

# Evaluación del estado del conocimiento oceánico a partir de una base de metadatos: fortalezas, limitaciones y aplicación a México

*Juliano Palacios-Abrantes, Andrés Cisneros-Montemayor, Miguel Ángel Cisneros-Mata, Laura Rodríguez, Francisco Arreguín-Sánchez, Santiago Domínguez-Sánchez, Verónica Aguilar, Raquel López-Sagástegui, Hector Reyes, Rocío Rivera, Silvia Salas, Stuart Fulton, Nuno Simoes y William Cheung*

5/2019

- **NOTA: esta es una traducción al español del artículo** *A metadata approach to evaluate the state of ocean knowledge: Strengths, limitations, and application to Mexico* escrito por Juliano Palacios-Abrantes, Andrés Cisneros-Montemayor, Miguel Ángel Cisneros-Mata, Laura Rodríguez, Francisco Arreguín-Sánchez, Santiago Domínguez-Sánchez, Verónica Aguilar, Raquel López-Sagástegui, Hector Reyes, Rocío Rivera, Silvia Salas, Stuart Fulton, Nuno Simoes y William Cheung. El artículo original se puede encontrar en el siguiente link.
- **Palacios-Abrantes J**, Cisneros-Montemayor, A. M., Cisneros-Mata, M. A., Rodríguez, L., Arreguín-Sánchez, F., Aguilar, V., Domínguez-Sánchez, S., Fulton, S., López-Sagástegui, R., Reyes-Bonilla, H., Rivera-Campos, R., Salas, S., Simoes, N., and Cheung, W. W. L., 2019. A metadata approach to evaluate the state of ocean knowledge: Strengths, limitations, and application to Mexico. PLoS ONE, 14 (6), e0216723.

Descargar traducción en PDF

## Resumen

Climate change, mismanaged resource extraction, and pollution are reshaping global marine ecosystems with direct consequences on human societies. Sustainable ocean development requires knowledge and data across disciplines, scales and knowledge types. Although several disciplines are generating large amounts of data on marine socio-ecological systems, such information is often underutilized due to fragmentation across institutions or stakeholders, limited standardization across scale, time or disciplines, and the fact that information is often not searchable within existing databases. Compiling metadata, the information which describes existing sets of data, is an effective tool that can address these challenges, particularly when metadata corresponding to multiple datasets can be combined to integrate, organize and classify multidisciplinary data. Here, using Mexico as a case study, we describe the compilation and analysis of a metadatabase of ocean knowledge that aims to improve access to information, facilitate multidisciplinary data sharing and integration, and foster collaboration among stakeholders. We also evaluate the knowledge trends and gaps for informing ocean management. Analysis of the metadatabase highlights that past and current research in Mexico focuses strongly on ecology and fisheries, with biological data more consistent over time and space compared to data on human dimensions. Regional imbalances in available information were also evident, with most available information corresponding to the Gulf of California, Campeche Bank and Caribbean and less available for the central and south Pacific and the western Gulf of Mexico. Despite existing knowledge gaps in Mexico and elsewhere, we argue that systematic efforts such as this can often reveal an abundance of information for decision-makers to develop policies that meet key commitments on ocean sustainability. Surmounting current cross-scale social and ecological challenges for sustainability requires transdisciplinary approaches. Metadatabases are critical tools to make efficient use of existing data, highlight and address strengths and deficiencies, and develop scenarios to inform policies for managing complex marine social-ecological systems.

## Introducción

El océano proporciona una variedad de bienes y servicios ecosistémicos como alimento, energía y rutas de transporte, además de ser una fuente de valores culturales y recreativos. Sin embargo, el cambio climático, la extracción excesiva de recursos marinos, la contaminación y demás productos de actividades humanas están afectando la biodiversidad marina mundial y los servicios ecosistémicos [3–6] impactando estructuras sociales y económicas de manera negativa [7]. La mitigación y gestión de dichos impactos humanos así como el desarrollo sostenible de los recursos marinos requieren datos de diferentes disciplinas, con largos intervalos de tiempo y que cubran diferentes escalas geográficas. Por ello, la elaboración de políticas públicas para la evaluación del estado de la investigación marina y el establecimiento de objetivos claros de manejo de recursos requieren diversas fuentes complementarias de información. Sólo con la inclusión de distintas fuentes y disciplinas es que el diseño e implementación de dichas políticas públicas serán efectivas [8]. La implementación de un enfoque multidisciplinario ha sido adoptado por alianzas internacionales que buscan alcanzar los Objetivos de Desarrollo Sostenible interdisciplinarios de las Naciones Unidas (UN-SDGs<sup>1</sup>) [9]. Sin embargo, a pesar de un llamado global de transición hacia una elaboración de “ciencia abierta” y los beneficios asociados donde datos y métodos son compartidos y las colaboraciones son más comunes [10], la identificación, el acceso y el intercambio de datos siguen siendo un desafío en todo el mundo [11].

