



UNIVERSIDAD DE LOS LAGOS

2023

Glosario de Gestión de Datos de Investigación y Publicaciones

DIRECCIÓN DE BIBLIOTECAS
PROYECTO INES CIENCIA ABIERTA INCA 210018

Elaborador por. Mg. Jocelyn Patterson
Coordinadora Ciencia Abierta
jocelyn.patterson@ulagoc.cl

Última actualización: 8 de agosto de 2023



GLOSARIO GESTIÓN DE DATOS DE INVESTIGACIÓN Y PUBLICACIONES

PRESENTACIÓN

A continuación, se presenta el Glosario de Gestión de Datos de Investigación y Publicaciones desarrollado en el contexto del desarrollo del Proyecto Ciencia Abierta (INCA 210018), con el objetivo de dar a conocer y socializar conceptos esenciales vinculados a la gestión de datos de investigación y publicación científica. El glosario comprende 17 conceptos organizados alfabéticamente¹.

CONCEPTOS

ACCESO ABIERTO (OPEN ACCESS)

Significa el acceso libre a la información y uso sin restricciones de recursos electrónicos para todos. Cualquier tipo de contenido puede ser open access, desde texto, datos de software, audio, vídeo y multimedia. Específicamente, se refiere al depósito en repositorios de acceso universal de copias digitales de dichos contenidos. Una publicación es considerada de acceso abierto si:

- Su contenido accesible de forma universal y gratuita, sin costo para el lector, vía internet, o de lo contrario;
- El autor o el propietario del derecho de autor irrevocablemente concede a todos los usuarios, por un periodo ilimitado, el derecho a usar, copiar o distribuir un artículo, bajo la condición de proporcionar la atribución correspondiente;
- Es depositada, inmediatamente, en formato completo y electrónico, en al menos un repositorio de acceso abierto (adaptado de la UNESCO y la ANID).

ADMINISTRACIÓN DE DATOS (DATA STEWARDSHIP)

Se refiere al conocimiento y habilidades requeridas para gestionar los activos de datos de forma efectiva. La administración de datos es a menudo descrita como la gobernanza de datos en acción. Esto incluye la supervisión de los datos para asegurar la adecuación al uso, la accesibilidad de los datos y el cumplimiento de las políticas, directrices y regulaciones (adaptado de Tri-Agency Research Data Management, Canadá).

ARTICLE PUBLISHING CHARGE (APC)

Son pagos solicitados por una editorial de acceso abierto que efectúa el autor o su institución para cubrir diversos costos de publicación.

CICLO DE VIDA DE LOS DATOS

Se refiere a todas las etapas de existencia de datos, desde su creación hasta su destrucción. El ciclo de vida de los datos provee una visión de alto nivel de las etapas involucradas en la gestión exitosa y en la preservación de datos para su uso y reutilización. De manera amplia, estas etapas son: planeación, creación, procesamiento, análisis, disseminación, preservación y reutilización.

¹ Algunas de las definiciones presentes en este glosario son parte de la “Política de acceso abierto a la información científica y a datos de investigación financiados por fondos públicos de la ANID. Disponible en: https://s3.amazonaws.com/documentos.anid.cl/estudios/Politica_acceso_a_informacion_cientifica_2022.pdf

CIENCIA ABIERTA (OPEN SCIENCE)

Es un movimiento que pretende hacer la ciencia más abierta, accesible, eficiente, transparente y beneficiosa para todas y todos. Impulsado por los avances sin precedentes del mundo digital, la transición hacia la ciencia abierta permite que la información, los datos y los productos científicos sean más accesibles y más fácilmente compartidos con la participación activa de todas las partes interesadas. Se caracteriza por la apertura no sólo de las publicaciones (lo que tradicionalmente se conoce como *open access*), sino también de los datos de investigación, las metodologías, los procesos, así como la vinculación de la ciudadanía en un entorno de investigación e innovación responsable. De forma sencilla, los procedimientos de ciencia abierta pretenden hacer accesibles, en formato digital, los resultados de la investigación financiada con fondos públicos, no sólo para la comunidad científica que los produce, sino para la sociedad en general que la financia, potenciando al tiempo la reproducibilidad y la reutilización de los resultados (UNESCO).

DATASET (CONJUNTO DE DATOS)

Es una colección de datos relacionados y codificados en una estructura definida. Los datasets pueden contener datos experimentales, datos geoespaciales, notas de laboratorio, datos genómicos, entre muchos otros.

DATOS DE INVESTIGACIÓN

Es material factual registrado, el cual es necesario para validar los descubrimientos científicos. Pueden ser mediciones, grabaciones, registros u observaciones que son recolectadas por investigadores u otros, y se generan a lo largo del desarrollo del proyecto de investigación. Presentan un mínimo de contexto para su interpretación. Los datos pueden tomar la forma de texto, números, símbolos, imágenes, grabaciones, grabaciones sonoras, reproducciones pictóricas, dibujos, diseños y otras representaciones gráficas, manuales de procedimiento, formas, diagramas, flujos de trabajo, descripciones de equipos, archivos de datos, algoritmos para el procesamiento de datos, software, lenguajes de programación, código o registros estadísticos. Sirven de soporte a los resultados de un proyecto de investigación (adaptado de Tri-Agency Research Data Management, Canadá).

