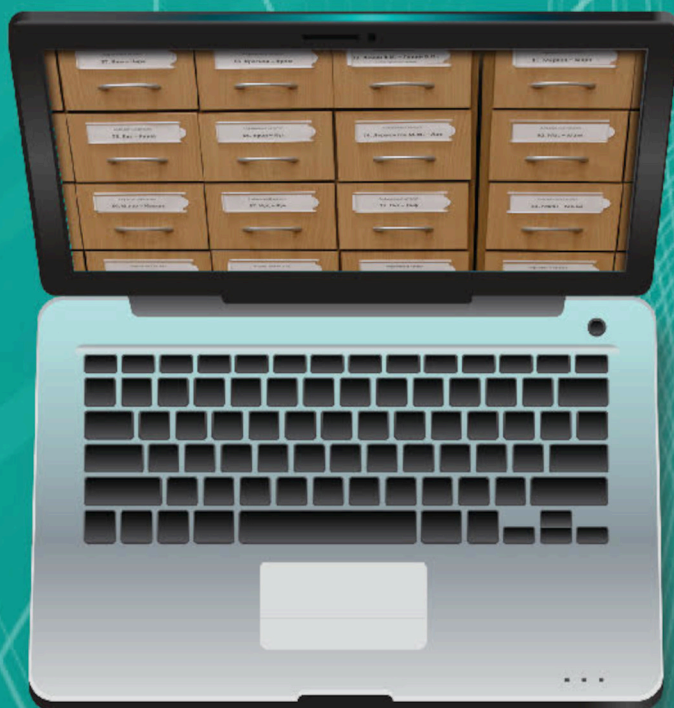


Áreas de oportunidad ante el nuevo entorno de la catalogación y los metadatos



Filiberto Felipe Martínez Arellano
Patricia de la Rosa Valgañón
Coordinadores



Z693

A74

Áreas de oportunidad ante el nuevo entorno de la catalogación y los metadatos / Coordinadores Filiberto Felipe Martínez Arellano, Patricia de la Rosa Valgañón. - México : UNAM. Instituto de Investigaciones Bibliotecológicas y de la Información, 2024.

Datos electrónicos (1 archivo PDF : ix, 245 p.). - (Organización de la información)

ISBN: 978-607-30-9813-7

1. Catalogación. 2. Metadatos. I Martínez Arellano, Felipe, coordinador. II. De la Rosa Valgañón, Patricia, coordinadora. III. ser.

Diseño de imagen: Dr. Bardo García Martínez

Diseño de cubierta: Mario Ocampo Chávez

Primera edición: 2024

D.R. © UNIVERSIDAD NACIONAL

AUTÓNOMA DE MÉXICO

Instituto de Investigaciones Bibliotecológicas
y de la Información

Circuito Interior s/n, Torre II de Humanidades,
pisos 11, 12 y 13, Ciudad Universitaria, C. P.
04510, Alcaldía Coyoacán, Ciudad de México

ISBN para PDF: 978-607-30-9813-7

Esta edición y sus características son propiedad
de la Universidad Nacional Autónoma de
México. Prohibida la reproducción total o
parcial por cualquier medio sin la autorización
escrita del titular de los derechos patrimoniales.

Publicación dictaminada

Impreso y hecho en México

Contenido

INTRODUCCIÓN	ix
Filiberto Felipe Martínez Arellano	
<i>RDA TOOLKIT</i> 2021: NUEVAS PERSPECTIVAS PARA LA CATALOGACIÓN	1
Filiberto Felipe Martínez Arellano	
ADAPTANDO RDA CON PERSPECTIVA DE DATOS ENLAZADOS. ALGUNOS PASOS DADOS EN LA BIBLIOTECA NACIONAL DE ESPAÑA	21
Ricardo Santos Muñoz	
DE LOS MODELOS CONCEPTUALES A LAS TAREAS DEL USUARIO: MÁS QUE UNA SECUENCIA DIDÁCTICA, UN MARCO DE TRABAJO PARA LA AUTOGESTIÓN DE LAS BIBLIOTECAS	41
Gerardo Belmont Luna	
HACIA UN CATÁLOGO MÁS INCLUSIVO: EL CASO DE <i>ILLEGAL ALIENS</i> EN LA UNIVERSIDAD DE MIAMI	57
José Antonio Vila Ruiz Margarita Pérez Martínez	
ORGANIZACIÓN DE LA INFORMACIÓN EN SERIUNAM: RETOS, OPORTUNIDADES Y APRENDIZAJES DURANTE LA PANDEMIA DE LA COVID-19	77
Hortensia Labra González Luis Javier Félix Acosta	
CATALOGACIÓN DE PERIÓDICOS EN LÍNEA: UN DESAFÍO ANTE LA NUEVA REALIDAD	95
María Gabriela García del CORRO	
ESQUEMAS DE METADATOS PARA REPOSITARIOS DE DATOS	109
Juan Miguel Palma Peña	

METADATOS PARA UN REPOSITORIO DE DATOS SOBRE CIENCIAS DE LA TIERRA	129
Minerva Castro Escamilla Saúl Armendáriz Sánchez Ricardo César Castro Escamilla Miguel Ángel Sosa Jiménez	
LOS ESTÁNDARES DE METADATOS COMO CRITERIOS DE CALIDAD DE LAS PUBLICACIONES PERIÓDICAS CIENTÍFICAS	157
Julio Zetter Patiño	
CONTEXTUALIZACIÓN DE RDA PARA ESTUDIANTES DE BIBLIOTECOLOGÍA, A PARTIR DE LOS ANTECEDENTES HISTÓRICOS Y MODELOS CONCEPTUALES	189
Argenis Rodríguez Salinas	
NORMALIZACIÓN Y CONTROL DE AUTORIDADES DE AUTORES COLOMBIANOS EN RDA: UN APOORTE AL CONTROL BIBLIOGRÁFICO NACIONAL	203
Yenny Alexsandra Carrillo Ibañez	
REGISTROS DE AUTORIDAD DE NOMBRE COLMEX PARA LA GESTIÓN Y CONTEXTUALIZACIÓN DE UNA IDENTIDAD ACADÉMICA	225
Guadalupe Vega Díaz Luz Elena Vergara Castro	

Metadatos para un repositorio de datos sobre Ciencias de la Tierra

MINERVA CASTRO ESCAMILLA

SAÚL ARMENDÁRIZ SÁNCHEZ

RICARDO CÉSAR CASTRO ESCAMILLA

MIGUEL ÁNGEL SOSA JIMÉNEZ

Universidad Nacional Autónoma de México, México

INTRODUCCIÓN

Para los repositorios institucionales (RI), la asignación de metadatos es fundamental para la creación, configuración y consulta de su contenido, bajo parámetros internacionales que brinden la posibilidad no sólo de realizar una organización técnica, sino además alternativas para el intercambio de registros y datos entre sus pares y la unificación de contenidos. Los metadatos tienen importantes variantes para los tipos de información, datos y textos completos que se manejan en repositorios como *Dublin Core*, *Open Aire*, *Geonetwork*, entre otros. En ese sentido, aquí se muestra la experiencia adquirida en la aplicación de dos esquemas distintos en un RI especializado en Ciencias de la Tierra que maneja datos para la investigación y literatura científica, tanto moderna como histórica, basándonos en lo establecido por la *American Geophysical Union* (AGU) de Estados Unidos y los Lineamientos Generales para el Repositorio Nacional (RN) y los Repositorios Institucionales (RI) del Conahcyt,¹ es decir, estructurado bajo

1 Conahcyt, *Lineamientos generales de ciencia abierta*.

diferentes normatividades con el fin de que la consulta por parte del usuario final sea de manera transparente y expedita.

