial (Gráfico 12).

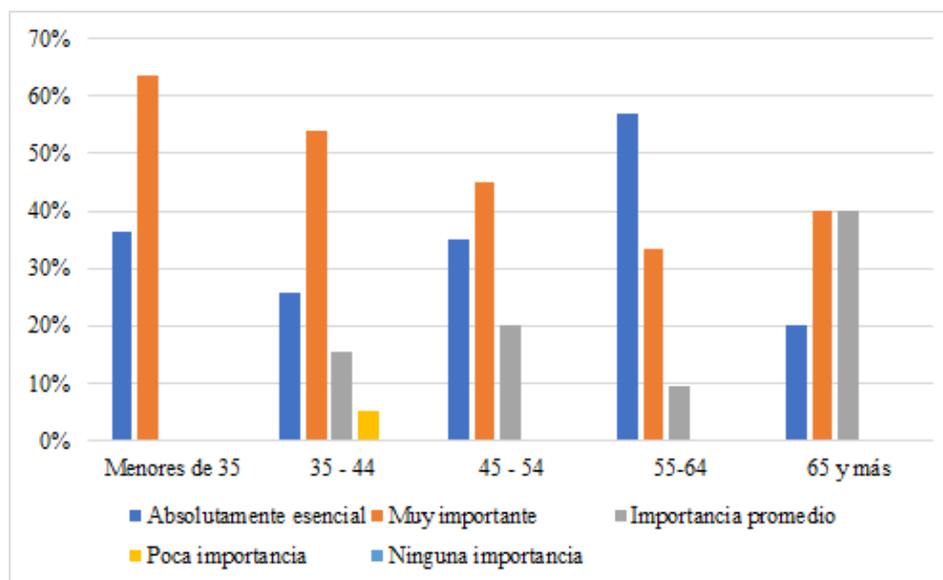


Gráfico 12. Importancia de contar con un repositorio de datos según grupos de edades

Es destacable el porcentaje de investigadores asistentes que consideran absolutamente esencial la disponibilidad de un repositorio de datos (56%), siendo también alto este

porcentaje en investigadores doctorales (42%). Los integrantes de equipo de investigación y los investigadores postdoctorales lo consideran muy importante: 69% y 64% respectivamente (Gráfico 13).

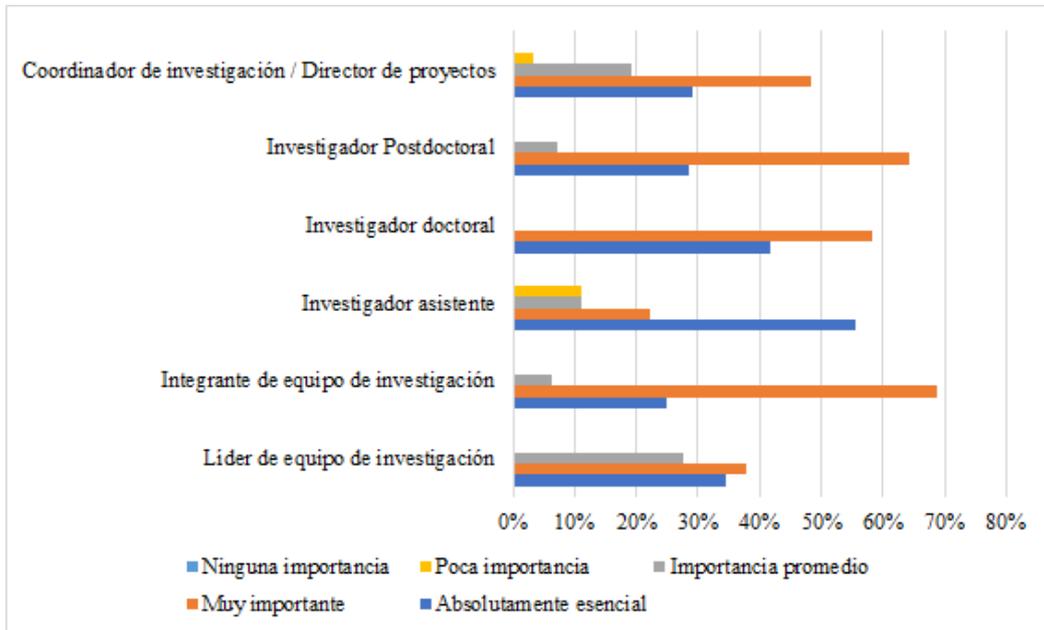


Gráfico 13. Importancia de contar con un repositorio de datos según la posición en el equipo de investigación

## A.2. Existencia de una política de intercambio de datos

La existencia de una política de intercambio de datos es más apreciada en los grupos etarios de 35-44 (18% lo considera absolutamente esencial y 61.5% muy importante) y 45-54 (22.5% lo considera absolutamente esencial y 57.5 muy importante), como se puede apreciar en el Gráfico 14. Además, dentro de los menores de 35 el 40% la considera absolutamente esencial y el 50% muy importante. En los siguientes grupos etarios, la consideración de esencial aumenta hasta los 64 años y el considerarla muy importante decrece respecto a la edad, a partir de los 35. Dentro de los mayores de 65 años hay un 22.2 % que manifiesta que la existencia de una política de datos no tiene ninguna importancia.

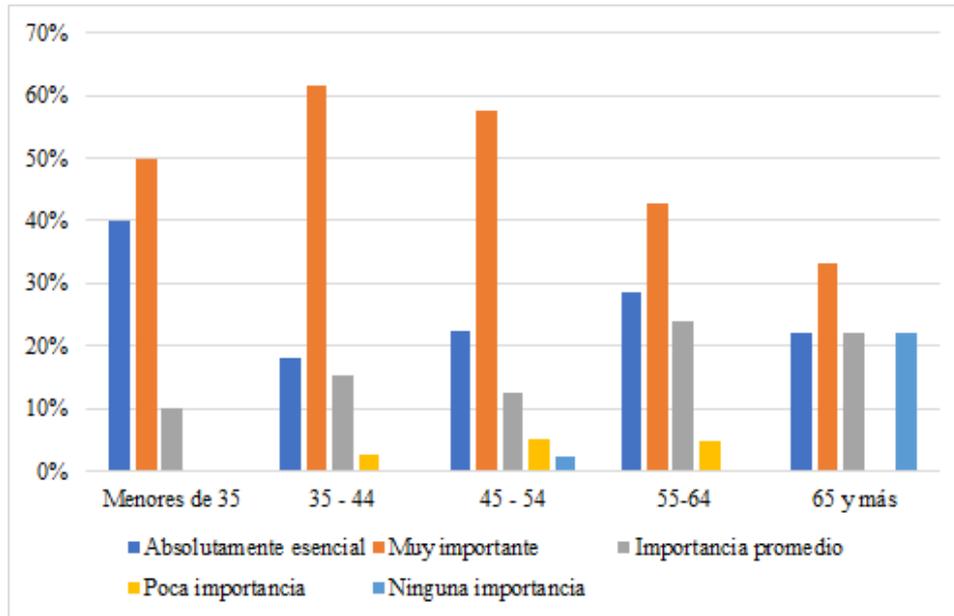


Gráfico 14. Importancia de la existencia de una política de intercambio de datos según grupos de edades

Los investigadores doctorales y posdoctorales son los que más importancia manifiestan cuando se pregunta por la existencia de una política de intercambio de datos (Gráfico 15).

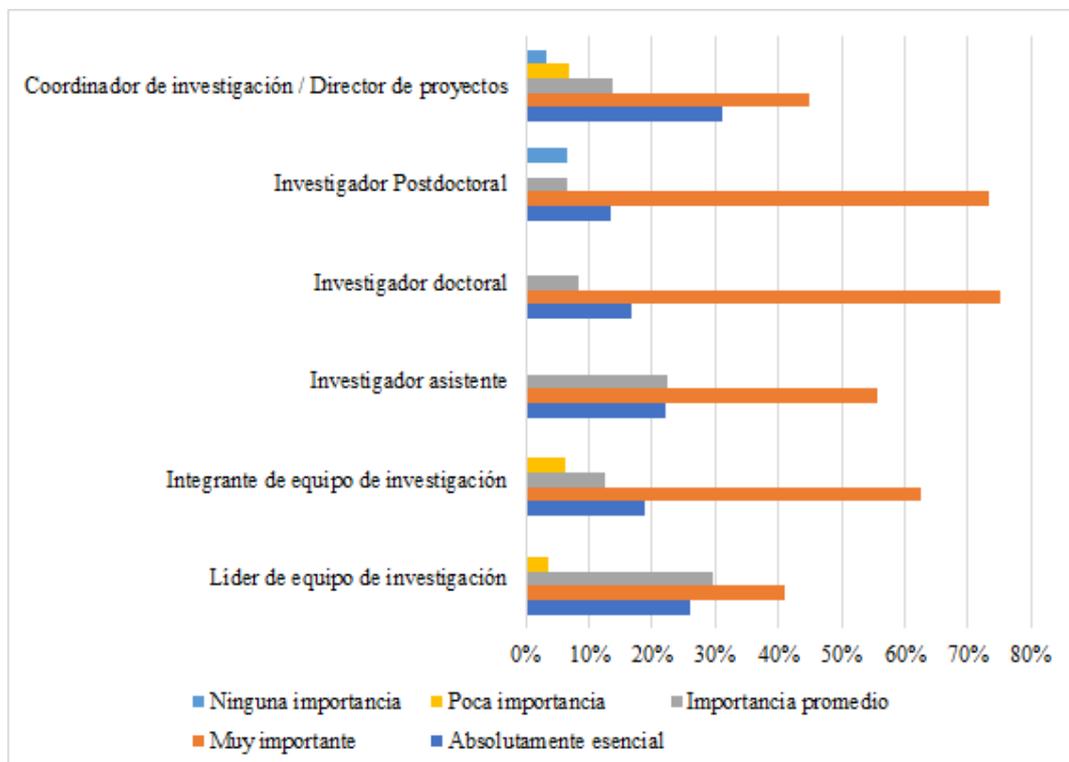


Gráfico 15. Importancia de la existencia de una política de intercambio de datos según posición en el equipo de investigación

### A.3. Utilización de estándares de biodiversidad

En los grupos de edades entre 35 a 64 es donde se constatan porcentajes más altos de aprecio por la utilización de estándares. Las opciones absolutamente esencial y muy importante sumadas, para cada uno de los grupos 35-44, 45-54 y 55-64, superan el 60%.

El análisis de los datos muestra que el considerarlo absolutamente esencial decrece con la edad del entrevistado, y el considerarlo muy importante aumenta con la edad (Gráfico 16).

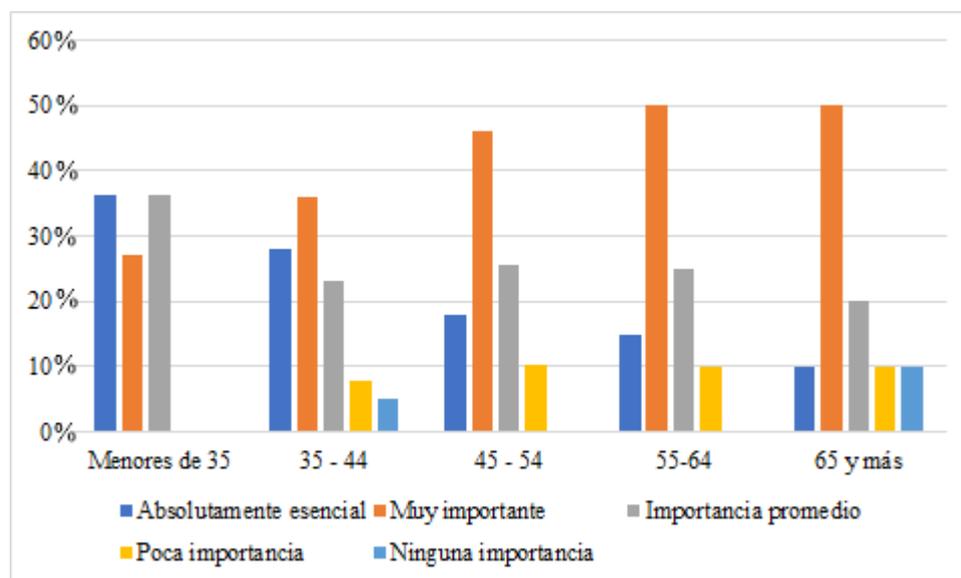


Gráfico 16. Importancia de la utilización de estándares según grupo de edades

Los que más importancia otorgan al empleo de estándares para datos sobre biodiversidad son los investigadores asistentes (absolutamente esencial 55.5%, muy importante 33.3%), los posdoctorales (absolutamente esencial 35.7%, muy importante 43%) y los integrantes de equipo de investigación (absolutamente esencial 35.7%, muy importante 43%) (Gráfico 17).

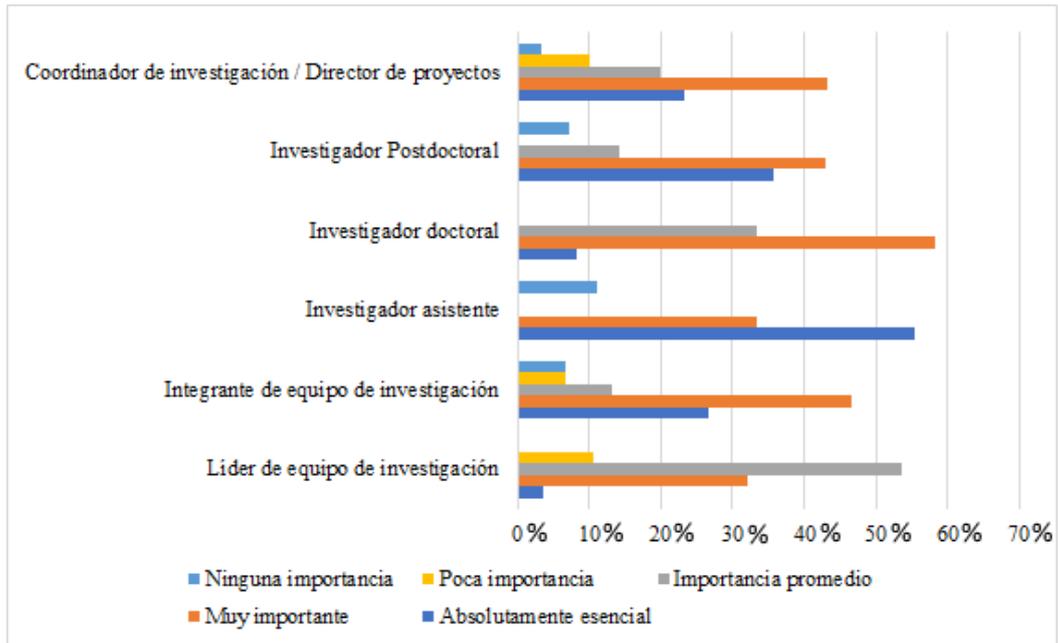


Gráfico 17. Importancia de la utilización de estándares según posición en el equipo de investigación

## B. Prácticas para obtener datos sobre biodiversidad marina

### B.1. Buscar datos en línea

La utilización de búsqueda en línea en motores de búsqueda de Internet como estrategia para localizar datos, muestra una tendencia decreciente en relación con la edad hasta los 54 años, si se observan los porcentajes de encuestados que mencionan utilizar esta estrategia siempre. Los mayores de 45 declaran utilizarla frecuentemente, resultando los mayores de 65 el grupo etario que indica utilizarla frecuentemente en mayor porcentaje que los demás grupos (Gráfico 18).

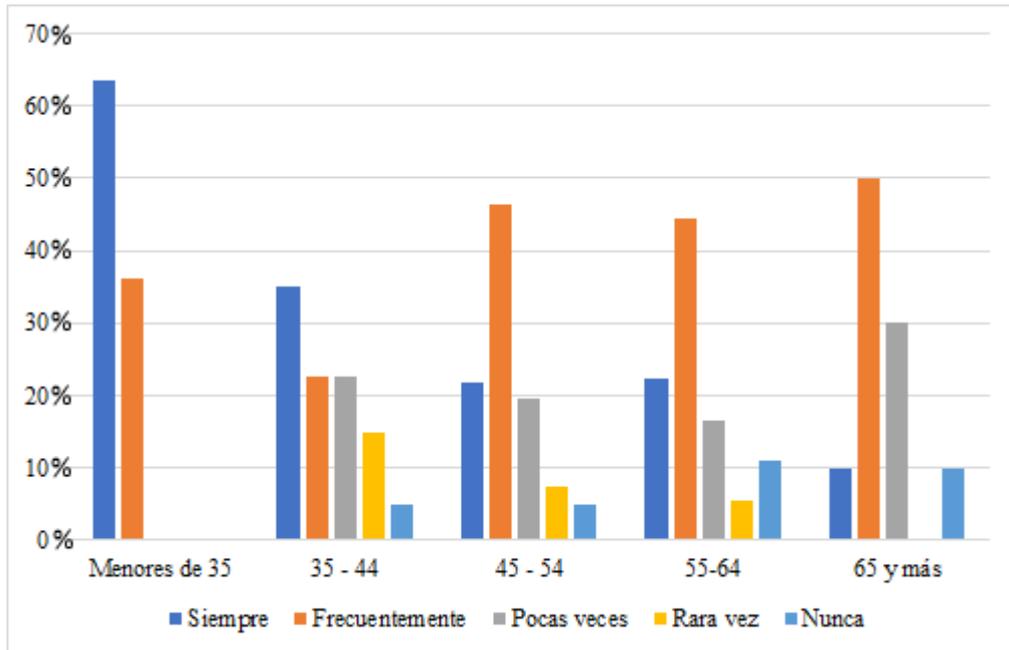


Gráfico 18. Frecuencia de búsqueda de datos sobre biodiversidad marina en línea según grupos de edad

Los investigadores doctorales resultaron ser los que más se inclinan por esta práctica (58% la utilizan siempre, 33% frecuentemente) y los líderes de equipo de investigación indican que la utilizan frecuentemente en un 59% y siempre un 17% (Gráfico 19).

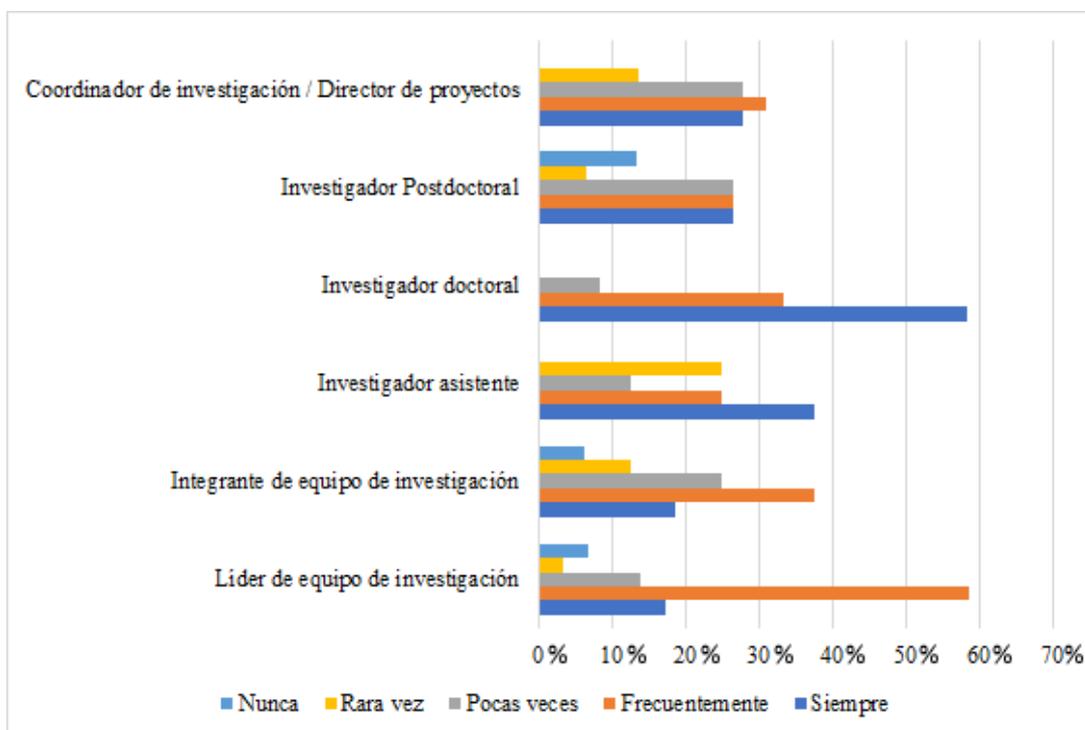


Gráfico 19. Frecuencia de búsqueda de datos sobre biodiversidad marina en línea según posición en el equipo de investigación

## B.2. Solicitudes personales

Es una práctica frecuentemente utilizada entre los menores de 45, los menores de 35 dicen utilizarla siempre en un 18.2% y 45.4% la utiliza frecuentemente. En el grupo de 35-44 años un 15% utiliza siempre la práctica de solicitar datos a otro colega y un 55% dice recurrir a los colegas frecuentemente. Se destaca en los resultados del análisis la frecuencia con la que los investigadores de entre 55-64 años solicitan siempre datos a sus colegas: 36.8 % (Gráfico 20).

