



Mejores prácticas en Gestión, Conservación y Uso de Datos

Esta guía proporciona consejos y ejemplos de las mejores prácticas en la gestión, conservación y uso de datos en el proceso de evaluación nacional de ecosistemas. El uso de datos de alta calidad por parte de los autores garantiza de manera crítica que la evaluación se base en evidencias sólidas, lo que aumenta la credibilidad de los resultados clave. Se deben considerar planes para la gestión, conservación y uso de datos durante la etapa de *análisis inicial* de la evaluación para garantizar la precisión, validación y disponibilidad de los datos desde el inicio del proceso. También debe resaltar cualquier vacío de datos que pueda abordarse durante la evaluación.

Gestión de datos

La gestión de datos implica la planificación de todas las etapas del ciclo de vida de los datos y se implementa durante todo el proceso de evaluación. Una buena gestión de datos fomenta el uso eficaz de los datos y puede aumentar el acceso y el uso de los datos tanto por parte de los desarrolladores como de los usuarios de la evaluación, aumentar la eficiencia operativa, reducir la probabilidad de errores o la pérdida de datos y ayudar a mantener los datos seguros. Las estrategias sólidas de gestión de datos para el almacenamiento, el uso, la eliminación y el intercambio de datos son fundamentales para que los procesos de uso de datos sean seguros y eficientes.

Se puede desarrollar un **plan de gestión de datos** que describa los procesos acordados en la etapa de *análisis inicial* de la evaluación, antes de recopilar o analizar los datos y los recursos. Esto puede garantizar la coherencia en toda la organización y reducir los costes relacionados con los datos al aumentar la eficiencia.

Considere su presupuesto : La gestión de datos acarreará ciertos costes en términos de almacenamiento de datos, seguridad de los datos y tiempo del personal. Estos costes deben considerarse al comienzo del proceso de evaluación nacional de ecosistemas.

Consideraciones importantes para un plan de gestión de datos:

Se pueden definir e implementar convenciones sólidas de denominación y catalogación de archivos en todo el equipo de evaluación. Esto permitirá a los autores y otros miembros del equipo encontrar, utilizar y actualizar datos fácilmente. Lo ideal es que tanto la nomenclatura como las convenciones de catalogación sean fáciles de usar y se puedan reconocer y actualizar con facilidad en el futuro. Entre los consejos para las mejores prácticas de nomenclatura de archivos se incluyen:

- Una jerarquía de carpetas que se ajuste a la estructura de la evaluación.
- Uso de nombres de archivos descriptivos pero cortos de manera coherente.
- Uso de un formato coherente para incluir fechas en los nombres de archivos. Se sugiere AAAAMMDD.
- Incluya un número de versión.
- Utilice CamelCase y guiones bajos en lugar de espacios.

La gestión de datos es responsabilidad de todos. La Unidad de Apoyo Técnico suele ser la responsable de la organización y el almacenamiento de los datos utilizados en el proceso de evaluación. Sin embargo, todos los autores son responsables de obtener datos y metadatos y de cumplir con el plan de gestión de datos. Las funciones y responsabilidades de la gestión de datos podrían asignarse claramente a personas y equipos específicos durante el proceso de planificación de la evaluación.

Registre y actualice los metadatos durante todo el ciclo de vida de los datos. Los metadatos son información sobre los datos que es crucial para mantener la calidad de los datos. Incluyen información sobre el contenido, la estructura y los permisos de los datos y ayudan a garantizar que los datos se utilicen y reutilicen correctamente. En condiciones ideales, los planes de gestión de datos alientan al equipo de evaluación a cumplir con los principios FAIR ¹ (Encontrable, Accesible, Interoperable, Reutilizable) para gestionar los metadatos. Un acuerdo sobre los estándares de metadatos también puede contribuir a un uso coherente dentro del equipo de evaluación.

Es necesario un plan de almacenamiento de datos para garantizar la seguridad, accesibilidad y disponibilidad de los datos tanto durante como después del proceso de evaluación. El plan debe abordar el almacenamiento, la copia de seguridad, el intercambio y la preservación, todos ellos aspectos interconectados del almacenamiento de datos.