DIGITAL OBJECT IDENTIFIER (DOI)

Es un identificador persistente que es usualmente asignado a ítems digitales, como por ejemplo, a artículos o a un *dataset* con el fin de que pueda ser encontrado o citado. Los DOIs pueden ser incorporados a una URL de manera tal que los usuarios siempre puedan acceder al contenido digital, incluso si estos han sido movidos de su ubicación en línea (adaptado de Research Data Oxford, Reino Unido).

FAIR (FINDABLE, ACCESIBLE, INTEROPERABLE, REUSABLE)

Los principios FAIR se utilizan para indicar que los datos cumplen con los requisitos de encontrabilidad, accesibilidad, interoperabilidad y reutilización. Los Principios FAIR ofrecen un conjunto de cualidades precisas y medibles que una publicación de datos. De forma general, estos son:

- Findable (Encontrables): Los datos y metadatos pueden ser encontrados por la comunidad después de su publicación, mediante herramientas de búsqueda.

- Accessible (Accesibles): Los datos y metadatos están accesibles y por ello pueden ser descargados por otros investigadores utilizando sus identificadores.
- Interoperable (Interoperables): Tanto los datos como los metadatos deben de estar descritos siguiendo las reglas de la comunidad, utilizando estándares abiertos, para permitir su intercambio y su reutilización.
- Reusable (Reutilizables): Los datos y los metadatos pueden ser reutilizados por otros investigadores, al quedar clara su procedencia y las condiciones de reutilización.

GESTIÓN DE DATOS DE INVESTIGACIÓN

Se refiere a los procesos aplicados al ciclo de vida de un proyecto de investigación para guiar la recolección, documentación, almacenamiento, intercambio y preservación de los datos de investigación. Una buena gestión de datos de investigación es un elemento crítico para la investigación de todas las disciplinas, particularmente en los casos donde las entidades financiadoras requieren de los investigadores que gestionen sus datos de investigación (adaptado de Tri-Agency Research Data Management, Canadá).

METADATOS

Consisten en información que caracteriza los datos. Describen el contenido, calidad, condiciones, historia, disponibilidad y otras características de los datos (ANID).

ORCID ID (OPEN RESEARCHER AND CONTRIBUTOR IDENTIFIER)

Es un identificador digital persistente que permite distinguir a los investigadores de forma unívoca. Resuelve el problema de identificar a autores a pesar de existir nombres idénticos o similares, cambios de nombre a lo largo de la vida, diferentes convenciones al escribir nombres, etc. Permite conectar de forma automática este iD con la información profesional, tales como la afiliación institucional, publicaciones, adjudicaciones y otros. El ORCID iD es administrado por el propio investigador y es de uso gratuito.

PLAN DE GESTIÓN DE DATOS DE INVESTIGACIÓN (PGDI)

Es un documento donde se establece cómo se gestionarán los datos (incluyendo códigos) durante la investigación y una vez finalizada la misma. El plan de GDI es un documento que debe ser actualizado durante la investigación.

PUBLICACIONES

La producción científico-tecnológica plasmada en artículos de revistas, capítulos de libros, libros, trabajos técnico-científicos, tesis académicas, entre otros que sean el resultado de la realización de actividades de investigación, que hayan atravesado un proceso formal de evaluación y que hayan sido estos publicados o no (ANID).

RUTA VERDE AL ACCESO ABIERTO (GREEN OPEN ACCESS)

El proceso mediante el cual el autor que ha publicado su artículo en una revista de suscripción deposita la versión aceptada del mismo en un sitio web o en un repositorio institucional o temático después de un periodo de embargo que se contabiliza desde la fecha en que el artículo se ha publicado en una revista, y que suele durar entre 6 y 24 meses (ANID).

RUTA DORADA AL ACCESO ABIERTO (GOLD OPEN ACCESS)

Implica el acceso público, inmediato, permanente y gratuito al artículo final por medio de un canal de publicación de acceso abierto de alta calidad, es decir una revista de acceso abierto. Ello se consigue mediante el pago al editor de una tarifa variable conocida como APC (Article Publishing Charge), que son los costos derivados del proceso de revisión, producción y publicación del artículo (adaptado de ANID).

REVISTAS HÍBRIDAS

Una revista de acceso abierto híbrida es una publicación que combina la publicación basada en suscripción y la publicación libre de pagos. Esto significa que el autor paga un monto (APC) determinado por el editor para dejar el artículo disponible de forma libre. De otra forma, la revista queda disponible a los lectores mediante el pago de una suscripción por el acceso.