La metodología empleada para la implementación de dos modelos de metadatos en un RI empleando software libre, es utilizar el Protocolo de Iniciativa de Archivos Abiertos para la Recolección de Metadatos *Open Archives Initiative Protocol for Metadata Harvesting* (OAI-PMH), el cual facilita el uso de varios prefijos que permiten acceder a los diferentes recursos, dependiendo de la naturaleza del objeto. Para ello es necesario conocer plenamente la información estructurada bajo un marcaje que posibilita la creación de los registros con el fin de ofrecer un producto normalizado y a disposición de cualquier acceso y de libre consulta.

Podemos señalar entonces que los RI pueden manejar diferentes conjuntos de metadatos de acuerdo con los requerimientos técnicos, al tipo de información incluida mediante la configuración en el software de administración a través de contextos y a la definición de la consulta del contenido libre.

LOS REPOSITARIOS

Se han desarrollado proyectos internacionales que apoyan el Acceso Abierto a la investigación. Tal es el caso de *Digital Repository Infrastructure Vision for European Research* (DRIVER), *Publisher and Institutional Repository Usage Statistics* (PIRUS) para unificar estadísticas y EuroCRIS, para sistemas de gestión de la investigación y los repositorios, los cuales sirven para mejorar la cooperación local, regional, nacional e internacional.

En México, en 2014, se crea la iniciativa que logra integrar en el capítulo 10 de Ciencia y Tecnología sobre acceso abierto a la información científica y tecnológica, que incluye la creación del Programa de Repositorios del Consejo Nacional de Humanidades, Ciencia y Tecnología (Conahcyt), con el fin de conformar el Repositorio Nacional (RN), y una serie de RI estructurados bajo una política pública con miras a un crecimiento científico nacional y a una apertura de la ciencia para todos los niveles, tanto de

literatura como de datos generados con recursos públicos. Se buscó así maximizar la disseminación de los recursos de información científica y tecnológica del país, y convertirse en un escaparate importante para toda la producción científica de las publicaciones académicas que se incluyan, así como de datos que producen las entidades que están inmersas en el proyecto nacional.

Bajo el contexto de la ciencia abierta, y con base en el desarrollo que tiene esta nueva infraestructura de servicios que ofrecen los repositorios, existen diferentes tipos de herramientas que recuperan información y datos disponibles en su mayoría en acceso abierto o, en algunos casos, con restricciones por las características de la publicación.

De acuerdo con Castro,² podemos encontrar diversos tipos de repositorios que apoyan y contribuyen en la investigación científica, entre los que podemos mencionar:

- Repositorios digitales temáticos o disciplinarios.
- Repositorios institucionales.
- Repositorios huérfanos (*homeless*).
- Repositorios de datos de investigación (RDR, *Research Data Repositories*).
- Repositorio de datos para la investigación (DRR, *Data Repositories for Research*).
- Los recolectores de repositorios.

A través de estas plataformas se ponen en acceso abierto publicaciones académicas que son dictaminadas y financiadas por convocatorias dirigidas a instituciones, así como los datos que producen las entidades que se dedican a la investigación científica y/o tecnológica. De esta manera, se logra que organismos que producen ciencia interoperen de forma transparente y dinámica.

Es por eso por lo que el impacto de un proyecto nacional en materia de acceso y ciencia abierta fomenta la difusión de

2 M. Castro, “Desarrollo de un Repositorio de datos científicos...”.

investigaciones en revistas, tesis, cuadernos de laboratorio, informes, entre otros, además el uso de datos, software y hardware que proveen los elementos para replicar (y también modificar) los resultados de recolecciones y/o experimentaciones. En cuanto a la publicación de los datos se pueden realizar en el repositorio y además vincularlo con todo un proceso científico con la finalidad de hacer sólidas estas iniciativas en nuestro país y aspirar a un ecosistema de innovación donde la generación del conocimiento científico supere las barreras.

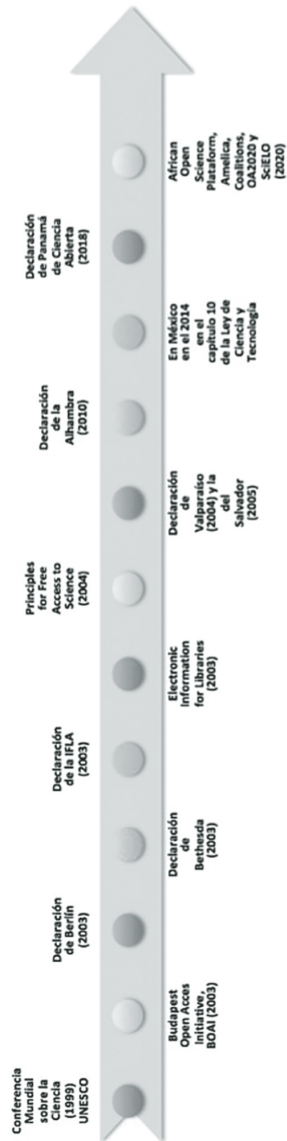
Actualmente en México se cuenta con un proyecto nacional que integra 106 (ciento seis)³ repositorios institucionales de literatura y de datos, en ambos casos interoperables con el Repositorio Nacional, por lo que se están generando enlaces y abriendo brechas en la ciencia para que sea consultada por todos los connacionales bajo una línea de libre acceso para la generación de nuevo conocimiento.

Un elemento importante tanto para la ciencia abierta como para el libre acceso son las declaraciones e iniciativas que hacen que de ellas partan los elementos para la consulta de datos y literatura científica, así como para el establecimiento, bajo diversas plataformas de metadatos de utilidad para la estructuración, de RI y otras fuentes de consulta. Cada una de ellas cuenta con elementos propios dependiendo de su origen institucional y fortalecen, apoyan y reforman el avance tanto en acceso como de la ciencia abierta en cuanto a su compartición, así como en el reúso de datos por parte de los investigadores.