Los metadatos, es decir la información requerida para entender los datos, como el tipo, contenido, fuente, calidad, formato, estructura y accesibilidad [10,12], son importantes para la armonización de datos a través de escalas, disciplinas y dominios. Los repositorios de metadatos (y su propio desarrollo) tienen la capacidad de mejorar el acceso a los datos, fomentar la colaboración entre las partes interesadas y facilitar los análisis posteriores así como el refinamiento de los datos [13,14]. Varios campos de investigación relacionados con sistemas marinos socio-ecológicos han generado grandes cantidades de datos. Sin embargo, dichos datos no están estandarizados y se encuentran dispersos, mantenidos por diferentes instituciones y su localización y acceso es limitado [6,8,15], por lo que a menudo son poco utilizados. Los metadatos son particularmente útiles para los países en desarrollo con capacidad de investigación limitada [11] y donde existen datos, pero se los considera limitados o no disponibles [16].

Repositorios de datos y metadatos nacionales para sistemas marinos se han creado para Australia [17], Canadá [13] y las Islas Canarias en España [18]. El Sistema Integrado de Observación Marina (IMOS) es un proyecto de investigación en Australia que opera a nivel nacional e incluye una base de datos que permite a los usuarios ver gráficos dinámicos, ingresar metadatos y acceder a datos [17]. Dicha base de datos ha dado lugar a cientos de publicaciones científicas, capítulos de libros y reportes [19]. En Canadá [13] se creó un repositorio de metadatos con el objetivo de identificar brechas temáticas y de información en la investigación marina para las regiones ártica, pacífica y atlántica. Posteriormente dicho repositorio se utilizó en la evaluación del progreso del país hacia el Convenio sobre la Diversidad Biológica: Objetivos de Aichi (CDB<sup>2</sup>) [14]. A su vez, el repositorio de datos marinos de las Islas Canarias (REDMIC) incluye datos, metadatos, documentos de investigación, mapas y gráficos interactivos relacionados con el ambiente marino, que han respaldado la toma de decisiones y la investigación regional [18]. Todas estas iniciativas tienen como objetivo aumentar el acceso a los datos, apoyar la investigación de metadatos y mejorar la toma de decisiones relacionadas con las políticas ambientales marinas a partir de un soporte científico.

En este estudio, desarrollamos un marco para la creación de una base interdisciplinaria de metadatos para sistemas marinos, con el objetivo de apoyar la toma de decisiones para el desarrollo sostenible de los

---

<sup>1</sup>United Nations Sustainable Development Goals (<https://www.undp.org/content/undp/es/home/sustainable-development-goals.html>)

<sup>2</sup>Aichi Biodiversity Targets - Convention on Biological Diversity (CBD) <https://www.cbd.int/doc/strategic-plan/2011-2020/Aichi-Targets-ES.pdf>

océanos mediante la evaluación del estado de la información existente, así como la determinación de vacíos y tendencias de investigación. Aplicamos este marco a México como ejemplo de una nación en desarrollo con extensas costas y area marina [15]. Al igual que en otras partes del mundo, múltiples organizaciones académicas (por ejemplo, instituciones públicas de investigación [20]), gobierno [21], organizaciones de la sociedad civil (OSC) [22], e instituciones privadas generan y albergan una gran cantidad de datos de múltiples campos de investigación. Sin embargo, la información sobre estos datos (y los datos en sí) no siempre está visible o accesible, ni se puede buscar en un formato estandarizado, por lo que personas que trabajan en campos específicos pueden no estar enterados de proyectos pasados o incluso actuales pertinentes a su campo de investigación. Así mismo, en varias ocasiones, los políticos encargados de la creación e implementación de políticas públicas se ven limitados por la disponibilidad de información (tanto temporal como espacial) para la implementación de políticas exitosas de manejo de recursos. Dichas limitaciones se pueden abordar mediante un esfuerzo colaborativo centrado en la creación y el mantenimiento de un repositorio de metadatos.

Este estudio describe los procesos de diseño, compilación y métodos para vincular y armonizar conjuntos de datos de diferentes escalas y dominios de una base de metadatos. Posteriormente, presentamos ejemplos de análisis de metadatos como tendencias históricas, regionales y temáticas. Crear y mantener un repositorio de metadatos de acceso abierto puede facilitar la interpretación de la información a través de consultas públicas y el intercambio de datos. Los análisis de metadatos son críticos para ayudar a identificar falta de datos y promover la creación de redes de colaboración entre una amplia gama de individuos, instituciones y organizaciones.

