

Gestión de los datos para la investigación en los repositorios del área de ciencias de la tierra de la UNAM

SAÚL ARMENDÁRIZ SÁNCHEZ

Instituto de Geofísica, Biblioteca Conjunta de Ciencias de la Tierra (UNAM)

asaul@atmosfera.unam.mx

MINERVA CASTRO ESCAMILLA

Instituto de Geofísica, Biblioteca Conjunta de Ciencias de la Tierra (UNAM)

minerva@igeofisica.unam.mx

EJE TEMÁTICO

Datos abiertos

RESUMEN

La Universidad Nacional Autónoma de México (UNAM) cuenta con seis entidades académicas dedicadas a la investigación que generan importantes volúmenes de datos para la investigación con base en diversos proyectos y considerando distintos equipos que se encargan de generarlos y procesarlos para ponerlos a disposición de los usuarios, de una manera abierta a través de sus Repositorios Institucionales, combinando ahora los datos con literatura científica para ampliar su posibilidad de visibilidad, impacto, consulta y recuperación, integrándose con ello al tema de los datos abiertos en México. Con base en ello, el presente documento se encamina a mostrar el proceso de gestión de los repositorios generados por el Instituto de Geofísica de la UNAM (SSN, Geotermia y Clima Espacial) estableciendo desde su planeación, estructuración de metadatos y liberación. Estas tres etapas son fundamentales para que un

repositorio de datos para la investigación se convierta en una herramienta clave para una especialidad, que en nuestro caso son las ciencias de la tierra. De igual manera la descripción de las etapas llevadas a cabo en su diseño, los procesos y el software empleado han sido dos de los elementos que permitieron su desarrollo, para lo cual se expone cada uno de ellos para entender la estructura del trabajo. Por ello, el objetivo principal del trabajo es mostrar todas estas actividades bajo proyectos reales y que se encuentran en operación en estos momentos, apoyando la ciencia no solo nacional sino también mundial.

PALABRAS CLAVE

Gestión de datos para la investigación; ciencias de la tierra; repositorios institucionales; ciencia abierta; literatura científica; marcaje de datos; conversión y migración de datos. Data management for research; Earth sciences; institutional repositories; open science; scientific literature; data marking; data conversion and migration.

Los datos en Ciencias de la Tierra en México

El aumento de la actividad científica produce cambios profundos a nivel mundial con sofisticados avances en Tecnologías de Información y Comunicación y en el incremento de la inversión, tanto pública como privada. El avance en áreas como las Ciencias de la Tierra, el Mar y la Atmósfera, así como en la Astronomía, la Física o las Ciencias Biomédicas y de la Vida está permitiendo la obtención y la recolección de datos, que, además, crecen de manera exponencial. Como consecuencia, tanto investigadores como organizaciones dedicadas a la investigación se encuentran cada vez con más dificultades para gestionar, analizar e interpretar esta producción de datos en grandes volúmenes.

Al hablar del desarrollo de la ciencia en México, y sobre todo al referirnos a las Ciencias de la Tierra en donde su producción científica nacional es de

aproximadamente 33.991¹ documentos indizados (que representa el 47,2 % de la producción que se realiza en América Latina), cuenta con difícil acceso a través de plataformas o estructuras propias o locales, aunándole a ello si el acceso es por suscripción, registro, pago o que sean financiados con recursos públicos.

Esta área de la ciencia 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. Un problema al que se enfrenta nuestro país es la falta de difusión de información científica y tecnológica que se deriva de una serie de investigaciones que se llevan a cabo tanto por entidades públicas, privadas y de investigación.

Bajo este panorama y debido al gran impacto de los diversos fenómenos naturales que han golpeado a México, así como el aumento de datos para la investigación e información académica, se considera necesario facilitar a la sociedad herramientas que permitan la consulta de estos datos de manera libre, oportuna e intuitiva para entender los cambios que sufre el orbe desde el punto de vista de las Ciencias de la Tierra. Temas como el calentamiento global, desarrollo sustentable, cambio climático, variación de oleaje en los mares, movimientos sísmicos, estados geotérmicos en el país, así como de clima espacial, entre otros, son estudiados por los científicos, y en ocasiones como resultado desarrollan infraestructuras tecnológicas en beneficio de la humanidad y las propias ciencias de la tierra.

