**PLANTILLA PARA EL PLAN DE GESTIÓN DE DATOS[[1]](#footnote-1)**

**INFORMACIÓN GENERAL**

|  |
| --- |
| **Nombre del proyecto:**  |
| **ID del proyecto:**  |
| **Fecha de inicio del proyecto:** |
| **Fecha de finalización del proyecto:**  |
| **Investigador Principal***El investigador principal es quien toma las decisiones con relación al uso de datos, acorde al proyecto de investigación. En caso de que se trate de un trabajo de grado/tesis, el investigador principal es el estudiante.*

|  |
| --- |
| **Nombre:**  |
| **Institución:** |
| **Cargo en la Institución:**  |
| **Correo Institucional:**  |
| **ORCID:**  |

 |
| **Equipo investigador**

|  |
| --- |
| **Nombre:**  |
| **Institución:** |
| **Cargo en la Institución:**  |
| **Rol en el proyecto:** |
| **Correo Institucional:**  |
| **ORCID:**  |

|  |
| --- |
| **Nombre:**  |
| **Institución:** |
| **Cargo en la Institución:**  |
| **Rol en el proyecto:** |
| **Correo Institucional:**  |
| **ORCID:**  |

 |
| **Descripción general del proyecto:** ***Descripción general del proyecto en la cual se incluya su objetivo y tipo de proyecto (Proyecto financiado por convocatoria interna/externa, proyecto de semillero, proyecto de trabajo de grado/tesis, entre otros).***  |
| **Nombre de las Instituciones financiadoras:**   |

1. **TIPOS DE DATOS**
	1. **Realizar un inventario de los datos de su proyecto**

*Los datos de investigación son datos que son recolectados, observados o creados para ser analizados y producir resultados de investigación originales (Biblioteca de la CEPAL, 2020). Los datos pueden ser de tipo cuantitativo o cualitativo, y que pueden venir en muchos formatos y soportes, ya sean físicos o digitales. Como los datos pueden tomar muchas formas y pueden cambiar durante el curso del proyecto, es importante primero crear un inventario de datos potenciales que serán generados para después planear la gestión de los mismos.*

* ¿Qué tipo de datos generó (o va a generar) su proyecto? \*

*Archivos de texto, notas, datos de encuestas, imágenes, cuadernos de laboratorio, software, materiales curriculares, etc.*

* ¿Por cuáles etapas pasaron (o van a pasar) sus datos? \*

*Explique de dónde fueron tomados (o serán tomados) los datos, de qué forma, y cómo se fueron transformando hasta tener su versión final. Datos de texto pueden ser considerados como “datos brutos”, se podría usar software para re-formatear estos datos para análisis, creando un set de “datos procesados”. Después, estos datos procesados podrían ser usados para crear gráficos o ecuaciones, conformando un set de “datos analizados”. Los sets de datos que están listos para ser publicados pueden ser considerados “finalizados”.*

* ¿Se va a realizar procesamiento de los datos? \*

En caso afirmativo por favor seleccione su escenario:

1. Requiere apoyo para el procesamiento o almacenamiento masivo de datos.
2. Requiere apoyo para el procesamiento, pero no masivo.
3. Requiere procesamiento, pero no masivo (procesamiento normal, del que se puede hacer un pc de escritorio) y no requiere de ningún tipo de apoyo al respecto.
* ¿Cómo se ven (o se verán) sus datos? ¿Cuáles son los aspectos técnicos de sus datos? \*

*Tipos de archivos, promedio de tamaño, volumen y/o número estimado de datos que serán generados.*

|  |  |  |
| --- | --- | --- |
| **Conjunto de datos** | **Tipo** | **Formato y tamaño** |
| Ejemplo: Resultado de encuestas de Etiopia | Audio | VMA – 1 GB |

* ¿Cuáles son las variables principales que caracterizan sus datos? Debe existir por lo menos un diccionario de datos. \*