En la imagen 1 se presenta una línea de tiempo sobre acceso abierto. Mientras que en la imagen 2 se enuncian los principios de la ciencia abierta.

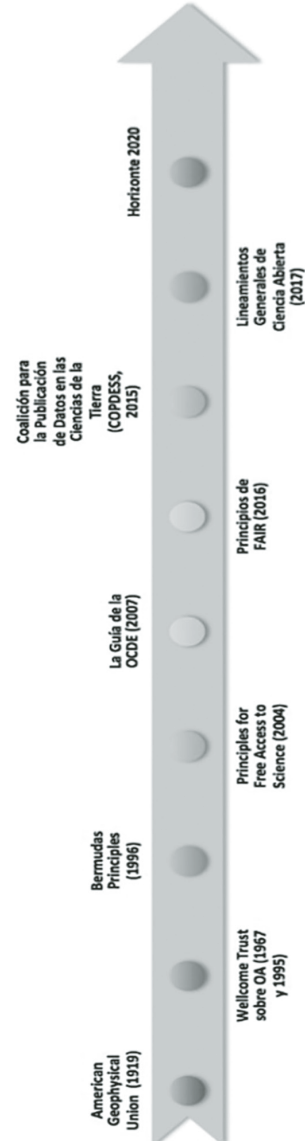
3 Conahcyt, *Repositorio Nacional*.

Imagen 1. Línea de tiempo en Acceso Abierto



Fuente: Castro, 2021.

Imagen 2. Principios de la Ciencia Abierta



Fuente: Castro, 2021.

LOS DATOS CIENTÍFICOS

Hablar de datos científicos en Ciencias de la Tierra y el Espacio para que sean conservados a largo plazo para el futuro uso, es referirnos a la Declaración de la Unión Geofísica Americana (AGU),⁴ la cual busca convertir a los datos en un patrimonio mundial que deben ser conservados, pues ayudan a los científicos a comprender los sistemas terrestre, planetario y heliofísico. Además, deben ser puestos a disposición de la comunidad científica y del público tan pronto como sea posible, de forma accesible, en formatos utilizables, con suficiente documentación y legible que permite su reutilización.

Los procesos de recopilación de información y análisis de los datos en las Ciencias de la Tierra resultan esenciales para la comprensión del entorno natural, pues cambian a cada momento. A diferencia de otros campos científicos que están muy alejados de la vida de la mayoría de las personas, en esta área se tratan los procesos que influyen todos los días en el mundo como las tasas de lluvia, tendencias en las pesquerías marinas, probabilidades de terremotos, potenciales de erupción volcánica, entre tantas otras.

Es por eso que la adhesión a esta política fomenta los avances científicos y produce beneficios económicos que repercuten en la toma de decisiones, a fin de mejorar la seguridad pública y el bienestar social, mismos que contribuyen a la seguridad nacional y mundial al mantener a la sociedad más informada. Sólo en situaciones limitadas de acceso por cuestiones éticas, pero no por propiedad, no podrán ser publicadas. Pero es importante documentar las tendencias y la producción de los datos a largo plazo para comprender muchos fenómenos naturales, debido a que el estado de los sistemas naturales nunca se repite, las pérdidas de datos o las oportunidades de recopilación no pueden ser corregidas.

La AGU es clara en su postura de considerar que el valor de los datos crece con el tiempo, llevando a cabo ventajas en la curación.

⁴ American Geophysical Union, *Declaration adopted by the American Geophysical Union*.

Por lo tanto, los datos y la documentación asociada deben depositarse en un repositorio de la especialidad, pues son parte integral de los programas operativos y de la investigación, y permitirán así que la comunidad científica tenga un amplio conjunto de datos para la realización de sus investigaciones en beneficio de la sociedad y de la ciencia. Uno de sus pilares es la apertura de los datos de investigación, lo que se está volviendo una demanda generalizada para los gobiernos y financiadores.

Se comenta que maximiza la transparencia y la fiabilidad de los datos, lo que favorece la verificación de los resultados de investigación y aumenta el impacto y la visibilidad de ésta. Ello da pie a promover proyectos de investigación de los que provienen datos nuevos y de algunas publicaciones, lo que reduce en gran medida los costos o inversiones, evitando con ello la duplicación de datos.

Al hablar del desarrollo de la ciencia en México, y sobre todo al referirnos a las Ciencias de la Tierra, su producción científica nacional es de aproximadamente 25 500 (veinticinco mil quinientos) documentos⁵ (equivalente al 42% de la producción que se realiza en América Latina), pero es difícil su acceso a ella, ya sea a través de plataformas o estructuras, aunándole si este acceso es por suscripción, registro, pago o que sea financiado con recursos públicos. Esta área científica ha crecido rápidamente en la última década, convirtiéndose en materia prima con un parámetro de desarrollo económico entre los países, en donde se invierten grandes cantidades de recursos para su desarrollo por ser prioritaria para la sociedad.

LOS PRINCIPIOS RECTORES DE FAIR

Hablar de estos principios es referirse a la gestión de datos científicos y la administración tal como lo establece la AGU en su

5 Elaboración de búsqueda de información en el Sistema de Administración del Conocimiento *Web of Science* sobre la producción científica en Ciencias de la Tierra, 30 de enero, 2020.

Declaración. Se requiere poner a disposición de forma accesible, en formatos utilizables, con suficiente documentación respaldando a la reutilización de datos académicos y el grupo de personas interesadas, que representan a la academia, la industria, las agencias de financiamiento y los editores académicos.

Se enfatiza en desarrollar estándares que conectan a investigadores, editores y repositorios de datos para habilitar bajo el Modelo de Datos FAIR⁶ (localizables, accesibles, interoperables y reutilizables) a gran escala. Apunta también a acelerar el descubrimiento científico y mejorar la integridad, transparencia y reproducción de estos.

Dentro de estos principios, están los habilitadores como la AGU, *re3data* y *DataCite*, quienes decidieron desarrollar una nueva herramienta que apoye a los investigadores a encontrar un repositorio adecuado para sus datos, como el *Repository Finder*. Este buscador de repositorios es un proyecto liderado por la Unión Geofísica Americana (AGU) en asociación con *DataCite* y la comunidad de Ciencias de la Tierra, el espacio y el medio ambiente.

Por ello los principios rectores de los datos deben ser accesibles, interoperables y reutilizables no sólo para las personas, sino también para las máquinas con identificadores y metadatos únicos y persistentes, apropiados para ayudar al descubrimiento y ser citados de manera equivalente a otros resultados académicos, mediante un protocolo estándar basado en la web, proporcionando la información de procedencia y el licenciamiento de uso y sean interoperables técnica y semánticamente. Los datos según su caso además deberán estar vinculados de forma segura a las publicaciones asociadas.