Ante la producción masiva de datos, se requiere facilitar el acceso, el intercambio, almacenamiento, disseminación, así como la difusión de información especializada en Ciencias de la Tierra, mediante herramientas funcionales que permitan a los investigadores y a la sociedad consultar de manera abierta sus líneas de acción y llevar a cabo un mejor estudio de las variantes que se presentan en México y en otras partes del mundo, a fin de

¹ Resultados de la búsqueda de información en el Sistema de Administración del Conocimiento Web of Science, sobre la producción científica en Ciencias de la Tierra en México. Mayo 11, 2022.

contar con información retrospectiva, actual y en tiempo real, producida en diversas entidades especializadas en Ciencias de la Tierra.

Estas herramientas deben apoyen el desarrollo y la preservación de la ciencia por medio de una serie de procesos y flujos de trabajo de la información (datos), generados por organismos, servicios nacionales, oficiales y de investigación, que abordan temas de prioridad y seguridad nacional, como los desastres naturales, cambio climático y temblores de tierra por ejemplo, los cuales se han incrementado en los últimos años y cada vez son más cercanos en tiempo.

Si consideramos que México cuenta con 28.578² investigadores, según el Portal de Datos del Sistema Nacional de Investigadores (SNI), y 27,09 % de la producción científica en Ciencias de la Tierra es generada por la UNAM, se necesitan mecanismos y recursos que puedan apoyar la difusión de toda esta producción mediante instrumentos que permitan el manejo, estudio, análisis y uso de los datos, lo que proporcionará un fortalecimiento de la información, la cual puede prevenirnos de cualquier eventualidad en un desastre natural.

La literatura científica en apoyo a la ciencia abierta

En los últimos años, el movimiento Open Access (OA) o Acceso Abierto ha tomado fuerza entre las instituciones académicas y científicas, por lo que este concepto se relaciona directamente a publicaciones de investigación, que buscan favorecer y promover el acceso libre y sin restricciones a los trabajos generados por la comunidad científica a nivel mundial. El Acceso Abierto a la literatura académica y documentación de investigación es vital para la comprensión de nuestro mundo y para la búsqueda de soluciones a los desafíos globales, en especial para reducir la desigualdad de la información, garantizando la integridad del sistema de comunicación académica, asegurando que toda investigación estará disponible a perpetuidad.

² Conacyt. Sistema Nacional de Investigadores (SIN). Tabla de Investigadores Vigentes 2018 por Grado, Nivel, Género, Adscripción, Entidad Federativa y Área de Conocimiento. Disponible en <https://datos.gob.mx/busca/dataset/sistema-nacional-de-investigadores> (Consultado: 10-05-2022).

En cuanto al panorama actual del estado que guarda el Acceso Abierto, se encuentran inmersos 158³ países, divididos en cinco regiones. Revisando estadísticas e informes, pocas naciones han tenido éxito, por su entorno y los factores que los rodean en la implementación; en otros casos, el avance ha sido bajo o se encuentran en sus primeras etapas de desarrollo.

Existen iniciativas regionales de Acceso Abierto, que ofrecen a los autores y usuarios finales publicaciones académicas y científicas de texto completo, en portales de revistas multidisciplinarios revisadas por pares. Plataformas patrocinadas principalmente por fondos públicos, buscando la visibilidad internacional, indicadores bibliométricos para una mejor calidad de los procesos editoriales, tales como DOAJ y SciELO. También hay portales como el de las revistas en Brasil –como SEER, a nivel institucional de la Universidad de Sao Paulo (USP)–. En México, el portal de revistas académicas y arbitradas de la Universidad Nacional Autónoma de México (UNAM), que tiene como objetivo facilitar la búsqueda y consulta de las revistas e incrementar su visibilidad, presencia e impacto. Por otro lado, la Universidad de Chile tiene un significativo número de revistas de Acceso Abierto, por dar algunos ejemplos.

Por ello, la literatura científica en apoyo a la ciencia abierta contribuye a mejorar la metodología y calidad de los procesos editoriales, la visibilidad de las revistas académicas editadas en América Latina y el crecimiento de la adopción del Acceso Abierto. Y, más recientemente, los Repositorios Institucionales de centros y universidades públicas y privadas. Si nos remontamos al año 2002, se ha fortalecido en gran medida la ciencia y su compartición, así como el reuso de datos por parte de los investigadores a nivel mundial. México, en el 2014, crea una iniciativa que logra integrar el Capítulo 10 de Ciencia y Tecnología sobre Acceso Abierto a la información científica y tecnológica, creando el Programa de Repositorios del Consejo Nacional de Ciencia y Tecnología (Conacyt)⁴, implementándose el Repositorio Nacional

³ The Global Open Access Portal (GOAP), UNESCO. Disponible en <http://www.unesco.org/new/en/communication-and-information/portals-and-platforms/goap/> (Consultado: 10-05-2022).