## Materiales y Métodos

El proceso de desarrollo de la base de metadatos de investigación marina para México (MDB) constató de cuatro etapas: (1) desarrollo de la estructura de la MDB; (2) identificación, comunicación y colecta de metadatos y repositorios disponibles; (3) elaboración de protocolos para la inclusión y distribución de metadatos; (4) publicación de la MDB como una plataforma accesible, de acceso abierto y estable a través del tiempo a través de una institución externa (En este caso la Comisión Nacional para el Conocimiento y uso de la Biodiversidad, CONABIO [23]). Posteriormente, presentamos algunos de los posibles análisis que se pueden realizar a partir de una base de metadatos mediante la identificación de tendencias y vacíos de información. La base de metadatos de investigación marina en México (*infoocéanos*) puede encontrarse en <https://www.infoceanos.conabio.gob.mx>.

### Estructura de la base de metadatos

Existen cinco niveles en la estructura de la MDB: Base de metadatos > Repositorio > Base de datos > Registro > Dato (Fig 1). La base de metadatos incluye los metadatos de los datos mientras que los repositorios son estructuras que contienen múltiples bases de datos. Los repositorios pueden estar en línea (e.g. Ocean Biogeographic Information System (OBIS) [24]), reportes de algún tema en específico (e.g. el anuario estadístico de pesca de SAGARPA-CONAPESCA [25]), o repositorios institucionales, de algún laboratorio o proyecto de investigación que comprenda múltiples bases (p.j. el catálogo de especies marinas en México del Instituto de Ciencias del Mar y Limnología de la UNAM [26]). Los registros de metadatos son individuales y describen cada dato dentro de un repositorio (p.j. “Captura de almejas en la región A” o “Captura de almejas en la región B”; Fig 1). Los metadatos contienen descripciones de los datos mas no los datos; en metadatos marinos, dicha información puede ser relacionada a captura pesquera, distribución de especies, o costo de pesca. Finalmente, un dato es un punto de información dentro de un record. Por ejemplo, un registro de abundancia anual de la población de cierta especie entre 2000 y 2003 incluye cuatro (promedio anual) datos de abundancia estimada. Así, los registros contemplan la escala espacial y temporal, por ejemplo, capturas pesqueras pueden ser registradas a nivel regional o nacional.

## Categorías de los Metadatos

Tener una estructura de metadatos estandarizada proporciona un marco para que la información incorporada sea consistente en sujetos (p.j. almejas, tunas, pulpos) y tipos de información (p.j. métodos, unidades de medida y detalles del diseño experimental) [12,17,27]. En el presente estudio asignamos campos de metadatos con el objetivo de maximizar la flexibilidad de la base para que pudiera acomodar datos multidisciplinarios y permitir una serie de meta-análisis. Inicialmente se adoptó la estructura de la base de metadatos marinos desarrollada para Canadá [13], ésta fue posteriormente modificada (principalmente para acomodar la geografía y clasificación de especie de acuerdo a estándares nacionales) y finalmente se incorporaron sugerencias proporcionadas por diversos expertos en temas marinos en el país. Dichas sugerencias fueron colectadas en diversas reuniones descritas en la sección *Colección de Metadatos*. La principal diferencia entre la estructura de la MDB y la preparada para Canadá es que cada registro de metadatos en la canadiense representa un repositorio específico (p.j. un reporte o una abse de datos), con un campo de la estructura de la base indicando el número de series de tiempo dentro de dicho registro. En la MDB, cada serie de tiempo es un registro de metadatos únicos y diferentes campos de la estructura denotan la base de datos y el repositorio a donde pertenecen. Si bien esta estructura requiere mayor esfuerzo de inclusión de registros, el producto final es una base de metadatos que se puede analizar de manera más fácil y que permite mayor detalle en cada registro. La estructura final de la MDB incluye 29 categorías desde información general como región o sujeto de estudio, hasta información más específica como número de puntos de datos en la base y/o disciplina (S1 Table).

## Colecta de metadatos

La primera etapa de colecta constó en capturar la información disponible en repositorios públicos en línea, como OBIS [28] o las estadísticas pesqueras de la Organización Mundial para la Alimentación (UN-FAO) [29], catálogos estadísticos nacionales como el de pesca [25] y bases de datos producidas y/o albergadas por universidades u OSC trabajando con el medio marino. Posteriormente, utilizamos esta primera versión de la base como plataforma de discusión en 20 talleres (~30 personas por taller), con grupos de investigación universitarios, gubernamentales y pertenecientes a ODC en ocho ciudades localizadas en distintas regiones de México (Fig 2). La última etapa constató en relajar reuniones en persona y virtuales así como presentaciones en conferencias nacionales e internacionales para demostrar el avance de la base y promover la colaboración entre distintas instituciones (S2 Table). Adicionalmente, nos reunimos con cuatro instituciones gubernamentales (CONACyT- Consejo Nacional de Ciencia y Tecnología [30], INAPESCA - Instituto Nacional de Pesca y Acuicultura [31], INECC - Instituto Nacional de Ecología y Cambio Climático [32] y CONABIO [23]), así como repositorios existentes (dataMares [33] y Monitoreo Noroeste [34]) para unificar esfuerzos. Cabe constatar que si bien esto representa un esfuerzo importante, no cubre todas las posibles fuentes de datos en México lo que resalta la importancia de continuar con dicho esfuerzo.