6 Simon Hodson, *Hacer realidad los datos de FAIR...*

METADATOS PARA DATOS Y LITERATURA

Para emplear metadatos en un repositorio institucional ya sea de datos y/o de literatura, deben discutirse normativas, técnicas y elementos que definirán el registro y la especificidad de cada uno de los objetos digitales albergados. Estos son indispensables para que los contenidos sean recuperables mediante herramientas de búsqueda en Internet.

Para ello, son necesarios metadatos estructurados que describan el contenido, la calidad, condición y características que expliquen y localicen el recurso de información, y, una vez recuperado, usarlo y administrarlo. Al trabajar con la estructura técnica del repositorio se establece una comunicación entre las diversas plataformas con las que decidamos interactuar.

Es importante dejar claro que se pueden emplear diferentes tipos de metadatos según la naturaleza del repositorio, ya sea de datos, literatura o mixto. En ellos se deben tomar en cuenta la descripción de contenido, calidad, condiciones, entre otros aspectos.

La tipificación de los elementos en los metadatos, generalmente se reconocen de acuerdo con la descripción del objeto digital: administrativos, descriptivos, de preservación, técnicos y de uso. Debemos asegurarnos de que los metadatos mantendrán su capacidad de interpretación y utilidad para que los registros puedan almacenarse y conservarse a lo largo del tiempo.

Las funciones de los metadatos facilitan la organización de los recursos basados en la audiencia permitiendo organizarlos mediante una identificación digital [*Uniform Resource Locator* (URL), *Persistent Uniform Resource Locator* (PURL) o un *Digital Object Identifier* (DOI)] y estructura, asegurando su preservación. Y pueden ser expresados en *Extensible Markup Language* (XML), lenguaje independiente del software e intercambiable entre distintos sistemas. Se expresa a partir de una Definición de Tipos de Documentos (DTD), especificando las etiquetas que se emplearán.

En cuanto a formatos de metadatos se presentan varios destinados a diversos objetos digitales tanto a publicaciones académicas como a datos de investigación, entre los que tenemos: *Dublin*

Core, Metadata Object Description Schema (MODS), Metadata Encoding and Transmission Standard (METS), Metadata for Images in XML (MIX Schema), Data Dictionary for Preservation Metadata (Premis), Learning Object Metadata (LOM), Open Microscopy Environment (OME-XML), Minimum Information for Biological and Biomedical Investigations (MIBBI), Agricultural Metadata Element Set (Agmes), DataCite, Common European Research Information Format (CERIF), Data Catalog Vocabulary (Dcat); en el área multidisciplinar, Ecological Metadata Language (EML), Geographic Information – Metadata (ISO 19115), Open Archives Initiative Object Reuse (OAI-Ore).

Ahora bien, los metadatos en el área de las Ciencias de la Tierra y del Espacio son:

- ISO 19139 e ISO 19139 GML 3.2., cumple el estándar de la información, así como también archivos exportados utilizan el espacio de nombre GML 3.2 y, por consiguiente, pueden validarse con las versiones de los Esquemas XML geográfica, Astronomy Visualization Metadata (AVM), Common Information Model (CIM), entre otros.
- Metadatos *Federal Geographic Data Committee - Content Standard for Digital Geospatial Metadata* (FGDC – CSDGM), permite ver y editar metadatos de acuerdo con las directrices Estándar de Contenidos Geoespaciales Digitales de CSDGM; permite exportar metadatos en el formato XML de esta norma y validarlos con CSDGM XML DTD [*CSDGM Document Type Declaration* (DTD) – usado para validar CSDGM XML].

La aplicación de estos puede ser en objetos digitales como videos, imágenes, mapas, recursos educativos, además de recursos geoespaciales, datos geográficos, entre otros.

En cuanto a los metadatos que empleamos en los RI en el área de Ciencias de la Tierra para los datos fueron los *Metadatos para datos*

para los repositorios institucionales ver. 3.0.80 (Lineamientos Específicos para Repositorios del Apéndice 2).⁷

Pero para las Ciencias de la Tierra los datos tienen condicionantes. Por ejemplo, se componen de varios archivos en diversos formatos, por lo que es necesario relacionar el dato que está siendo registrado, que en ocasiones, por nacimiento o naturaleza, se relaciona con otro recurso. Es decir, se tiene que llevar a cabo una asociación. Por lo que además consideramos para el RI las siguientes características:

- Descripción de los datos y formato y versión.
- Cómo se adquieren y procesan.
- Relación con otros datos.
- Metadatos de relación o con los que se requieran para que los datos posean sentido y guarden su contexto.
- Políticas para compartir, acceder y reusar los datos.
- Citación.
- Responsabilidad y roles.
- Derechos de autor.
- Almacenamiento.
- Licencia abierta para acceder libremente a la información.
- Los datos deberán estar estructurados en un formato legible por máquina.
- Se debe depositar en formatos no propietarios, es decir, datos abiertos, estructurados y sin propietarios y con nombres.
- Emplear *Uniform Resource Identifier* (URI) con otras fuentes de datos y aplicaciones, facilitando la identificación.

También pueden consultar el anexo 1, para la visualización de algunas otras características de los metadatos empleados.

Los metadatos que se utilizaron para la parte de literatura fue el modelo de *Dublin Core*, apegado a las estipulaciones de los

7 INFOTEC, *Interoperabilidad con el Metabuscador del Repositorio Nacional Versión 3.1. México*.

Lineamientos Específicos para repositorios, publicados en el sitio del Repositorio Nacional, haciendo uso de la notación de *Dublin Core* básica, no incluyendo calificadores así como también en las Directrices de OpenAIRE 3.0, que complementan a *Dublin Core* con la importancia de incluir en el encabezado del registro cualquier fecha relacionada a su gestión, no en la etiqueta `<dc:date>` y sobretodo utilizando los catálogos de autoridad proporcionados por el Conahcyt;⁸ ello con sus elementos opcionales y repetibles, codificados en *HyperText Markup Language* (HTML), *Extensible Markup Language* (XML) y *Resource Description Framework* (RDF), formatos más divulgados y aceptados para representación y recuperación de información en la web. Para visualizarlos consultar el anexo 2.

LAS DIRECTRICES DE OPENAIRE⁹

Se consideran estas directrices, por el apoyo que brindan a los administradores de los repositorios que difunden publicaciones, conjuntos de datos, *Current Research Information Systems* (sistemas CRIS) y repositorios de otros productos de investigación, a través del protocolo *Open Archive Initiative-Protocol for Metadata Harvesting* (OAI-PMH), para integrarse con la infraestructura de *Open Access Infrastructure for Research in Europe* (OpenAIRE) y se enlistan a continuación:

- Repositorio de literatura:
 - Perfil de aplicación y un esquema basado en *Dublin Core* y *DataCite* que incluye un nuevo OAI-metadata-Prefix (*Open Archives Initiative- metadataPrefix*).