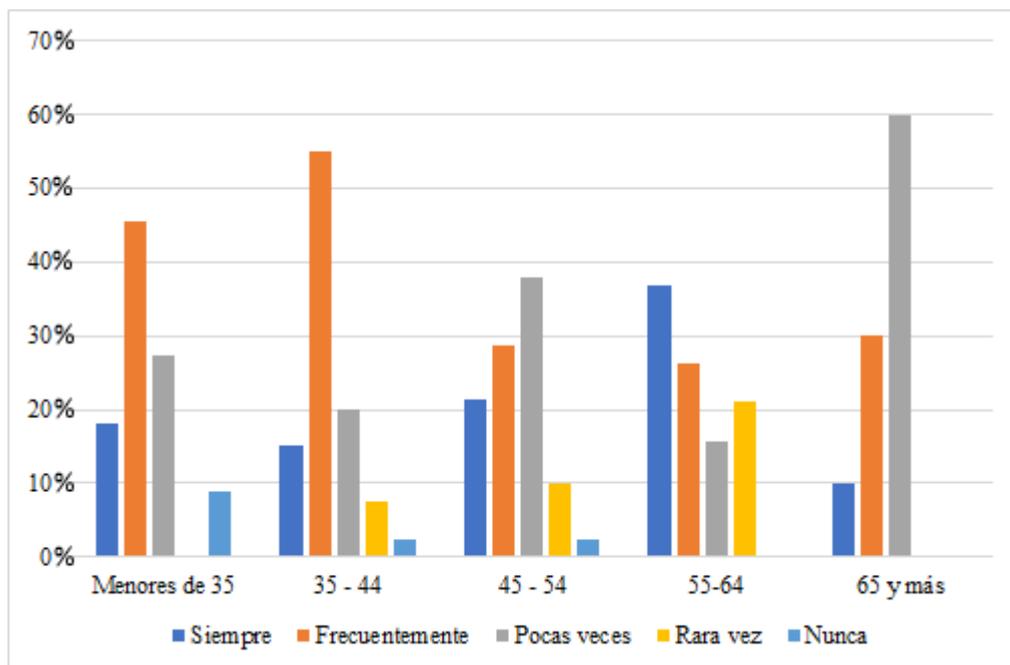


Gráfico 20. Solicitudes personales de datos sobre biodiversidad marina según grupos de edades

Los investigadores doctorales, investigadores asistentes y líderes de equipo de investigación resultaron ser los que más recurren a esta práctica (Gráfico 21).

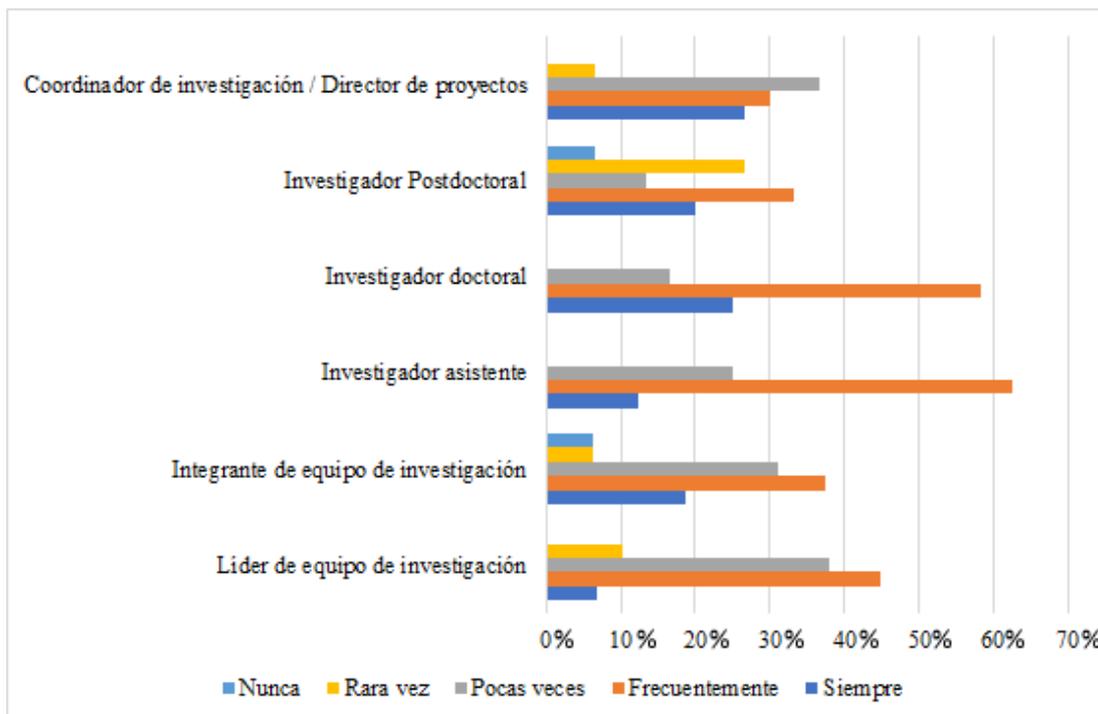


Gráfico 21. Solicitudes personales de datos sobre biodiversidad marina según posición en el equipo de investigación

### B.3. Buscar en un repositorio central o centro de datos

Los que más acuden a repositorios centrales o centros de datos para sus búsquedas son los menores de 35 (20% dice hacerlo siempre y 30% frecuentemente) y los del grupo etario entre 45-54 (14.6% dice hacerlo siempre y 36.6% frecuentemente). Los que menos recurren a esta opción son los mayores de 65 (Gráfico 22).

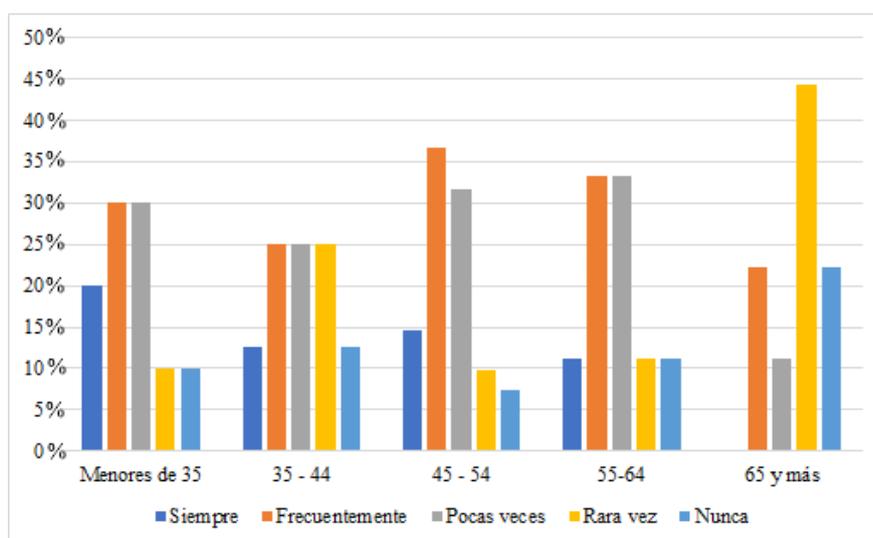


Gráfico 22. Frecuencia con que buscan en un repositorio central o centro de datos según grupo de edades

Los investigadores doctorales son los que más recurren a buscar en repositorios centrales o centros de datos (16.6% dice hacerlo siempre y 41.6% frecuentemente), seguidos por coordinadores de investigación / directores de proyectos (10.3% dice hacerlo siempre y 44.8% frecuentemente) y líderes de equipo de investigación (21.4% dice hacerlo siempre y 32.1% frecuentemente) (Gráfico 23).

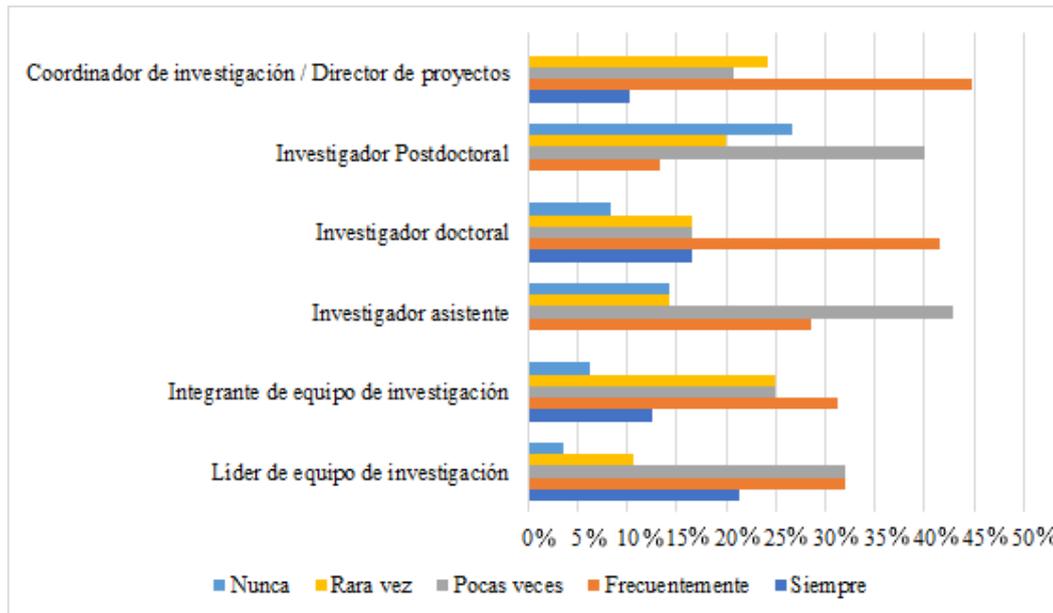


Gráfico 23. Frecuencia con que buscan en un repositorio central o centro de datos según posición en el equipo de investigación

## C. Factores que desalientan compartir datos sobre biodiversidad marina

### C.1. Preocupación por el uso inadecuado de los datos

El uso inadecuado preocupa menos a medida que aumenta la edad hasta los 54 años. En el grupo de 65 y más el 40% le atribuye mucha importancia al posible uso inadecuado de los datos (Gráfico 24).

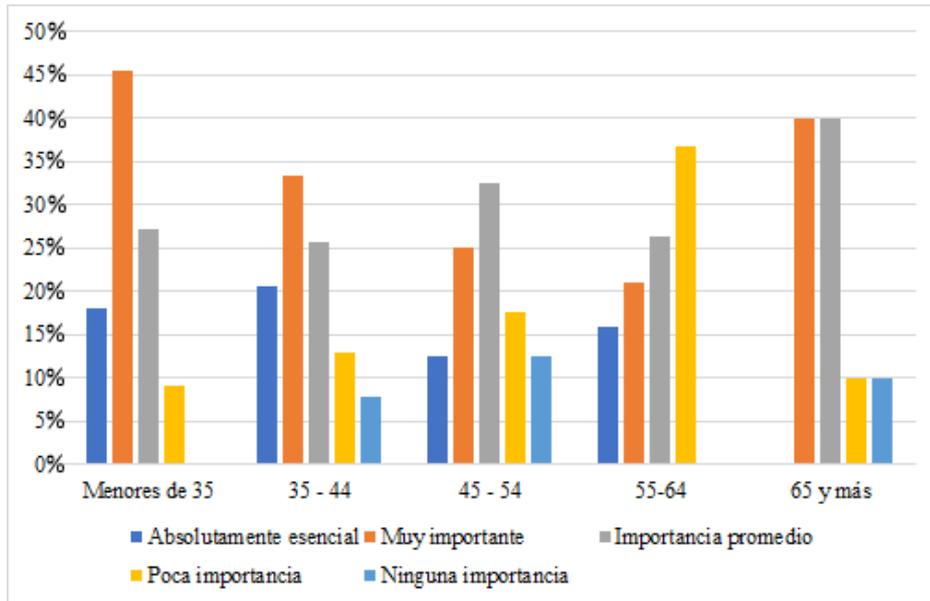


Gráfico 24. Preocupación por el uso inadecuado de los datos según grupos de edad

Si lo analizamos respecto a la posición en el equipo de investigación, los más preocupados por este factor son los investigadores asistentes y los integrantes de equipo de investigación (Gráfico 25).

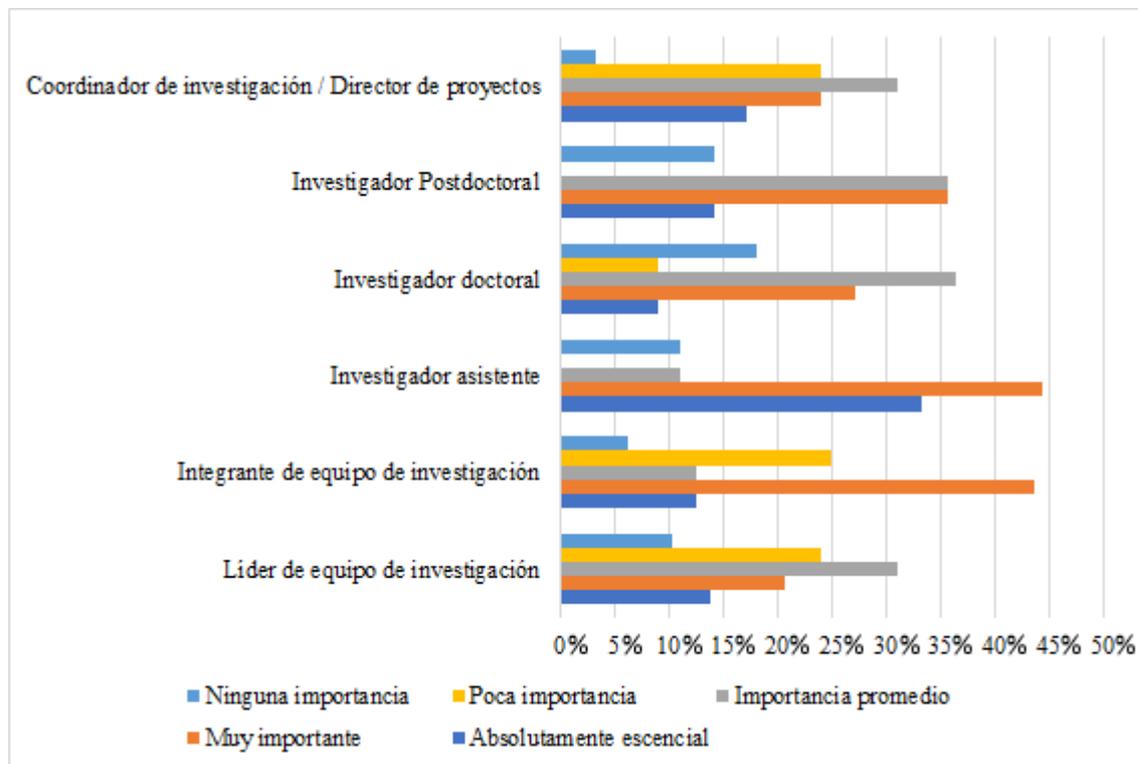


Gráfico 25. Preocupación por el uso inadecuado de los datos según posición en el equipo de investigación

## C.2. Preparar datos para compartirlos

El trabajo adicional de preparar los datos para hacerlos disponibles (ejemplo: ser depositados en un repositorio) es un factor que resulta muy importante en todos los grupos etarios. El grupo que más manifiesta la incidencia de este factor es el de 55-64 años (Gráfico 26).

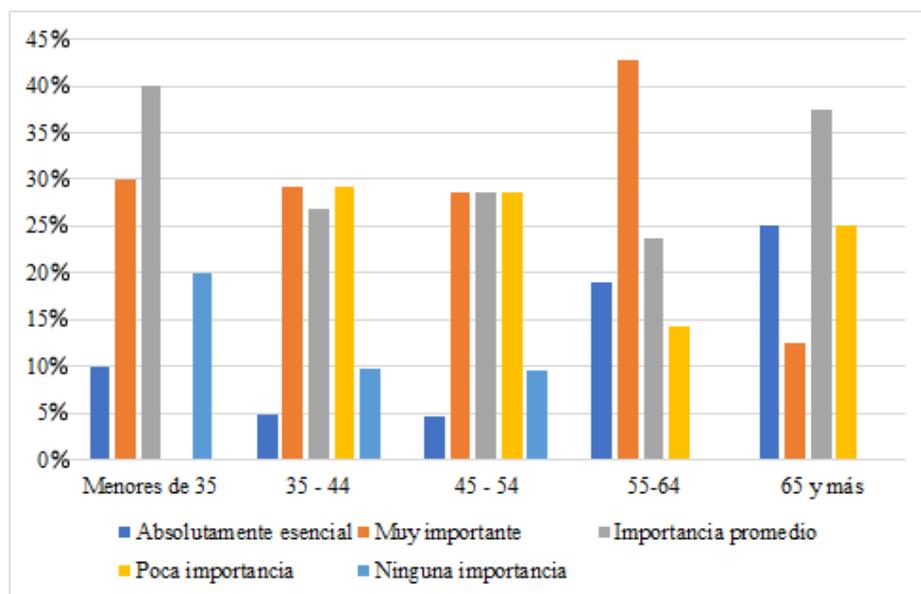


Gráfico 26. Importancia del trabajo adicional requerido para compartir los datos según grupos de edad

Principalmente los investigadores asistentes son los que mayor importancia le atribuyen (el 22.2% lo considera absolutamente esencial y el 33.3% muy importante), los integrantes de equipo de investigación (el 12.5% lo considera absolutamente esencial y el 31.2% muy importante) seguidos por los investigadores postdoctorales (40% lo considera muy importante) (Gráfico 27).

