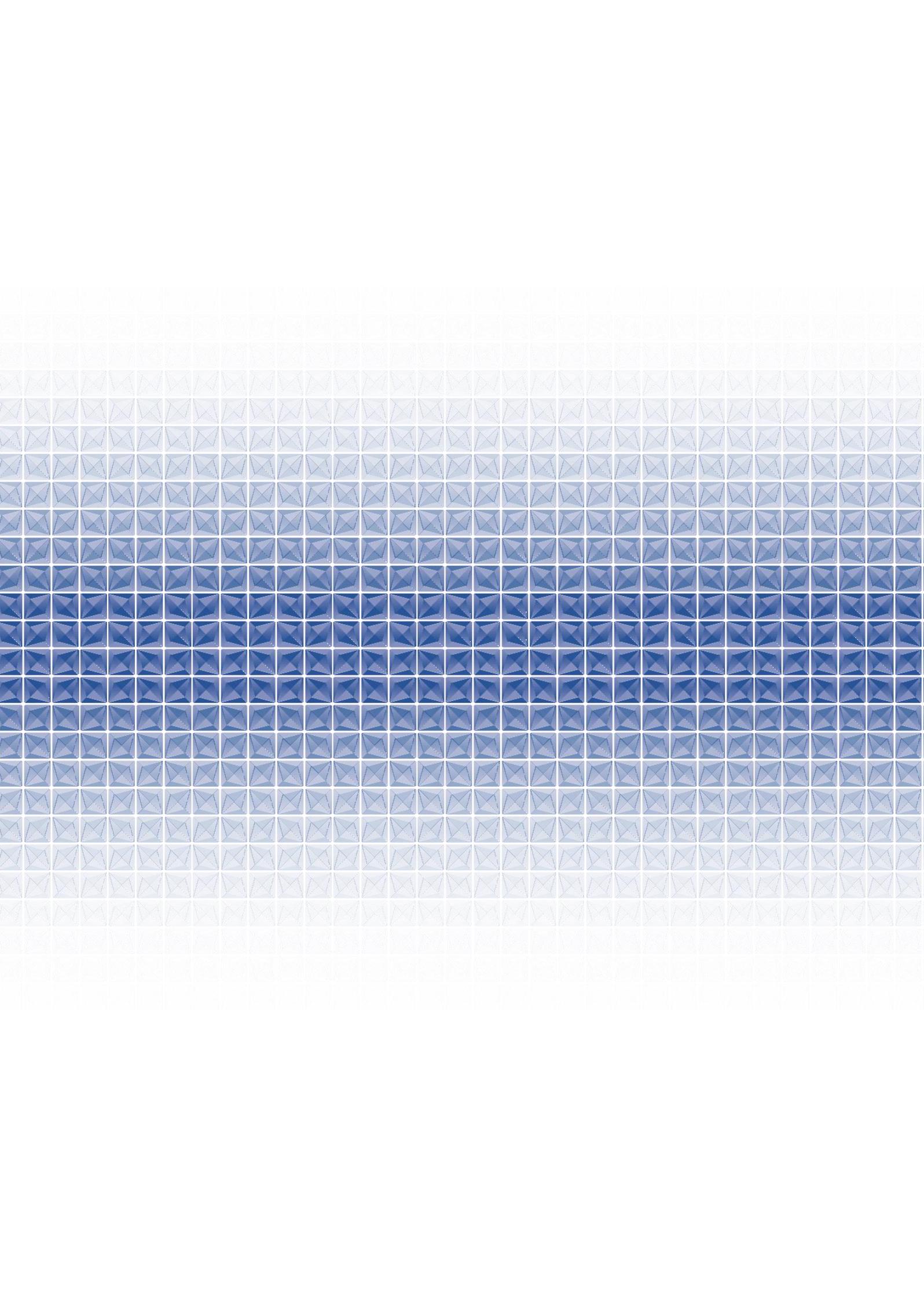


GUÍA PARA LA GESTIÓN DE DATOS EN INVESTIGACIÓN Y LA ELABORACIÓN DE UN PLAN DE GESTIÓN DE DATOS (PGD)



Vicerrectoría de
Investigación y Desarrollo
Universidad de Concepción

Dirección de Bibliotecas
Universidad de Concepción



"Esta guía tiene por objetivo ofrecer información y herramientas para la Gestión de Datos de Investigación (GDI). Está dirigida a la comunidad académica comprometida con las buenas prácticas en la gestión de los datos que generan las investigaciones, sin importar la disciplina o ámbito del conocimiento con el que se vinculen."