8 Simon Hodson, *Hacer realidad los datos de FAIR...*

9 J. Schirrwagen, *OpenAIRE Guidelines for institutional and thematic repository managers 4.0*.

- Soporte de esquemas de identificación para autores, organizaciones, financiadores, recursos académicos.
- Introducción de *COAR Controlled Vocabularies* (Vocabularios Controlados de COAR).
- Cumplimiento de la Política de Adquisición de Contenido de OpenAIRE, publicada el 5 de octubre de 2018 COAR.
- Repositorio para datos: El contenido del repositorio debe ser configurado bajo el *setSpec openaire_data*. El recolector usará el valor *setSpec* para realizar la recolección selectiva y determinará su comunicación, por lo que recolectará los conjuntos de datos, cumpliendo algunos de los siguientes criterios:
 - El conjunto de datos es el resultado de un proyecto de investigación ya sea financiado o no.
 - Los datos deberán vincularse a una publicación en el espacio de información de OpenAIRE.
 - Los dos anteriores puntos se deben o pueden determinar por la infraestructura de OpenAIRE.

Las versiones anteriores de las Directrices presentan cambios importantes como un perfil y esquema de aplicación basados en *Dublin Core* y *DataCite*, un nuevo OAI-metadataPrefix, apoyo de esquemas de identificación para autores, organizaciones, financiadores, recursos académicos e introducción de vocabularios controlados por COAR con el Cumplimiento de la Política de Adquisición de Contenido de OpenAIRE.¹⁰

INTEROPERABILIDAD

Para poder llevar a cabo este proceso se deben diseñar las estructuras técnicas bajo un estándar y deben normalizarse bajo su

10 INFOTEC, *Interoperabilidad con el Metabuscador del Repositorio Nacional Versión 3.1. México*.

función principal, que es mover metadatos de un servidor local a un agregador con la finalidad de reunirlos con otros metadatos y proporcionar un servicio de valor agregado. Es importante mencionar que se requiere de un acuerdo entre el proveedor de metadatos y el cosechador.

Ante esta situación, el repositorio institucional encargado de preservar la información debe estar configurado de tal manera que se pueda acceder a cualquier elemento sin importar la categoría a la cual pertenece, bien sea de datos o literatura.

Para cumplir con este objetivo en la plataforma *DSpace*, utilizando el Protocolo de Iniciativa de Archivos Abiertos para la Recolección de Metadatos, OAI-PMH (*Open Archives Initiative Protocol for Metadata Harvesting*), es posible hacer uso de prefijos que permitan acceder a los diferentes recursos, dependiendo de la naturaleza de la información, a través de los proveedores de datos que construyen servidores que atienden peticiones vía *Hypertext Transfer Protocol* (HTTP) y producen respuestas en XML.

De igual forma, se debe establecer una configuración de contestación con base en su prefijo de metadatos (debiendo ser *Datacite* para los datos y *Dublin Core* para la literatura, según sea el caso); no todos los elementos que integran el RI deben ser de la misma naturaleza, por ejemplo: revistas, artículos de investigación, boletines, monografías, libros, folletos, entre otros, son elementos que refieren a la producción. Por otro lado, existen conjuntos de datos que forman parte de la investigación de carácter científico, tales como: valores de temperatura, cantidad de sustancia, longitud de desplazamiento, entre otros, los cuales difícilmente aparecen bajo un mismo formato y con una adecuada indización, recuperación e interoperabilidad para su libre acceso y despliegue de contenidos.

De esta manera, para acceder a los metadatos de los recursos de literatura, se puede utilizar el prefijo *oai dc*, mientras que los registros de datos utilizan el prefijo *oai datacite*.

Es importante mencionar que el prefijo únicamente sirve para realizar la solicitud de los metadatos en el formato correspondiente, sin embargo, no existe validación alguna por parte de *DSpace* o del OAI-PMH, por lo que, al momento de solicitar información

para la cosecha de recursos, es indispensable emplear un *contexto*, uno para cada caso, es decir, un contexto para literatura y otro para los datos.

OAI-PMH por *default* incluye tres contextos:

- *Default Context*. Contexto por defecto que viene pre-cargado en la configuración del OAI-PMH de *DSpace*, que generalmente se utiliza para recursos de tipo literatura, pues responde utilizando el prefijo *oai_dc*.
- *Driver Context*. Se utiliza generalmente para exponer elementos del repositorio clasificados como de acceso abierto; no todos los elementos disponibles en un repositorio son de acceso abierto, además de ello, se deben cumplir con las siguientes etiquetas: *dc.date.issued*, que corresponde a la fecha de publicación del elemento, *dc.language*, que debe apegarse al estándar ISO639-3, y la etiqueta *dc.type*, cuyo valor debe coincidir con alguno de los 16 valores establecidos por la *Digital Repository Infrastructure Vision for European Research* (DRIVER).
- *OpenAIRE Context*. Destinado a la exposición de elementos que cumplen exclusivamente con los lineamientos establecidos por OpenAIRE.

Como se puede apreciar, los contextos provistos por *DSpace* tienen ciertas limitantes, que no necesariamente empatan con los recursos que se desean publicar. Por ello, la creación de uno o más contextos que se adapten a las necesidades de cada institución, se perfila como la opción más adecuada, pues ofrece múltiples beneficios, entre ellos, diseñar un esquema de metadatos que se adapte a los requerimientos solicitados, permitiendo aprovechar la misma infraestructura para reunir recursos que son de naturalezas diferentes. En otras palabras, el uso de diferentes contextos permite tener un repositorio híbrido, capaz de almacenar elementos tanto de literatura como de datos.

LA CONFIGURACIÓN DEL CONTEXTO EN *DSPACE* Y OAI-PMH

El archivo de configuración encargado de la gestión de contextos y esquemas de metadatos se ubica en *[dspace]/config/crosswalks/oai/xoai.xml*. Este archivo escrito en XML permite definir nuevos contextos, de igual manera, es posible definir el esquema de metadatos a utilizar en cada uno de ellos, así como personalizar los filtros y las reglas que se deberán aplicar en cada contexto. Es importante recurrir a la documentación oficial de OAI 2.0 Server,¹¹ pues la configuración del contexto requiere de cierta familiarización con los diferentes elementos que lo conforman, tales como el transformador, el formato de metadatos, filtros y conjuntos de datos.

Para desarrollar la configuración de DSpace y OAI-PMH, es necesario realizar los siguientes pasos, en donde se debe delimitar claramente el tipo de información que se incluye, ya sean datos o literatura:

1. Definición de un nuevo contexto utilizando el *prefijo oai_dc* para los registros de literatura.
2. Definición del esquema de metadatos.
3. Definición del filtro de literatura.
<Filter id="LiteFilter">
<Definition>

Con ayuda de una etiqueta auxiliar, se filtran todos aquellos elementos que cumplan con una serie de elementos, como:

```
<string name="field">dcterms.format</string>  
<string name="operator">contains</string>  
<string name="value">Literatura</string>
```

11 Atlassian Confluence Open Source Project, *OAI 2.0 Server*.

En la imagen 3, se presenta un listado de los contextos disponibles en el Repositorio Institucional de Datos del Servicio Sismológico Nacional (RDSSN).