## Tipos de fuentes de datos

Incluimos todo tipo de fuentes de datos en la MDB. En la primera etapa buscamos incluir toda la información públicamente disponible en línea relacionada con el medio marino mexicano. Éstas fuentes incluyeron datos de instituciones académicas, ambientales-OSCs, gubernamentales y privadas (personales no académicos o industria), tanto nacionales como internacionales. Asimismo, se buscó colectar información sobre datos aún no publicados o mantenidos en formatos no digitales. La siguiente sección señala algunas de las instituciones que contribuyeron información a la MDD.

### Academia

Las fuentes de datos académicos incluyeron bases de datos gerenciadas por instituciones académicas públicas y privadas de México. Algunas de las fuentes que contribuyeron muchos datos, comparativamente, incluyen

el Atlas Climático de México manejado por la UNAM [35] el cual posee una extensa base de variables fisicoquímicas usadas en modelos de cambio climático, entre otras cosas. La unidad académica de la UNAM en Sisal, Yucatán (UNAM-UAY) proporcionó información en diversos temas como oceanografía, ecología, pesquerías, biología y turismo [36]. Finalmente, el Centro de Investigación y Estudios Avanzados del Politécnico Nacional (CINVESTAV-IPN) contiene una extensa fuente de información relacionada, principalmente, a la pesca y el turismo en la península de Yucatán [37].

### **Instituciones gubernamentales**

El gobierno mexicano dió un pazo inpresedente hacia la transparencia de datos con el decreto de 2015 que establece la política nacional de datos abiertos a través del portal Datos Abiertos MX donde se encuentran miles de bases de datos institucionales [38,39]. Si bien el portal contiene sólo un porcentaje de la vasta información que distintas dependencias han colectado en décadas de programas públicos, representa una fuente de más de 500 bases de datos relacionadas con corrupción, desarrollo económico, servicios públicos, cambio climático y derechos humanos, entre otros [39]. A pesar de que algunos de estos datos no están relacionados directamente con el ambiente marino, siguen siendo fundamentales para la elaboración de políticas públicas relacionadas con sistemas socio-ecológicos marinos. La transición hacia el portal de Datos Abiertos MX no ha sido completa y diversas instituciones mantienen sus portales de datos. Este es el caso de la Secretaría de Economía [40], CONAPESCA [25] y CONABIO [41], entre otros. Toda la información de éstas y otras instituciones destacadas en la base de metadatos están públicamente disponible para consulta en forma de reportes, portales de internet y anuarios.

### **Organizaciones de la Sociedad Civil (OSCs)**

Las OSC en México tienen datos de pesquerías, conservación, oceanografía y sociología. Comunidad y Biodiversidad, A.C. (COBI) representa el mayor repositorio de las OSC en la base de metadatos. Dicha organización tiene como objetivo conservar los ecosistemas marinos que se están deteriorando debido a la sobre-explotación de recursos y cuenta con varios programas de monitoreo que se pueden trazar más de dos décadas atrás [22]. El proyecto *Monitoreo Noroeste* del Fondo Mexicano para la Conservación de la Naturaleza es el segundo repositorio con mayor cantidad de registros en la MDB y a su vez representa un repositorio con más de 1,000 bases de datos de monitoreo en el noroeste mexicano proveniente de 20 OSCs [34].

### **Fuentes académicas internacionales**

Grupos de investigación internacionales contienen una gran variedad de datos para México sobre todo como parte de estudios a nivel global siendo dataMares y OBIS los principales repositorios internacionales en la MDB. dataMares es una plataforma de datos abiertos con base en el instituto de oceanografía *Scripps* de la Universidad de California, San Diego [33]. Dicho programa tiene como objetivo albergar y facilitar el acceso a datos científicos relacionados con las costas mexicanas. A su vez, OBIS es una plataforma global de datos abiertos relacionados con biodiversidad (en este caso marina) [24]. IN ADDITION, el museo del desierto de Arizona-Sonora contiene un listado extensivo de invertebrados del Golfo de California, la Universidad de British Columbia en Canadá tiene más de tres mil registros relacionados con economía pesquera, modelos de cambio climático para los océanos, y proyecciones futuras de abundancia y captura para al rededor de 200 especies que ocurren a nivel nacional. Dichos datos están concentrados en las unidades de investigación *Changing Ocean Research Unit* [42] y *Fisheries Economic Research Unit* [43]. Finalmente, FishBase [44] y SeaLifeBase [45] son bases de datos en línea que proporcionan diversa información ecológica (p.j. historia de vida, ecología trófica, dinámicas poblacionales) para más de dos mil especies que ocurren en México.