Fuente: Basada en Comisión Económica para América Latina y el Caribe (CEPAL). (2020). Biblioguías: Gestión de datos de investigación. Recuperado de <https://biblioguías.cepal.org/c.php?g=495473&p=3390849>

Contenido

Introducción.....	1
1. ¿Qué es un Plan de Gestión de Datos de investigación?.....	2
2. Actores de un proceso de gestión de datos.....	3
3. Responsabilidades de cada área en el PGD.....	4
4. Guía para la gestión de datos	5
5. Beneficios de la Gestión de Datos de Investigación (GDI)	6
6. ¿Cómo identificar tipos de datos de investigación?	7
7. Proceso de depósito de datos	10
8. Bibliografía	11

INTRODUCCIÓN

Los datos “son considerados como la nueva materia prima del siglo 21, sin embargo, en la actualidad los datos de investigación quedan encerrados en las oficinas de los investigadores, o en las instituciones y empresas que los crearon. Sin embargo, la mayoría de las instituciones académicas, organizaciones de investigación, las principales editoriales requieren que los datos se presenten adecuadamente para poder utilizarlos, que estén disponibles para cualquier propósito y puedan reutilizarse para nuevas investigaciones”(Alonso-Arévalo, 2019).

La adecuada gestión de datos es fundamental para maximizar la utilidad y el valor de los datos de investigación de alta calidad y, por tanto, la investigación de excelencia. La organización y archivo de los datos también es crucial para facilitar el intercambio de datos y asegurar la sostenibilidad y la accesibilidad de los datos a largo plazo, y por lo tanto, su reutilización para la ciencia futura. En este contexto, es determinante la labor de la biblioteca en el apoyo a la investigación, de cara a gestionar y compartir datos a través del **Repositorio de Datos de Investigación de la Universidad de Concepción**, así como el apoyo y orientación a los investigadores en torno al tema; facilitar la citación y vinculación de los datos con publicaciones, con el objetivo de proporcionar mayor visibilidad y accesibilidad de los datos y la investigación misma (Erway, Horton, Nurnberger, Otsuji, & Rushing, 2016).

Los desafíos para las bibliotecas universitarias son amplios, desde el diagnóstico, pasando por diversas áreas, pero podemos recalcar que para avanzar es fundamental la colaboración con investigadores, la construcción de conocimientos técnicos entre el personal de biblioteca y el establecimiento de directrices en cuanto a la descripción de los metadatos, rol que recae por la expertise en la curatoría desde los inicios de la bibliotecología.

La transición a los formatos digitales conlleva una avalancha de datos que necesitan ser curados, desde las descargas, citas, redes sociales, medios de comunicación etc. Esto ha llevado a que instituciones, investigadoras e investigadores junto con profesionales de la información, entre otros, incorporen nuevas competencias y conocimientos, dando como resultado la adquisición de nuevas habilidades para la gestión de grandes cantidades de datos.

Para que el modelo basado en la tríada educación, experticia y curaduría funcione, es importante que las instituciones cuenten con una política explícita y clara en cuanto a qué, cómo, para qué conservar y compartir los datos. El modo de organizar en llevar a cabo los tres servicios no responde a un esquema universal o rígido, sino que estará fuertemente asociado a la política particular de cada institución y variará según cada contexto.

Esta guía tiene por objetivo ofrecer información y herramientas para la Gestión de Datos de Investigación (GDI). Está dirigida a la comunidad académica comprometida con las buenas prácticas en la gestión de los datos que generan las investigaciones, sin importar la disciplina o ámbito del conocimiento con el que se vinculen. (CEPAL, 2020).

Es en este marco, que la Vicerrectoría de Investigación y Desarrollo (VRID) junto a la Dirección de Bibliotecas de la Universidad de Concepción, ponen a disposición este documento que se enmarca en la ejecución del INCA 210004“Ciencia Abierta en la Universidad de Concepción: Creando Cultura y Democratizando la Ciencia”.

	Elaboración	Revisión	Aprobación
Fecha	Enero- 2023	Mayo- 2023	Junio- 2023
Cargo	Unidad de Apoyo a la Investigación - DIRBIB	Dirección de Bibliotecas Dirección de Investigación y Creación Artística	Vicerrectoría de Investigación y Desarrollo

1. ¿Qué es un Plan de Gestión de Datos de Investigación?