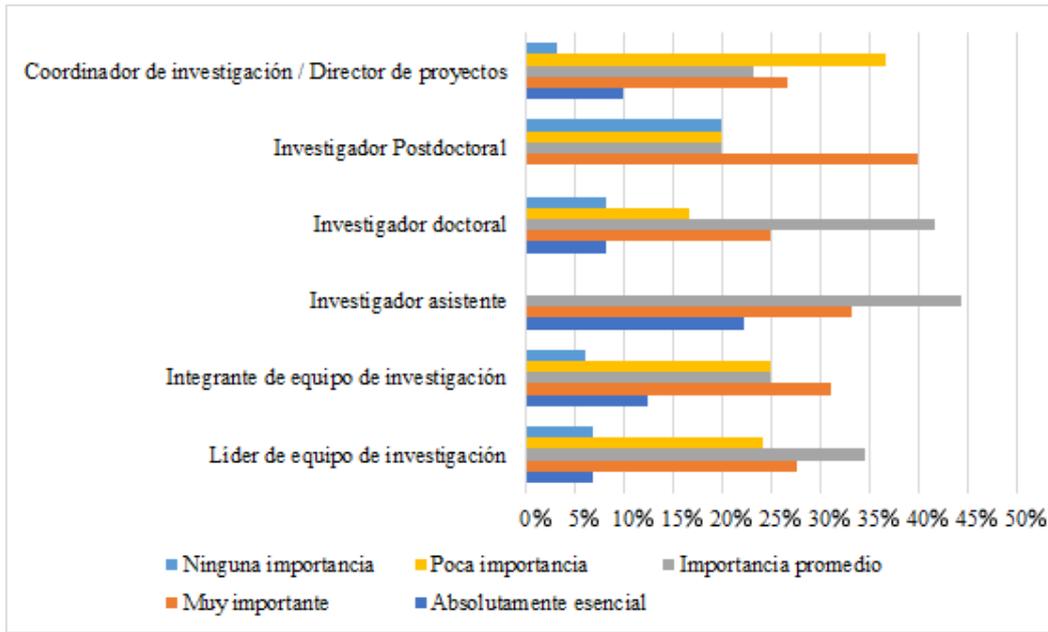


Gráfico 27. Importancia del trabajo adicional requerido para compartir los datos según posición en el equipo de investigación

### C.3. Riesgo de perder oportunidades de publicación

Los más preocupados por perder oportunidades de publicación a causa de compartir datos son los investigadores entre 35-44 años: el 25.6 % lo considera absolutamente esencial y el 20.5 % muy importante. A los grupos de edades de 45-54 también les preocupa el riesgo de perder oportunidades de publicación, pero en menor medida que a los del grupo de 55-64. Los menores de 35 consideran este factor absolutamente esencial en un 9% y muy importante en un 27.3 %. Los menos preocupados son los mayores de 65.

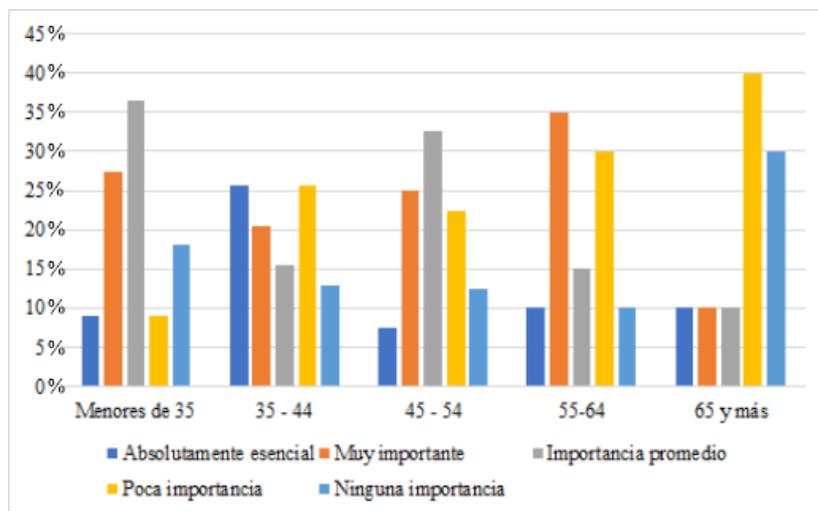


Gráfico 28. Importancia conferida a perder oportunidades de publicación por compartir datos según grupo de edades

Los investigadores postdoctorales son los más preocupados por el riesgo de perder oportunidades de publicación a causa de compartir datos: 21.4 % consideran el riesgo de perder oportunidades de publicación absolutamente esencial y 28.6 % muy importante, seguidos por los investigadores doctorales, quienes lo consideran absolutamente esencial en un 8.3 % y muy importante en un 41.6 %. Los investigadores asistentes creen que es absolutamente esencial en un 33.3 % y muy importante en un 11.1 % (Gráfico 29).

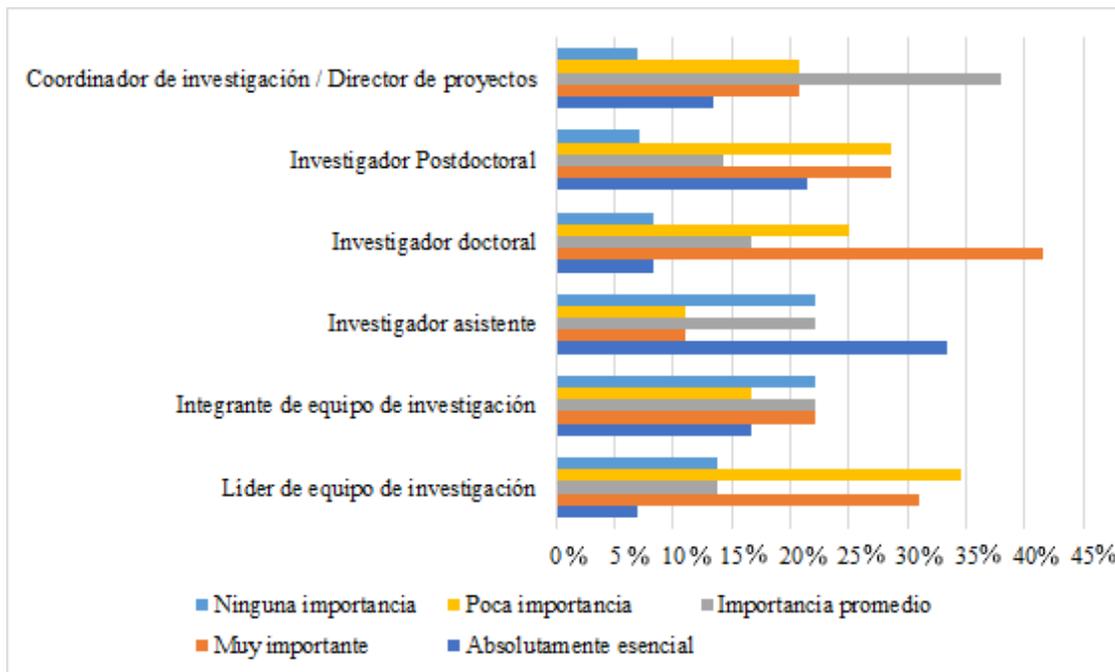


Gráfico 29. Importancia conferida a perder oportunidades de publicación por compartir datos según posición en el equipo de investigación

## 5.2. Análisis de contenido de los documentos de las reuniones del Grupo Directivo de OBIS

El sistema de información OBIS tiene una forma de gobierno en la cual el órgano de conducción es el Grupo Directivo. El grupo Directivo (*Steering Group*, SG) está integrado por representantes de los nodos (*Node managers* / gestores, coordinadores de nodo) y cuenta con un *Chair* y un *Co-Chair* elegidos entre sus miembros, que presiden las reuniones anuales, dirigen y representan al SG. El SG cuenta con varios grupos de trabajo, (*Task Teams*, TT) a los cuales delega tareas específicas. Los TT pueden ser permanentes o transitorios según demande la tarea o actividad. La Secretaría permanente del sistema es desempeñada por IODE (UNESCO).

En los documentos de las reuniones del SG se reportan actividades desarrolladas por los nodos, pero también por los grupos de trabajo. Son las actividades de los grupos de trabajo las que resultan de mayor interés para esta investigación, pues es en esa colaboración donde pretendemos identificar ciertas prácticas al compartir datos e información, que presumimos muy amalgamadas con el compartir capacidad técnica, científica, recursos e infraestructura de información.

Las actividades de colaboración son muy importantes para el Sistema, tanto así que la participación activa de los nodos en estas actividades es fundamental para el éxito del mismo: “*success of OBIS depends on the active role of the OBIS nodes*”, “*international collaboration is essential to achieve the objectives of OBIS*” (IODE Steering Group for OBIS (SG-OBIS), 2017).

Los resultados del AC permiten identificar actividades de colaboración muy variadas, desde la participación voluntaria de los integrantes del Sistema en tareas propuestas por el SG o por alguno de los nodos, hasta colaboración con otros sistemas de información o proyectos con objetivos comunes o bien objetivos complementarios a los de OBIS (*to encourage the sharing, to document new datasets, sharing planned and current activities*).

La colaboración con otros editores y proveedores de datos es vista de forma muy positiva. Las expresiones acuerdos de colaboración (*collaboration agreement, letter agreement*), colaboración con (*collaboration with*) y más involucramiento con otros programas (*more involved with other programmes*) se repiten como un punto de interés constante en los documentos de las reuniones, siempre asociados a expresiones positivas como fomentar el amplio intercambio de datos o alcanzar formas más eficientes de combinar datos (*fostering wider data sharing, more efficient ways to combine data*).

Las actividades orientadas al mantenimiento y crecimiento del sistema que aparecen mencionadas con más frecuencia son las relacionadas con fortalecer la red de proveedores de datos (*strengthen the network of data providers*). Hay varias actividades colaborativas en curso que apuntan a este objetivo. Las que más ocupan la atención están centradas en desarrollar las capacidades técnicas del sistema (*new online tools and*

*applications, new OBIS infrastructure, new OBIS website, technology stack, real-time data integration, upscale the system, real-time analytics capability, prototype harvester, OBIS tools package based on R for data cleaning*) y en desarrollar nuevos productos y servicios (*develop information products and services, OBIS data exploration portal, OBIS R-package, more indicators and products*).

Las actividades colaborativas para brindar capacitación son otro de los focos que ocupan la atención del SG. Las frecuencias (F) de palabras relacionadas a actividades de capacitación en los reportes de las reuniones anuales de 2016 (*teaching* F: 1, *training* F: 63, *teacher* F: 20, *train* F: 11, *guidelines* F: 20, *manual* F: 20, *tutorial* F: 3, *lectures* F: 2) y 2017 (*training* F: 60, *lessons* F: 3, *teacher* F: 8, *train* F: 13, *guidelines* F: 4, *manual* F: 6, *tutorial* F: 4) son un claro ejemplo de la importancia que tienen estas actividades para el sistema. Las expresiones relacionadas con estas actividades son: *teaching material, training course/s, provide training, training workshop, workshops, training activities, training, training strategy, training needs, lessons, lectures, tutorial videos, tutorials, online training materials, Ocean Teacher, webinars, Train, training opportunities, OBIS Manual, share information about local training, share links to their guidelines*.

La capacitación se brinda en varios niveles, adaptada a los diferentes grupos de interés (gestores de datos, científicos, académicos, proveedores de datos, profesionales). Los temas prioritarios para la capacitación son: uso de estándares y mejores prácticas, acceso, procesamiento, control de calidad y manejo de diferentes tipos de datos.

*“ ... training the OBIS network of data managers and scientists in data standards and best practices remains a top priority, in order to increase data quality and ensure equitable participation of all Member States in OBIS”* (IODE Steering Group for OBIS (SG-OBIS), 2016).

Los nodos contribuyen de diferentes maneras con actividades relacionadas a capacitación, por ejemplo, desarrollando módulos para aprendizaje en línea que luego son compartidos.

Las actividades de capacitación se posicionan cumpliendo un rol muy importante en asegurar la calidad de los datos, una preocupación que vuelve a aparecer en otras actividades del sistema destinadas a mejorar este aspecto.

Las actividades relacionadas con promover la colaboración que se mencionan en los documentos analizados tienen dos focos principales: el fomento de los vínculos y la comunicación y extensión.

Los vínculos de colaboración aparecen de forma implícita en diferentes acciones del sistema como establecer colaboraciones, expandir el intercambio de datos, desarrollar redes regionales o fortalecer los nodos locales (*set up a collaboration International Network, expand the data exchange network on marine biodiversity, development of a regional network, strengthen the local OBIS nodes*).

La comunicación y la extensión ocupan espacio importante dentro de las actividades del SG. El grupo de trabajo de comunicación y extensión es el que coordina esas actividades. Los miembros del grupo colaboran a través de diferentes canales, incluidas las redes sociales (*Newsletter, OBIS twitter account, OBIS Facebook page, LinkedIn group*) en difundir qué es OBIS, qué puede hacer y que no.

### **5.2.1. Factores que desalientan las actividades de compartir datos e información**

Los aspectos legales aparecen insertos en actividades como incrementar el escrutinio en cuanto a las formas de indicar que los datos son adecuados o inadecuados para determinados usos y vinculados a las tareas de actualización con licencias CC que aplican a los conjuntos de datos (*flag data appropriate or inappropriate, consultation with legal experts, datasets updated with CC-BY licenses*).

El factor baja participación de los nodos se manifiesta fundamentalmente con expresiones de preocupación por la falta de involucramiento activo de los nodos en el Sistema o en los Grupos de Trabajo (*concern over the lack of active engagement, node exit strategy, nodes that are no longer active, transition strategy for inactive nodes*).

Las dificultades técnicas son vistas como problemas que dificultan el procesamiento de los datos. Las dificultades en publicar datos en línea aparecen asociadas a nodos específicos y a diversas situaciones que van desde problemas técnicos hasta problemas en procedimientos que causan demoras en la publicación de los datos (*data harvesting*

*and indexing process is not modularized, datasets are processed in bulk, it causes delays for all).*

La falta de consistencia de los datos e información está muy vinculada a la preocupación por la calidad de los datos. El no uso del nombre científico (taxonómico) aceptado de las especies en los conjuntos de datos es un problema que requiere colaboración de los nodos para resolverlo (*non-matching names, dealing with the scientific names, names that do not match with the WoRMS*). De acuerdo a los resultados del AC, la calidad de los datos es el aspecto más recurrente luego de la capacitación y muy ligada a esta. Las expresiones y los términos más frecuentes referidos a este factor son: *quality control, quality controlled datasets, improve data quality, best available quality datasets, QC summaries, summary of the QC results, data quality issues, data errors corrected, node activities related to data quality, compliance, taxonomic catalogue, taxonomic QC, new OBIS QC, validation against WoRMS, QC reports, QC flag, non-matching names, taxon names that do not match, WoRMS taxon match, unsolvable names, OBIS standards, high quality.*

### **5.2.2. Factores propicios a las actividades de compartir datos e información**

La actualización de la variedad de herramientas, productos y servicios que el sistema OBIS puede desarrollar y proveer es una preocupación constante y las expresiones para referirse a este tema así lo corroboran: *data enhancement tools, new website will be multilingual, processing and visualizing OBIS data using R package, development of an OBIS R-package, new OBIS portal interface, building of the OBIS database and information products, new technology stack, increase the web performance, mapping server, greater future collaboration on software development.*

Los nodos mantienen un relacionamiento (*data sharing relationship*) que hace posible el compartir datos y se sustenta en la compatibilidad para cosechar de conjuntos de datos que poseen sus sistemas.

La compatibilidad de los formatos es un tema muy ligado a los estándares que se utilicen. Es un aspecto muy mencionado en los documentos analizados, pues el sistema está desarrollando una extensión del estándar Dwc denominada OBIS-ENV-DATA

que integra datos biogeográficos con datos ambientales. Las referencias a la implementación de OBIS-ENV-DATA son numerosas (*full and timely implementation, 7 required terms from the Darwin Core, detailed specifications, support data types beyond species occurrences, OBIS Darwin Core Manual, compatibility with [...]*).

En cuanto a las políticas de datos, la política actual fue adoptada en la reunión del SG de 2015. Además, el reconocimiento de la necesidad de desarrollar guías y procedimientos para gestionar y compartir conjuntos de datos está muy presente en los documentos de las dos reuniones analizadas (*procedures and guidelines for managing and sharing datasets*).

El uso de estándares es uno de los temas que más se mencionan, juntamente con la capacitación y la calidad de los datos. Aparecen variadas expresiones relacionadas con estándares en los documentos analizados: *fitness for use of marine biodiversity data, OBIS-ENV-DATA standard, Ocean Data Standards, track and resolve issues related to OBIS-ENV-DATA, test OBIS-ENV-DATA dataset, assisting nodes with using the OBIS-ENV-DATA standard, Data transformation into the Darwin Core*.

### **5.3. Entrevistas a coordinadores de nodos OBIS**

De los 33 nodos existentes en abril de 2018, seis coordinadores han aceptado concedernos la entrevista. Las preguntas, primero que nada, buscan corroborar las principales prácticas de datos y los factores que según el resultado de la encuesta condicionan más esas prácticas, y en segundo lugar buscan profundizar en los motivos por los cuales se producen esos comportamientos a través de la opinión de los entrevistados.

#### **5.3.1. Principales problemas al compartir datos**

La opinión de la mayoría de los entrevistados cuando enumeran los problemas al compartir datos es coincidente con los resultados de la encuesta en varios aspectos, pero también aparecen en las entrevistas aspectos novedosos que en la encuesta no han sido mencionados.

Dentro de los aspectos coincidentes están la importancia del uso de estándares o vocabularios estandarizados (E01), la existencia de políticas o planes de manejo de datos (E02), y el tiempo requerido para preparar los datos y hacerlos disponibles (E01, E03).

La existencia de políticas de gestión de datos es fundamental para E02. Este entrevistado es de la opinión que compartir datos en cualquier entorno requiere un enfoque inequívoco, bien planificado y sólido, y esto también aplica a los estudios sobre biodiversidad marina. Allí radica la importancia de emplear un plan completo de gestión de datos.

Respecto a los motivos por los cuales el tiempo requerido para completar el proceso de compartir datos es referido específicamente como problema, E01 opina que “la carga [de datos] en sistemas especializados usualmente no está contabilizada dentro de los tiempos de los proyectos, por lo que representa una tarea adicional que muchos investigadores no están dispuestos a asumir” (mi traducción). También E03 coincide en este aspecto. Refiriéndose a la preparación de los metadatos, menciona que la descripción incluye muchas características diferentes, que requieren tiempo y una capacitación adecuada para llevarla a cabo. En muchos casos, los metadatos son muy escasos debido a diferentes razones, dentro de las cuales especifica la falta de información, la falta de comunicación con el autor de los datos, o simplemente pereza en completar esos metadatos.

El tiempo que transcurre entre la recolección y la liberación de los datos aparece mencionado por un entrevistado (E01) como problemático y también estuvo presente en las respuestas a preguntas abiertas de la encuesta (R016).

Otro factor que aparece en encuestas (R016, R087, R112) y entrevista (E06) es la escasa cultura relacionada a compartir datos. En las encuestas se mencionaba no querer compartir datos por tratar los datos de investigación con recelo gestionándolos como propiedad personal. En la entrevista, E06 expresa que “El científico considera que los conjuntos de datos son propiedad privada y el biólogo tiene una tradición débil respecto a compartir datos” (mi traducción).

Además, el trabajo que implica para los biólogos marinos recolectar los conjuntos de datos, tiene para E06 consecuencias al compartirlos.

A menudo, un conjunto de datos biológicos (identificación de especies) representa muchas horas de trabajo para un individuo, a diferencia de los datos físicos que frecuentemente son recogidos de forma semiautomática (E06) (mi traducción).

Dentro de los factores que no aparecen en la encuesta y sí en las entrevistas, está la falta de capacitación en el manejo de datos por parte de los investigadores (tanto en lo que refiere a metadatos como al proceso de carga de datos a los sistemas). E04 opina que el único problema en términos de facilitarle a los proveedores el compartir datos puede deberse a veces a complicaciones y dificultades para comprender *Darwin Core* y el proceso de gestión y envío de los datos. Esta falta de capacitación se puede apreciar, en opinión de la entrevistada, en las generaciones más antiguas de científicos o bien en los taxónomos que no estén muy familiarizados con la gestión y el intercambio de datos.

Para E05, el único problema es cuando los datos no se comparten, o se comparten parcialmente. Debido a esto último, la historia completa que cuentan esos datos no puede ser totalmente evaluada/analizada y allí radica el problema para los potenciales usuarios de esos datos.