Almacenamiento de datos: El almacenamiento de datos: El almacenamiento de datos puede incluir datos digitales o no digitales. El equipo podría determinar de antemano los requisitos sobre qué información se retendrá y qué se destruirá. Las opciones de almacenamiento incluyen almacenamiento local (portátiles, unidades de red, almacenamiento óptico y dispositivos de almacenamiento externo) y almacenamiento en la nube. El almacenamiento en la nube, como Dropbox, Sharepoint y Onedrive, ha ganado popularidad porque puede reducir las inversiones de capital, el espacio de almacenamiento físico, y proporcionar múltiples niveles de seguridad de datos. El almacenamiento de datos puede incluir datos digitales o no digitales. El equipo podría determinar de antemano los requisitos sobre qué información se retendrá y qué se destruirá. Las opciones de almacenamiento incluyen almacenamiento local (portátiles, unidades de red, almacenamiento óptico y dispositivos de almacenamiento externo) y almacenamiento en la nube. El almacenamiento en la nube, como Dropbox, Sharepoint y Onedrive, ha ganado popularidad porque puede reducir las inversiones de capital, el espacio de almacenamiento físico, y proporcionar múltiples niveles de seguridad de datos.

Mejores prácticas en almacenamiento de datos digitales²:

- Utilice sistemas de almacenamiento de alta calidad.
- Utilice formatos objetivos
- Realice una migración de los archivos de datos cada dos a cinco años a un nuevo lugar de almacenamiento
- Compruebe periódicamente los datos almacenados para garantizar que no se ha perdido nada
- Utilice diferentes formas de almacenamiento para los mismos datos
- Etiquete y organice archivos almacenados de forma lógica
- Considere el cifrado

Se deben evaluar los problemas de privacidad, la facilidad de accesibilidad según la ubicación y los costes al elegir opciones de almacenamiento.

Copia de seguridad: Las mejores prácticas destacan la necesidad de realizar copias de seguridad de los datos durante y después (consulte “Preservación” a continuación) de los procesos de evaluación. Los datos se pueden perder debido a errores de hardware, problemas con el software, virus, cortes de energía, errores humanos, robo o pérdida y daños al hardware³. Una estrategia de respaldo ideal podría incluir la protección contra todos estos riesgos, priorizando ante todo la protección contra los riesgos más probables en el contexto específico. El plan para realizar copias de seguridad podría incluir qué datos incluir en las copias, con qué frecuencia se realizan, dónde guardarlas, quién es el responsable de realizarlas y si se necesitan copias de seguridad completas o incrementales. Una práctica recomendada para realizar copias de seguridad y almacenar datos es la regla 3-2-1, que recomienda tener tres copias de datos, en dos medios de almacenamiento diferentes, uno de los cuales debe almacenarse de manera externa, es decir, fuera del sitio, oficina, edificio, etc. Probar el sistema de copias de seguridad durante la configuración inicial y posteriormente a intervalos regulares ayuda a garantizar que los archivos guardados se puedan recuperar y leer.

Intercambio: Las necesidades de intercambio de datos son una consideración importante al desarrollar un

¹ Principios FAIR <https://www.go-fair.org/fair-principles/>

² Sociedad Ecológica Británica. 2018. Gestión de datos. Sociedad Ecológica Británica. Disponible en: https://www.britishecologicalsociety.org/wp-content/uploads/2019/06/BES-Guide-Data-Management-2019.pdf?utm_source=web&utm_medium=web&utm_campaign=mejor-ciencia

³ Sociedad Ecológica Británica. 2018. Gestión de datos. Sociedad Ecológica Británica. Disponible en: https://www.britishecologicalsociety.org/wp-content/uploads/2019/06/BES-Guide-Data-Management-2019.pdf?utm_source=web&utm_medium=web&utm_campaign=mejor-ciencia

plan de almacenamiento de datos para una evaluación nacional de ecosistemas. Un plan de almacenamiento de datos bien pensado puede simplificar el intercambio de datos y la colaboración con los miembros del equipo para su uso actual y futuro. Muchas opciones de almacenamiento en la nube también sirven como servicios para compartir archivos, lo que permite a los usuarios compartir y colaborar directamente con los datos, documentos y otros archivos. Es importante determinar quién necesitará acceder a los datos y desde dónde necesitarán acceder a ellos al diseñar el plan de gestión de datos, ya que esto puede ser importante para seleccionar la mejor manera de almacenarlos y compartirlos. Un paso fundamental es el desarrollo de acuerdos para compartir datos y posibles restricciones en el uso, sobre todo antes de compartir la publicación públicamente. Si los datos se van a compartir antes de la publicación de la evaluación, será necesario firmar documentos de permiso y confidencialidad. Este paso es imprescindible para garantizar la protección de los datos y sus fuentes para la evaluación.

Preservación: Contar con un plan de preservación de datos es esencial para garantizar que los datos puedan usarse más allá del proceso de evaluación. Supone el registro de datos para garantizar la viabilidad y accesibilidad futuras. Durante el proceso de planificación se podría elaborar un plan sobre cómo y dónde se archivarán los datos, que se incluiría en el plan de gestión de datos. Los datos archivados podrían incluir archivos de datos con nombres de archivo descriptivos en una estructura bien organizada, metadatos fiables y documentación y archivos complementarios que ayudarán a los futuros usuarios a utilizar y comprender los datos⁴. Por otro lado, los datos pueden depositarse en un archivo o repositorio al finalizar la evaluación para preservarlos y facilitar su uso para futuros procesos de toma de decisiones, investigaciones u otras actividades.