Para la Comisión Económica para América Latina y el Caribe, CEPAL, un **Plan de Gestión de Datos** (PGD) es un documento formal elaborado por el investigador o grupo de investigación, que se desarrolla al inicio de un proyecto de investigación. Describe todos los aspectos de la gestión de los datos, es decir, qué se hará con los datos durante y después del proyecto de investigación.

Un PGD es un paso vital en el proceso de investigación y que no debe ser obviado. Este instrumento ayuda a asegurar que los datos de investigación sean precisos, completos, confiables y seguros tanto durante como después de completar una investigación.

El PGD tiene como propósito proporcionar un análisis de los principales elementos de una política de gestión de datos, y describe el tratamiento que van a recibir los conjuntos de datos de investigación recopilados o generados en el curso de un proyecto de investigación.

La Agencia Nacional de Investigación y Desarrollo (ANID) establece en su **Política de Acceso Abierto** que el plan de gestión de datos es un documento escrito, que describe los datos que se espera adquirir o generar durante el curso de un proyecto de investigación, cómo se gestionarán, describirán, analizarán y almacenarán y qué mecanismos se utilizarán para compartirlos y preservarlos (ANID, 2022). Es decir, el plan de gestión de datos evoluciona a medida que avanza el proyecto. Por lo tanto, su finalidad es servir de herramienta de apoyo durante el ciclo de vida de aquellos datos que se recopilen, se procesen o se generen en el ámbito de un proyecto e investigaciones.

Por ello, Bibliotecas UdeC asume el desafío de apoyo a la investigación tomando como referencia el modelo entregado por la ANID y la CEPAL.

Para ver modelos de PGD de distintas disciplinas, puede revisar en los siguientes sitios:

- [DPM Digital Curation Center](#)
- [DMP Tool - University of California](#)

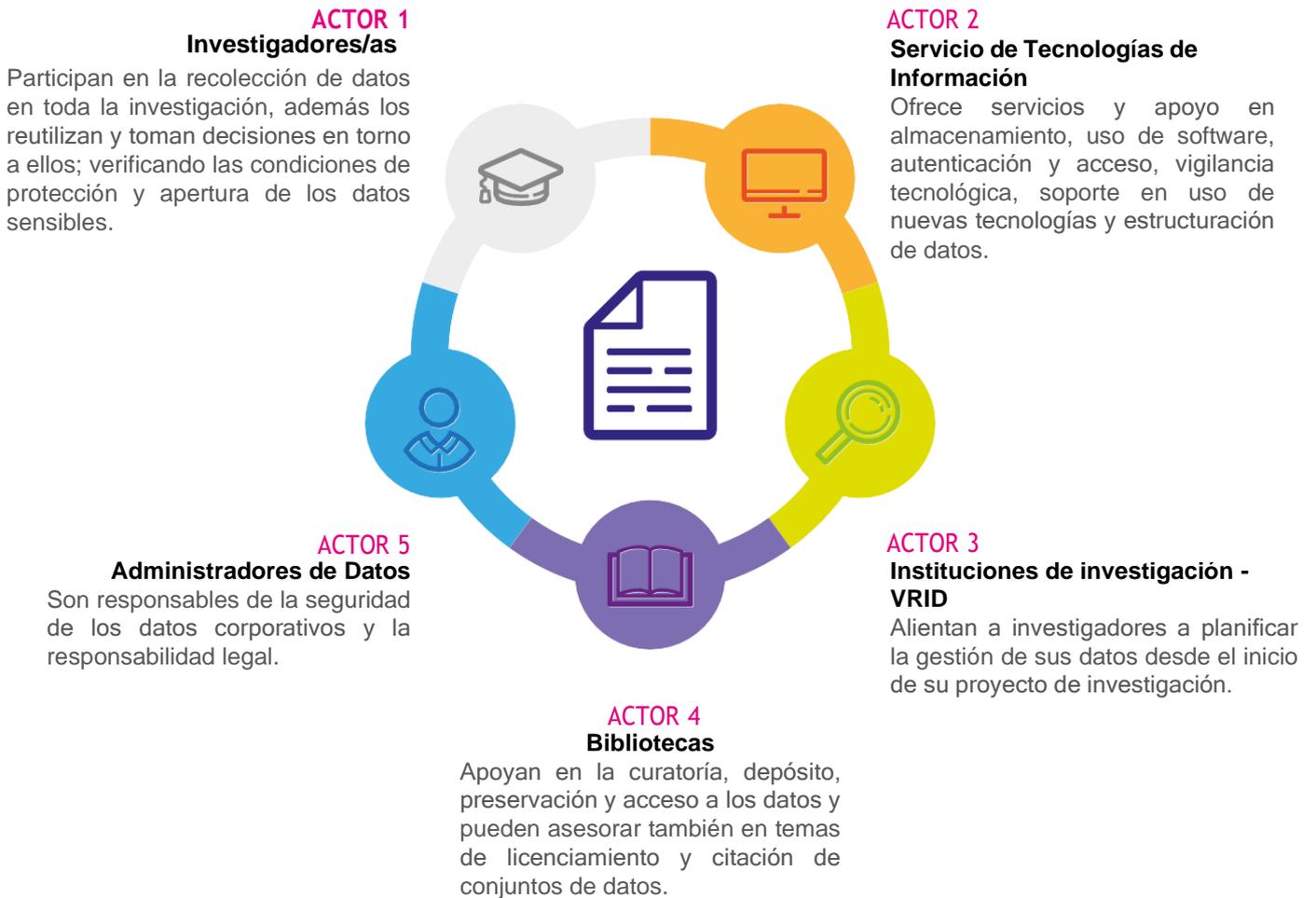
Para ver los ítems a considerar en un PGD puede revisar:

- [Plan de Gestión de Datos ANID](#)
- [Plan de Gestión de Datos de la Unión Europea](#)

	Elaboración	Revisión	Aprobación
Fecha	Enero- 2023	Mayo- 2023	Junio- 2023
Cargo	Unidad de Apoyo a la Investigación - DIRBIB	Dirección de Bibliotecas Dirección de Investigación y Creación Artística	Vicerrectoría de Investigación y Desarrollo

2. Actores de un Proceso de Gestión de Datos

Este apartado, describe a los agentes o actores involucrados en el proceso de elaboración, depósito y difusión de un Plan de Gestión de Datos (PGD). Los aspectos éticos, como la protección de los datos, son descritos en el cuarto apartado del presente documento que ofrece una breve guía al respecto.



Fuente: Basada en Comisión Económica para América Latina y el Caribe (CEPAL). (Dec 18, 2020). Biblioguias: Gestión de datos de investigación: Módulo 2 - Plan de Gestión de Datos (PGD). Recuperado de <https://biblioguias.cepal.org/gestion-de-datos-de-investigacion/PGD>

	Elaboración	Revisión	Aprobación
Fecha	Enero- 2023	Mayo- 2023	Junio- 2023
Cargo	Unidad de Apoyo a la Investigación - DIRBIB	Dirección de Bibliotecas Dirección de Investigación y Creación Artística	Vicerrectoría de Investigación y Desarrollo

3. Responsabilidades de cada área en el PGD

El presente apartado, describe en términos generales las principales responsabilidades en la elaboración de un Plan de Gestión de Datos (PGD) tanto de parte de investigadores e investigadoras, como las entidades que organizan y apoyan la investigación, en el contexto académico:

INVESTIGADORAS/ES:

- Generan los datos (crean, recolectan, recogen)
- Diseñan el PGD de acuerdo con las políticas de las instituciones y agentes de financiamiento;
- Deciden grado de apertura, confidencialidad, restricciones de acuerdo con el tipo de estudio, por medio de la verificación de condiciones de protección y apertura de datos sensibles, por ejemplo, mediante el Comité de Ética de su facultad.
- Describen y archivan los datos en los repositorios;
- Reutilizan datos generados por otros investigadores.

INSTITUCIONES DE INVESTIGACIÓN, VICERRECTORÍA DE INVESTIGACIÓN Y DESARROLLO UDEC:

- Establecen las políticas generales y flexibles de apertura y compartición de datos;
- Armonizan políticas con entes financiadores e intereses de los investigadores;
- Sugieren aspectos éticos en el uso de datos de investigación.
- Proveen infraestructura tecnológica.
- Proveen apoyo y capacitación a los investigadores/as;
- Asumen todo o parte de los costos de archivo, preservación y acceso;
- Promueven la cultura de datos abiertos.

