Generalist Repository Comparison Chart

This chart is designed to assist researchers in finding a generalist repository should no domain repository be available to preserve their research data. Generalist repositories accept data regardless of data type, format, content, or disciplinary focus. For this chart, we included a repository available to all researchers specific to clinical trials (Vivli) to bring awareness to those in this field.

https://fairsharing.org/collection/GeneralRepositoryComparison

TOPIC	HARVARD DATAVERSE REPOSITORY	DRYAD	FIGSHARE	MENDELEY DATA	OSF	VIVLI	ZENODO
Brief Description	Harvard Dataverse Repository is a free data repository open to all researchers from any discipline, both inside and outside of the Harvard community, where you can share, archive, cite, access, and explore research data.	Dryad is an open data publishing platform and community committed to the open availability and routine re-use of all research data. Dryad fully curates all data and metadata and publishes exclusively under a Creative Commons Public Domain License (CC0).	Figshare is a freely available open data publishing platform for all researchers where they can share and get credit for all types of scholarly output including any file type from any research discipline. The Figshare+ repository supports sharing of larger datasets.	Mendeley Data is a free repository specialized for research data. Search more than 20+ million datasets indexed from 1000s of data repositories and collect and share datasets with the research community following the FAIR data principles.	OSF is a free and open source project management tool that supports researchers throughout their entire project lifecycle in open science best practices.	Vivli is an independent, non-profit organization that has developed a global data-sharing and analytics platform. Our focus is on sharing individual participant-level data from completed clinical trials to serve the international research community.	Powering Open Science, built on Open Source. Built by researchers for researchers. Run from the CERN data centre, whose purpose is long term preservation of digital objects. CERN maintains one of the largest scientific datasets in the world for high-energy physics.
Size limits	No byte size limit per dataset. Harvard Dataverse Repository currently sets a file size limit of 2.5GB.	300GB per dataset through browser submission system and up to 1TB with assistance from help@datadryad.org.	20GB for free figshare.com accounts. Figshare+ offers storage in tiers beginning at 100GB up to 10TB+ per dataset. System limit of 5TB/file.	10GB per dataset	Projects and child/sub projects currently have a 50GB storage limit if they are public, and 5GB limit if they are private. There is a 5GB/file upload limit for native OSF Storage. There is no limit imposed by OSF for the amount of storage used across add-ons connected to a given project.	If more than 1TB of study data, reach out to us at support@vivli.org so we can help transfer your data.	50GB per dataset, contact us via https://zenodo.org/support for higher limits
Storage space per researcher	1 TB	No limit	No limit	No limit	No limit	No limit	No limit
Persistent, Unique Identifier Support	DOI	DOI	DOI	DOI	DOI	DOI	DOI

TOPIC	HARVARD DATAVERSE REPOSITORY	DRYAD	FIGSHARE	MENDELEY DATA	OSF	VIVLI	ZENODO
Licensing Options	By default, datasets are published in the public domain (CC0). Depositors can change this to apply their own licenses.	CC0	Figshare default licenses supported: CC0 1.0 CC BY 4.0 MIT Apache 2.0 GPL v3 GPL v2 Figshare+ also supports: CC BY-SA 4.0 CC BY-ND 4.0 CC BY-NC 4.0 CC BY-NC-SA 4.0 CC BY-NC-ND 4.0 BSD 3-Clause	Default licenses supported: CC0 1.0 CC BY 4.0 CC BY NC 3.0 MIT Apache 2.0 BSD 3-Clause BSD 2-Clause GPL v3 GPL v2 LGPL MPL-2.0 CeCILL CeCILL-B CERN OHL TAPR OHL	The following 14 licenses are available: No License - is a copyright license for the project authors and contributors CC0 1.0 CC-By 4.0 MIT Apache 2.0 BSD 2-Clause BSD 3-Clause GPL 3.0 GPL 2.0 Artistic 2.0 Eclipse 1.0 LGPL 3.0 LGPL 2.1 Mozilla 2.0 Other- user defines a license in a .txt file and uploads to the project (not available on registrations or collections)	https://vivli.org/resources/vivli-data-use-agreement/	Content is available publicly under any one of 400 open licences (from opendefinition.org and spdx.org). Restricted and Closed content is also supported.
Costs to the researcher	Harvard Dataverse Repository is free for all researchers worldwide (up to 1 TB)	Costs covered by institutional, publisher, and funder members, otherwise a one-time fee of $150 for authors to cover cost of curation and preservation.	Free Figshare+ has a one time fee based on storage size for sharing larger datasets and data review.	Free	Free	$4,000 for datasets <500GB. $10,000 for datasets >500GB. Fees waived for academic researchers from member institutions. https://vivli.org/resources/sharedata/	Free
Equitable free and ongoing access to data	Use of CC0 and free data access highly encouraged. Support for broad indexing and long-term preservation strategies	Required use of CC0, publicly available, broadly indexed research data with long-term preservation strategy	Metadata is licensed CC0. All files and metadata are open access unless embargoed and can be accessed from Figshare's Open API. Figshare has a long-term preservation plan to restore public user content from an archive if needed.	Metadata is licensed CC0. Datasets are and will continue to be free access. Long-term access in the event of cease of operations granted by DANS. Accesso to archived datasets will be provided for free in perpetuity.	Open, public API to support broad indexing, partnership with Internet Archive for long-term preservation, $250k preservation fund.	Access to data is via managed access as Vivli focuses on sharing anonymized individual participant data from completed clinical research. No charge period for data that is accessible only with a research environment. Costs after no charge time period ends. https://vivli.org/resources/requestdata/	Metadata is licensed CC0. Content is both online on disk and offline on tape as part of long-term preservation policy