De igual manera, es posible realizar cualquier ajuste o modificación al contexto por *default* de DSpace y OAI-PMH, “request”, el cual, en el caso particular del RDSSN, se utiliza para poder servir todos los recursos que son catalogados como datos.

4. Definición del contexto predeterminado “request” para datos
5. Definición del esquema de metadatos para el contexto por default “request”
<Format id=”datacite”>
6. Definición del filtro para el contexto por default “request”
<Filter id=”defaultFilter”>
</Definition>
7. Regla auxiliar para la discriminación de elementos de tipo “evento”
<!-- Para filtrar los sismogramas (Event). -->
<CustomCondition id=”ExcluyeEvents”>

Estos elementos señalados tienen que llevarse a cabo bajo una división clara de ¿qué estamos trabajando? y ¿qué tipo de datos estamos incluyendo en nuestro repositorio (de investigación o para la investigación)?, para con ello hacer la estructura de cosecha, consulta y despliegue de información.

CONSIDERACIONES Y REFLEXIONES FINALES

Para finalizar, podemos comentar que los metadatos para un repositorio en Ciencias de la Tierra se deben llevar a cabo mediante una normatividad, directrices y políticas, así como con estándares internacionales que cubran el valor tanto de los datos como de la producción científica y ambos puedan ser compartidos bajo ciencia abierta.

Imagen 3. Listado de los contextos disponibles en el RDSSN

Available Contexts

Default Context This is the default context of the DSpace OAL-PMH data provider.				
Identify	List Sets	List Metadata Formats	List Identifiers	List Records
Driver Context This context complies with Driver rules.				
Identify	List Sets	List Metadata Formats	List Identifiers	List Records
OpenAIRE Context This context complies with OpenAIRE Guidelines for Literature Repositories v3.0.				
Identify	List Sets	List Metadata Formats	List Identifiers	List Records
Contexto de Literatura. Contexto de tipo documental del RDSSN.				
Identify	List Sets	List Metadata Formats	List Identifiers	List Records

Fuente: Universidad Nacional Autónoma de México.
Servicio Sismológico Nacional, 2019.

Para llevar a cabo esta tarea, es imprescindible analizar estructuras técnicas correspondientes para asignar los metadatos adecuados a los datos y estos puedan contar con un contexto e información relevante y se puedan fusionar con más fuentes de información generando nuevo conocimiento en otros ámbitos y a su vez crear valor adicional y, por otro lado, en cuanto a la literatura ofrecer producción científica en abierto al alcance de la sociedad.

Como se puede ver a lo largo del trabajo es evidente que al contribuir en la línea de lo abierto se relacionan cuatro factores clave que se requieren tanto para generar ciencia como para incrementarla y difundirla, los cuales son: los RI tanto de literatura y datos, los investigadores y una comunidad que requiere consultar información sobre esta temática y generar nuevo conocimiento.

Es importante mencionar que, dentro de las Ciencias de la Tierra, el Mar y la Atmósfera se generan una gran cantidad de información en lapsos muy cortos, por lo que se necesitan infraestructuras en donde los metadatos sirvan para suministrar información sobre lo que se produzca en datos y en la parte de la literatura, lo cual facilitará la descripción de los contenidos, asegurará la calidad, y ofrecerá condiciones de uso. Con ello se mostrará el contexto en donde se encuentra la información, garantizando su disponibilidad y preservación.

Es necesario trabajar junto con otras comunidades de partes interesadas (investigadores, universidades y financiadores de la ciencia) con el fin de continuar con esta misión de ponerlo en abierto y valorar aún más inclusión, diversidad, confianza e innovación. Asimismo, los metadatos deben estar eficientemente conformados y establecidos bajo parámetros internacionales que ofrezcan una garantía de cosecha e intercambio de información y datos.

El uso de los metadatos dentro de estas herramientas, como son los repositorios, apoya en la representación de los proyectos de investigación y ofrece un dominio específico o un ámbito que se requiere comprender para un proceso de investigación que necesitan los especialistas. Los investigadores no conocen los esquemas de metadatos ni la lógica del acceso, por lo que nuestra contribución como profesionales de la información es primordial

dentro de estas actividades tanto académicas como científicas para seguir en las líneas de la ciencia abierta.

BIBLIOGRAFÍA

- American Geophysical Union. *Declaration adopted by the American Geophysical Union*. 2019. <https://ethics.agu.org/>.
- Atlassian Confluence Open Source Project. *OAI 2.0 Server*. <https://wiki.lyrasis.org/display/DSDOC6x/OAI+2.0+Server>.
- Castro Escamilla, M. “Las declaraciones, principios e iniciativas de la Ciencia Abierta: una visión”. En *Ciencia Abierta Opciones y experiencias para México y Latinoamérica*. Coordinado por Monserrat García Guerrero *et al.* Barcelona: Octaedro. Colección: Universidad. 2021.
- . “Desarrollo de un Repositorio de datos científicos de apoyo a la investigación: el caso de las ciencias de la Tierra”. Tesis de maestría. México: UNAM. 2020. <http://132.248.9.195/ptd2020/agosto/0802846/Index.html>.
- Conahcyt. *Lineamientos generales de Ciencia Abierta*. México: Conahcyt. 2017.
- . *Repositorio Nacional*. México: Conahcyt. 2021. [https:// 2021. www.repositorionacionalcti.mx/](https://2021.www.repositorionacionalcti.mx/).
- Hodson, Simon. *Hacer realidad los datos de FAIR... y los desafíos de la interoperabilidad y la reutilización*. s. f. <http://www.eddi-conferences.eu/ocs/index.php/eddi/eddi18/paper/view/383>.
- INFOTEC. *Interoperabilidad con el Metabuscaador del Repositorio Nacional Versión 3.1*. México. 2018.

Schirrwagen, J. *OpenAIRE Guidelines for institutional and thematic repository managers 4.0*. 2018.

Universidad Nacional Autónoma de México. Servicio Sismológico Nacional. *Repositorio Institucional de Datos del Servicio Sismológico Nacional*. 2019. <http://rigeofisica.ssn.unam.mx/jspui/>.