AGENCIAS DE FINANCIAMIENTO:

- Establecen las políticas de apertura y compartición de datos;
- Proveen modelos de PGD para ser empleados por los investigadores;
- Pueden proveer repositorios;
- Promueven la cultura de datos abiertos.

CENTROS DE APOYO A LA INVESTIGACIÓN: (BIBLIOTECAS UDEC):

- Apoyo a las instituciones e investigadores en la curaduría, administración y gestión de datos y archivos (descripción, metadatos, acceso, expurgo, estándares);
- *Alfabetización en datos;*
- *Proveen servicios de búsqueda de datos;*
- *Promueven la reutilización de los datos;*
- *Promueven la cultura de datos abiertos.*

Desde el ámbito de la gestión de información, para proteger y preservar sus datos de investigación, hemos implementado un repositorio que permite el depósito y acceso a los datos de las investigaciones en servidores institucionales que garantizan su preservación y seguridad.

- [Repositorio de Datos de Investigación UdeC](#)

	Elaboración	Revisión	Aprobación
Fecha	Enero- 2023	Mayo- 2023	Junio- 2023
Cargo	Unidad de Apoyo a la Investigación - DIRBIB	Dirección de Bibliotecas Dirección de Investigación y Creación Artística	Vicerrectoría de Investigación y Desarrollo

4. Guía para la Gestión de Datos: Aspectos y preguntas en torno a la Gestión de Datos. Responsabilidades del/de la Investigador/a



1. Recopilación de datos

Lo primero es detallar qué datos se van a crear y determinar los formatos, estándares y metodologías para compartirlos.



2. Documentación y metadatos

Posteriormente se deben identificar los datos y establecer su organización, definiendo estándares y metadatos para describirlos, como por ejemplo: [DataCite Metadata Schema](#) o [DublinCore Schema](#)



3. Aspectos éticos

En esta etapa se debe explicar cualquier restricción, como una solicitud de patente pendiente, períodos de embargo o acceso restringido a datos o nombres por ser datos sensibles o personales, por ejemplo.



4. Almacenamiento

En esta fase se debe preparar una estrategia de almacenamiento, definir el repositorio donde se almacenarán los datos y definir su uso y propiedad intelectual. Como, por ejemplo: [Repositorio de Datos UdeC](#).



5. Selección y conservación

En este paso se debe seleccionar una base de datos disciplinaria apropiada, un centro de datos o un repositorio institucional, asegurándose que los datos permanezcan accesibles. Como, por ejemplo: [Repositorio de Datos](#)



6. Compartir datos

En esta fase se debe proporcionar detalles específicos sobre el acceso, es decir dónde, cuándo y cómo se pondrán a disposición sus datos. Las agencias de financiamiento suelen indicar los plazos esperados para su publicación, o si requiere imponer restricciones.



7. Responsabilidades y recursos

Finalmente, se deben definir los roles que tendrán los diferentes actores en la gestión de datos y la curación de datos dentro de los equipos de investigación.

En el siguiente enlace podrás identificar los componentes básicos de un Plan de Gestión de Datos:

<https://biblioguias.cepal.org/gestion-de-datos-de-investigacion/ComponentesPGD>

En el siguiente enlace encontrarás algunas técnicas de anonimización que le ayudarán en la protección de datos sensibles:

<https://biblioguias.cepal.org/c.php?g=495473&p=4961125>

En el siguiente enlace podrás descargar la lista de verificación o checklist, para un Plan de Gestión de Datos.

https://biblioguias.cepal.org/ld.php?content_id=34263381

(*) Traducción literal de: DCC. (2013). Checklist for a Data Management Plan. v.4.0. Edinburgh: Digital Curation Centre. Disponible en línea: http://dmponline3.dcc.ac.uk/files/DMP_Checklist_2013.pdf

Elaboración		Revisión	Aprobación
Fecha	Enero- 2023	Mayo- 2023	Junio- 2023
Cargo	Unidad de Apoyo a la Investigación - DIRBIB	Dirección de Bibliotecas Dirección de Investigación y Creación Artística	Vicerrectoría de Investigación y Desarrollo

5. Beneficios de la Gestión de Datos de Investigación (GDI)

Ahorrar tiempo y hacer uso eficiente de los recursos disponibles:
Podrás encontrar, comprender y utilizar los datos en el momento que necesites.

