

Data Archive (Publication) Options Comparative Overview

DOI Options	Dryad	Hugging Face	Zenodo
Cost	\$150 for a 50GB dataset, then an additional \$50 per 10GB of data (different for members, eg., OSU)	Free	Free
Size Limits	300GB per data publication (flexible upon contact) recommend 10GB per file limitation	None, though <i>individual</i> files should be < 46.6GB	50GB per dataset, willing to be flexible on request, <i>IF</i> the dataset will be included in a peer-reviewed publication.
Storage	Extra copy stored with Zenodo	??? DataCite member so this should be available info?	CERN Data Centres (mostly Geneva), duplicates in Budapest, nightly tape backups
File types			No restriction
Link to Code	Through Zenodo	Dataset Card with information, part of repo as well, can archive github repo(s) with it	Stores this data too, can archive associated GitHub Repos (also get DOI)
Repository Focus	General Science	NLP, Computer Vision, and Audio	General Science
Licensing	CC0	Flexible	Required, type flexible
Time to Publication	After their review	Immediate	Immediate
Discoverability	Indexed by the Thomson-Reuters Data Citation Index, Scopus, and Google Dataset Search, search and browse on site	Well-known/utilized model and dataset repository/archive, on our org to link to	Indexed by Google Scholar, search and browse on site, can make our own community to link to
Link to Imageomics	Through metadata	On our organization page	Through metadata, Could make our own community in Zenodo
Publication links		arXiv tag to connect to papers (included in dataset card) and filter other models citing same paper	
Access		Web Interface, integration with their Spaces, Hub API Endpoints	Dev : REST API, OAI-PMH, Web Interface
Upload Process	From local computer, cloud, or remote servers (by URL), NOT GoogleDrive	Many options (UI, push, git commit, web), see adding data	In process of updating their API to take up to 50GB/file (currently 100MB/file)
Guides	README template (required), Guide for metadata available	Guide to Dataset Card Generation, allows editing through site	
Extras		Ability to host demos (eg, smaller butterfly models) Also functional as Model repository (one-stop-shop)	
Color Key	No Fill is Neutral		
Negative	Potential Negative	Potential Positive	Positive