

Repositorios de Datos

Luis Norberto Zúñiga Morales

25 de enero de 2022

- 1 Objetivo de un Repositorio de Datos
- 2 Características de un Repositorio de Datos
- 3 Repositorios Gratuitos para Empezar

Objetivo de un Repositorio de Datos

Toda base de datos que se utiliza para una investigación o artículo científico debe ser almacenada para su posterior uso:

- Ya sea por el investigador o grupo de investigación o,
- Para que investigadores ajenos al trabajo puedan **verificar** resultados o **expandir** el alcance del proyecto.

Características de un Repositorio de Datos

- Asegurar **la persistencia y preservación a largo plazo** de los conjuntos de datos (mínimo de 5 años después de la publicación).
- Ser **apoyado por una comunidad** de investigación o institución de investigación.
- Proporcionar conjuntos de datos depositados con **identificadores estables y persistentes**, e.g., identificadores de objetos digitales (DOI) como DataCite, ISTIC o JaLC.
- Permitir el acceso a los datos **sin restricciones innecesarias**.
- **Proporcionar términos claros** para el uso y accesos de datos (o especificar una licencia).
- Facilitar el acceso de revisores anónimos a **datos embargados**.

Repositorios Gratuitos para Empezar

Nombre	Comentario
Google Dataset Search	Fuera de beta desde el 2020. Google para todos.
Kaggle	Sitio de competencias que también aloja datos de diversas áreas.
UCI Machine Learning Repository	Una de las bases de datos más respetada y de libre acceso, con 622 bases mantenidas al momento.
Data World	Repositorio abierto que contiene datos contribuidos por usuarios y. empresas alrededor del mundo

Repositorios Gratuitos para Empezar

Nombre	Comentario
Awesome Public Datasets	Repositorio de Github enfocado en información extraída de blogs, respuestas o comentarios en la web sobre diversos temas.
Data is Plural	Sitio que envía datos como un servicio tipo <i>Newsletter</i> .