*Con el objetivo de que los datos sean reproducibles, se debe generar una documentación asociada que permita que estos sean claros para cualquier usuario que los requiera. Es importante considerar quién va a realizar esta documentación, la cual consiste en una explicación simple de los nombres de cada columna de la base de datos.*

|  |  |
| --- | --- |
| **Variable** | **Definición** |
| Ejemplo: Individuo (Ind.) | Esta columna hace referencia a la identificación del individuo que se encuestó, la cual se incluye de manera codificada para que no sea posible identificar la identidad del encuestado. |
| Ejemplo 2: Necesidad | Esta columna hace referencia a la principal necesidad que el individuo identificó dentro de su comunidad o grupo social. |

* 1. **Políticas de selección y evaluación.**

*No todos los datos requerirán su almacenamiento por siempre, y después de su manipulación y procesamiento, puede ser necesario únicamente guardar las versiones de los datos brutos y finales. Tomar decisiones de qué conjuntos de datos mantener y subsecuentemente compartir es muy importante dado que múltiples versiones de los datos son más difíciles y costosas de migrar, preservar y proveer acceso, comparado con únicamente seleccionar los datos principales para su inclusión dentro del plan de gestión de datos.*

* ¿Cuál es el valor de los datos de su proyecto para otros? \*

*Considere qué componentes de sus datos pueden ser valiosos, para el público en general y/o para audiencias específicas, en términos de re-uso y combinación con datos existentes.*

* ¿Cuál (es) comunidad(es) podría(n) estar interesada en sus datos? \*

*¿Sólo los integrantes de la disciplina específica estarán interesados en los datos? ¿O sus datos tendrán un valor más interdisciplinario y general? ¿Sus datos son de interés local, nacional, o internacional?*

* ¿Por cuánto tiempo sus datos serán vigentes? \*

*Considere la naturaleza de su disciplina y campo de investigación y la vigencia de la utilidad de sus datos: 5, 10, 15 años, indefinidamente, etc. Entiéndase “vigencia”, como el tiempo por el cual los datos estarán actualizados y serán valiosos para otros.*

1. **METADATOS**
	1. **Estándares de metadatos**

*Metadato puede ser definido como “información estructurada que describe, explica, localiza y/o hace más fácil la recuperación, uso o gestión de un recurso de información” tal como un conjunto de datos.*

*Por tanto, la asociación adecuada del metadato a los datos de investigación que se generan permitirá al investigador y a otros localizar, compartir, re-usar sus datos siempre y cuando cada dato esté documentado adecuadamente.*

* ¿Dónde serán almacenados los conjuntos de datos? \*

Los datos de investigación deben ser almacenados en una plataforma diferente al equipo de cómputo del investigador, que asegure su preservación a través del tiempo. Se debe especificar concretamente en qué plataforma serán almacenados los datos de investigación.

\*Recuerde que la Universidad del Rosario cuenta con el [Repositorio de datos de investigación](https://research-data.urosario.edu.co/), una plataforma diseñada especialmente para sus investigadores y aliados.

* ¿Cómo espera que sus conjuntos de datos sean descritos y almacenados? \*

|  |
| --- |
| **Nombre del Dataset 1** |
| Encuesta Etiopia 2019 |
| **Descripción**  |
| (Español) |
| **Description** |
| (Inglés) |
| **Archivos que contiene** |
| **Nombre** | **Descripción (Español)** | **Description (Inglés)** |
| Cuestionario Etiopia |   |   |
| Datos recolectados Etiopia |   |   |
| Reporte de resultados Etiopia |   |   |

|  |
| --- |
| **Nombre del Dataset 2** |
| Encuesta Vietnam 2019 |
| **Descripción**  |
| (Español) |
| **Description** |
| (Inglés) |
| **Archivos que contiene** |
| **Nombre** | **Descripción (Español)** | **Description (Inglés)** |
| Cuestionario Vietnam |   |   |
| Datos recolectados Vietnam |   |   |
| Reporte de resultados Vietnam |   |   |