Facilitar la reutilización de los datos que has generado o recopilado:

La correcta gestión y documentación de los datos a lo largo de su ciclo vital permitirá que estos se mantengan exactos, completos, auténticos y fiables. Estos atributos permitirán que puedan ser comprendidos y utilizados por otras personas.

Cumplir con las exigencias de agencias de financiamiento:

Cada vez son más las agencias que exigen la presentación de planes de gestión de datos y/o el depósito de datos en repositorios como requisitos para el financiamiento de investigaciones.

Proteger y preservar los datos:

Al gestionar y depositar los datos en repositorios apropiados podrás resguardarlos de forma segura a lo largo del tiempo, protegiendo tu inversión de tiempo y recursos y permitiendo que puedan servir a nuevas investigaciones y descubrimientos en el futuro.



Para proteger y preservar sus datos, hemos implementado un repositorio que permite el depósito y acceso a los datos de las investigaciones en servidores institucionales que garantizan su preservación y seguridad.

- [Repositorio de Datos de Investigación UdeC](#)

Fuente: Basada en Comisión Económica para América Latina y el Caribe (CEPAL). (Dec 18, 2020). Biblioguias: Gestión de datos de investigación: Módulo 2 - Plan de Gestión de Datos (PGD). Recuperado de <https://biblioguias.cepal.org/gestion-de-datos-de-investigacion/PGD>

	Elaboración	Revisión	Aprobación
Fecha	Enero- 2023	Mayo- 2023	Junio- 2023
Cargo	Unidad de Apoyo a la Investigación - DIRBIB	Dirección de Bibliotecas Dirección de Investigación y Creación Artística	Vicerrectoría de Investigación y Desarrollo

6. ¿Cómo identificar tipos de datos de investigación?

Datos según su nivel de procesamiento en el curso de la investigación:

Datos primarios o sin procesar:

Datos originales que han sido recolectados, pero aún no han sido procesados o analizados. Algunos ejemplos son dos registros sonoros, observaciones, notas de campo o datos de experimentos.

Datos procesados:

Datos que han sido digitalizados, traducidos, transcritos, limpiados, validados, verificados y o anonimizados.

Datos analizados:

Modelos, gráficos, tablas, textos u otros, que han sido creados a partir de los datos primarios y procesados, y que se pretende sean de ayuda en el descubrimiento de información útil, la presentación de conclusiones y la toma de decisiones.

Datos según la fuente de la que provienen:

Datos canónicos o de referencia:

Conjunto de datos que pueden ser utilizados para validación, comparación o búsqueda de información (por ejemplo, secuencias de genoma humano, estructuras químicas o portales de datos espaciales).

Datos experimentales:

Datos generados en experimentos científicos. Generalmente son reproducibles y pueden ser generados por equipos de laboratorio.

Modelos o simulaciones:

Datos generados en computadores por algoritmos, modelos matemáticos, o simulaciones de experimentos.

Datos derivados:

Conjunto de datos creados al tomar datos ya existentes y realizar algún tipo de manipulación sobre los mismos.

Observaciones:

Datos generados al registrar observaciones de un evento específico y posiblemente irrepitable, en un lugar y tiempo dados.

En el siguiente enlace encontrarás algunas técnicas de anonimización que le ayudarán en la protección de datos sensibles:

<https://biblioguias.cepal.org/c.php?g=495473&p=4961125>

Fuente: Basada en Comisión Económica para América Latina y el Caribe (CEPAL). (Dec 18, 2020). Biblioguias: Gestión de datos de investigación: Módulo 2 - Plan de Gestión de Datos (PGD). Recuperado de <https://biblioguias.cepal.org/gestion-de-datos-de-investigacion/PGD>

	Elaboración	Revisión	Aprobación
Fecha	Enero- 2023	Mayo- 2023	Junio- 2023
Cargo	Unidad de Apoyo a la Investigación - DIRBIB	Dirección de Bibliotecas Dirección de Investigación y Creación Artística	Vicerrectoría de Investigación y Desarrollo

Datos según su forma o tipo:

- Documentos de texto electrónicos.
- Planillas de datos.
- Cuadernos de campo o anotaciones de laboratorio.
- Cuestionarios o transcripciones.
- Fotografías o películas.
- Registros sonoros.
- Muestras, artefactos, especímenes.
- Objetos digitales. Modelos, algoritmos, scripts.
- Bases de datos.
- Metadatos.
- Esquemas de metadatos.
- Configuraciones de software.
- Archivos de pre o post procedimiento de software.