⁴ Conacyt (2022). Repositorio Nacional. <https://www.repositorionacionalcti.mx/>

(RN), y a finales de ese mismo año se crean los Lineamientos Generales para el Repositorio Nacional (RN) y los Repositorios Instituciones (RI)⁵.

Poco después, en 2019, se reunieron los representantes de la African Open Science Platform, Amelica, Coalitions, OA2020 y SciELO, cinco de las principales iniciativas mundiales de Acceso Abierto en la reunión anual de la Declaración de São Paulo, con el fin común de hacer que el conocimiento esté disponible y sea accesible en cualquier lugar, pues este ayuda a resolver los desafíos de la humanidad.

Proceso de datos

Este proceso 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 (tiempo). 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 consideramos y partimos de la declaración de la American Geophysical Union (AGU),^{6 7} en donde convierte a los datos en un patrimonio mundial, que deben ser conservados a largo plazo para futuro uso, pues ayudarán a los científicos a comprender los sistemas terrestre, planetario y heliofísico. Y se tendrá la obligación a ponerlos a disposición de la comunidad científica y del público tan pronto como sea posible, debiendo ser accesibles, en formatos utilizables, con suficiente documentación legible que permita su reutilización.

⁵ Conacyt (2017). Lineamientos generales de Ciencia Abierta: Conacyt. México.

⁶ La American Geophysical Union, establecida en 1919 por el Consejo Nacional de Investigación, asociación científica internacional sin fines de lucro –con 60.000 miembros en 137 países– crea su Declaración.

⁷ Declaración adoptada por la American Geophysical Union el 29 de mayo de 1997; y fue refirmado en mayo de 2001, mayo de 2005, mayo de 2006 y finalmente revisado y refirmado en mayo de 2009, febrero de 2012 y septiembre de 2015.

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.

Asimismo, busca ser lo más abierta posible para maximizar el rendimiento científico, económico y social de la obtención de los datos, ya sea con fondos públicos y/o privados. Solo existirán situaciones limitadas de acceso por cuestiones éticas, pero no por propiedad. Estas deben ser justificadas a un nivel de protección de seguridad, la confidencialidad para así poder evitar daños a entidades protegidas. Lo importante es 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 tanto en su postura, considerando el valor de los datos que crece con el tiempo, llevando a cabo ventajas en la curación de datos. Por lo tanto, los datos y la documentación asociada deben depositarse en un repositorio de la especialidad, siendo parte integral de los programas operativos y de la investigación, por lo que la comunidad científica tiene un amplio conjunto de datos para la realización de sus investigaciones en beneficio de la sociedad y de la ciencia.

Con el desarrollo de iniciativas internacionales de investigación a gran escala, el apoyo a los datos tomó una variedad de formas. Se establecieron una serie de instalaciones de datos, por ejemplo, los Centros de Datos Mundiales y la Federación de Servicios de Análisis de Datos Astronómicos y Geofísicos⁸; surgió una visión revolucionaria en el área científica de las Ciencias de la Tierra, centrando la atención de los geocientíficos en conjunto en métodos científicos, mediciones y datos.

En cuanto a recursos nacionales mediante infraestructuras tecnológicas para los investigadores. Estos centros se convirtieron en depósitos de datos

⁸ Surge en 1957 el Año Geofísico Internacional, en inglés International Geophysical Year (IGY), el cual ha evolucionado con una visión revolucionaria en el área científica de las Ciencias de la Tierra.

junto con otras instituciones de preservación, como archivos con colecciones de fotos y manuscritos, museos con objetos físicos y bibliotecas con libros y revistas.

Actualmente, hay debates científicos sobre la producción de grandes datos y de infraestructuras que se centran en investigaciones, reavivando el interés social y cultural, transformándose en nuevas líneas de investigación y de conocimiento.

La gestión de contenidos

Para trabajar en la gestión de estos contenidos, tanto la información como los datos ambos deben ser accesibles, interoperables y reutilizables no sólo para las personas, sino también para las máquinas con identificadores, incluyendo 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, así como reutilizables.