E04 enfoca su respuesta en los datos de acceso abierto. No ve problemas en compartir datos en bases de acceso abierto siempre y cuando exista un acuerdo entre el proveedor de los datos y la base.

### **5.3.2. Retención de datos**

La pregunta sobre los tipos de datos más probables de ser retenidos suscitó tres tipos de respuesta: los que responden tipificando los datos (E01, E03, E04), los que opinan que el problema de la retención no está en el tipo de datos sino en cómo se sienta (“*feelings regarding data sharing*”) quien genera o recoge esos datos respecto a la posibilidad de perjuicios que alguien podría ocasionarle (E05), y los que opinan que la causa más probable de retención de datos se debe a fallas en el proceso de hacer disponibles esos datos. “No es debido a la falta de voluntad, sino a la falta de conciencia y la subestimación del cómo hacer los datos disponibles” (E02) (mi traducción).

Las tipologías de datos más probables de ser retenidos según los diferentes entrevistados son: datos relacionados con investigaciones en curso/ trabajos inéditos (E02, E04), datos que están siendo comercializados (ejemplo: datos para consultorías ambientales) y ciertos datos sensibles (ejemplo: datos de pesquerías) (E01), datos recolectados recientemente (E03), datos de cruceros en curso, datos de científicos de la vieja generación, datos en monografías y archivos impresos antiguos, datos en otros idiomas diferentes del inglés, datos relacionados con programas de investigación realizados en colaboración (E04), datos de proyectos de investigación llevados a cabo por investigadores individuales o en instituciones con una política de no compartir datos (E06).

Para E02 es claro que hay una tendencia en las personas a retener los datos de su investigación hasta que finalmente se publique el artículo o libro. La falla más grande de este proceso, en opinión de E02, es que los conjuntos de datos en los que se basan esos documentos, artículos o libros, rara vez se publican de manera útil o significativa después de la publicación de la investigación asociada. Esos datos a menudo se pierden o nunca se ponen a disposición del público. “Tengo varios ejemplos de investigadores que han realizado un trabajo excelente, pero sus datos se perdieron porque nunca se procesaron y se pusieron a disposición” (E02) (mi traducción).

En opinión de E05, el hecho de que los datos sean compartidos o publicados no radica en el tipo de datos, sino que está totalmente relacionado con los sentimientos del investigador principal respecto a compartir datos. Es una posición más centrada en los temores por posibles daños y consecuencias de que otras personas perjudiquen su investigación.

### **5.3.3. Prácticas para obtener datos**

La pregunta número 3 de la pauta de entrevista está orientada a profundizar un poco más en la descripción de las prácticas de los investigadores para obtener datos ya generados en otras investigaciones sobre biodiversidad marina. Las respuestas van desde menciones a herramientas y sitios específicos a los que suelen recurrir, incluyendo motores de búsqueda en Internet (aparece mencionado *Google Search*),

paquetes R y herramientas específicas del sistema OBIS (E01, E03, E04), hasta aspectos más ligados a los formatos que suelen preferir.

E02 manifiesta abiertamente que “¡Encontrar datos es a menudo una experiencia frustrante y desconcertante!” (mi traducción). Esto en su opinión se debe a que comúnmente los datos disponibles están en formatos diversos y presentaciones muy variables. Comenta a modo de ejemplo que en muchas ocasiones los datos se proporcionan en formatos inutilizables, porque el software que se utilizó para generarlos está desactualizado o ya no funciona. Pasa a menudo que los formatos que se utilizan son muy específicos de una disciplina en particular, y eso también limita que los datos sean utilizables por todos. Apuntando al tema de la usabilidad de los datos, afirma que el formato de salida de sus datos es .CSV (valores separados por comas), para garantizar que cualquier persona pueda usarlos de varias maneras.

Las estrategias para localizar datos que se mencionan en las entrevistas comprenden desde utilizar términos amplios en motores de búsqueda generales de Internet para encontrar nuevas fuentes de datos hasta aprender el funcionamiento de diversas plataformas especializadas. Para E02, el enfoque básico es aprender sobre las muchas plataformas existentes (como OBIS) que brindan posibilidades de búsqueda y acceso a conjuntos de datos. En su opinión, es necesario aprender a usar esas plataformas antes de que se pueda hacer un uso real de los datos.

E04 explica que su metodología consiste en buscar continuamente utilizando términos amplios en Google, para encontrar fuentes nuevas y confiables. Comenta que es importante utilizar varios motores de búsqueda ya que todos tienen capacidades variables para ubicar lo que está disponible. También recurre a archivos, libros y artículos publicados para localizar datos que a menudo están contenidos en forma de tablas o de información textual. Esto último coincide con otras menciones a estas mismas prácticas que aparecen en los resultados de la encuesta.

#### **5.4. Discusión de los resultados de las tres técnicas aplicadas**

Se presenta a continuación una síntesis del análisis de los resultados de la encuesta, las entrevistas y el análisis de contenido. La discusión está planteada en el mismo orden que se propusieron los objetivos específicos de la investigación y enfatiza en aspectos

que están presentes de forma coincidente en los resultados de la aplicación de las tres técnicas.

#### **5.4.1. Motivos y factores favorables a compartir datos**

Los motivos para compartir datos que más opiniones a favor obtuvieron fueron “Ayuda a evitar la duplicación de esfuerzos en la recogida de datos”, “Mejora la eficiencia en el uso de los fondos” y “es útil para entrenar nuevos investigadores”.

En opinión de los encuestados, el factor que más favorece al momento de compartir datos de investigación sobre biodiversidad marina es la disponibilidad de un repositorio, el segundo es la existencia de una política de intercambio de datos y el tercero la utilización de estándares de biodiversidad para el intercambio de datos.

La disponibilidad de un repositorio donde alojar los datos es muy apreciada por investigadores de todas las edades, resultando los menores de 35 años los que más relevancia le dan a este factor, con un 56% de investigadores asistentes que consideran absolutamente esencial disponer de un repositorio de datos.

La existencia de una política de intercambio de datos es más apreciada en los grupos etarios de 35-44 y 45-54. Las políticas y los planes de manejo de datos aparecen mencionados en las entrevistas como importantes y también es un tema relevante en las reuniones anuales del Grupo Directivo de OBIS (SG-OBIS), resultando muy ligado a la calidad de los datos.

La utilización de estándares de biodiversidad es muy apreciada por los grupos de edades entre 35 y 64 años. El considerar muy importante la utilización de estándares aumenta con la edad del entrevistado. Los que más importancia otorgan al empleo de estándares para datos sobre biodiversidad son los investigadores asistentes, los posdoctorales y los integrantes de equipo de investigación. La importancia del uso de estándares o vocabularios estandarizados también aparece en las entrevistas (E01) y en es el tercer tema más mencionado en las reuniones del SG-OBIS, luego de la capacitación y de la calidad de los datos, de acuerdo a los resultados del análisis de contenido de los documentos dichas reuniones.

### **5.4.2. Motivos y factores que desalientan compartir datos**

Dentro de los motivos para no compartir (retener) datos, aparece en primer lugar el no tener tiempo suficiente para formatear y preparar los datos para compartirlos, luego motivos de derechos de autor, y en tercer lugar porque los patrocinadores de la investigación no aceptan hacer disponibles los datos.

En cuanto a los factores que desalientan compartir datos, las preocupaciones sobre el uso inadecuado de los datos es la primera causa que desalienta compartir datos, la segunda es el trabajo adicional requerido para poner los datos a disposición en repositorios o centros de datos, y la tercera el riesgo de perder oportunidades de publicación.

Los factores desfavorables fueron los más mencionados en la pregunta abierta de la encuesta: la confianza es el factor más mencionado muy vinculada a experiencias previas de apropiación indebida de los datos. Otro factor que se menciona repetidas veces son las actitudes de resistencia a compartir datos.

La preocupación por el uso inadecuado es menor a medida que aumenta la edad hasta los 54 años, siendo los menos preocupados por este factor los investigadores de entre 55-64 años. En el grupo de 65 y más el 40% le atribuye mucha importancia al posible uso inadecuado de los datos. Los más preocupados por este factor son los investigadores asistentes y los integrantes de equipo de investigación. Esta preocupación también aparece en las entrevistas, donde se mencionan sentimientos de temor o desconfianza frente a la posibilidad de perjuicios para quien comparte datos (E05).

El trabajo adicional de preparar los datos para hacerlos disponibles (ejemplo: ser depositados en un repositorio) es un factor que resulta muy importante en todos los grupos etarios. Los investigadores asistentes son los que mayor importancia le atribuyen, seguidos por los integrantes de equipo de investigación. Preparar los datos para hacerlos disponibles también es mencionado como problemático por dos de los coordinadores de nodo entrevistados (E01, E03), siendo vivenciado como una carga adicional para el investigador.

Otros factores desfavorables que aparecen en el AC, las entrevista y la encuesta son:

- La falta de capacitación en el manejo de datos por parte de los investigadores (tanto en lo que refiere a metadatos como al proceso de carga de datos a los sistemas). E02 menciona en la entrevista la falta de conciencia y la subestimación del cómo hacer los datos disponibles. Del resultado del AC podemos concluir que la prioridad que se le da a la capacitación por parte del sistema OBIS es alta, pues se quiere mejorar justamente este aspecto.
- La no existencia de un acuerdo entre proveedor de los datos y la base / repositorio en aspectos de licencia de uso aparece como problemático en la entrevista y en el AC.

### **5.4.3. Prácticas para obtener datos**

Buscar en línea resultó ser la práctica más utilizada para conseguir datos, la segunda, las solicitudes personales a los colegas y la tercera buscar en un repositorio central o centro de datos.

La utilización de búsqueda en línea en motores de búsqueda de Internet como estrategia única para localizar datos muestra una tendencia decreciente con la edad. Los mayores de 35 suelen combinar estrategias en sus prácticas de localización de datos.

La solicitud de conjuntos de datos a colegas es una práctica frecuentemente utilizada entre los menores de 45. Investigadores doctorales, investigadores asistentes y líderes de equipo de investigación, en ese orden, resultaron ser los que más recurren a esta práctica.

Todos los grupos etarios recurren a repositorios centrales o centros de datos como práctica para buscar datos, siendo los que menos los utilizan los mayores de 65 años. Los investigadores doctorales son los que más recurren a buscar en repositorios centrales o centros de datos, seguidos por coordinadores de investigación / directores de proyectos y líderes de equipo de investigación.

En cuanto a las condiciones para compartir datos, el 96.7% de los encuestados pone alguna condición. Las más mencionadas fueron, en primer lugar, el permiso y/o el reconocimiento de la autoría, en segundo lugar, que los datos no sean utilizados con

finés comerciales y en tercer lugar solicitar coautoría en publicaciones que surjan de los datos.

## Capítulo 6. Conclusiones y proyecciones

### 6.1. Conclusiones

Estas conclusiones, cuestionan sobre todo el enfoque individualista respecto del usuario de información que impera en los enfoques tradicionales, y lo hacen desde una postura construccionista social. El paradigma construccionista pone énfasis en las prácticas sociales, o sea las actividades concretas y situadas de las personas que interactúan. Al centrarnos en el estudio de las prácticas de información, el análisis hace énfasis en lo social, por lo tanto, el foco de las conclusiones de esta investigación se sitúa en el comportamiento de las personas interactuando con la información y los datos en su contexto social.

Centrarse en el estudio de cómo los datos se comparten en la comunidad OBIS ha permitido identificar y explorar prácticas, así como también descubrir puntos de vista coincidentes respecto de los motivos más habituales que suscitan ciertas prácticas de información, o más específicamente “prácticas de datos” en el área temática de la biodiversidad marina.

Estudiar estas actividades socialmente mediadas respecto a la información y los datos en ambientes de investigación, ya sea que estas se realicen en el ambiente material (físico) o en el digital, es avanzar en comprender los comportamientos informativos actuales en ambientes de colaboración. Los factores estudiados, tanto los dados por el contexto local como por el global juegan un papel importante en las prácticas de los investigadores al compartir datos.

En el contexto de un sistema de información global de datos sobre biodiversidad marina como lo es OBIS, era de esperar que la receptividad a la encuesta y las entrevistas presentara diferencias según el área geográfica. Investigadores de Europa y Sudamérica fueron los que más respondieron, con un porcentaje un poco mayor de hombres (61%) que de mujeres participando de la encuesta, y un promedio de edad de 47 años para los investigadores encuestados.

La problemática planteada en este estudio ha resultado de mucho interés a los investigadores con puestos de trabajo en niveles de dirección y conducción: directores

de proyectos de investigación, coordinadores de investigación y líderes de equipos de investigación.

Las motivaciones expresadas por los investigadores relacionadas con compartir datos sobre biodiversidad marina permiten llegar a interesantes conclusiones sobre el comportamiento informativo colaborativo respecto a los datos de investigación en este ámbito, básicamente porque las motivaciones tienen que ver con el por qué compartir datos e información sobre los datos.

La posición en el equipo de investigación incide de diferentes maneras en los comportamientos al compartir datos según las prácticas que se consideren. La posición en el equipo determina ciertos roles que tienen que ver con las relaciones sociales dentro de ese grupo humano. Como se puede apreciar en análisis de los resultados de esta investigación, las opiniones, comentarios y reflexiones respecto de los motivos y los factores (favorables o desfavorables) sobre compartir datos surgen de la experiencia personal, pero además son brindadas desde la perspectiva socio-cultural del integrante del equipo que brinda su opinión o comentario.

En primer lugar se presentan las conclusiones sobre motivaciones y factores propicios o favorables a compartir datos, y en segundo lugar las conclusiones sobre los motivos y factores que desalientan o no favorecen compartir datos, para luego pasar a concluir respecto a las prácticas de datos más habituales y cómo varía la utilización de unas u otras, resultando en comportamientos informativos diferentes según edad y posición en el equipo de investigación.

### **6.1.1. Motivaciones y factores que propician compartir datos sobre biodiversidad marina**

Los motivos por los cuales la comunidad estudiada comparte datos aparecen muy claros en los resultados del análisis y permite concluir que los principales motivos para compartir datos de investigación sobre biodiversidad marina son prácticos, relacionados con la consecución de objetivos y con el logro de mayor eficiencia: evitar la duplicación de esfuerzos, mejorar la eficiencia en el uso de los fondos y entrenar nuevos investigadores. También aparecen motivos de tipo altruista como beneficiar a otros

investigadores (reutilización de los datos), beneficiar a la comunidad científica y alentar la colaboración, pero son los menos mencionados en la encuesta (9, 4 y 3 menciones respectivamente en un total de 123 encuestados).

La disponibilidad de repositorios de datos, la existencia de una política de intercambio de datos y la utilización de estándares de metadatos sobre biodiversidad son las principales influencias en comportamientos favorables a compartir datos. Estos tres factores se repiten en los resultados de las tres técnicas aplicadas en la investigación: encuestas, entrevistas y análisis de contenido.

El pertenecer a una red o un proyecto sobre biodiversidad marina es también un factor positivo, favorable a compartir datos, apareciendo en cuarto lugar en los resultados de la encuesta. El primer factor favorable (existencia de un repositorio de datos) coincide justamente con uno de los productos de información principales que proporciona el sistema OBIS: un repositorio central de datos. Los otros dos factores que resultaron en los primeros tres lugares (existencia de una política para intercambio de datos y utilización de estándares) están ambos entre los temas que el sistema OBIS considera importantes y así se recoge de los resultados del análisis de contenido de los documentos de las reuniones del Grupo Directivo de los años 2016 y 2017.

La capacitación que brinda el sistema a sus proveedores de datos está muy enfocada actualmente al uso de estándares y a mejorar la calidad de los datos, tanto en manejo y procesamiento de diferentes tipos de datos (*best practices*) como en el control de calidad de estos. El uso de estándares es uno de los temas que más se mencionan en los documentos de las reuniones. A su vez, está muy relacionado con otro de los factores favorables a compartir datos identificado en los resultados del AC: la compatibilidad de los datos.

Si bien en opinión de Costello y Wieczorek (2014, p. 68) existe una creciente presión de la comunidad científica, incluidos los organismos de financiación, las revistas y los pares, para que los autores publiquen datos sobre biodiversidad que hayan utilizado en artículos y otra literatura científica, en esta investigación ese factor no aparece como una causa principal para compartirlos.

### **6.1.2. Motivos y factores que desalientan compartir datos sobre biodiversidad marina**

Las actitudes individuales no favorables a compartir datos que aparecen en la encuesta, asociadas con la práctica de tratar los datos como propios, se indagaron posteriormente en las entrevistas. Son actitudes que se manifiestan en comportamientos de no compartir y de gestionar los datos como propiedad personal. Los motivos por los cuales se da esta práctica, según la opinión de los investigadores entrevistados, se relacionan con que los conjuntos de datos sobre biodiversidad marina representan muchas horas de trabajo para quienes los generan, a diferencia los datos de oceanografía física que se recogen generalmente de forma automatizada.

De acuerdo a los resultados del análisis de los datos recogidos en esta investigación, se concluye que los motivos más habituales por los cuales se retienen datos sobre biodiversidad marina son, en primer lugar, no tener tiempo suficiente para preparar y formatear los datos (datos relacionados con investigaciones en curso o recientemente completadas), en segundo lugar que los datos estén sujetos a derechos de autor y en tercer lugar porque los patrocinadores de la investigación no permiten que los datos estén disponibles (datos sensibles, datos que están siendo comercializados, datos generados por proyectos llevados a cabo en instituciones con una política de no compartir datos de investigación). En cuarto lugar, aparece el no contar con apoyo de la organización donde se trabaja para gestionar los datos generados. Estas dos últimas conclusiones son coincidentes además con uno de los resultados de la investigación de Kim y Burns (2016, p. 230) respecto a que el entorno institucional (incluidos los organismos que financian la investigación y las editoriales) es un factor importante determinando las actitudes y comportamientos al compartir datos.

Los factores más habituales incidiendo en comportamientos no favorables a compartir datos identificados en esta investigación son, en primer lugar las preocupaciones por el uso inadecuado de los datos, basadas principalmente en motivos que tienen que ver con un sentimiento de desconfianza originado por haber tenido experiencias previas de apropiación indebida luego de compartir datos, o bien por sentimientos de inseguridad del generador de los datos respecto a los perjuicios que podría ocasionarle el hecho de compartir los datos. El segundo factor más habitual es el trabajo requerido

para poner los datos a disposición en repositorios o centros de datos, y el tercero la posibilidad de perder oportunidades de publicación. Esto nos permite coincidir con Tenopir et al. (2011, p. e0021101) cuando aseguran que las "barreras" para compartir datos están profundamente arraigados en las prácticas y en la cultura del proceso de investigación y en los investigadores.

En el análisis de las preguntas abiertas de la encuesta, así como también en pasajes de las entrevistas, se encuentran opiniones coincidentes acerca de una falta de cultura de compartir datos sobre biodiversidad marina. La falta de cultura de compartir datos es percibida como un factor desfavorable, manifestándose además una necesidad no satisfecha de compartir datos de forma permanente más allá de la pertenencia a una red o sistema de información sobre biodiversidad marina. Aparecen menciones a la importancia de la reciprocidad como valor que debe formar parte de esa cultura de compartir datos, lo que da la razón a Bao y Bouthillier (2007, p. 12) cuando afirman que "compartir información es un tipo de comportamiento colaborativo basado en la comunidad, impulsado por beneficios mutuos". La reciprocidad es muy importante para una cultura donde compartir datos de forma permanente sea predominante.

El enfoque construccionista social destaca la dimensión social del usuario y las relaciones del usuario con su contexto histórico-cultural al estudiar sus prácticas de información. Eso hace que tener en cuenta los significados atribuidos por los sujetos durante las acciones de buscar, usar y compartir informaciones y datos en los espacios en los que se dan tales prácticas sea muy importante para el estudio de las prácticas de información.