Datos según su forma de representación electrónica:

- Textuales (Microsoft Word, PDF, RTF, ODT, etc.)
- Numéricos (Excel, CSV, etc.)
- Multimedia (JPG, MPEG, WAV, etc.)
- Estructurados (XML, bases de datos MySQL, etc.)
- Código de software (Java, C, etc.)
- Específicos de un software (Mesh, 3D CAD, modelo estadístico, etc.)
- Específicos de una disciplina (FITS en astronomía, CIF en química, etc.)
- Específicos de un instrumento (Olympus Confocal Microscope Data Forma, Carl Zeiss Digital Microscopic Image Format)

Fuente: Basada en Comisión Económica para América Latina y el Caribe (CEPAL). (Dec 18, 2020). Biblioguias: Gestión de datos de investigación: Módulo 2 - Plan de Gestión de Datos (PGD). Recuperado de <https://biblioguias.cepal.org/gestion-de-datos-de-investigacion/PGD>

	Elaboración	Revisión	Aprobación
Fecha	Enero- 2023	Mayo- 2023	Junio- 2023
Cargo	Unidad de Apoyo a la Investigación - DIRBIB	Dirección de Bibliotecas Dirección de Investigación y Creación Artística	Vicerrectoría de Investigación y Desarrollo

7. Proceso de depósito de datos

Bibliotecas UdeC pone a disposición de su comunidad académica el repositorio institucional que permite el depósito y acceso a los datos de las investigaciones, en servidores institucionales que garantizan su preservación y seguridad.

- [Repositorio de Datos de Investigación UdeC](#)

A continuación, se describen una serie de pasos a seguir. Sin embargo, es importante señalar que solo los administradores del repositorio, después de revisar los datos enviados y generar el respectivo contacto entre las partes involucradas, aplicarán los permisos requeridos, es decir, después de una inducción y los términos de publicación.

Para iniciar el proceso deberá:

Para ver modelos de PGD de distintas disciplinas, puede revisar en los siguientes sitios:

- [DPM Digital CurationCenter](#)
- [DMP Tool - University of California](#)

Para ver los ítems a consideraren en un PGD puede revisar:

- [Plan de Gestión de Datos ANID](#)
- [Plan de Gestión de Datos de la Unión Europea](#)



Fuente: Elaboración propia Bibliotecas UdeC – Vicerrectoría de Investigación y Desarrollo UdeC. (2023). Flujos de depósito de datos Bibliotecas UdeC – VRID.

	Elaboración	Revisión	Aprobación
Fecha	Enero- 2023	Mayo- 2023	Junio- 2023
Cargo	Unidad de Apoyo a la Investigación - DIRBIB	Dirección de Bibliotecas Dirección de Investigación y Creación Artística	Vicerrectoría de Investigación y Desarrollo

8. Bibliografía

Alonso-Arévalo, J. (2019). **La gestión de datos de investigación en el horizonte de las bibliotecas universitarias y de investigación**. Cuadernos de Documentación Multimedia, 30(0). doi:10.5209/cdmu.62806.

Angelozzi, S. M. (2020). **La gestión de datos de investigación en abierto: introducción al rol emergente para las bibliotecas universitarias y científicas argentinas**. Palabra Clave (La Plata), 9(2). doi:10.24215/18539912e091.