Anexo 1.
Ejemplo de etiquetado xml para datos

```
<oai_datacite xmlns:dc="http://purl.org/dc/elements/1.1/"
xmlns:doc="http://www.lyncode.com/xoai
xmlns="http://schema.datacite.org/oai/oai-1.1"
xsi:schemaLocation="http://schema.datacite.org/oai/oai-1.1
http://schema.datacite.org/oai/oai-1.1/oai.xsd">
<schemaVersion>2.1</schemaVersion>
<datacentreSymbol>TIB.WDCC</datacentreSymbol>
<payload>
<resource xmlns="http://datacite.org/schema/kernel-2.1"
xsi:schemaLocation="http://datacite.org/schema/kernel-2.1
http://schema.datacite.org/meta/kernel-2.1/metadata.xsd">
<identifier identifierType="Handle">http://rigeofisica.ssn.unam.mx/jspui/handle/ssn-
mx/892</identifier>
<creators>
<creator>
<creatorName>Servicio Sismológico Nacional (IGEF-UNAM)</creatorName>
<nameIdentifier nameIdentifierScheme="RNCTIMX"
schemeURI="http://repositorionacionalcti.mx/">
info:eu-repo/dai/mx/rn/3190557
</nameIdentifier>
</creator>
</creators>
<contributors>
<contributor contributorType="ProjectLeader">
<contributorName>XYOLI PEREZ CAMPOS</contributorName>
<nameIdentifier nameIdentifierScheme="RNCTIMX"
schemeURI="http://repositorionacionalcti.mx/">
info:eu-repo/dai/mx/or-
cid/0000-0001-8970-7966
</nameIdentifier>
</contributor>
<contributor contributorType="ProjectMember">
<contributorName>SAÚL ARMENDÁRIZ SÁNCHEZ</contributorName>
<nameIdentifier nameIdentifierScheme="RNCTIMX"
schemeURI="http://repositorionacionalcti.mx/">
info:eu-repo/dai/mx/orcid/0000-0001-5229-0642
</nameIdentifier>
</contributor>
<contributor contributorType="ProjectMember">
<contributorName>Jesus Antonio Perez Santana</contributorName>
```

```
<nameIdentifier nameIdentifierScheme="RNCTIMX"
  schemeURI="http://repositorionacionalcti.mx/">
  info:eu-repo/dai/mx/or-
cid/0000-0002-0261-0836
</nameIdentifier>
</contributor>
<contributor contributorType="ProjectMember">
<contributorName>Caridad Cardenas Monroy</contributorName>
<nameIdentifier nameIdentifierScheme="RNCTIMX"
  schemeURI="http://repositorionacionalcti.mx/">
info:eu-repo/dai/mx/orcid/0000-0001-7268-979X
</nameIdentifier>
</contributor>
<contributor contributorType="ProjectMember">
<contributorName>Miguel Ángel Sosa Jiménez</contributorName>
<nameIdentifier nameIdentifierScheme="RNCTIMX"
  schemeURI="http://repositorionacionalcti.mx/">
info:eu-repo/dai/mx/orcid/0000-0002-2618-2151
</nameIdentifier>
</contributor>
<contributor contributorType="Producer">
<contributorName>UNAM, IGEF, Grupo de Trabajo del SSN</contributorName>
<nameIdentifier nameIdentifierScheme="RNCTIMX"
  schemeURI="http://repositorionacionalcti.mx/">
info:eu-repo/dai/mx/rn/3190471
</nameIdentifier>
</contributor>
<contributor contributorType="Producer">
<contributorName>Servicio Sismológico Nacional,
Estación de Red Convencional Acapulco,
Guerrero,
ACX</contributorName>
<nameIdentifier nameIdentifierScheme="RNCTIMX"
  schemeURI="http://repositorionacionalcti.mx/">
info:eu-repo/dai/mx/rn/3192706
</nameIdentifier>
</contributor>
</contributors>
<identifier identifierType="URL">http://www.sismoteca.unam.mx/sn/pdf/8143.pdf</
identifier>
<subjects>
<subject subjectScheme="LCC"
schemeURI="http://id.loc.gov/authorities/subjects/sh85119634">
Sismología</subject>
```

```
<subject subjectScheme="GeoRef"
schemeURI="https://pubs.geoscienceworld.org/georef/term/5/30320/seismology">
Sismología</subject>
<subject subjectScheme="UNESCO"
schemeURI="http://vocabularies.unesco.org/thesaurus/concept8601">
Sismología</subject>
</subjects>
<sizes>
<size>1 MB</size>
</sizes>
<dates>
<date dateType="Created">2011-03-08</date>
</dates>
<descriptions>
<description descriptionType="Methods">Entrada marzo 08, 2011 at 13:00 p.m.
(GMT) / Salida marzo 09, 2011 at 12:52 p.m. (GMT)</description>
</descriptions>
<titles>
<title>Sismograma de la estación Acapulco.</title>
</titles>
<publisher>UNAM, IGEF, SSN, Grupo de Trabajo</publisher>
<subjects>
<subject subjectScheme="CTI">1</subject>
</subjects>
<subjects>
<subject subjectScheme="CTI">25</subject>
</subjects>
<subjects>
<subject subjectScheme="CTI">2507</subject>
</subjects>
<subjects>
<subject subjectScheme="CTI">250705</subject>
</subjects>
<publicationYear>2011</publicationYear>
<resourceType resourceTypeGeneral="Event">Sismogramas</resourceType>
<relatedIdentifier relatedIdentifierType="DOI"
relationType="IsSupplementTo">10.21766/SSNMX/SP/MX</relatedIdentifier>
<rightsList>
<rights rightsURI="info:eu-repo/semantics/openAccess" />
<rights rightsURI="http://creativecommons.org/licenses/by-nc-nd/4.0">
CC BY-NC-ND
</rights>
</rightsList>
<geoLocations>
```

```
<geoLocation>
<geoLocationPoint>-99.9167 16.8667</geoLocationPoint>
<geoLocationBox>-99.8667 16.8167 -99.9667 16.9167</geoLocationBox>
<geoLocationPlace>Guerrero</geoLocationPlace>
</geoLocation>
</geoLocations>
<language>spa</language>
<formats>
<format>application/pdf</format>
</formats>
<descriptions>
<description descriptionType="Abstract">Presenta registro de movimiento sísmico.</description>
</descriptions>
<descriptions>
<description descriptionType="Abstract">Estación telemática que envía su señal a la estación central localizada en el Instituto de Geofísica de la UNAM y pertenece a la red sísmológica convencional.</description>
</descriptions>
<descriptions>
<description descriptionType="Abstract">Componente Z.</description>
</descriptions>
</resource>
</payload>
</oai_datacite>
Anexo 2 ejemplo de etiquetado xml para literatura
<oai_dc:dc xmlns:dc="http://purl.org/dc/elements/1.1/"
xmlns:doc="http://www.lyncode.com/xoai"
xmlns:oai_dc="http://www.openarchives.org/OAI/2.0/oai_dc/"
xsi:schemaLocation="http://www.openarchives.org/OAI/2.0/oai_dc/
http://www.openarchives.org/OAI/2.0/oai_dc.xsd">
<dc:title>Intraslab earthquake of 16 june 2013 (M-w 5.9), one of the closest such
events to Mexico City</dc:title>
<dc:creator id="info:eu-repo/dai/mx/orcid/0000-0002-0219-9813">Shri Krishna
Singh</dc:creator>
<dc:subject>FUENTE SPECTRA</dc:subject>
<dc:subject>TERREMOTOS INSLAB</dc:subject>
<dc:subject>NORTEAMÉRICA</dc:subject>
<dc:subject>FUERTE MOVIMIENTO</dc:subject>
<dc:subject>ZONA</dc:subject>
<dc:subject>AMPLIFICACIÓN</dc:subject>
<dc:subject>ATENUACIÓN</dc:subject>
<dc:subject>SUBDUCCIÓN</dc:subject>
<dc:subject>GEOMETRÍA</dc:subject>
```

```
<dc:subject>info:eu-repo/classification/LCC/Seismology</dc:subject>
<dc:subject>info:eu-repo/classification/cti/1</dc:subject>
<dc:subject>info:eu-repo/classification/cti/25</dc:subject>
<dc:subject>info:eu-repo/classification/cti/2507</dc:subject>
<dc:subject>info:eu-repo/classification/cti/250705</dc:subject>
<dc:subject>info:eu-repo/classification/GeoRef/Seismology</dc:subject>
<dc:subject>info:eu-repo/classification/UNESCO/concept8601</dc:subject>
<dc:date>info:eu-repo/date/embargoEnd/2025-03-01</dc:date>
<dc:date>2014-03-01</dc:date>
<dc:type>info:eu-repo/semantics/article</dc:type>
<dc:type>info:eu-repo/semantics/publishedVersion</dc:type>
<dc:identifier>10.1785/0220130179</dc:identifier>
<dc:identifier>http://rigeofisica.ssn.unam.mx/jspui/handle/ssnmx/43</dc:identifier>
<dc:language>eng</dc:language>
<dc:relation>SEISMOLOGICAL RESEARCH LETTERS</dc:relation>
<dc:relation>0895-0695</dc:relation>
<dc:relation>Singh, S. K.; X. Perez-Campos; V. H. Espindola; V. M. Cruz-Atienza and
A. Iglesias. 2014. "Intraslab Earthquake of 16 June 2013 (M-W 5.9), One of the Clo-
sest Such Events to Mexico City." Seismological Research Letters, 85(2), 268-277.</
dc:relation>
<dc:rights>info:eu-repo/semantics/embargoedAccess</dc:rights>
<dc:rights>http://creativecommons.org/licenses/by-nc-nd/4.0</dc:rights>
<dc:format>pdf</dc:format>
<dc:publisher>Seismological Society of America</dc:publisher>
<dc:source>https://pubs.geoscienceworld.org/ssa/srl/article/85/2/268/315346</
dc:source>
</oai_dc:dc>
```