Las políticas, normas y demás aspectos relacionados con gestión de los datos e información sobre los datos, forman parte de un contexto histórico-cultural que puede ser más o menos propicio a compartir datos. En el caso de la comunidad estudiada, subyace un comportamiento de recelo respecto de compartir datos, que según los resultados de este estudio se puede deber a distintas causas, como desconfianza de que los datos sean objeto de apropiación indebida o de utilización en una forma que no esté autorizada, inseguridad y temor por posibles perjuicios si los datos se comparten, temor a perder oportunidades de publicación, así como también a la falta de una cultura de compartir datos.

El trabajo adicional de preparar los datos para compartirlos es un factor que preocupa a todos los grupos etarios, siendo los investigadores asistentes e integrantes de equipo de investigación los que más importancia le dan a este factor. La falta de capacitación en el manejo de datos, metadatos y en el proceso de carga de los datos en los sistemas informáticos es otro de los factores detectados, el cual está muy asociado a la responsabilidad y el compromiso de preparar los datos para compartirlos. Además, la falta de capacitación también es reconocida por los investigadores como factor que disminuye la calidad de los datos que se comparten, y cuando la calidad es baja, se dificulta el aprovechamiento de los datos.

Para la gran mayoría de los investigadores en biodiversidad marina la importancia del uso de estándares para datos no es cuestionada. Sin embargo, el tiempo requerido para preparar los datos y sus metadatos si es percibido como un obstáculo a la hora de compartir datos. Preparar los datos para compartirlos puede convertirse en un trabajo adicional que requiere una capacitación adecuada para llevarlo a cabo. Ese tiempo requerido en muchos casos no está previsto dentro de la duración de los proyectos.

Se puede concluir luego del análisis de resultados de encuestas, entrevistas y análisis de contenido que el factor mala calidad de los datos influye de forma negativa en las actividades relacionadas a compartir datos sobre biodiversidad marina. Los aspectos en relación con la calidad que influyen de forma desfavorable al compartir datos van desde errores en los datos y duplicación de registros hasta paquetes de datos incompletos, o cuyos metadatos están incompletos.

### **6.1.3. Principales prácticas para compartir datos sobre biodiversidad marina**

Colaborar es un componente esencial de la forma de trabajar de los integrantes de OBIS, y esa forma de trabajo también se extiende a las prácticas al compartir datos e información sobre los datos.

En toda práctica de compartir datos hay alguien (persona, organización o institución) que los brinda y alguien que utiliza esos datos. Es de destacar que se encontraron actitudes diferentes cuando se ponen datos a disposición de otros que cuando se trata de conseguir datos generados por otros investigadores. Cuando comparten, la gran mayoría

pone alguna condición, generalmente reconocimiento de la autoría o restricciones al uso que se le puede dar a los datos. Cuando necesitan conseguir datos generados por otros, son varias las estrategias utilizadas, que derivan en diferentes prácticas.

Las principales prácticas de los investigadores en biodiversidad marina para obtener datos identificadas por esta investigación son: en primer lugar, buscar en línea, en segundo lugar, realizar solicitudes personales a colegas y en tercer lugar buscar en un repositorio central o centro de datos. Esas prácticas tienen estrategias asociadas como utilizar términos amplios en motores de búsqueda generales de Internet para encontrar nuevas fuentes de datos o aprender el funcionamiento de diversas plataformas especializadas.

Las prácticas al compartir datos se ven influidas por el tipo de datos que se comparte: hay datos que se comparten inmediatamente sin restricciones y otros datos más específicos que se comparten luego de publicado el trabajo de investigación, o no se comparten por diferentes motivos.

Existe una propensión a retener datos de investigación hasta que se publique el artículo o libro. Haciendo referencia a los antecedentes de investigaciones previas sobre compartir datos de investigación, esta es una práctica extendida en la comunidad científica y no solo entre los investigadores que estudian la biodiversidad marina.

Preocupa a los investigadores de la biodiversidad marina que luego de la publicación del artículo o del libro esos datos se pierdan a causa de que no se publiquen de manera útil. Las razones argumentadas están vinculadas a la usabilidad de los datos, pues las dificultades aparecen cuando los formatos utilizados son específicos de un determinado software o el software con el que se generaron ha quedado obsoleto. En consecuencia, no tener en cuenta la usabilidad de los datos tiene consecuencias negativas al momento de compartirlos y es un factor a tener muy en cuenta en la planificación y el diseño de productos y servicios de información que involucren datos.

#### **6.1.4. Edad y posición en el equipo de investigación y comportamiento al compartir datos**

A partir de los resultados obtenidos en el análisis cuantitativo de los datos de la encuesta, se puede concluir que la edad, así como también la posición en el equipo de investigación, son variables que resultan en preferencias por diferentes prácticas respecto a los datos, y en distintas percepciones sobre los factores que inciden al compartir datos sobre biodiversidad marina.

Buscar datos en línea es una práctica muy habitual entre los investigadores doctorales y frecuente entre los líderes de equipo de investigación. Las solicitudes personales a colegas son una práctica utilizada frecuentemente por los menores de 45, investigadores doctorales e investigadores asistentes. Los investigadores doctorales son los que más utilizan repositorios centrales o centros de datos para la búsqueda de conjuntos de datos, seguidos por los coordinadores de investigación / directores de proyectos y los líderes de equipo de investigación.

La disponibilidad de un repositorio donde alojar datos sobre biodiversidad marina es muy apreciada por investigadores de todas las edades, pero es el grupo de los menores de 35 años el que más relevancia le da a este factor. En cambio, la existencia de una política de intercambio de datos es más apreciada en los grupos etarios de 35 a 54 años y la utilización de estándares de biodiversidad entre los investigadores de 35 a 64 años.

El posible uso inadecuado de los datos al momento de compartirlos preocupa menos a los investigadores de más edad. El grupo de 55 a 64 años resultó ser el menos preocupado por el uso inadecuado de los datos de investigación y los más preocupados por este factor son los investigadores asistentes y los integrantes de equipo de investigación.

El factor perder oportunidades de publicación preocupa más al grupo etario de 35 a 44 años al momento de compartir datos, siendo los investigadores doctorales y postdoctorales los que más importancia le dan a este factor. En este sentido los resultados de esta investigación coinciden con las conclusiones a las que arribaron Piwowar y Vision (2013, p. e175) respecto a identificar como "barrera" la preocupación de que otros investigadores puedan realizar publicaciones a partir de los datos que se

compartan. Dicho de otra manera, la preocupación de los generadores de los datos por perder oportunidades de publicación y en consecuencia perder oportunidades competitivas, se convierte en un factor desfavorable al momento de compartir datos de investigación sobre biodiversidad marina, constituyendo en estos casos un motivo determinante de su comportamiento respecto a compartir los datos que generan.

## 6.2. Proyecciones

Está prevista la publicación de la tesis en el repositorio institucional Colibrí, en acceso abierto, bajo la licencia *Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International*.

Se continúa trabajando en el análisis de los datos recabados en las encuestas y entrevistas con el objetivo de elaborar un artículo a presentar para revisión en una revista arbitrada.

Además, se realizará la presentación de los resultados y conclusiones de esta investigación en conferencias y reuniones relacionadas con información y ciencias marinas: *Annual Conference of IAMSLIC*<sup>15</sup>, *ASFA*<sup>16</sup> *Board Meeting*, etcétera, así como también ponencias en congresos relacionados con prácticas de información y comportamiento humano informativo.

---

<sup>15</sup> IAMSLIC: *International Association of Aquatic and Marine Science Libraries and Information Centres*

<sup>16</sup> ASFA: *Aquatic Sciences and Fisheries Abstracts*

## Referencias Bibliográficas

- AnthroDPA Metadata Working Group. (2009). *Report of the AnthroDPA MetaData Working Group*. Retrieved from <http://anthrodatadpa.org/Media/AnthroDataDPAReport.pdf>
- Bao, X., & Bouthillier, F. (2007). Information Sharing: As a Type of Information Behavior. In *Proceedings of the Annual Conference of CAIS / Actes du congrès annuel de l'ACSI* (pp. 1–14). Retrieved from <https://www.cais-acsi.ca/ojs/index.php/cais/article/view/604>
- Bardin, L. (1986). *Análisis de contenido*. Madrid: Akal.
- Bates, M. J. (2005). An Introduction to Metatheories, Theories, and Models. In K. E. Fisher, S. Erdelez, & L. McKechnie (Eds.), *Theories of Information Behavior* (p. 431). Medford, NJ: Information Today. Retrieved from [https://is.muni.cz/el/1421/jaro2011/VIKBA12/um/Bates\\_2005\\_intro.pdf](https://is.muni.cz/el/1421/jaro2011/VIKBA12/um/Bates_2005_intro.pdf)
- Bernete, F. (2013). Análisis de contenido (cuantitativo y cualitativo). In A. L. Marín & A. Noboa (Eds.), *Conocer lo social: estrategias y técnicas de construcción y análisis de datos* (pp. 221–261). Madrid. Retrieved from <http://analisisdecontenido.pbworks.com/w/file/attach/110765119/Bernete2013b.pdf>
- Blumenthal, D., Campbell, E. G., Gokhale, M., Yucel, R., Clarridge, B., Hilgartner, S., & Holtzman, N. A. (2006). Data withholding in genetics and the other life sciences: prevalences and predictors. *Academic Medicine : Journal of the Association of American Medical Colleges*, 81(2), 137–145. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/16436574>
- Borgman, C. L. (2012). The conundrum of sharing research data. *Journal of the American Society for Information Science and Technology*, 63(6), 1059–1078. <https://doi.org/10.1002/asi.22634>
- Borgman, C. L., Wallis, J. C., & Mayernik, M. S. (2012). Who's Got the Data? Interdependencies in Science and Technology Collaborations. *Computer Supported Cooperative Work (CSCW)*, 21(6), 485–523. <https://doi.org/10.1007/s10606-012-9169-z>
- Byström, K., & Lloyd, A. (2012). Practice theory and work task performance: How are they related and how can they contribute to a study of information practices. *Proceedings of the American Society for Information Science and Technology*, 49(1), 1–5. <https://doi.org/10.1002/meet.14504901252>
- Chapman, A. D., & Grafton, O. (2008). *Guide to Best Practices for Generalising Primary Species-Occurrence Data*. Copenhagen: Global Biodiversity Information Facility. Retrieved from [https://assets.contentful.com/uo17ejk9rkwj/6ie7n89wYMA6IcGKyoqW2/46d527fcd192ac18ec6c0be909bb8f20/gbif\\_Sensitive\\_Data\\_guide\\_en\\_v1.pdf](https://assets.contentful.com/uo17ejk9rkwj/6ie7n89wYMA6IcGKyoqW2/46d527fcd192ac18ec6c0be909bb8f20/gbif_Sensitive_Data_guide_en_v1.pdf)

- 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 & Evolution*, 28(8), 454–461. <https://doi.org/10.1016/j.tree.2013.05.002>
- Costello, M. J., & Wieczorek, J. (2014). Best practice for biodiversity data management and publication. *Biological Conservation*, 173, 68–73. <https://doi.org/10.1016/j.biocon.2013.10.018>
- Cox, A. M. (2012). An exploration of the practice approach and its place in information science. *Journal of Information Science*, 38(2), 176–188. <https://doi.org/10.1177/0165551511435881>
- Cragin, M. H., Palmer, C. L., Carlson, J. R., & Witt, M. (2010). Data sharing, small science and institutional repositories. *Phil. Trans. R. Soc. A*, 368, 4023–4038. <https://doi.org/10.1098/rsta.2010.0165>
- Engeström, Y. (2001). Expansive learning at work: toward an activity theoretical reconceptualization. *Journal of Education and Work*, 14(1), 133–156.
- Hyldegård, J. (2009). Beyond the search process – Exploring group members’ information behavior in context. *Information Processing & Management*, 45(1), 142–158. <https://doi.org/10.1016/j.ipm.2008.05.007>
- IODE Steering Group for OBIS (SG-OBIS). (2016). *Fifth Session, 25-27 May 2016. Reports of Meetings of Experts and Equivalent Bodies, UNESCO (English)*. Retrieved from [https://www.iode.org/index.php?option=com\\_oe&task=viewDocumentRecord&ocID=18012](https://www.iode.org/index.php?option=com_oe&task=viewDocumentRecord&ocID=18012)
- IODE Steering Group for OBIS (SG-OBIS). (2017). *Sixth Session, 1-3 February 2017. Reports of Meetings of Experts and Equivalent Bodies, UNESCO (English)*. Retrieved from [https://www.iode.org/index.php?option=com\\_oe&task=viewDocumentRecord&ocID=19894](https://www.iode.org/index.php?option=com_oe&task=viewDocumentRecord&ocID=19894)
- Kim, Y., & Adler, M. (2015). Social scientists’ data sharing behaviors: Investigating the roles of individual motivations, institutional pressures, and data repositories. *International Journal of Information Management*, 35, 408–418. <https://doi.org/10.1016/j.ijinfomgt.2015.04.007>
- Kim, Y., & Burns, C. S. (2016). Norms of data sharing in biological sciences: The roles of metadata, data repository, and journal and funding requirements. *Journal of Information Science*, 42(2), 230–245. <https://doi.org/10.1177/0165551515592098>
- Kratz, J. E., & Strasser, C. (2015). Researcher perspectives on publication and peer review of data. *PLoS ONE*. <https://doi.org/10.1371/journal.pone.0117619>
- Krippendorff, K. (2008). *Content Analysis: an introduction to its methodology* (2nd Ed.). Sage.

- Nguyen, V. M., Brooks, J. L., Young, N., Lennox, R. J., Haddaway, N., Whoriskey, F. G., ... Cooke, S. J. (2017). To share or not to share in the emerging era of big data: perspectives from fish telemetry researchers on data sharing. *Canadian Journal of Fisheries and Aquatic Sciences*, 74(8), 1260–1274. <https://doi.org/10.1139/cjfas-2016-0261>
- Perez, M. C., & Sabelli, M. (2003). Uso de información electrónica por los académicos de la Universidad de la República (Uruguay). *Información, Cultura y Sociedad*, (9), 29–52. Retrieved from <http://www.scielo.org.ar/pdf/ics/n9/n9a03.pdf>
- Pilerot, O., & Limberg, L. (2011). Information sharing as a means to reach collective understanding: A study of design scholars' information practices. *Journal of Documentation*, 67(2), 312–333. <https://doi.org/10.1108/00220411111109494>
- Pinto, F. V. M., & Araújo, C. A. A. (2012). Contribuição ao campo de usuários da informação: em busca dos paradoxos das práticas informacionais. *TransInformação*, 24(3), 219–226. Retrieved from <http://www.scielo.br/pdf/tinf/v24n3/a06v24n3.pdf>
- Piwozar, H. A., & Chapman, W. W. (2008). A Review of Journal Policies for Sharing Research Data. In *Proceedings ELPUB 2008 Conference on Electronic Publishing. Toronto, Canada*. Retrieved from [proceedings.nature.com/documents/1700/version/1](http://proceedings.nature.com/documents/1700/version/1)
- Piwozar, H. A., Day, R. S., & Fridsma, D. B. (2007). Sharing Detailed Research Data Is Associated with Increased Citation Rate. *PLoS ONE*, 2(3). <https://doi.org/10.1371/journal.pone.0000308>
- Piwozar, H. A., & Vision, T. J. (2013). Data reuse and the open data citation advantage. *PeerJ*, 1, e175. <https://doi.org/10.7717/peerj.175>
- Poisot, T., Mounce, R., & Gravel, D. (2013). Moving toward a sustainable ecological science: don't let data go to waste! *Ideas in Ecology and Evolution*, 6(2). <https://doi.org/10.4033/iee.2013.6b.14.f>
- Poteri, E. (2007). *Collaborative information behaviour: a case study of two research groups*. Abo Akademi University.
- Reddy, M. C., & Jansen, B. J. (2008). A model for understanding collaborative information behavior in context: A study of two healthcare teams. *Information Processing and Management*, 44(1), 256–273. <https://doi.org/10.1016/j.ipm.2006.12.010>
- Renear, A. H., Sacchi, S., & Wickett, K. M. (2010). Definitions of dataset in the scientific and technical literature. *Proceedings of the ASIST Annual Meeting*, 47, 3–6. <https://doi.org/10.1002/meet.14504701240>
- Rodríguez, R. G. (2017). Aplicabilidad de la Teoría de la Actividad Histórico-Cultural en los estudios de comportamiento informacional. *Biblios*, (67), 69–83.

<https://doi.org/10.5195/biblios.2017.336>

- Sandoval Moya, J. (2010). Construcciónismo, conocimiento y realidad: una lectura crítica desde la Psicología Social. *Revista MAD*, (23), 31–37.  
<https://doi.org/10.5354/0718-0527.2010.13633>
- Savolainen, R. (2007). Information Behavior and Information Practice: Reviewing the “Umbrella Concepts” of Information-Seeking Studies. *The Library Quarterly*, 77(2), 109–132. <https://doi.org/10.1086/517840>
- Savolainen, R. (2008). *Everyday information practices: a social phenomenological perspective*. Lanham, MD: Scarecrow Press.
- Savolainen, R. (2012). Conceptualizing information need in context. *Information Research*, 17(4). Retrieved from <http://www.informationr.net/ir/17-4/paper534.html#.WmIL7qinHIU>
- Silvello, G. (2018). Theory and practice of data citation. *Journal of the Association for Information Science and Technology*, 69(1), 6–20.  
<https://doi.org/10.1002/asi.23917>
- Simons, N., Visser, K., & Searle, S. (2013). Growing institutional support for data citation: Results of a partnership between Griffith University and the Australian National Data Service. *D-Lib Magazine*, 19(11/12).  
<https://doi.org/10.1045/november2013-simons>
- Soulier, L., Tamine, L., & Bahsoun, W. (2014). On domain expertise-based roles in collaborative information retrieval. *Information Processing and Management*, 50(5), 752–774. <https://doi.org/10.1016/j.ipm.2014.04.002>
- Talja, S. (1997). Constituting “information” and “user” as research objects: a theory of knowledge formations as an alternative to the Information Man-Theory. In P. Vakkari, R. Savolainen, & B. Dervin (Eds.), *ISIC '96 Proceedings of an international conference on Information seeking in context Tampere, Finland* (pp. 67–80). London, UK: Taylor Graham Publishing.
- Talja, S. (2002). Information sharing in academic communities: Types and levels of collaboration in information seeking and use. *New Review of Information Behavior Research*, 3(1), 143–159.
- Talja, S. (2005). The Domain Analytic Approach to Scholar’s Information Practices. In *Theories of information behavior*. Information Today.
- Talja, S., & Hansen, P. (2006). Information Sharing. In *New Directions in Human Information Behavior* (pp. 113–134). Berlin/Heidelberg: Springer-Verlag.  
[https://doi.org/10.1007/1-4020-3670-1\\_7](https://doi.org/10.1007/1-4020-3670-1_7)
- Talja, S., & Nyce, J. M. (2015). The problem with problematic situations: Differences between practices, tasks, and situations as units of analysis. *Library and Information Science Research*, 37, 61–67.