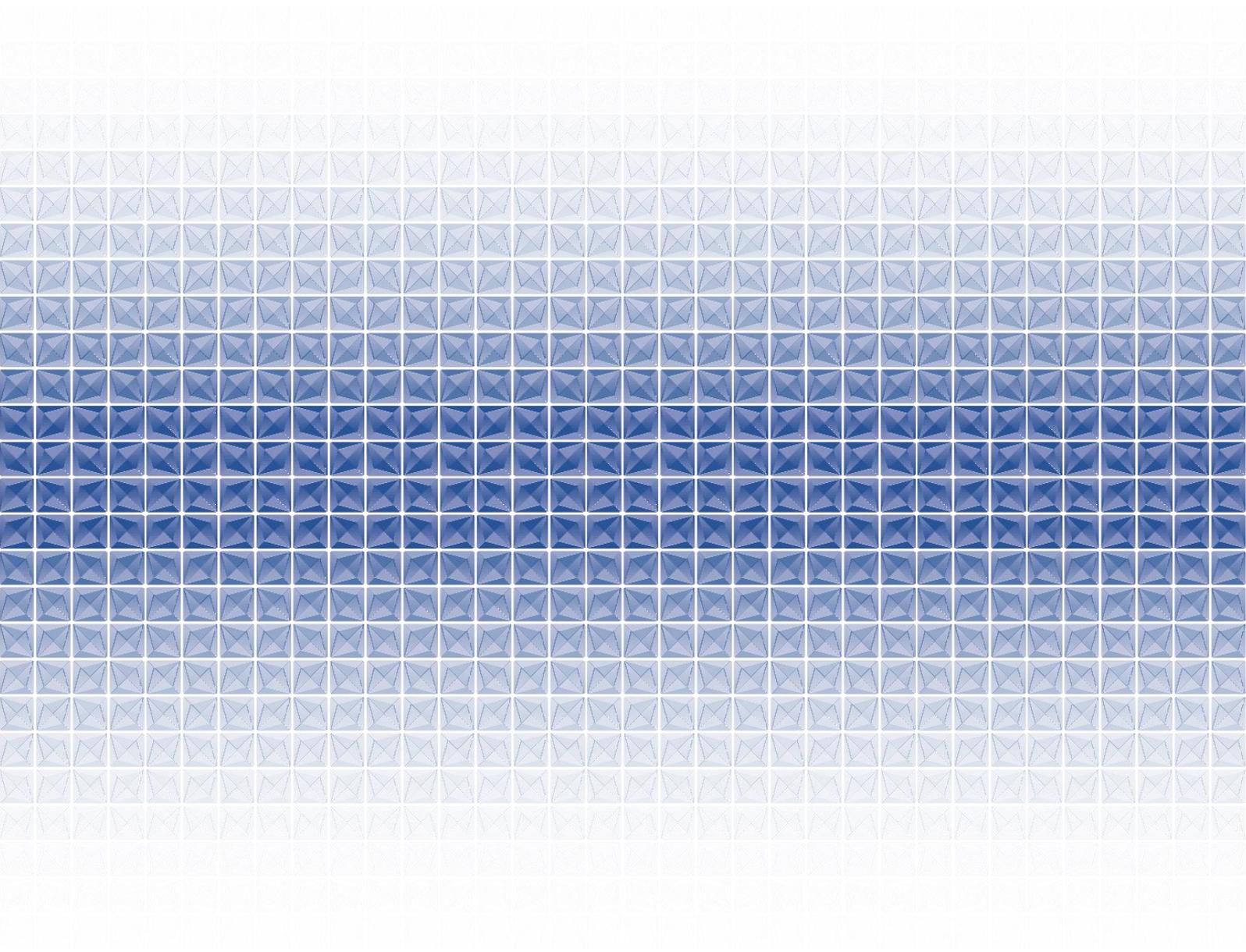
Comisión Económica para América Latina y el Caribe (CEPAL). (2020, Dec 18, 2020). **Biblioguías: Gestión de datos de investigación: Módulo 2 - Plan de Gestión de Datos (PGD)**. Recuperado de <https://biblioguías.cepal.org/gestion-de-datos-de-investigacion/PGD>

Comisión Nacional de Ciencia y Tecnología (CONICYT). (2014). **Datos científicos abiertos: La ciencia la hacemos entre todos**. Recuperado de <http://datoscientificos.cl/files/manual-2014.pdf>.

Erway, R., y Rinehart, A. (2016). **If You Build It, Will They Fund? Making Research Data Management Sustainable**. Recuperado de <http://www.oclc.org/content/dam/research/publications/2016/oclcresearch-making-research-datamanagement-sustainable-2016.pdf>

Strasser, C., Cook, R., Michener, W., y Budden, A. (2012). **Primer on Data Management: What you always wanted to know. In U. O. o. t. President (Ed.)**. Recuperado de <https://escholarship.org/uc/item/7tf5q7n3>
doi:10.5060/D2251G48

	Elaboración	Revisión	Aprobación
Fecha	Enero- 2023	Mayo- 2023	Junio- 2023
Cargo	Unidad de Apoyo a la Investigación - DIRBIB	Dirección de Bibliotecas Dirección de Investigación y Creación Artística	Vicerrectoría de Investigación y Desarrollo





Universidad de Concepción