TOPIC	HARVARD DATAVERSE REPOSITORY	DRYAD	FIGSHARE	MENDELEY DATA	OSF	VIVLI	ZENODO
Version Support	Yes, including version comparison and W3C provenance support	Yes, versioning is supported	Yes, versioning is supported and DOIs are versioned for certain changes	Yes, including version comparison, to easily see what has changed from version to version	Yes for OSF Storage, when supported for integrated storage providers	Yes, DOIs are updated when new versions of data are added.	Yes, with "Concept" DOI to represent "all versions"

Characteristics - Standards

TOPIC	HARVARD DATAVERSE REPOSITORY	DRYAD	FIGSHARE	MENDELEY DATA	OSF	VIVLI	ZENODO
Supported Metadata Schemas	Dublin Core, Data Documentation Initiative Codebook, DataCite, OpenAIRE, Schema.org, Open Archives Initiative Object Reuse and Exchange (OAI-ORE)	DataCite, schema.org	Dublin Core (oai_dc), Datacite (oai_datacite), RDF (rdf), CERIF XML (cerif), Qualified Dublin Core (qdc) (hasPart support), Metadata Encoding and Transmission Standard (mets) and UKETD_DC (uketd_dc),,schema.org for datasets	Dublin Core, DataCite, Schema.org	Datacite, Crossref for preprint	DataCite, Schema.org	DataCite
Formats Supported for Export	JSON, XML	Fully documented API available for direct integration. Exports available in JSON, XML, schema.org	JSON, XML, DataCite, Dublin Core and other metadata schemas. Open API.	JSON, XML	JSON, API, JSONAPI	All formats are supported to export results.	DataCite, Dublin Core, DCAT-AP, JSON, JSON-LD, GeoJSON, MARCXML, Citation Style Language JSON + support for custom metadata formats
Supported Community Vocabulary or Taxonomy	ISO 3166-1 CV for geospatial metadata, ISO 639-1 CV for languages, DataCite's dataset contributor vocab, and subsets of the OBI Ontology and NCBI Taxonomy for Organisms	ORCID - Authors ROR - Institutions Open Funder Registry - Funding	ORCID - Authors GRID and ROR - Institutions Dimensions funder and grant IDs - Funding ANZSRC Field of Research (FoR) codes - Research category	Omniscience taxonomy for subject categories and keywords, and additional custom vocabularies can be loaded on institutional version	BePress 3 tier taxonomy for preprints and custom taxonomy mapped to Bepress on preprints	Cochrane Linked Data Vocabulary	ORCID - Authors, FundRef + OpenAIRE Projects Database - Funding

Characteristics Useful for Linking data to other relevant digital information

TOPIC	HARVARD DATAVERSE REPOSITORY	DRYAD	FIGSHARE	MENDELEY DATA	OSF	VIVLI	ZENODO
Support for creators/ authors identifiers	ORCID, ISNI, LCNA, VIAF, GND, DAI, ResearcherID and ScopusID	ORCID required for corresponding author, co-author ORCID supported	ORCID integration	ORCID - Mendeley ID - Scopus Author ID	ORCID, ResearcherID	ORCID	ORCID, GND, ISNI