<https://doi.org/10.1016/j.lisr.2014.06.005>

- Talja, S., Tuominen, K., & Savolainen, R. (2005). “Isms” in information science: constructivism, collectivism and constructionism. *Journal of Documentation*, 61(1), 79–101. <https://doi.org/10.1108/00220410510578023>
- Tenopir, C., Allard, S., Douglass, K., Aydinoglu, A. U., Wu, L., Read, E., ... Frame, M. (2011). Data sharing by scientists: Practices and perceptions. *PLoS ONE*. <https://doi.org/10.1371/journal.pone.0021101>
- Tenopir, C., Dalton, E. D., Allard, S., Frame, M., Pjesivac, I., Birch, B., ... Dorsett, K. (2015). Changes in Data Sharing and Data Reuse Practices and Perceptions among Scientists Worldwide. *PLOS ONE*, 10(8), e0134826. <https://doi.org/10.1371/journal.pone.0134826>
- Tuominen, K., & Savolainen, R. (1997). A social constructionist approach to the study of information use as discursive action. In P. Vakkari, R. Savolainen, & B. Dervin (Eds.), *ISIC 1 Conference Proceedings* (pp. 81–96). Retrieved from [http://www.informationr.net/isic/ISIC1996/96\\_Tuominen.pdf](http://www.informationr.net/isic/ISIC1996/96_Tuominen.pdf)
- Tuominen, K., Talja, S., & Savolainen, R. (2002). Discourse, cognition and reality: toward a social constructionist metatheory for library and information science. In H. Bruce (Ed.), *Emerging frameworks and methods. CoLIS 4 : proceedings of the Fourth International Conference on Conceptions of Library and Information Science, Seattle, WA, USA, July 21-25, 2002* (pp. 271–284). Washington DC: Libraries Unlimited.
- UNESCO. IODE. (2016). Manual. OBIS Nodes. Retrieved July 21, 2018, from <http://iobis.org/manual/nodes/>
- United Nations Environment Programme. (1992). Convention on Biological Diversity. UN. Retrieved from <https://www.cbd.int/doc/legal/cbd-en.pdf>
- Wallis, J. C., Rolando, E., & Borgman, C. L. (2013). If We Share Data, Will Anyone Use Them? Data Sharing and Reuse in the Long Tail of Science and Technology. *PLoS ONE*, 8(724), e67332–e120778. <https://doi.org/10.1371/journal.pone.0067332>
- Widén, G., & Hansen, P. (2012). Managing collaborative information sharing: bridging research on information culture and collaborative information behaviour. *Information Research*, 17(4). Retrieved from [http://www.informationr.net/ir/17-4/paper538.html#.WLWLV\\_nyvIU](http://www.informationr.net/ir/17-4/paper538.html#.WLWLV_nyvIU)
- Zhao, M., Yan, E., & Li, K. (2018). Data set mentions and citations: A content analysis of full-text publications. *Journal of the Association for Information Science and Technology*, 69(1), 32–46. <https://doi.org/10.1002/asi.23919>

## **Anexos**

## Anexo 1. Formulario de la encuesta

### Data sharing practices in marine biodiversity

Thank you in advance for your interest in this study. Your participation will contribute to the growth of knowledge on collaborative information behavior and information sharing practices.

Please indicate which of the following statements in the first person most apply to you.

#### 1. I am willing to share my data (with or without restrictions) because:

*Please choose one or several of the following options.*

- Helps avoiding duplicate data collection.
- Discourages fraud.
- Increases efficient use of funding.
- It is useful for training new researchers.
- Implies benefits for my career.
- Helps to identify errors in the data.
- Other:

#### 2. In my opinion, factors that positively influence marine biodiversity data sharing are:

	Not important at all	Of little importance	Of average importance	Very important	Absolutely essential
Being a member of a marine biodiversity project/network.	<input type="checkbox"/>				
Availability of a data repository.	<input type="checkbox"/>				

	Not important at all	Of little importance	Of average importance	Very important	Absolutely essential
Pressure to publish the data by funding agencies.	<input checked="" type="radio"/>				
Pressure to publish the data by journals.	<input checked="" type="radio"/>				
Existence of a data sharing policy.	<input checked="" type="radio"/>				
Usage of biodiversity standards for data exchange (i.e. Darwin Core, ABCD).	<input checked="" type="radio"/>				

### 3. In order to share my datasets ...

- I do not put any conditions.
- I agree to release or make my data available with conditions that restrict their use or distribution.

### 4. Regarding the use or distribution of my data I usually require that:

*Please choose one or several of the following options.*

- The data is not being used without permission and acknowledge of the author/s.
- The data is not to be used for commercial purposes.
- Any use includes my co-authorship on publications that arise from the data.
- Other:

**5. When I need to use data generated by other marine biodiversity scientists, I ...**

	Never	Rarely	Sometimes	Very often	Always
Do online searching (Internet Search Engines).	<input type="checkbox"/>				
Send a request to a colleague (personal requests).	<input type="checkbox"/>				
Search in my institution repository.	<input type="checkbox"/>				
Search in cloud-based repositories for research data (i.e. Mendeley Data).	<input type="checkbox"/>				
Search in a central repository or data center.	<input type="checkbox"/>				

**6. Other data resources or strategies that you use and may like to add:**

**7. When I had to withhold data in the past it was because:**

*Please choose one or several of the following options.*

- Data were copyrighted and cannot be freely shared.

- I didn't have support from my organization to manage my data.
- I didn't have enough time to format and prepare my data for sharing.
- I felt that sharing the data will decrease my own competitive advantage.
- My conclusions based on the data may be challenged if the data were available.
- De-identification of sensitive data was complex.
- The sponsors of my research did not agree to release the data.
- Other:

**8. Things that discourage me from sharing data with other researchers are:**

	Not important at all	Of little importance	Of average importance	Very important	Absolutely essential
The required additional work to submit content to repositories or data centres (creating metadata, etc.).	<input type="checkbox"/>				
The risk of losing publication opportunities.	<input type="checkbox"/>				
Worries about misuse of my data.	<input type="checkbox"/>				
Concerns about misunderstanding or misinterpretation of the data.	<input type="checkbox"/>				

**9. Please feel free to include any other comments you would like to add about your experience on data sharing and/or other factors not mentioned on this form that encourage or discourage you for sharing your data.**



**10. Region**

- Africa
- Asia
- Central America & The Caribbean
- Europe
- North America
- Oceania
- South America
- Other:

**11. Position**

- Team member
- Team leader
- Assistant researcher
- Doctoral researcher
- Postdoctoral researcher
- Research manager/project manager
- Other:

**12. Gender**

- Male
- Female

**13. Your age:**

## Anexo 2. Lista de Nodos OBIS

**Antarctica:** AntOBIS. Biodiversity and Ecosystems Data and Information Centre (BEDIC), Royal Belgian Institute of Natural Sciences.

**Arctic:** Arctic OBIS. Conservation of Arctic Flora and Fauna (CAFF), Iceland.

**Argentina:** ArOBIS. Centro Nacional Patagónico (CENPAT) CONICET.

**Australia:** OBIS Australia. Commonwealth Scientific and Industrial Research Organisation (CSIRO).

**Black Sea:** OBIS Black Sea. Ukrainian Scientific Centre of Ecology of the Sea.

**Canada:** OBIS Canada. Centre of Marine Biodiversity, Bedford Institute of Oceanography.

**Caribbean:** Caribbean OBIS. Universidad Simón Bolívar - Centro de Biodiversidad Marina.

**China:** OBIS China. Department of Marine Organism Taxonomy & Phylogeny, Institute of Oceanology, Chinese Academy of Sciences.

**Colombia:** OBIS Colombia. The Marine and Coastal Research Institute of Colombia (INVEMAR).

**CPPS - South Pacific:** OBIS CPPS. Comisión Permanente Pacifico Sur.

**Deep-Sea:** Deep-Sea OBIS. Senckenberg Research Institute, Germany and Natural History Museum & Marine Institute, Plymouth University.

**Europe:** EurOBIS. Vlaams Instituut voor de Zee (VLIZ).

**Fish Data:** FishBase.

**Harmful Algal Blooms:** OBIS HAB. IOC Science and Communication Centre on Harmful Algae.

**Historical Research:** Oceans Past Initiative (OPI). Global research network for marine historical research

**Indian Ocean:** IndOBIS. Centre for Marine Living Resources and the Environment, Ministry of Earth Sciences, India.

**iOBIS.** Intergovernmental Oceanographic Commission of UNESCO, IOC Project office for IODE, Oostende, Belgium.

**Japan:** OBIS Japan. Japan Agency for Marine-Earth Science and Technology.

**Korea:** KOBIS. Korea Institute of Ocean Science and Technology.

**Malaysia:** OBIS Malaysia. Institute of Oceanography and Environment (INOS) of the University of Malaysia Terengganu.

**Marine Megavertebrates:** OBIS SEAMAP. Marine Geospatial Ecology Lab, Nicholas School of the Environment, Duke University.

**Marine Microbes:** MicrOBIS. The Josephine Bay Paul Center, Marine Biological Laboratory, University of Chicago.

**Mediterranean:** MedOBIS. Hellenic Centre for Marine Research.

**Persian Gulf and Gulf of Oman:** PEGO-OBIS. Iranian National Institute for Oceanography (INIO).

**Senegal:** OBIS Senegal. Centre De Recherches Océanographiques Dakar-Thiaroye (C.R.O.D.T).

**South-East Asia:** SEAOBIS. ASEAN Centre for Biodiversity.

**South-West Pacific:** South Western Pacific OBIS. National Institute of Water & Atmospheric Research.

**Sub-Saharan Africa:** AfrOBIS. Southern African Data Centre for Oceanography.

**Tracking Data:** Ocean Tracking Network. Dalhousie University.

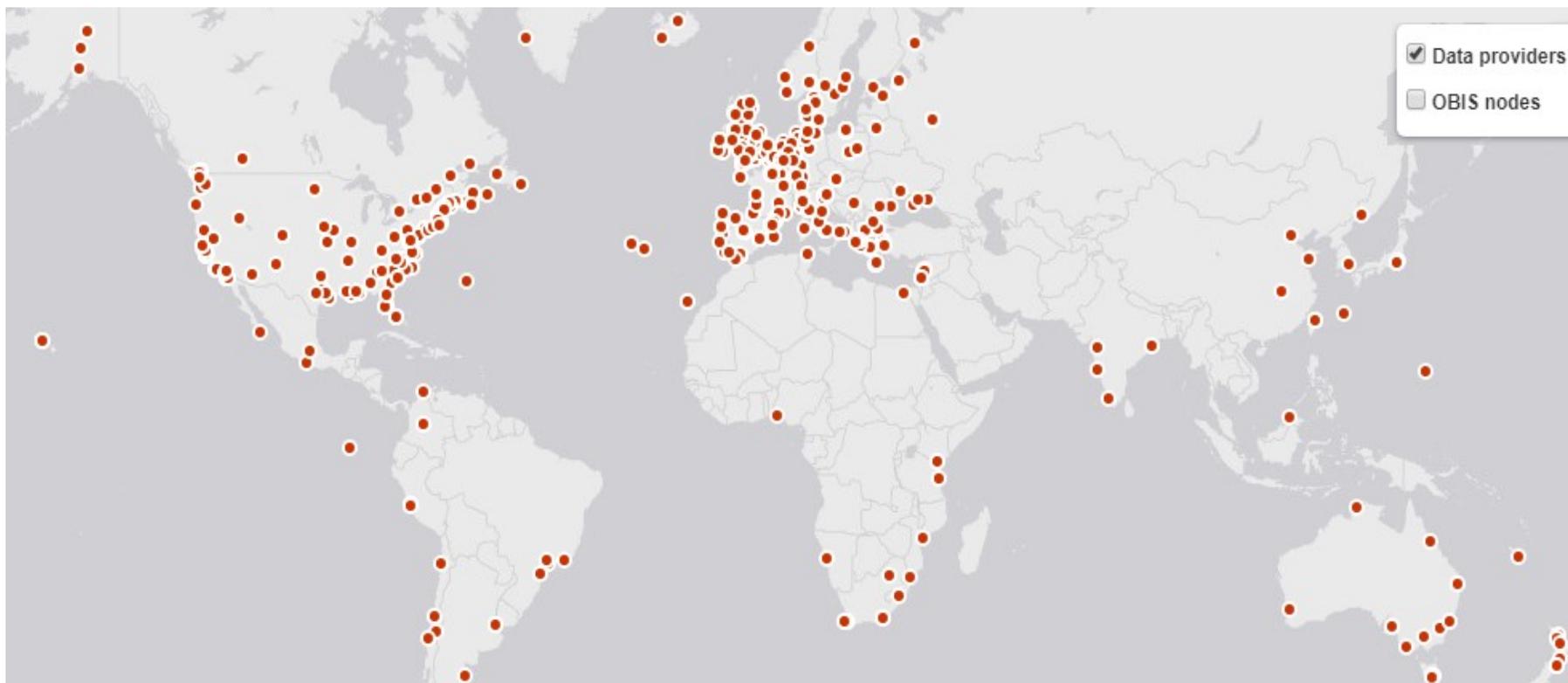
**Tropical and Subtropical Eastern South Pacific:** ESPOBIS. FONDAP COPAS.

**Tropical and Subtropical Western South Atlantic:** WSAOBIS. University of São Paulo.

**United Kingdom:** OBIS UK. The Marine Biological Association of the United Kingdom.

**USA:** OBIS-USA. US Geological Survey.

### Anexo 3. Mapa de instituciones proveedoras de datos a OBIS



Fuente: Sitio web de OBIS <<http://www.iobis.org/contact/>>

### Anexo 4. Datos recabados en la encuesta

Date	ID	1	2.1	2.2	2.3	2.4	2.5	2.6	3	4	5.1	5.2	5.3	5.4	5.5	6	7	8.1	8.2	8.3	8.4	9	10	11	12	13
2/12/18 19:06:06	R001	1.7; 1.1; 1.3; 1.4; 1.5; 1.6;	1	3	2	2	3	1	3.2	4.3;	3	2	0	1	1	7		3	1	1	2		10.5	11.6	12.1	47
2/13/18 9:55:04	R002	1.1; 1.3; 1.4; 1.6;	3	3	1	1	3	3	3.2	4.1; 4.2; 4.3;	2	3	3	1	1	7.1; 7.2;		3	1	3	2		10.8	11.7	12.1	39
2/13/18 10:45:26	R003	1.2; 1.3; 1.4; 1.6;	2	4	3	3	3	3	3.2	4.1;	3	3	4	3	4	7.2; 7.3; 7.6;		1	1	2	2		10.8	11.6	12.1	53
2/13/18 11:13:57	R004	1.1; 1.2; 1.4; 1.5; 1.6;	4	4	3	3	3	4	3.1	4.4;	3	0	0	0	4	7.1;		2	0	0	3		10.4	11.1	12.1	41
2/13/18 15:51:06	R005	1.1; 1.2; 1.4; 1.6;	3	4	3	3	3	2	3.1	4.1;	4	4	3	0	3	7.2; 7.3; 7.7;		4	3	4	4		10.2	11.4	12.2	34
2/13/18 17:34:32	R006	1.1; 1.3; 1.4;	3	4	2	2	3	4	3.2	4.1;	4	2	3	2	2	7.1; 7.3;		3	3	2	2		10.6	11.5	12.2	52
2/13/18 17:34:33	R007	1.1; 1.3; 1.6;	2	3	0	2	3	3	3.2	4.1; 4.3;	3	3	3	0	0	7.1; 7.7;		1	1	1	2		10.2	11.4	12.1	35
2/13/18 17:43:33	R008	1.1; 1.4; 1.5;	2	3	2	4	2	4	3.2	4.3;	4	3	0	1	1	7.2;		3	3	2	0		10.2	11.5	12.1	40
2/13/18 17:58:50	R009	1.1; 1.3; 1.4; 1.5; 1.6;	2	4	4	4		4	3.2	4.1;	4	0	0	0	0	7.3; 7.6;			0	3	4		10.4	11.7	12.1	26
2/13/18 19:10:49	R010	1.1;	1	2	2	0	3	2	3.2	4.1; 4.2; 4.3;	0	2	1	0	0	7.3; 7.4;		1	3	4	4		10.7	11.2	12.1	46
2/14/18 0:45:05	R011	1.1;	1	4	3	2	3	2	3.2	4.1;	3	1	2	2	2	7.2; 7.4;		3	2	2	1		10.2	11.5	12.1	46
2/14/18 1:31:56	R012	1.1;	2	4	3	3	3	3	3.2	4.3;	3	2	3	3	2	7.7;		3	2	3	3		10.2	11.1	12.2	40
2/14/18 7:37:09	R013	1.4;							3.2	4.1;	2	2	2	2	2								10.2	11.6	12.1	36
2/14/18 7:51:29	R014	1.1;	1	3	2	3	3	3	3.2	4.1; 4.2;	2	2	3	1	1	7.2; 7.3;		2	0	3	3		10.4	11.1	12.1	72
2/14/18 9:19:27	R015	1.1; 1.3; 1.5;	3	3	2	1	3		3.1	4.2;	3	2	1	1	3	7.2; 7.3;		1	1	1	1		10.7	11.1	12.1	62
2/14/18 11:21:53	R016	1.7; 1.1; 1.2; 1.3; 1.4; 1.5; 1.6	3	4	2	4	4	4	3.1	4.4; 4.1;	4	4	3	1	4	7.8;		0	0	2	2		10.5	11.7	12.2	28
2/14/18 11:43:09	R017	1.2;	3	4	3	3	3	3	3.2	4.1;	3	4	4	3	3	7.2; 7.7;		2	2		1		10.7	11.4	12.2	57
2/14/18 12:10:37	R018	1.2; 1.3;	1	3	3	2	3	4	3.2	4.1; 4.2; 4.3;	3	1	3	1	1	7.2;		1	2	2	2		10.7	11.6	12.2	41
2/14/18 12:33:24	R019	1.7; 1.1; 1.2; 1.5;	3	3		2	3	2	3.2		3	3	3	1	1	7.1; 7.4; 7.7;		3	4	3	3		10.5	11.4	12.1	24
2/14/18 16:28:02	R020	1.7; 1.3;	3	4	3	3	3	3	3.2	4.4; 4.1;	2	2	3	2	2	7.1; 7.3; 7.7;		2	3	4	3		10.6	11.6	12.1	48
2/14/18 18:19:09	R021	1.1; 1.3;	3	3	2	1	3	2	3.2	4.1; 4.2; 4.3;	3	2	1	0	3	7.4; 7.6; 7.7;		2	3	3	3		10.5	11.2	12.1	46
2/15/18 11:03:20	R022	1.7; 1.1; 1.4;	4	4	2	2	3	4	3.2	4.4; 4.1; 4.2; 4.3;	4	3	1	2	4	7.8; 7.2; 7.3;		3	2	3	4		10.8	11.7	12.2	39
2/14/18 19:00:42	R023	1.1; 1.2; 1.3;	3	3	2	2	3	3	3.1	4.1;	4	3	3	1	4	7.2; 7.5;		2	2	2	2		10.7	11.4	12.2	32
2/15/18 7:27:46	R024	1.3;	3	3	1	0	3	3	3.1	4.1; 4.2;	4	4	4	4	4	7.1;		2	0	0	0		10.7	11.2	12.1	50
2/15/18 8:07:51	R025	1.1; 1.2; 1.3; 1.4; 1.6;	2	3	4	4	2	2	3.1		4	2	0	0	2	7.3;		2	3	2	3		10.4	11.7	12.2	35
2/15/18 9:02:06	R026	1.1; 1.3; 1.5; 1.6;	2	3	0	0	4	2	3.1	4.1;	3	2	4	3	3	7.2; 7.7;		1	1	3	0		10.7	11.2	12.1	65
2/15/18 9:37:33	R027	1.1; 1.6;	0	3	3	3	3	3	3.2	4.1; 4.2;	2	1	1	3	3	7.8;		1	1	3	3		10.7	11.5	12.2	36
2/15/18 9:45:15	R028	1.7; 1.1; 1.3; 1.6;	3	3	2	2	4	3	3.1	4.1; 4.2;	3	2	1	3	3	7.2;		2	3	3	3		10.7	11.4	12.2	30
2/15/18 10:03:50	R029	1.6;	3	2	1	1	1	1	3.2	4.1; 4.3;	0	2	0	0	0	7.4;		2	4	3	2		10.7	11.1	12.1	44
2/15/18 10:15:12	R030	1.2; 1.5;	2	1	3	1	3	0	3.2	4.1;	1	3	0	0	0	7.3;		4	1	3	1		10.7	11.3	12.1	44
2/15/18 10:38:05	R031	1.7; 1.3;	3	4	2	2	4	2	3.2	4.1;	3	3	0	4	4	7.8; 7.2;		1	3	1	1		10.7	11.2	12.2	62
2/15/18 10:22:45	R032	1.1; 1.2; 1.3; 1.4; 1.5; 1.6;	3	4	2	2	3	3	3.2	4.4;	4	3	0	3	4	7.1; 7.3;		3	0	0	1		10.7	11.2	12.1	45