Anexo 2.

Ejemplo de etiquetado XML para literatura

```
<oai_dc:dc xmlns:dc="http://purl.org/dc/elements/1.1/"
  xmlns:doc="http://www.lyncode.com/xoai"
  xmlns:oai_dc="http://www.openarchives.org/OAI/2.0/oai_dc/"
  xsi:schemaLocation="http://www.openarchives.org/OAI/2.0/oai_dc/
    http://www.openarchives.org/OAI/2.0/oai_dc.xsd">
  <dc:title>Intraslab earthquake of 16 june 2013 (M-w 5.9), one of the closest such
```

```
events to Mexico City</dc:title>
<dc:creator id="info:eu-repo/dai/mx/orcid/0000-0002-0219-9813">Shri Krishna
Singh</dc:creator>
<dc:subject>FUENTE SPECTRA</dc:subject>
<dc:subject>TERREMOTOS INSLAB</dc:subject>
<dc:subject>NORTEAMÉRICA</dc:subject>
<dc:subject>FUERTE MOVIMIENTO</dc:subject>
<dc:subject>ZONA</dc:subject>
<dc:subject>AMPLIFICACIÓN</dc:subject>
<dc:subject>ATENUACIÓN</dc:subject>
<dc:subject>SUBDUCCIÓN</dc:subject>
<dc:subject>GEOMETRÍA</dc:subject>
<dc:subject>info:eu-repo/classification/LCC/Seismology</dc:subject>
<dc:subject>info:eu-repo/classification/cti/1</dc:subject>
<dc:subject>info:eu-repo/classification/cti/25</dc:subject>
<dc:subject>info:eu-repo/classification/cti/2507</dc:subject>
<dc:subject>info:eu-repo/classification/cti/250705</dc:subject>
<dc:subject>info:eu-repo/classification/GeoRef/Seismology</dc:subject>
<dc:subject>info:eu-repo/classification/UNESCO/concept8601</dc:subject>
<dc:date>info:eu-repo/date/embargoEnd/2025-03-01</dc:date>
<dc:date>2014-03-01</dc:date>
<dc:type>info:eu-repo/semantics/article</dc:type>
<dc:type>info:eu-repo/semantics/publishedVersion</dc:type>
<dc:identifier>10.1785/0220130179</dc:identifier>
<dc:identifier>http://rigeofisica.ssn.unam.mx/jspui/handle/ssnmx/43</dc:identifier>
<dc:language>eng</dc:language>
<dc:relation>SEISMOLOGICAL RESEARCH LETTERS</dc:relation>
<dc:relation>0895-0695</dc:relation>
<dc:relation>Singh, S. K.; X. Perez-Campos; V. H. Espindola; V. M. Cruz-Atienza and
A. Iglesias. 2014. "Intraslab Earthquake of 16 June 2013 (M-W 5.9), One of the Clo-
sest Such Events to Mexico City." Seismological Research Letters, 85(2), 268-277.</
dc:relation>
<dc:rights>info:eu-repo/semantics/embargoedAccess</dc:rights>
<dc:rights>http://creativecommons.org/licenses/by-nc-nd/4.0</dc:rights>
<dc:format>pdf</dc:format>
<dc:publisher>Seismological Society of America</dc:publisher>
<dc:source>https://pubs.geoscienceworld.org/ssa/srl/article/85/2/268/315346</
dc:source>
</oai_dc:dc>
```

Áreas de oportunidad ante el nuevo entorno de la catalogación y los metadatos. Instituto de Investigaciones Bibliotecológicas y de la Información/UNAM. La edición consta de 100 ejemplares. Coordinación editorial, Sergio J. Sepúlveda H.; revisión especializada: Marcos Emilio Bustos Flores; corrección de pruebas: Carlos Ceballos Sosa y Marcos Emilio Bustos Flores; formación editorial: Mario Ocampo Chávez. Fue impreso en papel cultural de 90 g en los talleres Gráfica Premier, 5 de Febrero núm. 2309, San Jerónimo Chicahualco, C.P. 52170, Metepec, Estado de México. Se termino de imprimir en 2024.