1. **ACCESO Y DIFUSIÓN**
	1. **Componentes de las políticas de acceso y difusión**
* ¿Cómo se van a asignar los permisos de acceso a los miembros del proyecto? \*

| Rol | Descripción | Dataset 1 | Dataset 2  | Dataset N |
| --- | --- | --- | --- | --- |
| **Admin.** | Persona que tiene todos los permisos para todos los dataverses, datasets y archivos.  |   |   |   |
| **Curator** | Para datasets, persona que puede editar, editar permisos y publicar datasets |   |   |   |
| **Contributor** | Para datasets, persona que puede editar después enviar para revisión |   |   |   |
| **Dataverse + Dataset creator** | Persona que puede agregar subdataverses y datasets dentro de un dataverse |   |   |   |
| **Dataset creator** | Persona que puede agregar datasets dentro de un dataverse |   |   |   |
| **Dataverse creator** | Persona que puede agregar subdataverses dentro de un dataverse |   |   |   |
| **File downloader** | Persona que puede descargar un archivo publicado.  |   |   |   |
| **Member** | Persona que puede ver dataverses y datasets no publicados |   |   |   |

* ¿Qué datos se pondrán disponibles al público? \*

Se debe explicar claramente cuáles datos serán públicos, cuáles quedarán cerrados temporalmente y cuáles quedarán cerrados definitivamente.

* ¿Cómo se podrán disponibles los datos? \*

Para los datos que quedarán disponibles al público, se debe explicar claramente:

* Si se va a restringir la descarga de algún set de datos.
* Si se permitirá la descarga inmediata de los datos, o si esta quedará condicionada al diligenciamiento de los datos del visitante. Esto, a través de un formulario en el que se debe especificar determinada información previa descarga. Si este es el caso, indicar en el presente formato para qué se solicitará este formulario (por ejemplo, para hacer seguimiento de quiénes harán uso de los datos y/o para futuras investigaciones).
* ¿Cuándo se pondrán disponibles los datos? \*

Ej. *Los datos serán compartidos con otros inmediatamente después de que los resultados derivados de estos datos sean publicados, para dar oportunidad a los investigadores del proyecto de beneficiarse directamente de su trabajo.*

* 1. **Privacidad y confidencialidad/políticas**

*Algunos conjuntos de datos (datasets) son generados por procesos de investigación que involucran seres humanos y pueden contener información sensible que no quiera ser revelada públicamente.*

*Los investigadores en la UR deben seguir los lineamientos que hayan sido indicados por el Comité de ética, antes de que los datos estén accesibles a otros. Un análisis y consideración cuidadosa debe darse a la información que será directa o indirectamente transmitida por los datos en el momento de su publicación, y sí esta información necesita o no el desarrollo de medidas de protección.*

* ¿Cómo se realizará la protección la identidad de los seres humanos de ser revelada cuando los datos se hagan accesibles? ¿Cómo se demostrará el cumplimiento de las regulaciones y guías públicas e institucionales que cubren la investigación conducida y los conjuntos de datos obtenidos como resultado? \*

*Esta pregunta es obligatoria en caso de que existan datos sensibles. Para la publicación de los datos, se debe tener en cuenta no solo la remoción de información que identifique directamente seres humanos, sino también información indirecta que pueda ser agrupada y usada para identificar a los sujetos.*

* 1. **Aspectos de seguridad/políticas**

*Es importante tener en cuenta que la gestión de diferentes tipos de datos podría requerir diferentes niveles de seguridad. Por ejemplo, datos que contienen información confidencial y sensible podrían requerir medidas de seguridad adicionales que los datos que no tienen esta información. Una forma de determinar el nivel de seguridad necesario para un conjunto de datos particular será determinar qué consecuencias tendría la divulgación accidental o la pérdida de los datos.*

* ¿Quién será designado para garantizar la seguridad de los datos? \*

*A pesar de que el investigador principal del proyecto será probablemente el responsable de planear, implementar y asegurar la seguridad de los datos, los roles individuales y responsabilidades particulares del personal del proyecto deberían ser identificadas.*