Los datos también deberán estar bien conservados, ser accesibles y vinculados de forma segura a las publicaciones asociadas, recursos o facilitar el reuso.

Existen estándares para el uso y tratamiento de los datos y se deben contar con las siguientes características:

1. Para publicar los datos en cualquier formato se debe contar con una licencia abierta, con finalidad de que el público que los encuentre tenga el beneficio de acceder libremente a la información bajo una licencia de uso.
2. También los datos que sean publicados bajo una estructura con licencia abierta deberán estar estructurados en un formato legible por máquina, por lo que el usuario podrá reutilizar los datos publicados y hacer su propio análisis.

3. Se deberán usar formatos no propietarios, es decir, datos abiertos, estructurados, y sin propietarios, con el objetivo de que ellos se consoliden y ofrezcan potencial para que su reutilización sea fácil, productiva, además de que genera impacto.
4. Se deben emplear URI con la finalidad de integrarse con otras fuentes de datos y aplicaciones, pasando de datos abiertos a datos vinculados, por lo que se facilita su identificación y localización mediante “identificador uniforme de recursos” o URI.
5. Finalmente los datos deberán enviar a otros datos para proveer contexto. Por ello, se vincularán y se relacionaran con otros datos de información relevante, la fusión de estas dos o más fuentes de información pone todo este conocimiento en un nuevo contexto, creando valor adicional.

Es importante mencionar que para la gestión de los datos para la investigación en los repositorios del área de Ciencias de la Tierra, se debe considerar además un elemento esencial de apoyo, refiriéndonos a las publicaciones académicas debido a que es una fuente utilizada por la comunidad científica para dar a conocer los avances de sus investigaciones, siendo por lo regular, que los datos se complementan, respaldan y/o vinculan con la publicación (artículo, informe, libro, capítulo, tesis, etc.) o un proyecto ofreciendo como resultado su presentación.

Con el fin de llevar a cabo el repositorio para manejar y hospedar datos, se pueden considerar una serie de buenas prácticas en su desarrollo, dentro de los cuales podemos destacar:

- Citación y procedencia en publicaciones académicas –cuando se considere necesario–. Cuando los investigadores utilizan datos creados por otros, estos deben ser citados con referencia a su autor, a su procedencia y a un identificador digital permanente.
- Interoperabilidad en la mayor medida posible. Tanto los datos de investigación como los metadatos que permiten la evaluación y reutilización, todo ello con el objetivo de que el intercambio y el enlace de plataformas e información permita un flujo transparente para el usuario.

- Reutilización no restrictiva. Los datos para investigación que están en el dominio público deben ser etiquetados como reutilizables por medio de una renuncia a los derechos o una licencia no restrictiva, mediante la cual se deje en claro que los datos pueden ser reutilizados sin mayor requisito que el reconocimiento al autor.
- La información debe llevar el dataset o el dato del lugar donde se hospedará –que se incluirá en el repositorio dentro del archivo llamado *readmefile* con la siguiente información y descripción:
 - I. Proporcionar la información sobre los datos y/o el set de datos para que sea correctamente interpretado tanto por personas como máquinas y se puedan realizar procesos con otras investigaciones.
 - II. Cuando el registro (unidad de descripción) se compone de varios sets de datos, se debe vincular a un archivo que se llame *readmefile* (archivo que contendrá la descripción de los datos), de tal manera que pueda ser asociado al set de datos y se especifique en el archivo de formato de fácil consulta como csv o txt.
 - III. Debe incluir una breve descripción de los datos según sea el caso: contacto del investigador principal, fecha de recolecta de datos y de creación del set de datos; información geográfica de los datos; metodología y enlace a publicaciones y otra documentación; unidades de medida, protocolos, abreviaciones, códigos, símbolos asociados a los datos; licencia de uso; citación recomendada.
- Capacidad de vinculación de los datos abiertos, siempre que sea posible estar vinculados con otros datos basados en su contenido y contexto, con el fin de maximizar el valor semántico.

Los principios rectores de FAIR para la gestión y la administración de los datos científicos

La necesidad urgente de mejorar la infraestructura que respalda 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 unieron para diseñar y respaldar los Principios (HODSON, 2021) que referimos, con la intención de utilizarlos como una guía para aquellos interesados que deseen mejorar la reutilización de sus datos. A diferencia de las iniciativas de pares –que se centran en especialistas–, estos principios ponen especial énfasis en mejorar la capacidad de las máquinas para encontrar y utilizar los datos automáticamente, además de respaldar su reutilización por parte de los individuos.