## Análisis de metadatos

El análisis de la MDB fue realizado con el programa *R-Studio (R) Version 1.1.463* y paquetes *data.table* [46] y *tidyverse* [47]. Las categorías de metadatos fueron comparadas por números totales y porcentajes disponibles por disciplina e incluyeron la distribución espacial y temporal de los registros, la cantidad de registros por taxa, disciplina y tipo de fuente de datos así como la relación socio-ecológica de los metadatos. Todas las figuras fueron producidas utilizando los siguientes paquetes de *R*: *ggplot2* [48], *cowplot* [49], *ggpubr* [50], *ggrepel* [51], *gridExtra* [52] y *wesanderson* [53].

Para el análisis espacial se utilizaron los paquetes *ggplot2* [48] y *sf* [54] y el mapa de México se obtuvo de *Natural Earth* (<http://naturalearthdata.com>). Existen diversas categorías espaciales para los océanos mexicanos, por ejemplo la CONABIO reconoce cinco eco-regiones marías en México mientras que la CONAPESCA establece seis regiones de pesca nacionales por lo que tuvimos que estandarizar la zona económica exclusiva (EEZ) de México para armonizar datos de distintas fuentes (Fig 2). También se estandarizaron los nombres en la categoría de Sujeto de investigación (taxa). En el caso de nombres comunes, éstos se modificaron para reflejar el objeto principal (p.j. “camarón” se utilizó para incluir “*shrimp*”, “camarones”, “camarón sin cabeza”, etc.) y los nombres científicos fueron homologados y corregidos con el paquete *taxize* [55].

El análisis de tendencias temáticas fue hecho a partir del número de registros en la MDB así como la cantidad de puntos de datos (años de datos) disponibles durante los años de colecta de cada registro. Todos los registros fueron catalogados de acuerdo a su interacción socio-ecológica de acuerdo al marco de DPSIR (*Drivers*; Causantes, *Pressures*; Presión, *State*; Estado, *Impacts*; Impactos, y *Response*; Respuestas) [56]. De acuerdo al marco, los Beneficios representan beneficios sociales provenientes de los sistemas naturales (p.j. capturas pesqueras), Presión (que juntamos como causantes) representan cualquier presión al sistema natural por parte del social (p.j. esfuerzo de pesca), Respuesta considera cualquier acción que reduce dicha presión (p.j. limitar el esfuerzo pesquero) y Estado se refiere al estado actual de los sistemas naturales (p.j. *stock assessments*). Utilizamos el paquete *networkD3* [57] para analizar la relación entre registros, instituciones y disciplinas y DPSIR. Finalmente, utilizamos el método estadístico de *Chi-Square* [58] para investigar diferencias significativas en el número de registros entre distintas variables.

Cabe recalcar que es posible que existan bases de datos repetidas en la MDB. Sin embargo, no esperamos que la repetición de bases de datos sea un problema debido al tamaño de la MDB y el esfuerzo que se hizo para identificar dichas repeticiones. Utilizamos *R* para automatizar e identificar fuentes de información redundante (p.j. instituciones que tienen la misma base de datos) y se consultó, cuando posible, a curadores y dueños si alguna de sus bases de datos estaban publicadas en otros repositorios. Los registros que presentan la misma base de datos (p.j. Capturas pesqueras de CONAPESCA y dataMares) pero con distinto nivel de procesamiento (p.j. distintos niveles taxonómicos o anuales) no fueron considerados como duplicados y se mantuvieron como registros diferentes.

## Resultados

La base de metadatos de investigación marina en México incluye 114,237 registros en su versión de Octubre de 2018. Dichos registros provienen de 216 repositorios albergados por instituciones académicas ( $n = 19$ ), gubernamentales ( $n = 22$ ), inter-gubernamentales ( $n = 2$ ) e internacionales ( $n = 29$ ), así como OSC ( $n = 21$ ). Los registros no están distribuidos de forma equitativa entre disciplinas catalogadas ( $X^2 = 337060$ , d.f. = 10,  $p < 0.001$ ) ya que Ecología (45%) y Pesquerías (38%) representan la mayoría de los registros (Fig 3).

Fuentes internacionales (p.j. Global Biodiversity Information Facility-GBIF, dataMares, OBIS) contribuyen la mayor cantidad de registros para México (49%), aun que cabe mencionar que éstas fuentes incluyen datos colectados por investigadores mexicanos, en instituciones mexicanas, o fueron apoyadas económicamente por el gobierno mexicano [59,60]. En general, predominan los metadatos provenientes de instituciones académicas (sin importar la disciplina) y gubernamentales (principalmente para pesquerías). Si bien existe una variedad en el tipo de instituciones que colaboraron, dataMares (52 bases principalmente de “pesquerías” representando más de 22,000 registros), Datos Abiertos MX (90 datos pertenecientes a 9 dependencias) y

OBIS (19,000 registros de más de 13,000 especies) representan el 46% de todos los registros. Sólo 20 bases de datos están clasificadas como “privadas” lo que sugiere que, virtualmente, todos los datos (metadatos) analizados son de libre acceso y los autores/dueños están dispuestos a nuevas colaboraciones.