1. **PROPIEDAD INTELECTUAL Y RE-USO DE LOS DATOS**

*Las agencias financiadoras pueden tener diferentes enfoques hacia la propiedad intelectual, los derechos de autor y aspectos relacionados. Se deben verificar los aspectos específicos del financiador y las políticas de la Universidad.*

* 1. **Aspectos de propiedad intelectual y derechos de autor/políticas**

*Compartir datos efectivamente requiere una cuidadosa consideración de aspectos de derechos de autor y propiedad intelectual. A pesar de que los hechos y el conocimiento común no puede ser protegido a través de propiedad intelectual, la consolidación, interpretación y expresión de datos está protegida. Los datos pueden ser licenciados, dando a los investigadores la posibilidad de establecer las condiciones bajo las cuales los datos serán usados por otros. A pesar de que los agentes financiadores generalmente reconocen los derechos de propiedad intelectual de los datos y otros resultados de proyectos financiados, cada vez hay una mayor expectativa de que los investigadores compartirán sus datos con otros dentro de un tiempo razonable y un costo mínimo.*

* ¿Quién es el titular de los derechos morales[[2]](#footnote-2) de los datos? \*
* ¿Quién es el titular de los derechos patrimoniales[[3]](#footnote-3) de los datos? \*
	1. **Re-uso y propiedad intelectual**
* ¿Se permitirá el re-uso de los datos, con o sin condiciones? \*
* ¿Se permitirá la re-distribución de los datos, con o sin condiciones? \*
* ¿Se permitirá la creación y publicación de elementos derivados de los datos, con o sin condiciones? \*
* ¿Se permitirá a otros usar los datos para el desarrollo de productos comerciales o en formas que produzcan beneficios financieros para otros, con o sin condiciones? \*
* ¿Cómo se licenciarán los datos? \*

*Condiciones para el re-uso de los datos, condiciones para la creación y publicación de datos derivados, y condiciones para propósitos comerciales.*

* ¿Cómo se realizará la atribución del trabajo a las personas que generaron los datos? \*
1. **ARCHIVO**
	1. **Estrategias de preservación**

*La preservación activa de los datos es necesaria para asegurar su disponibilidad y utilidad a largo plazo. El archivo y preservación efectiva de los datos va más allá de elementos de almacenamiento. Se centra en desarrollar estrategias para salvaguardar los datos y para asegurar que sirvan para propósitos contemporáneos. A pesar de que la preservación se realiza al final del ciclo de los datos, la preservación efectiva de datos requiere planificación y preparación, los cuales deben iniciar en la etapa de desarrollo del proyecto.*

* ¿Por cuánto tiempo considera que se deben preservar los datos? \*
	+ Se debe tener en cuenta que, en caso de que se haya decidido almacenar los datos en el Repositorio de Datos de Investigación de la Universidad del Rosario, estos se preservarán de forma indefinida y cualquier cambio sobre ellos será consultado con los autores.
* ¿Quién será el responsable de verificar que los datos estén abiertos o restringidos por el tiempo que se determine? \*
1. Plantilla adaptada teniendo como base la guía creada por la Biblioteca de la Universidad de Purdue de acuerdo con los requerimientos de planeación de gestión de datos del National Science Foudation (NSF). [↑](#footnote-ref-1)
2. Son derechos personalísimos, a través de los cuales se busca salvaguardar el vínculo que se genera entre el autor y su obra, en tanto ésta constituye la expresión de su personalidad. Surgen en cabeza del autor por el sólo hecho de creación de la obra. Son inalienables, inembargables, intransferibles e irrenunciables. [↑](#footnote-ref-2)
3. Son prerrogativas de naturaleza económico - patrimonial, con carácter exclusivo, que permiten a su titular controlar los distintos actos de explotación de que la obra puede ser objeto. Son otorgados al financiador del proceso de investigación. [↑](#footnote-ref-3)