Se desarrollan bajo estándares que conectan a investigadores, editores y repositorios de datos para habilitar los datos (localizables, accesibles, interoperables y reutilizables) a gran escala. Apunta también a acelerar el descubrimiento científico y mejorar la integridad, transparencia y su reproducción.

El marcaje de datos y la literatura

Para llevar a cabo el desarrollo de un Repositorio Institucional de datos, o en su caso literatura, deben discutirse una serie de normativas, técnicas y elementos –como son los metadatos que definirán el registro y la especificidad de cada uno de los objetos digitales albergados en el RI–. Estos son indispensables para que los contenidos sean recuperables mediante herramientas de búsqueda en Internet.

A través de los metadatos se crean datos estructurados que describen el contenido, calidad y condición y otras características de un dato y además expliquen y localicen un recurso de información, y, una vez recuperado, usarlo y administrarlo. Al trabajar con la estructura técnica del repositorio se debe hacer mediante estándares internacionales con el fin de poder establecer una comunicación entre las diversas plataformas con las decidamos interactuar.

Es importante dejar claro que se pueden emplear diferentes tipos de metadatos según la naturaleza del Repositorio ya sea de Datos o Literatura. En los cuales se deben tomar en cuenta la descripción de contenido, calidad, condiciones, entre otros, así como 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.

La tipificación de los elementos en los metadatos, generalmente se reconocen de acuerdo a la descripción del objeto digital: administrativos, descriptivos, de preservación, técnicos y de uso. Debemos asegurarnos 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 del marcaje facilitan la organización de los recursos basados en la audiencia permitiendo organizarlos mediante una identificación digital (URL, PURL o DOI) y estructura, asegurando su preservación. Y pueden ser expresados en 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.

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

Por lo que es necesario una interconexión para mantener una sincronización de la información, accesible para el usuario mediante los llamados *web services* que manejan servicios como:

- Interoperabilidad entre aplicaciones de software, independientemente de sus propiedades o de las plataformas sobre las que se instalen.
- Fomentan los estándares y protocolos basados en texto, que hacen más fácil acceder a su contenido y entender su funcionamiento.
- Apoyan los servicios eeb (HTTP).
- Permiten que los servicios y el software de diferentes compañías se combinen fácilmente para proveer servicios necesarios para el tratamiento de los datos.

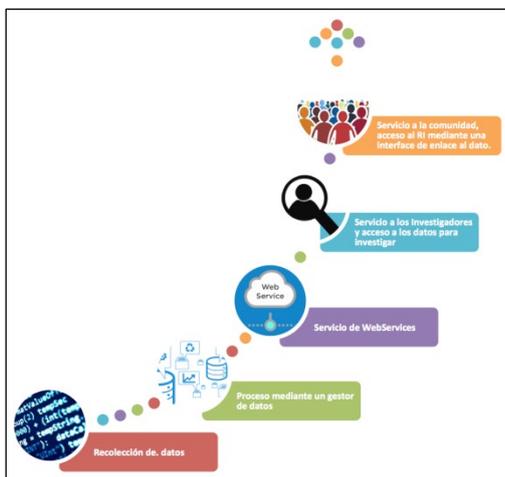


FIGURA 1. Estructura del software con el dato. Elaboración propia

Para las Ciencias de la Tierra los datos tienen condicionantes, por mencionar algunas de ellas, es que se componen de varios archivos en diversos formatos, por lo que es necesario relacionar el dato que está siendo registrado, 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 se debe considerar 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.

necesario contar con la identificación digital, es decir, utilizar varios esquemas de metadatos contienen elementos para cargar números normalizados que permitan identificar unívocamente la obra o el recurso mediante el URL, PURL o DOI.

Con ellos se podrán generar archivos con información que facilitará la preservación, por lo regular, es probable que muchos recursos digitales no puedan ser usados en el futuro y se busca mantenerlos y facilitar su acceso. La información electrónica puede ser alterada por el cambio de la tecnología del hardware o el software. La migración de formatos y la emulación de programas son estrategias para evitar los inconvenientes.

Es importante que en el marcaje los metadatos se pueden expresar en 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 el documento.

