COAR Community Framework for Good Practices in Repositories

Draft for COAR Community Input - June 16, 2020 - provide input here: https://comments.coar-repositories.org/

Objective	Essential Characteristics	Desired Characteristics
Discoverability	The repository supports high quality metadata (discipline-based, regional or general metadata schema such as Dublin Core) The repository supports harvesting of metadata using OAI-PMH Metadata in the repository are accessible, even in cases when the resource is no longer available The repository assigns a citable, persistent identifier (PID) that points to the landing page of the resource (even in cases where the resource is not available) and supports PIDs for externally housed versions of related resources (e.g. published article) The repository offers a search facility or the repository metadata and/or is indexed in an external discovery service or aggregation The repository is included in one or more disciplinary or general registries of resources (e.g. Re3data, OpenDOAR or other national, regional or domain registries)	The repository supports persistent identifiers for data at multiple levels of granularity, where appropriate. The repository supports linking between related materials such as articles, data and software HTTP link headers are supported to inform about content negotiation options The repository provides support for PIDs for authors (e.g., ORCID), funders, funding programmes and grants, institutions, and other relevant entities. Metadata are available for download in a standard bibliographic format

Access	There is no cost to the user for accessing data published in the repository The repository ensures ongoing access to data for a publicly stated time frame	The repository supports access to its records and documentation for persons with disabilities The repository provides a mechanism to transfer large datasets to users
Reuse	The repository supports the use of licenses (e.g. Creative Commons Licenses) The repository provides citable PIDs[1] The repository provides documentation and support for data citation	The repository has open API's to support full text harvesting and/or text and data mining The repository provides a list of preferred, non-proprietary formats Landing pages include the metadata required for citation (including license and PID), and ideally also metadata facilitating discovery, in human-readable and machine-readable format Resources are stored in a machine-readable community standard formats
Integrity and authenticity	The repository provides documentation of its practices that prevent unauthorized access/manipulation of resources The repository provides a record of all changes to metadata and resources in the repository The repository supports versioning of metadata and resources after deposit	The repository provides information about the content provider(s) including the name of the person(s) or institution(s) responsible for the resource

[1] A citable PID would involve the persistent identifier expressed as an URL resolving to a landing page specific for that dataset, and that landing page must contain machine readable metadata describing the dataset. We recommend the use of signposting protocol to support this.

Quality assurance	The repository provides documentation about what curation processes are applied to the resources and metadata	The repository undertakes basic curation of metadata (and data if applicable)[2] The repository supports annotation, commenting or reviewing of resource and metadata
Privacy of sensitive data (e.g. human subjects, etc.)	In cases where the repository is collecting sensitive research data, the repository has mechanisms that allow data owners to limit access to authorized users only	In cases where the repository is collecting sensitive research data, the repository provides tiered access based on the different levels of security requirements of data
Sustainability and preservation	The repository (or organization that manages repository) has a long term plan for managing and funding the repository The repository has a public resource retention policy that defines the duration of time the resources will be managed over the long term and documentation about preservation practices	Repository has a documented approach to preservation, that adopts widely accepted preservation practice The agreement between depositor and repository provides for all actions necessary to meet preservation responsibilities - e.g. rights to copy, transform, and store the items Archiving takes place according to defined workflows from ingest to dissemination
Other	The repository has a contact point or helpdesk to assist depositors and users The repository responds to queries within a reasonable time The repository provides documentation about the scope of content accepted into the repository	The repository collects and shares usage information using a standard method (e.g. number of views, downloads) The repository functions on well-supported operating systems and other core infrastructural software and is using hardware and software technologies appropriate to the services, including web readiness The submission / deposit system supports both individual creator uploads and bulk uploads of records.

[2] As defined by the CORE Seal of Approval, basic level of curation involves brief checking and addition of basic metadata or documentation where needed.