Date	ID	1	2.1	2.2	2.3	2.4	2.5	2.6	3	4	5.1	5.2	5.3	5.4	5.5	6	7	8.1	8.2	8.3	8.4	9	10	11	12	13
2/15/18 10:38:13	R033	1.7; 1.1; 1.3;	3	3	4	4	3	2	3.1	4.4;	3	1	0	3	3	7.8;	7	1	1	1	1		10.7	11.2		50
2/15/18 10:55:46	R034	1.1; 1.3; 1.4; 1.6;	3	3	2	2	3	3	3.2	4.1;	3	2	2	2	2	7.6; 7.7;	7	3	2	3	3		10.7	11.1	12.1	54
2/15/18 11:06:53	R035	1.3; 1.4;	3	4	3	3	4	3	3.2	4.1; 4.2;	3	1	0	2	2	7.2; 7.3;	7	1	1	3	2		10.7		12.1	64
2/15/18 13:15:23	R036	1.1;	4	4	2	2	2	3	3.2	4.1;	4	4	3	3	3	7.1;	7	4	4	4	3		10.6	11.6	12.2	56
2/16/18 8:24:29	R037	1.4; 1.6;	4	1	3	2	4	2	3.2	4.1; 4.2; 4.3;	1	3	3	1	2	7.2; 7.4; 7.7;	7	1	3	4	0		10.1	11.6	12.1	38
2/16/18 8:34:53	R038	1.4;	3	2	1	1	0	1	3.2	4.1; 4.3;	1	3	3	0	1	7.1; 7.3; 7.4;	7	1	1	3	3		10.8	11.6	12.1	38
2/16/18 9:25:56	R039	1.1; 1.3;	3	3	2	2	1	2	3.2	4.1; 4.2;	2	3	3	2	3	7.1;	7	2	3	3	3		10.3	11.6	12.2	53
2/16/18 10:36:54	R040	1.1; 1.3;	3	4	1	1	3	3	3.1	4.2;	3	4	2	0	0	7.4;	7	3	1	2	1		10.4		12.1	56
2/16/18 11:13:06	R041	1.7; 1.4;	2	2	0	0	0	0	3.2	4.1;	4	2	0	0	0	7.8;	7		3	2	1		10.1	11.7	12.1	69
2/16/18 11:26:05	R042	1.1; 1.2; 1.3; 1.4; 1.6;	2	4	3	1	3	3	3.2	4.2;	4	3	0	1	3	7.3; 7.5;	7	1	1	2	0		10.4	11.4	12.1	48
2/16/18 12:03:15	R043	1.1; 1.3; 1.4; 1.6;	3	3	1	2	3	1	3.2	4.1; 4.2;	2	3	3	0	3	7.3; 7.4;	7	3	2	2	2		10.4	11.2	12.1	46
2/16/18 12:53:44	R044	1.1; 1.3;	2	3	3	3	3	2	3.1	4.4;	2	1	1	2	2	7.4;	7	2	3	2	2		10.5	11.2	12.1	63
2/16/18 13:42:08	R045	1.1; 1.2; 1.3; 1.4; 1.5; 1.6;	2	3	2	3	3	3	3.2	4.1; 4.2;	3	3	3	3	3	7.1; 7.3;	7	1	1	2	1		10.3	11.1	12.1	39
2/16/18 15:24:54	R046	1.5;	3	3	1	3	2	2	3.2	4.1; 4.2; 4.3;	4	4	4	4	4	7.7;	7	0	0	2	2		10.7	11.4	12.1	38
2/16/18 15:48:49	R047	1.4;	3	3	4	4	3	3		4.1;	0	4	4	1	1	7.2;	7	0	4	3	3		10.1	11.5	12.1	43
2/16/18 17:09:13	R048	1.1; 1.3; 1.6;	3	4	4	3	4	4	3.1	4.1; 4.2;	4	2	2	2	3	7.1; 7.3; 7.7;	7	2	4	3	3		10.7	11.3	12.2	40
2/16/18 23:39:29	R049	1.1; 1.4; 1.6;	2	4	3	3	2	2	3.2	4.1; 4.2;	3	3	2	1	2	7.8;	7	2	4	4	4		10.6	11.2	12.2	39
2/17/18 10:30:14	R050	1.1; 1.3; 1.6;	3	4	2	3	4	4	3.2	4.1; 4.2; 4.3;	2	2	4	1	3	7.1; 7.3; 7.7;	7	1	2	2	1		10.7	11.6	12.2	45
2/18/18 4:34:08	R051	1.1; 1.3; 1.6;	2	3	4	3	3	3	3.2	4.1; 4.3;	2	3	2	0	1	7.7;	7	3	3	2	1		10.1	11.4	12.2	35
2/19/18 7:58:24	R052	1.1; 1.2; 1.3; 1.4; 1.5; 1.6;	3	3	2	2	3	3	3.2	4.1; 4.2;	4	2	2	2	2	7.1; 7.2; 7.3; 7.6;	7	3	2	3	3		10.4	11.3	12.2	27
2/19/18 8:04:42	R053	1.7; 1.1; 1.3; 1.6;	2	4	4	4	4	2	3.2	4.1; 4.2;	4	3	2	1	2	7.8; 7.1;	7	2	1	2	2		10.4	11.4	12.2	27
2/19/18 8:13:00	R054	1.3; 1.6;	3	4	2	2	3	4	3.2	4.1; 4.2;	1	2	3	3	3	7.1; 7.2; 7.6; 7.7;	7	3	2	3	3		10.4	11.6	12.2	54
2/19/18 8:08:00	R055	1.1; 1.3;	3	3	1	0	3	1	3.2	4.1;	2	3	1	2	3	7.1; 7.3; 7.4;	7	3	2	1	1		10.4	11.6	12.2	39
2/19/18 8:48:34	R056	1.1; 1.3;	4	4	1	0	2	2	3.1	4.1; 4.2;	4	4	0	0	4	7.3; 7.4;	7	4	4	1	0		10.4	11.2	12.1	47
2/19/18 8:59:37	R057	1.1; 1.2; 1.6;	3	2	3	3	2	3	3.1	4.1; 4.2;	3	3	1	2	2	7.8; 7.3; 7.4; 7.6; 7.7;	7	2	1	2	3		10.4	11.2	12.1	65
2/19/18 8:55:48	R058	1.2;	2	3	0	0	0	3	3.2	4.1; 4.3;	3	1	3	3	0	7.2; 7.4;	7	1	3	4	1		10.7	11.5	12.1	52
2/19/18 8:55:48	R059	1.2; 1.3;	0	4	3	3	2	2	3.1								7	2	2	1	2		10.4	11.6	12.1	59
2/19/18 9:48:46	R060	1.6;	3	3	2	2	4	3	3.2	4.1; 4.2;	3	4	3	3	3	7.1; 7.2; 7.7;	7	2	2	2	2		10.4	11.6	12.2	51
2/19/18 9:47:50	R061	1.1;	2	3	0	0	1	0	3.1	4.2;	2	3	1	1	1	7.3;	7	3	1	1	1		10.4	11.6	12.1	62
2/19/18 10:40:32	R062	1.1; 1.3; 1.6;	3	3	3	2	2	3	3.2	4.1; 4.2;	4	1	2	1	2	7.2; 7.3;	7	4	3	3	3		10.4	11.1	12.2	57
2/19/18 11:03:07	R063	1.7; 1.1; 1.5;	2	3	4	3	4	3	3.2	4.4; 4.1; 4.2;	3	4	0	1	2	7.1; 7.2; 7.4; 7.7;	7	3	2	1		10.4	11.7	12.2	61	
2/19/18 11:00:00	R064	1.1; 1.3; 1.4;	3	4	2	2	3	3	3.2	4.2;	0	2	2	2	2	7.1; 7.3;	7	3	1	3	3		10.4	11.2	12.2	55

Date	ID	1	2.1	2.2	2.3	2.4	2.5	2.6	3	4	5.1	5.2	5.3	5.4	5.5	6	7	8.1	8.2	8.3	8.4	9	10	11	12	13
2/19/18 11:13:58	R065	1.4;	2	4	0	0	4	4	3.1	4.1;4.2;	4	4	0	4	4		7.1;	0	0	0	0		10.7	11.6	12.1	51
2/19/18 12:21:40	R066	1.7;	3	4	4	2	3	4	3.1	4.4;	3	3	4	0	3		7.1;	4	4	4	1		10.4	11.3	12.1	37
2/19/18 12:57:57	R067		3	3	3	3	3	0	3.2	4.1;	3	1	0	3	1		7.3;7.4;	2	4	3	2		10.4	11.5	12.1	37
2/19/18 17:53:41	R068	1.1;1.3;	3	3	3	3	3	4	3.2	4.1;4.3;	1	3	0	1	0		7.4;	3	4	3	4		10.1	11.5	12.1	42
2/19/18 19:14:10	R069	1.1;1.4;1.6;	3	3	0	0	4	4	3.2	4.1;4.2;4.3;	3	2	2	3	2		7.3;7.4;	1	3	3	3		10.6	11.6	12.1	50
2/19/18 19:59:53	R070	1.1;1.3;1.5;	3	4	2	1	3	4	3.2	4.1;4.3;	1	3	1	2	1		7.2;7.4;	2	3	4	3		10.7	11.3	12.1	40
2/20/18 10:16:12	R071	1.1;1.3;1.6;	3	3	3	3	4	3	3.2	4.1;4.2;	4	4	4	4	4		7.1;	1	3	4	3		10.4	11.6	12.1	42
2/20/18 10:24:44	R072	1.1;1.3;1.6;	2	4	2	2	3	3	3.1	4.4;4.1;	2	3	1	1	2		7.8;	2	0	0	0		10.7	11.3	12.1	38
2/20/18 10:33:27	R073	1.1;1.3;	4	3	3	3	3	4	3.2	4.3;	1	3	4	4	1		7.1;7.2;7.7;	1	4	4	2		10.7	11.1	12.1	39
2/20/18 10:28:49	R074	1.1;1.3;1.6;	3	3	3	2	4	3	3.2	4.1;4.2;	1	3	2	1	2		7.2;	1	1	2	2		10.1	11.2	12.2	49
2/20/18 10:42:00	R075	1.3;	2	3	2	2	3	3	3.2	4.1;	3	2	2	1	2		7.3;7.4;	2	3	3	2		10.7	11.5	12.1	50
2/20/18 11:25:15	R076	1.1;1.3;1.4;	3	2	3	3	3	1	3.1	4.1;	3	2	2	2	3			2	2	2	1		10.4	11.2	12.1	45
2/20/18 12:24:39	R077	1.1;	2	4	1	1	3	3	3.2	4.1;	3	2	2	3	2		7.7;	2	2	2	1		10.7	11.2	12.1	51
2/20/18 14:20:43	R078	1.1;	1	2	1	1	2	2	3.2	4.1;4.2;4.3;	2	3	0	1	1		7.1;7.4;	2	3	4	2		10.7	11.2	12.1	41
2/20/18 16:08:16	R079	1.3;1.5;	3	4	3	2	3	3	3.2	4.1;4.2;	4	3	1	3	3		7.3;	2	3	1	1		10.6	11.4	12.2	57
2/20/18 17:19:54	R080	1.1;1.2;1.3;1.6;	2	4	2	0	4	4	3.1	4.2;	4	4	1	2	3		7.2;7.3;	4	0	2	3		10.8	11.6	12.2	56
2/21/18 1:10:55	R081	1.1;1.3;1.4;1.5;1.6;	3	3	3	3	3	3	3.2	4.1;4.2;4.3;	2	1	1	1	2		7.2;7.6;	2	2	3	2		10.2	11.6	12.1	54
2/21/18 3:22:42	R082	1.1;1.4;	4	2	2	1	4	3	3.2	4.1;4.2;4.3;	4	4	4	0	1		7.3;7.4;	4	4	3	3		10.1	11.6	12.2	46
2/21/18 8:00:57	R083	1.3;	1	3	3	0	3	2	3.2	4.1;4.2;4.3;	2	3	0	0	2		7.2;7.3;	3	2	4	4		10.4	11.5	12.2	43
2/21/18 8:24:12	R084	1.1;1.2;1.3;1.4;1.5;1.6;	3	3	3	3	3	4	3.1	4.3;	4	2	3	3	3		7.1;7.7;	3	2	2	2		10.2	11.6	12.1	40
2/21/18 9:04:52	R085	1.1;1.2;1.3;1.4;1.5;1.6;	3	2	3	3	4	3	3.2	4.1;4.2;	4	4	3	3	2		7.1;7.2;7.4;7.5;7.7;	3	4	3	2		10.7	11.3	12.1	42
2/21/18 8:57:34	R086	1.1;1.2;1.3;1.4;1.6;	3	3	0	0	2	2	3.1	4.1;4.2;	3	3	1	2	1		7.1;7.3;7.7;	3	2	1	1		10.1	11.6	12.1	49
2/21/18 10:00:55	R087	1.1;1.3;1.4;1.5;1.6;	3	3	2	2	4	4	3.2	4.1;4.2;4.3;	2	3	1	0	1		7.2;7.3;	3	4	3	4		10.5	11.1	12.1	78
2/21/18 10:56:50	R088	1.7;1.1;1.3;	2	2					3.1	4.4;	3	2	0	0	2		7.3;7.7;	2	0	0	0		10.5	11.2	12.1	48
2/21/18 11:38:13	R089	1.7;	2	3	2	2	4	2	3.2	4.1;	3	3	1	1	3		7.1;7.2;	0	2	1	1		10.7	11.2	12.2	32
2/21/18 13:44:49	R090	1.4;	4	3	3	3	4	3	3.1	4.2;	2	2	2	2	2		7.3;7.5;	3	3	3	3		10.3	11.2	12.1	57
2/21/18 17:48:21	R091	1.7;1.3;1.5;	3	2	1	1	3	3	3.2	4.1;4.2;	4	3	2	2	2		7.8;7.1;7.3;7.4;	2	2	4	4		10.4	11.6	12.1	48
2/22/18 2:11:30	R092	1.1;1.2;1.3;1.4;1.5;1.6;	3	3	2	2	3	4	3.2	4.1;	4	2	4	3	2		7.7;	3	3	3	2		10.2	11.4	12.2	32
2/22/18 4:04:28	R093	1.1;1.3;1.4;1.6;	4	4	3	3	3	4	3.2	4.1;4.2;	2	4	3	2	4		7.8;7.1;7.3;7.6;7.7;	2	2	2	1		10.1	11.1	12.2	45
2/22/18 7:33:57	R094	1.1;1.6;	2	3	4	3	2	3	3.1	4.2;	3	2	3	1	3		7.3;7.7;	3	4	1	2		10.2	11.6	12.1	48
2/22/18 10:49:20	R095	1.3;1.4;1.6;	2	4	3	3	2	4	3.2	4.1;							7.3;	3	0	2	3		10.5	11.3	12.2	61
2/22/18 11:57:35	R096	1.1;1.2;1.3;1.4;1.5;	2	2	1	0	1	1	3.2	4.1;4.3;	3	2	2	2	2			1	1	2	1		10.3	11.2	12.1	53

Date	ID	1	2.1	2.2	2.3	2.4	2.5	2.6	3	4	5.1	5.2	5.3	5.4	5.5	6	7	8.1	8.2	8.3	8.4	9	10	11	12	13
2/22/18 13:07:16	R097	1.2; 1.4;	3	2			3	2	3.1	4.1; 4.2;	4					7.1;	3						10.4	11.6	12.1	55
2/22/18 13:16:37	R098	1.1; 1.3; 1.4; 1.6;	2	4	2	2	3	3	3.2	4.1; 4.2;	3	3	3	3	3	7.1;	1	1	2	2			10.7	11.2	12.2	45
2/22/18 13:51:25	R099	1.1; 1.4;	2	3	2	2	2	4	3.2	4.1; 4.2;	3	3	1	3		7.1; 7.3; 7.7;	2	2	4	4			10.7	11.3	12.1	31
2/22/18 14:30:56	R100	1.3;	3	3	3	3	3	3	3.2	4.2; 4.3;	4	4	3	3	3	7.3;	2	1	2	1			10.3	11.5	12.2	41
2/22/18 14:47:12	R101	1.7;	2	3	2	3	2	2	3.1	4.4;	3	1	1	1	3	7.8; 7.3;	2	3	2	2			10.5	11.2	12.1	57
2/25/18 18:07:15	R102	1.7; 1.1; 1.3;	2	2	3	4	2	2	3.1	4.4;	4	2	1	0	4	7.8; 7.3;	3	1	1	1			10.4	11.2	12.1	42
2/26/18 8:10:24	R103	1.1; 1.3; 1.6;	3	4	1	1	4	4	3.2	4.1; 4.2; 4.3;	3	3	4	4	4	7.1; 7.7;	3	3	4	3			10.7	11.2	12.1	61
2/26/18 9:20:23	R104	1.1; 1.3; 1.4; 1.5; 1.6;	3	4	4	4	3	3	3.2	4.1; 4.2; 4.3;	2	3	3	3	2	7.1;	1	1	2	2			10.4	11.5	12.1	38
2/26/18 9:36:23	R105	1.1; 1.3; 1.6;	2	2	0	0	0	1	3.1	4.1; 4.2;	0	2	0	0	0	7.8;	4	0	0	0			10.1		12.1	65
2/26/18 10:29:38	R106	1.1;	3	3	3	3	2	2	3.2	4.1; 4.2;	2	3	3	2	3	7.1; 7.7;	3	2	3	2			10.7	11.1	12.2	47
2/26/18 10:51:21	R107	1.3; 1.4;	2	3	1	2		2	3.2	4.1; 4.2;	3	3	3	1	1	7.1; 7.3;	3	1	3	3			10.4	11.2	12.1	67
2/27/18 3:22:59	R108	1.7; 1.1; 1.3;	3	3	2	2	4	3	3.2	4.1; 4.2; 4.3;	3	3	2	2	3	7.8; 7.7;	1	3	3	3			10.4	11.6	12.2	43
2/27/18 8:30:05	R109	1.1; 1.3; 1.4;	3	3	2	0	3	3	3.2	4.4; 4.1; 4.2;	1	2	2	0	3	7.8;	1	1	1	2			10.1	11.6	12.1	53
2/27/18 8:51:39	R110	1.1; 1.3;	1	2	1	0	3	3	3.2	4.1;	2	4	3	3	3	7.1; 7.7;	1	1	2	1			10.1	11.6	12.1	69
2/27/18 9:29:46	R111	1.2; 1.3; 1.4; 1.5; 1.6;	3	3	3	4	4	3	3.2	4.1; 4.2; 4.3;	3	4	2	2	3	7.1; 7.4; 7.7;	1	3	4	1			10.7	11.1	12.2	45
2/27/18 9:38:37	R112	1.1; 1.3; 1.4; 1.5;	3	2	2	2	2	2	3.2	4.1; 4.2;	3	3	2	1	1	7.8;	0	1	1	1			10.4	11.2	12.1	53
2/27/18 10:57:06	R113	1.1;	3	3	2	3	4	2	3.2	4.1;	4	4	4	2	3	7.2; 7.3;	0	0	1	1			10.3	11.1	12.2	39
2/27/18 13:08:16	R114	1.3; 1.5;	2	2	1	1	3	3	3.2	4.1; 4.3;	0	4	3	0	0	7.8; 7.1; 7.3; 7.4;	3	4	4	4			10.5	it in cl	12.2	59
2/27/18 15:24:57	R115	1.4;	3	4	3	3	3	3	3.2	4.1; 4.2; 4.3;	3	2	3	1	1	7.2;	2	2	2	3			10.6	11.6	12.1	76
2/27/18 16:01:40	R116	1.1; 1.3; 1.4;	3	3	2	2	3	3	3.2	4.1; 4.2; 4.3;	4	3	2	3	2	7.3;	3	0	1	1			10.7	11.1	12.1	40
2/27/18 22:39:41	R117	1.1; 1.3; 1.4; 1.6;	3	2	3	0	3	2	3.2	4.4; 4.2;	4	2	2	3	3	7.1; 7.4; 7.7;	1	4	2	1			10.5	11.6	12.1	44
2/28/18 3:45:57	R118	1.4; 1.5;	3	4	4	3	4	2	3.2	4.1; 4.2;	4	3	4	3	3	7.1; 7.3; 7.5;	3	3	3	2			10.4	11.2	12.1	54
2/28/18 19:42:02	R119	1.1; 1.3; 1.5;	0	4	4	4	4	4	3.2	4.1;	4	3	0	0	0	7.2; 7.7;	0	1	0	0			10.5	11.5	12.2	39
3/1/18 11:14:00	R120	1.3;	4	4	2	2	2	3	3.2	4.1; 4.2;	3	2				7.1; 7.2; 7.3;	4	0	1	1			10.7	11.2	12.2	68
3/6/18 10:28:12	R121	1.4; 1.5; 1.6;	0	2	0	4	3	4	3.1	4.1; 4.2;	2	4	3	2	2	7.7;	3						10.1	11.5	12.2	50
3/6/18 12:13:59	R122	1.3; 1.4;	4				4		3.2	4.1; 4.3;	0	0	0	0	0	7.7;	0	0	0	0			10.4	11.5	12.2	46
3/12/18 11:33:32	R123	1.1; 1.4;	3	4	3	3	3	0	3.2	4.3;	1	3	1	0	1	7.3;	4	1	1	0			10.6	11.1	12.2	59

 Respuestas a las preguntas abiertas. Pregunta 6, p. 17-18.