Considerar además que las versiones anteriores de las directrices, presentan los cambios como un perfil y esquema de aplicación basados en Dublín Core y DataCite o con un nuevo OAI-metadataPrefix, apoyándose en los esquemas de identificación para autores, organizaciones, financiadores, recursos académicos introducción de vocabularios controlados por COAR con el Cumplimiento de la Política de Adquisición de Contenido de OpenAIRE.¹¹

Conclusiones

Para llevar a cabo esta tarea, es imprescindible analizar estructuras técnicas correspondientes para asignar los metadatos adecuados a los datos y que 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.

¹¹ México. Infotec. Interoperabilidad con el Metabuscador del Repositorio Nacional Versión 3.1., publicado el 05-Oct-2018.

Como se puede ver a lo largo del trabajo es evidente que se debe llevar a cabo un proceso y análisis de los contenidos, además para un correcto marcaje es importante mencionar que dentro de las Ciencias de la Tierra, el Mar y la Atmósfera se genera una gran cantidad de información en lapsos muy cortos, por lo que se necesitan infraestructuras en donde la gestión de los datos como de contenidos sirvan para suministrar información sobre lo que se produzca en datos y en la parte de la literatura, lo cual facilitará su descripción y accesos, asegurando la calidad y el ofrecimiento de las 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 este misión de ponerlo en abierto y valorar aun más inclusión, diversidad, confianza e innovación. Así mismo, los metadatos deben estar eficientemente conformados y establecidos bajo parámetros internacionales que ofrezcan una garantía de cosecha e intercambio de información y datos.

El llevar cabo una gestión correcta y adecuada en los repositorios, beneficia en gran medida la representación de los proyectos de investigación y ofrecen 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 esta actividades tanto académicas como científicas para seguir en las líneas de la ciencia abierta.

Bibliografía

AMERICAN GEOPHYSICAL UNION (2019a). Declaration adopted by the American Geophysical Union. <https://ethics.agu.org/>

AMERICAN GEOPHYSICAL UNION (2019b). Repository Finder. <https://repositoryfinder.datacite.org/>

ARMENDÁRIZ SÁNCHEZ, S.; SOSA ZARAGOZA, P.; SOSA JIMÉNEZ, M. (2021). Los datos para la investigación en los repositorios institucionales y su impacto científico y social: los RI en Ciencias de la Tierra. Barcelona: Octaedro. Colección: Universidad; pp. 153-164 .

ATLASSIAN CONFLUENCE OPEN SOURCE PROJECT.
<https://wiki.lyrasis.org/display/DSDOC6x/OAI+2.0+Server>

CASTRO ESCAMILLA, M. (2020). Desarrollo de un Repositorio de datos científicos de apoyo a la investigación: el caso de las ciencias de la Tierra [tesis de maestría]. Universidad Nacional Autónoma de México.
<http://132.248.9.195/ptd2020/agosto/0802846/Index.html>

CASTRO ESCAMILLA, M. (2021) Las declaraciones, principios e iniciativas de la Ciencia Abierta: una visión. En: Ciencia Abierta Opciones y experiencias para México y Latinoamérica. Barcelona: Octaedro. Colección: Universidad; pp. 79-90.

CONACYT (2017). Lineamientos generales de Ciencia Abierta: Conacyt.

CONACYT (2019). El Repositorio Nacional de México.
<http://www.repositorionacionalcti.mx/>

DATA CITE (2021). DataCite Metadata Schema. <https://schema.datacite.org/>

HODSON, S. (2021). Turning FAIR into Reality Report and Action Plan.
https://www.eoscpilot.eu/sites/default/files/18-11-22-hodson-turning_fair_into_reality.pdf

LYRASIS. DSpace: resource. <https://duraspace.org/dspace/resources/>

MÉXICO. INFOTEC. Interoperabilidad con el Metabuscaador del Repositorio Nacional Versión 3.1., Publicada el 05-Oct-2018.

SCHIRRWAGEN, J. (2018). OpenAIRE Guidelines for institutional and thematic repository managers 4.0 November 30, 2018.

TORRES-SALINAS D., NICOLÁS; ROBINSON GARCÍA; ALVARO CABEZAS-CLAVIJO (2012). Compartir los datos de investigación en ciencia: introducción al datasharing. *El Profesional de la Información*, 2(12).

UNAM. IGEF (2019). Repositorio Institucional de Recursos Geotérmicos para México.
<http://rigeotermia.geofisica.unam.mx/jspui/>

UNAM. SSN (2019). Repositorio Institucional de Datos del Servicio Sismológico Nacional.
<http://rigeofisica.ssn.unam.mx/jspui/>