El análisis histórico de la recopilación de metadatos puede informarnos sobre tendencias históricas de investigación nacional (Fig 4). Los primeros registros en la MDB fueron colectados en 1791 (muestras de plancton) y los datos ecológicos han estado bien representados históricamente por diversos períodos de muestreo. Por otra parte, la mayoría de los datos pesqueros fueron inicialmente colectados al inicio de los 50's con una expansión paulatina a través del tiempo acompañando un incremento en la investigación regional y conservación presentá una expansión considerable a partir de la primera década del siglo 21. El análisis también muestra una tendencia negativa en número total de registros a partir de 2010 seguido de una caída abrupta en 2015 (Fig 4). Creemos que dicha tendencia (especialmente a partir de 2015) es debido al tiempo que se demora entre la recopilación, procesamiento y publicación de datos.

Existen 24,083 objetos de estudio (taxa objetivo de la recolecta de datos) representados en la MDB de los cuales 97% representan especies específicas (p.j. Pulpo maya, *Epinephelus spp.*) y tan sólo el restante 3% son categorías comunes como “pulpo” o “manglar”. Los registros que no son especie-específicos son catalogados como “Multiple especies” y comprenden tan solo el 3% del total de registros. La distribución de los datos resultó ser desigual entre especies, a pesar del largo listado en la MDB: 3.7% objetos de estudio comprenden 52.29% de todos los registros colectados siendo *Carcharhinus porosus* y *C. falciformis* los más altos con 1,200 registros *C. u* y *C. limbatus* con casi 1,000 registros.

Encontramos una diferencia significativa ( $X^2 = 93114$ , d.f. = 6,  $p < 0.001$ ) en términos de la distribución de los registros siendo el océano Pacífico el más abundante (49% de los registros casi todos concentrados en una zona) seguido por el Atlántico (37%) y finalmente los registros a nivel nacional (14%). Diferencias regionales también fueron significativas ( $X^2 = 63175$ , d.f. = 3,  $p < 0.001$ ) con una mayoría de registros en la zona del Golfo de California y Pacífico Norte (43% de los registros y 77% de los registros del Pacífico), seguido por la región del Banco de Campeche y Caribe (27%) (Fig 5). Para México, la mayor parte de los datos académicos fueron catalogados como *Estado* (p.j. listados de especies) mientras que los datos gubernamentales reportaron principalmente *Beneficios* (p.j. ganancia por turismo) y *Presiones* (p.j. subsidios que incrementan pesca en vez de disminuirla). finalmente, los registros provenientes de OSC (tanto nacionales como internacionales) se relacionana principalmente al *Estado* de los recursos naturales (e.g. monitoreo de arrecifes) y *Beneficios* sociales (e.g. empleo del sector pesquero) (Fig 6). Una cantidad limitada de información relacionada con conservación fue relacionanda con *Beneficios* y una cantidad comparativamente inferior de investigación pesquera y sobre acuicultura relacionada con *Presiones* vs *Beneficios*.

## Discusión

El análisis de la base de metadatos de investigación marina en México nos ayudó a entender la disponibilidad de información multi-disiplinaria relacionados a los océanos de México, identificar el estado y tendencias de la investigación y los datos marinos en México e identificar vacíos de datos; todo esto con el objetivo de informar políticas públicas relacionadas a los océanos mexicanos. Particularmente, la construcción de una base como esta permite realizar una evaluación general del tipo de investigación y datos disponibles proporcionando información fundamental para el proceso de toma de decisiones [13].

El presente estudio reveló el largo historial de investigación marina en México con una substancial cantidad de datos relacionados con ecología y pesca elaborados principalmente por instituciones académicas y de gobierno. Sin embargo

Sin embargo, identificamos la necesidad de incorporar y/o invertir en monitoreo ecológico a largo plazo, otros aspectos además de desembarques pesqueros y demás temas como conservación y oceanografía como los vistos en iniciativas como FMCN-Monitoreo Noroeste y la Red de Investigación Ecológica a Largo Plazo (LTER-Mex) [61]. Tales esfuerzos ciertamente apoyarán el progreso de las políticas públicas hacia objetivos de sostenibilidad tales como los objetivos del CBD-Archi [14]. También identificamos una distribución regional de registros sesgada hacia el Golfo de California y el Pacífico Norte y casi inexistente en otras áreas del

Pacífico. Esto destaca una posible falta de datos en otras regiones, o que los datos disponibles son menos accesibles en las áreas mal representadas. Este resultado puede ayudar a crear conciencia de que hacen falta recursos en estas regiones para apoyar más investigación marina y/o mejorar la colaboración y el intercambio de investigación entre instituciones.