TOPIC	HARVARD DATAVERSE REPOSITORY	DRYAD	FIGSHARE	MENDELEY DATA	OSF	VIVLI	ZENODO
Support linking to related publications	Yes	DataCite relation types	DataCite relation types	Yes, DataCite and Scholix relation types	Preprints link to related Peer-reviewed publications using Crossef 'isPreprint of', Registrations can associate with papers (and other data/code/materials) using Datacite relation types	Yes	DataCite relation types
Linking of derived products from another	Yes	DataCite relation types	DataCite relation types	Yes, DataCite and Scholix relation types	Datacite relation types	Yes	DataCite relation types
Grant ID(s)	Yes	Yes	Yes via Dimensions.ai	Yes	Yes	Yes	Yes
Grant ID affiliation(s)	Yes	CrossRef Funder Registry	Yes via Dimensions.ai	Yes	CrossRef Funder Registry	Yes	Yes
Creator/author affiliations(s)	Yes	ROR for all authors	GRID and ROR	Yes, Mendeley Institution IDs (Ringgold)	Yes	Yes	Yes
Linking to related software	Yes	DataCite relation types	DataCite relation types	Yes with semantic links	Yes	Yes	DataCite relation types
Linking to related research product	Yes	DataCite relation types	DataCite relation types	Yes with semantic links	Yes	Yes	DataCite relation types
Characteristics supporting Metrics							
Supported Data Use Metrics	Downloads, explorations, data volume	Investigations (Views), Requests (Downloads), citations	View, downloads, citations, Altmetrics	Views, downloads, altmetrics	Downloads (per version), Links, Forks, views, referrer	Views, Downloads and Citations and additional metrics https://vivli.org/resources/platform_metrics/	Views, Downloads, Data Volume, Citations, Altmetrics
Make Data Count	No	Following standards: Counter Code of Practice for Reseach Data (Make Data Count) in both standardizing and reporting usage to DataCite	Following standards: Counter Code of Practice for Reseach Data Usage (Make Data Count)	Following standards: Counter Code Of Practice for Research Data Usage Metrics (Make Data Count)	No	Following standards: Counter Code Of Practice for Research Data Usage Metrics (Make Data Count)	Following standards: Counter Code Of Practice for Research Data Usage Metrics (Make Data Count)
Characteristics supporting protection and related journal articles							
Support embargo	Yes	So long as the related journal requires or allows this	Yes	Yes	Yes for registrations	Yes	Yes

TOPIC	HARVARD DATAVERSE REPOSITORY	DRYAD	FIGSHARE	MENDELEY DATA	OSF	VIVLI	ZENODO
Supports peer review during embargo (e.g. "temporary share link")	Yes	Yes	Yes Private link feature	Yes	View only link with ability to anonymize contributor list	Yes	Yes
Support managed access	Yes, with ability to request access and grant access	No	Yes with private link and embargo features	Yes	Yes with request access and private sharing setting	Yes	Yes, with access request workflow
Preservation	Files are preserved at Harvard FAS Research Computing; archive-friendly versions of certain tabular files are also created and preserved	Core Trust Seal Certified Merritt repository with storage in US and EU, at San Deigo Supercomputing Center, DANS, and Zenodo	BagIt bags preserved in CLOCKSS	DANS. Storage in perpetuity	$250k Preservation fund, IMLS grant for transfer to Internet Archive (In progress)	Secure storage using Microsoft Azure. Vivli preserves the data, where possible, for its useful lifetime, which we define as a minimum of 10 years.	CERN Tape Archive
Human subject research data at the individual participant level	No support for preserving/publishing data with identifying human information	No support for identifying human information. Dryad does support anonymous or deidentified human subject data.	No support for identifying human information	No support for identifying human information	No	Yes, anonymized individual-level patient data in clinical research made available to qualified researchers.	Any anonymous or anonimised data
Other Information							
Business Model	Support from Harvard University, Public and Private Grants, and an emergent Consortium model	Non-profit with community memberships (institutions, publishers, funders), direct funder support through grants, and data publishing charges	Institutional, Publisher, Funder, Government agency, and Corporate SaaS model. Also one-time data publishing charges for researchers for large datasets	Subscription model for Academic & Government entities	Non-profit with direct funder support through grants, government contracts, and community memberships.	Funded via grants and member fees.	Base infrastructure by CERN, a non-profit IGO. Projects through Grants
Links to Registries for more information							
FAIRsharing record	https://doi.org/10.25504/FAIRsharing.t2e1ss	https://doi.org/10.25504/FAIRsharing.wkggtx	https://doi.org/10.25504/FAIRsharing.drtwnh	https://doi.org/10.25504/FAIRsharing.3epmpp	https://doi.org/10.25504/FAIRsharing.g4z879	https://doi.org/10.25504/FAIRsharing.uovQrT	https://doi.org/10.25504/FAIRsharing.wy4egf
re3data record	http://doi.org/10.17616/R3C880	http://doi.org/10.17616/R34S33	http://doi.org/10.17616/R3PK5R	http://doi.org/10.17616/R3DD11	http://doi.org/10.17616/R3N03T	http://doi.org/10.17616/R3SB9S	http://doi.org/10.17616/R3QP53
RRID record	https://n2t.net/RRID:SCR_001997	https://n2t.net/RRID:SCR_005910	https://n2t.net/RRID:SCR_004328	https://n2t.net/RRID:SCR_002750	https://n2t.net/RRID:SCR_003238	https://n2t.net/RRID:SCR_018080	https://n2t.net/RRID:SCR_004129