 Respuestas a las preguntas abiertas. Pregunta 9, p. 18-21.

## Anexo 4 - Parte 2. Respuestas a la opción “Otros” de las preguntas 1, 4 y 7

### Pregunta 1. Opción Otros (codificada como 1.7)

*R001* Most of my data is collected to improve management of resources and I would like the data to be used as widely as possible.

*R016* Benefits conservation.

*R019* Encourages collaborations.

*R020* Benefits other researchers, government agencies and regulatory authorities.

*R022* Makes data useful for other analyses.

*R028* Alienta al trabajo en equipo.

*R031* Contribuye al avance del conocimiento.

*R033* Es una exigencia para la publicación de *papers*.

*R041* Helps the research of colleagues.

*R053* I am funded by a public research organization so it is only fair that my data are then freely available. I also hope it will help more research get done, more efficiently and collaboratively.

*R063* Helps data visibility.

*R066* Increases reuse in experiments by other people.

*R088* I would also like to have access to other's data.

*R089* Helps to progress advancing research.

*R091* Enables development of cooperation, allows new use and new view on the results, allows development of novel analysis etc.

*R101* Maximize the benefit of the data for the scientific community.

*R102* Enables it to be added to other datasets with potential for unforeseen applications.

*R108* Expands the amount of data available for researchers.

### Pregunta 4. Opción Otros (codificada como 4.4)

*R004* No requirements.

*R016* May inquire about co-authorship but wouldn't require it.

- R020* Any data that was generated through contract with another agency (e.g., government, regulatory, commercial business) must also be authorized for sharing by that agency. Contractual and Intellectual Property agreements must be adhered to.
- R022* Data are used appropriately for the questions needed to be addressed.
- R032* Make it available only after publishing.
- R033* Su utilización luego de la publicación de los resultados por nuestro grupo.
- R044* None of the above;
- R063* Any use includes proposing joint collaboration towards publications that include the data.
- R066* Acknowledgment to the data publishers.
- R072* No solicito coautoría pero si tal vez algún apartado de agradecimientos para el proyecto que lo financió.
- R088* When collected by a student, have him/her included in the new study.
- R101* Acknowledgment.
- R102* Acknowledgement / citation is sufficient - no need for permission.
- R109* Institution and funders are acknowledged.
- R117* Ya hayan sido usados en una publicación de mi autoría.

**Pregunta 7. Opción Otros** (codificada como 7.8)

- R016* N/A - I'm not an owner of any large datasets, so am unable to answer this question. Answers to previous questions were based on the hypothetical situation in which someone asked me for data. They reflect what I WOULD do... not what I HAVE done.
- R022* Permissions from colleagues can take time to obtain and not worth the effort if the requester is not using it for an approved use.
- R027* No fue planteado.
- R031* Los datos estaban siendo usados y aún no se habían completado las publicaciones.
- R033* Para datos moleculares, no es posible retener los datos si la intención es una publicación científica. Para aplicaciones biotecnológicas retengo los datos.

- R041* My own paper was not yet finished.
- R049* My students had not yet published the data.
- R053* It could not be released until my thesis was finished.
- R057* Colleagues contributing data refused further sharing of the data.
- R072* Nunca retuve datos.
- R091* I did not trust fair use by those requesting the data.
- R093* Not being aware of the importance of sharing data.
- R101* Still working on publications.
- R102* Often I work on compiled data, and so require the permission of all original data providers before sharing.
- R105* Not applicable.
- R108* So commercial businesses could not make a profit from my institute's data.
- R109* Question is not applicable.
- R112* I suspect most of the reasons above have been behind unsuccessful requests for data by me.
- R114* Distrust of the potential users of the data.

## Anexo 4 - Parte 3. Respuestas a las preguntas abiertas de la encuesta (preguntas 6 y 9)

### Pregunta 6. Otros recursos para conseguir datos u otras estrategias que utilice.

- R001* Reach out to organizations sponsoring research to see if they can point me to datasets/sources perhaps not yet advertised/completed.
- R010* It starts with conferences or discussions among scientists around specific questions, or ideas.
- R012* -
- R022* Read current literature related to data collection effort
- R033* Bases de datos de genes, genomas, metagenomas como GenBank, IMG, etc.
- R034* Sistemas de información institucionales.
- R041* ASKING colleagues! (probably the most simple and most common method - why is it missing here ?)
- R046* Contacto directo con especialistas. Involucrar a gente que posea datos al proyecto o grupo de trabajo.
- R047* National Central database.
- R050* Publicaciones.
- R058* Institutions that collect data for other purposes than research (ie. Meteorology, water quality, etc)
- R069* Web of Knowledge.
- R073* Convenios de cooperación entre Instituciones e Investigadores afines.
- R084* By google.
- R086* None.
- R090* En la Web o página sencilla de abrir en donde podamos encontrar la base de datos.
- R094* Direct sampling and interaction with locals.
- R098* Convenios, trabajo colaborativo.
- R100* Normally search databases and contact colleagues
- R104* -

*R111* Informes técnicos de acceso público. Ejemplo, Estudios de línea base ambiental, informes finales de Fondo de Investigación Pesquera. También publicaciones seriadas.

*R112* Collaborative projects specifically designed to bring together data (or metadata) for analysis.

*R117* Datos en paqueterías de R.

*R121* I ask colleagues for helping.

### **Pregunta 9. Comentarios sobre experiencias en compartir datos y otros factores que alienten o desalienten el compartir datos no mencionados en el cuestionario**

*R001* Mostly the issue is time. Many of my data sets come from citizen scientists or broad collaboration and require a good amount of QA/QC. That is the biggest barrier to more widely distributed data.

*R009* Not for marine, but still issues: -privacy issue in citizen-based dataset. - Military areas observations and the possibility to publish observations in private areas that shouldn't be covered by operators

*R012* -

*R016* I'm actually a person who requests a lot of data for global efforts to aggregate tracking datasets... I've encountered some resistance to sharing data, particularly tracking data. Many of the potential contributors (many of whom were funded by tax payer dollars and are technically required to make those data publicly available under U.S. law) are hesitant or unwilling to share data that they've been sitting on for ~15-20 years that is still unpublished. Reasons range from concern about missing publishing opportunities to having previous bad experiences with data sharing. I understand these concerns but I think it is counter to science to sit on data for that long, especially data funded by the public. I don't blame the people, I blame a system that rewards publication records over data collection and all the hard work that goes into the HUGE effort that is data collection. I think for our attitudes towards data sharing to change we need to think critically about making published datasets carry just

as much "academic weight" as a publication. And how we can create a better system to reward data collection.

*R031* Las respuestas dependen del tipo de datos. Hay datos de base (por ejemplo, registros de temperaturas en el mar) que he compartido inmediatamente y sin restricciones. Otros datos biológicos más específicos (ej., resultados de censos de abundancia animal) que se pueden hacer públicos una vez que se han publicado.

*R033* Es esencial disponer de sistemas de almacenamiento de datos fáciles de utilizar, preferentemente nacionales. En nuestro caso, utilizamos bases de datos internacionales, pero los metadatos están disponibles en el sistema nacional de datos genómicos.

*R034* Entregar datos brutos o muestrales, pueden ser mal interpretados, por lo tanto, se sugiere depurar y realizar el análisis exploratorio previo antes de entregar o publicar.

*R035* no additional comments

*R046* En algunos casos, debería priorizarse el uso compartido de datos. Por ejemplo, institutos de investigación nacional, financiados por el estado, tienen más obligaciones prioritarias con el público en general.

*R047* I would like to know if all scientists of the same specialty will have to be networked in order to exchange data on a permanent basis.

*R049* It is essential to also involve the communities where the data was collected in this process (and in this survey). We have to provide agency to the local stakeholders who facilitate and enable the research and not focus on just researcher to researcher exchange

*R053* Encourage:

- easy to use platforms for sharing
- easy to use metadata forms
- help from collaborators
- errors to be fed back on by collaborators
- users to send any issues/requests back to the owner, so the data can be updated and cleaned accordingly

Discourage:

- people using the data without permission

*R057* I have previously found errors in biodiversity data (omission of zero observations, duplication of records, insufficient metadata describing the sampling methods, taxonomic errors, etc.) that make it hard to use the data for serious work.

*R058* Data appropriation without consent (stealing data) is a common issue in the Brazil, and as is a problem in sharing valuable information with colleagues.

*R069* In past experience with other types of data, some unscrupulous researchers used the data as their own without any form of acknowledgement of the data producer or project/institution. Their argument for doing so was that the data was "freely available". This highlights the fact that clear conditions for the use of the data need to be in place.

*R073* En lo personal creo que compartir y requerir datos es esencial para fortalecer líneas de investigación y generar redes de apoyo entre distintos investigadores con el fin de mejorar el conocimiento actual que se puede tener sobre un tema.

*R084* publish all data

*R086* None.

*R087* It is with some hesitation that I comply with the overall intent of this philosophy. I believe that misuse of data, extra work providing it to a repository, or the possibility of misuse are all real. However, in order to bring about change it all of these issues are critical in order to advance the field of marine biodiversity.

*R088* Reciprocity is essential, clear data publishing rules by funding agencies is needed and, when existent must be enforced. At least in Mexico we have a huge lack of data sharing rules and culture.

*R094* No comment.

*R098* Patentes.

*R100* I do not have any issue sharing my data is for the benefit of all

*R104* -

*R105* I am very willing to share data, but situations where my data are relevant to others are rather rare and generally I work only with my own data, but I have no objections at all to the idea of sharing data widely.

*R111* La confianza en que el investigador usará los datos apropiadamente y compartirá autoría en publicaciones es fundamental.

*R112* I am always happy to share my data, but much of my work involves using data that I have not generated. I share my data freely as it gives my work exposure, offers opportunities to develop publications, and increasingly because it is required by funding organizations anyway. It is a positive thing to do. For certain datasets, especially time-series, I understand some reluctance to share data. I believe this reluctance is misguided and I know of cases where funding has been withdrawn because the data were not being made available. In most cases ownership of data belongs to organizations (or funding agencies) not individual scientists, but this does not seem to stop some scientists treating data as their own.

*R114* I was much more open to freely sharing data in the past but am now extremely cautious, due to three bad experiences - one outright theft and publication of data by someone I considered a friend and colleague; have also shared data in two cases where it was referred to in two publications which I didn't see before they were published, and in both cases it was described incorrectly as supporting conclusions which it absolutely didn't support (which was personally embarrassing as anyone familiar with the field would immediately see it was wrong, and was stated as a pers. comm. as if I had given them the conclusions). These experiences have totally discouraged me from sharing data outside of a close circle of trusted colleagues. I will not put data on any publicly accessible site where I have no control over its use.

*R120* It lacks support of my institution for the preparation of the data and also lacks the culture of sharing data in a constant flow.

## Anexo 5. Ficha de registro - Análisis de contenido

No. de referencia: (1) \_\_\_\_\_

Título del *report*: (2)

---



---

Año: (3) \_\_\_\_\_

### 1. Actividades de colaboración

	Cód.	F.	Expresión/es
1.1. Colaboración con OBIS	(4)		
1.2. Colaboración con otros editores y proveedores de datos	(5)		
1.3. Actividades orientadas al mantenimiento y crecimiento de OBIS	(6)		
1.4. Actividades para brindar capacitación	(7)		
1.5. Promover la colaboración	(8)		

## 2. Factores que desalientan las actividades de compartir datos e información

	Cód.	F.	Expresión/es
2.1. Consideraciones legales	(9)		
2.2. Baja participación de los nodos	(10)		
2.3. Dificultades técnicas	(11)		
2.4. Falta de consistencia de los datos	(12)		
2.5. Calidad de los datos	(13)		
2.6. Otros factores que dificultan	(14)		

### 3. Factores propicios a las actividades de compartir datos e información

	Cód.	F.	Expresión/es
3.1. Variedad de herramientas, productos y servicios	(15)		
3.2. Interoperabilidad de los sistemas	(16)		
3.3. Compatibilidad de los metadatos	(17)		
3.4. Fortalecimiento de redes regionales	(18)		
3.5. Políticas de datos	(19)		
3.6. Uso de estándares	(20)		
3.7. Actualización	(21)		
3.8. Otros factores que favorecen	(22)		

## Anexo 6. Textos correspondientes a pasajes de las entrevistas mencionados en esta tesis

### 1. What are, in your opinion, the main issues when sharing marine biodiversity data nowadays?

#### E01

a) El no uso de estándares para organizar los datos o de vocabularios estandarizados; b) el tiempo requerido para el proceso de compartir los datos, desde el formato, hasta la carga en sistemas especializados, usualmente no está contabilizado dentro de los tiempos de los proyectos, por lo que representa una carga adicional que muchos investigadores no están dispuestos a asumir; c) el tiempo que toman los investigadores en hacer públicos sus datos.

#### E02

Sharing data in any environment requires an unambiguous approach that is well planned and robust. This is also true for Marine Biodiversity studies such as those deployed in OBIS. I always employ a full Data Management Plan (I can provide examples). This enables users (projects or individuals) to remain aware of the issues that may confront them in terms of managing the processes of data management. These include security, reliability and integrity.

#### E03

Metadata preparation. The metadata is the description of your data-set. In my opinion when you have more detailed description, more credited your data. The description includes a lot of different features that take time and proper training to complete. When you are publishing your data, you need to prepare your metadata. But sometime the metadata are very scarce due to the different reasons (lack of the information, lack of communication with the originator of the data, or just laziness).

#### E04

I do not see any issues in sharing the data as long as the data will be shared through open-access databases which have the data sharing agreement between the data provider and open access database (e.g., copy right and citation). If there is no such an agreement, then it could be an issue.

The only issue in terms of ease of data sharing for the data providers might be sometimes complication and difficulties in understanding the Darwin Cores and the data management and submission process. Mostly older generations of scientists and pure taxonomists are not very familiar with the data mobilization and management and they need to be trained to be able to share their data easily.

**E05**

Not sure about this question, my thoughts on “issues” when sharing marine biodiversity data is when someone is NOT sharing and/or partially sharing so the entire story cannot be evaluated/analyzed.

**E06**

The scientist consider data sets private property and that biologist have a weak tradition for sharing data. Often a biological data set (species identification) represents many working hours for an individual, unlike physical data often collected semi-automated.

**2. In your experience, what kind of data are more likely to be withheld, even if there are no impediments or restrictions for those datasets to be shared?**

**E01**

a) datos que no han sido publicados; b) datos de colecciones que están siendo comercializadas como servicios de consultoría ambiental; c) ciertos datos sensibles para algunos sectores, por ejemplo, datos asociados con pesquerías.

**E02**

In my view, data is a universal and necessary shared commodity for everyone. I am NOT in favour of academics or scientists withholding data. However, I am aware that this is an issue and primarily it relates to academics who are in the process of writing up their research for publication. There is a tendency for these people to hold onto their research until it is finally published in a journal or book. The Biggest failure of this process is that the datasets that inform these papers, articles or books are often left as an afterthought and are rarely published in a usable or meaningful manner after the publication of the associated research. These data are often either lost, or never made available to the public. I have several examples of researchers who have done excellent work, but their data was lost because it was never properly

processed and made available. Importantly, any European research that is funded through the European Union, for example, MUST be published - and the associated datasets must also be published in the public domain. In my view, the real obstacle is NOT unwillingness to make data available, but the lack of awareness and underestimating in HOW to make it available.

### **E03**

The data collected in the most recent time, new, modern data.

### **E04**

Data from unpublished papers, big and on-going cruises, old generation of scientists, old monographs, old hard print archives, data in other languages than English, and vast collaborative research programs. In this case having workshops, lectures, educational programs are necessary to be done by the data scientists or the open-access databases' staff to explain the data sharing process and make it clear to the data providers that their data will still be cited under their names even if they publish their dataset before journal publications.

### **E05**

I find the measure of whether or not data will be shared/published is not the type of data but rather it is totally related to the principal investigator's feelings regarding data sharing and possibilities and associated damages of someone else scooping or harming their research.

### **E06**

Data from research projects carried out by single investigators or at institutions with a data policy of not sharing but only depositing data in e.g. PANGEA.

## **3. When searching for datasets or data sources generated by others, what are the most usual strategies and/or tools you use to locate the data?**

### **E01**

Fuentes de datos: OBIS. Herramientas: paquetes R que interactúan con esas bases de datos, ej robis (<https://github.com/iobis/robis>)

### **E02**

Locating data is often a frustrating and bewildering experience! Based on what I said in question 2 about researchers often nor being aware of how to publish their

data in the public domain, it seems that whatever data is available is often in a wildly variable series of formats and output methods. For example, data are often provided in unusable formats because the digital software that was used is now outdated, or corrupted. Often, the formats (layouts) that are used are very specific to a particular discipline rather than usable by everyone. I have many examples of how data was produced in excellent research software packages, but the outputs are limited for use by those who are also licensed to use that software. My solution is to provide data to the world in formats and methods that are most widely accessible to all. Typically, I output my data as .CSV (Comma Separated Value) files to ensure that anyone can use it in a variety of ways.

In terms of data search strategies, the basic approach is to learn about the many existing platforms (such as OBIS) that provide access. It is often necessary to learn how to use these platforms before any real use can be made of the data.

### **E03**

<http://iobis.org/mapper/>. Select taxa first, then zoom in to the area of interest.

### **E04**

My methodology is to continually search in the broadest terms using, for example, Google Search to find new and reliable sources. It is important to use several search engines as they are all variable in their ability to locate what is available. Of course, I also use archives and published books/articles to locate data which is often contained in the form of tables or textual information.

### **E05**

I actually do not search for other datasets though I have in the past extracted other data from OBIS but only for checking out what is available.

### **E06**

N/A.