Las tendencias históricas en los datos disponibles, como se refleja en los metadatos, pueden atribuirse a iniciativas tanto nacionales como internacionales. La disponibilidad de datos nacionales en la década de los 1950's es una respuesta a la solicitud de la FAO a los países en desarrollo de recopilar y reportar datos sobre el estado de la pesca nacional [29,62]. En todo el mundo, este aumento en la disponibilidad de datos permitió nuevas iniciativas de investigación para complementar la información relevante para crear políticas públicas a nivel local, regional y global (por ejemplo, *Sea Around Us* [63], el *Ocean Health Index* [5] y *Too Big To Ignore* — Sistema de Información Sobre Pesca en Pequeña Escala (TBTI-ISSF) [64].

Los esfuerzos del gobierno desde principios de la década de 2000 han mejorado drásticamente la disponibilidad de datos de pesca [62] como la disponibilidad de los anuarios de pesca de CONAPESCA en formato de base de datos [65] y el portal de Datos Abiertos MX [39]. Los metadatos relacionados con ecología y de conservación también aumentaron durante este período, principalmente a través de programas de monitoreo ejecutados por centros académicos y OSCs. Particularmente, el repositorio de la UNAM-UAY en la península de Yucatán y el de COBI en el Caribe, los cuales tienen políticas de datos abiertos (Fig. 3).

Al mismo tiempo, estatutos federales a principios de la década de 2000 permitió que OSC se establecieran impulsando el estudio sistemático de sistemas socio-ecológicos marinos del Golfo de California [66]. La primera década fue dedicada a la organización, posteriormente, los primeros programas sobre pesca y biodiversidad se establecieron una vez que las OSC, agencias gubernamentales y académicos desarrollaran una relación más formal. Las colaboraciones inter-institucionales resultaron en abundante información disponible al público que ha sido empleada recientemente en iniciativas de conservación [67] e investigación además de los propios resultados productos de investigación científica aplicada [68,69]. La disminución aparente en la disponibilidad de datos en los últimos años puede ser debido a varios factores. Muy probablemente se deba por un retraso entre la recopilación de datos y la disponibilidad (debido a los tiempos de procesamiento o publicación) [13], así como la falta de financiamiento para la recopilación de datos sobre temas específicos que pudieron haber proporcionado más datos a través del tiempo [62, 70, 71].

Es interesante que varias de las tendencias generales encontradas en la MDB de México son similares a la versión canadiense que utilizó un enfoque similar con categorías casi idénticas, lo que facilita intercomparaciones [13]. Por ejemplo, alrededor de 60% de los registros de Canadá correspondieron a pesquerías, mismas que contribuyen el mayor número de registros en la MDB de México (Figs. 3 y 4) y ecología como el principal y segundo mayor contribuyente en México y Canadá, respectivamente [13].

También existe una fuerte tendencia hacia la investigación a nivel de especie en ambos países (por ejemplo, captura, rasgos de la historia de vida y datos de presencia / ausencia) representando alrededor del 70% de los registros de Canadá [13] y más del 90% en México. Al mismo tiempo, la investigación a nivel ecosistémico ha aumentado en ambos países desde fines de la década de 1990, probable debido a la consolidación del manejo ecosistémico como método principal de manejo de recursos marinos [72,73] alineado con una capacidad de investigación relativamente extensa en México, a pesar de ser una nación en desarrollo. A pesar de esto, actualmente la información sobre temas más allá de la pesca, o del uso de recursos como tal, está subrepresentada en la MDB. Destaca especialmente la necesidad de una mayor atención a la investigación social sobre las dimensiones humanas de los sistemas marinos para generar evaluaciones marinas integradas y ayudar a que la toma de decisiones sea un proceso inclusivo. Esto no se limita a América del Norte, ya que bases de metadatos en otros lugares como Australia [17] o las Islas Canarias [18] contienen vasta información sobre especies y ecosistemas, mas carecen de información socio-económica de los usuarios dependientes de los recursos marinos.

Si bien el largo historial de colecta de datos ecológicos en aguas mexicanas ha producido varios catálogos de especies desde invertebrados marinos hasta peces y mamíferos [74], existe una inconsistencia importante entre los metadatos de especies comerciales y no comerciales. Los datos ecológicos tienden a ser registros de observaciones esporádicas ya que la mayoría de los proyectos no consisten en monitoreos a largo plazo debido a limitaciones en el financiamiento del proyecto o limitaciones temporales de los mismos [13]. En



contraste, los datos de pesca tienen series temporales más consistentes, con más registros a largo plazo en comparación con otros datos ecológicos, y por esa razón representan el mayor número de puntos de datos en los metadatos (Fig. 3). Por lo tanto, una especie importante para la pesca puede tener más de 50 años de datos de captura, mientras que las especies no comerciales a menudo tienen un solo registro de observación durante el mismo período de tiempo. La abrumadora cantidad relativa de información sobre especies de peces no es algo sorprendente y ni exclusivo de México [75], sin embargo, se requiere una gama mucho más amplia de datos para aplicar un manejo basado en los ecosistemas adecuado y tener en cuenta los impactos ecosistémicos de las pesquerías [76]. Además, la investigación no relacionada con el uso de recursos marinos es crucial para evaluar las interacciones, las externalidades y las posibles respuestas futuras a impactos del sistema.