La seguridad y protección de los datos es esencial para garantizar su privacidad, disponibilidad e integridad. Se podría planificar e implementar una infraestructura de seguridad sólida para reducir el riesgo de filtraciones de datos. Por ejemplo, se podrían instalar un software antivirus y herramientas antispyware en todas las estaciones de trabajo. Además, el cumplimiento de las normas de privacidad podría ser esencial y deberá considerarse desde la etapa de determinación del análisis inicial de la evaluación.

Identificación de fuentes de datos, adquisición y selección datos.

Identifique las necesidades de datos

Antes de recopilar datos, es importante considerar las preguntas clave que se centrarán en la necesidad existente de información sobre el estado y las tendencias de la biodiversidad y los servicios de los ecosistemas, que da cuenta de sus contribuciones al bienestar humano en la toma de decisiones en todos los sectores. Para identificar la necesidad de una evaluación, es importante comprender el contexto ambiental, social y económico nacional actual, por ejemplo, a través del marco conceptual y los desafíos relacionados a los que se enfrentan los responsables de la toma de decisiones. Identificar estas filtraciones en la etapa del análisis inicial ayuda a los autores a entender cuáles son los datos específicos necesarios para la evaluación.

Calidad y validación de datos.

Antes de recopilar datos, también es importante considerar el nivel aceptable de calidad de los datos. El uso de datos de alta calidad garantiza que la evaluación se base en evidencias sólidas, lo que aumenta la credibilidad de los resultados clave. Se podrían utilizar métricas de calidad de los datos para ayudar a decidir si un conjunto de datos tiene la calidad suficiente. Las métricas clave de calidad de los datos por considerar incluyen los términos de uso, la accesibilidad, la frecuencia de actualización, la cobertura geográfica, la escala, la inclusión de la metodología y los metadatos de recopilación de datos, así como la autenticidad del proveedor de datos.

Consejos básicos para recopilar y seleccionar datos

- Mantenga un registro de todas las fuentes de datos, incluida la información de contacto de los proveedores de datos individuales y los sitios web donde se encontraron conjuntos de datos disponibles públicamente. Tenga en cuenta que los datos de los proveedores de datos pueden estar

⁴ Ciencia de datos de Axiom. 2017. Metadatos y documentación. Disponible en: <https://www.axiomdatascience.com/best-prácticas/MetadatayDocumentación.html#metadata-metadata>

- sujetos a acuerdos formales para compartirlos.
- Informe acerca del uso previsto de los datos a los proveedores de datos. Así se aclararán los derechos que tiene para utilizar los datos.
 - Utilice identificadores únicos para cada punto de datos dentro de un conjunto de datos. Esto ayudará a realizar un seguimiento de los datos a lo largo del tiempo, aunque cambie otra información descriptiva sobre el punto de datos.
 - Use versiones. Si un conjunto de datos se actualiza o cambia con el tiempo, cada nueva versión podría guardarse como un archivo nuevo con la misma convención de nomenclatura y una nueva marca de tiempo.
 - Se recomienda un sistema de coordenadas coherente cuando se trata de datos espaciales.
 - Se podrían utilizar títulos y valores coherentes cuando se actualicen los conjuntos de datos.

Use datos de manera efectiva

Todos los aspectos de la gestión de datos llevan al uso y reutilización de los datos. Una vez que se han identificado las fuentes de datos, es importante considerar cómo utilizar los datos de manera efectiva en los análisis.

Consejos de autores para utilizar los datos de forma eficaz

- Vaya a su propio ritmo. Puede resultar útil seguir la regla 70-20-10 para trabajar con datos. El 70 % de su tiempo debería dedicarse a buscar y catalogar datos, el 20 % a limpiar y preparar datos y el 10 % a los análisis reales.
- La organización es fundamental y tener un plan para cada paso del proceso podría ayudar a mantener la concentración. Los consejos incluyen un sistema de nombres para los resultados y la estructuración de estos. También puede tomar notas detalladas de cada paso de los análisis.
- Haga que su trabajo se pueda rastrear. Esto será útil para el uso futuro de la información y los datos recopilados como parte del proceso de evaluación. Github es una plataforma potente y útil para compartir código de análisis con socios internos o externos.
- Es esencial que se sigan los términos de uso de los conjuntos de datos y que esta información se registre en metadatos.
- Se deben leer y respetar los acuerdos de uso de datos, como los derechos de propiedad intelectual, la información de permisos y las licencias. Se recomienda usar un software de referencia para citar fuentes.