Las diferencias regionales en la disponibilidad de datos reflejaron las tendencias de investigación subyacentes, pero también las diferencias en la capacidad regional de las instituciones y los ecosistemas y los patrones socioeconómicos [77]. La región del Golfo de California, entre las áreas con mayor biodiversidad del mundo [78] y de gran importancia para la pesca mexicana, se ha convertido en un centro de investigación académica y de conservación y de iniciativas relacionadas con la pesca. Estas instituciones de investigación proporcionan la infraestructura para generar posteriormente grandes cantidades de datos [77]. En contraste, el Pacífico centro-sur de México y el oeste del Golfo de México tienen muchos menos centros de investigación pesquera, OSC e instituciones educativas que el resto del país [77]. Como era de esperar, estas áreas también son las menos representadas en la metadatabase y deben priorizarse en la recopilación de metadatos futuros.

En el Golfo de México, el catastrófico impacto ambiental y económico causado por la explosión del pozo Deepwater Horizon en 2010 [79] destacó los limitados datos ecológicos disponibles para evaluar los impactos y provocó una mayor investigación científica apoyada en México por las agencias federales. La mayoría de los datos producidos a partir de estas nuevas investigaciones aún no están disponibles debido a los continuos litigios entre gobiernos, asociaciones de pesca y turismo y productores de petróleo, pero esto eventualmente proporcionará información importante para la región. Además, el desarrollo de importantes iniciativas interinstitucionales como el Consorcio de Investigación del Golfo de México (CIGoM) con sede en el CICESE, CINVESTAV [80] y el Instituto de Investigación Harte [81], y el proyecto de Biodiversidad Marina del Sur del Golfo de México liderado por el Laboratorio de Biodiversidad Marina (BDMY) [82] ayudará a sentar las bases de un observatorio marino en la región.

Regional differences in data availability reflected underlying research trends, but also differences in the regional capacity of institutions, and ecosystem and social-economic patterns [77]. The Gulf of California region, among the most biodiverse areas of the world [78] and of paramount importance for Mexican fisheries, has become a hub for academic research and conservation and fisheries-related initiatives. These research institutions provide the infrastructure to subsequently generate large amounts of data [77]. In contrast, the south-central Pacific of Mexico and the western Gulf of Mexico have far fewer fisheries research centers, CSOs, and education institutions than the rest of the country [77]. Unsurprisingly, these areas are also the least represented in the metadatabase and should be prioritized in future metadata collection. In the Gulf of Mexico, the catastrophic environmental and economic impact caused by the Deepwater Horizon well blowout in 2010 [79] highlighted the limited ecological data available to evaluate impacts and prompted increased scientific research supported in Mexico by federal agencies. Data produced from these new research are mostly not available yet due to ongoing litigation between governments, fishing and tourism associations, and oil producers, but this will eventually provide important information for the region. In addition, the development of important inter-institutional initiatives such as The Gulf of Mexico Research Consortium (CIGoM) based at the CICESE, CINVESTAV [80], and the Harte Research Institute [81], and the project of Marine Biodiversity of the South of the Gulf of Mexico led by the Marine Biodiversity Lab (BDMY) [82] will help lay the foundations for a marine observatory in the region.

## Conclusión

La base de metadatos desarrollada en el presente proyecto tiene el objetivo de ser una manera efectiva (en términos de costo y tiempo) de identificar información, tendencias y vacíos de investigación, así como

proporcionar un vehículo de comunicación y colaboración a investigadores e instituciones nacionales e internacionales. La incorporación de diversos grupos de investigación e instituciones en conjunto con el uso de nuevas tecnologías puede mejorar este tipo de proyectos tanto en México como en otros lugares. Consideramos que este esfuerzo puede y debe de ser repetido en otras regiones y países. En nuestro caso, el objetivo final de la base de metadatos es facilitar métodos multidisciplinarios que apoyen de manera ecológica, económica y social la creación de políticas públicas sustentables que sean inclusivas y efectivas a través del tiempo y espacio. La versión más actualizada de la base de metadatos de investigación marina en México puede ser encontrada en el portal de CONABIO <https://www.infoceanos.conabio.gob.mx> (*Infocéanos*).

- **Forma de Citar (Artículo original):** Palacios-Abrantes, J., Cisneros-Montemayor, A. M., Cisneros-Mata, M. A., Rodríguez, L., Arreguín-Sánchez, F., Aguilar, V., Domínguez-Sánchez, S., Fulton, S., López-Sagástegui, R., Reyes-Bonilla, H., Rivera-Campos, R., Salas, S., Simoes, N., and Cheung, W. W. L., 2019. A metadata approach to evaluate the state of ocean knowledge: Strengths, limitations, and application to Mexico. PLoS ONE, 14 (6), e0216723.

Traducido por Juliano Palacios Abrantes. Revisión técnica de\_\_\_\_\_

## Referencias