



## COAR Community Framework for Good Practices in Repositories

Public version 1 - October 8, 2020

### Purpose

The purpose of the framework is to assist repositories to evaluate and improve their current operations based on a set of applicable and achievable good practices.

Currently, there are a number of existing frameworks and evaluation criteria that were developed to assist repositories in assessing certain facets of their operations (such as discovery, access, reuse, integrity, quality assurance, preservation, privacy, and sustainability), but these criteria are spread across different organizations and are often relevant for only one region or one type of repository.

The aim of this work was to bring together relevant criteria into a **global, multidimensional framework for assessing best practices** that can be adopted and used by different types of repositories (publication, institutional, data, etc.) and in different geographical and thematic contexts.

### Process

The COAR Working Group reviewed existing frameworks, identified gaps, and assessed their level of importance, relevance and feasibility of implementation, and categorized each characteristic as either essential or desired.

The framework was disseminated to COAR members in June-August 2020 for further community feedback related to the following issues:

- Ease of Adoption - Are any of the essential characteristics too difficult to implement?
- Gaps - Are there any important criteria missing?
- Clarity - Are any of these criteria difficult to interpret?
- Relevance - Are any of these criteria not appropriate?

In the coming months, COAR will be working to provide links and citations to examples, guides and instructions that will assist the community in adopting these best practices.

The COAR Repository Assessment Working Group will review the framework on an annual basis (in July/Aug each year) to ensure it reflects current best practices and is relevant for the diverse, international repository community.

## Relationship with other assessment tools and frameworks

The following frameworks were taken into consideration for this work: [Data Citation Roadmap for scholarly data repositories](#), [Core Trust Seal](#), [FAIR data principles](#), [PLOS "Criteria that Matter"](#), [TRUST Principles for Digital Repositories](#), [COAR Next Generation Repositories Technologies Plan S](#)

## Definitions

**Content resource:** This refers to the object or item that has been deposited / uploaded into the repository (e.g. manuscript, dataset, video)

**Metadata record:** This refers to the collection of metadata elements for a given resource

**Landing page:** This describes the repository page which represents a particular content resource in the repository



Objective	Essential Characteristics	Desired Characteristics
<b>Discoverability</b>	<ul style="list-style-type: none"> <li>1.1 The repository supports quality metadata and controlled vocabularies (discipline-based, regional or general metadata schema such as Dublin Core)</li> <li>1.2 The repository supports harvesting of metadata using OAI-PMH</li> <li>1.3 The metadata in the repository are available, even in cases when the resource is no longer available</li> <li>1.4 The repository assigns a persistent identifier (PID) that points to the landing page of the resource, even in cases where the resource is not available</li> <li>1.5 The repository offers a search facility and/or the metadata is indexed by external discovery services and/or aggregators</li> <li>1.6 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)</li> </ul>	<ul style="list-style-type: none"> <li>1.7 The repository supports linking between related resources such as articles, data and software (e.g. including PIDs for related resources held elsewhere)</li> <li>1.8 The repository supports HTTP link headers to provide automated discovery of metadata records and content resources associated with repository items. We recommend <u>Signposting</u> typed links to support this.</li> <li>1.9 The repository supports PIDs for authors, funders, funding programmes and grants, institutions, and other relevant entities</li> <li>1.10 The metadata in the repository are available under a Creative Commons Public Domain License and are available for download in a standard bibliographic format</li> <li>1.11 The metadata in the repositories are available in human-readable and machine-readable formats</li> <li>1.12 In the case of data, the repository supports PIDs for data at multiple levels of granularity, where appropriate (for example, if there is research using a subset of the full dataset, a citation of the data subset will be needed)</li> </ul>
<b>Access</b>	<ul style="list-style-type: none"> <li>2.1 There is no cost to the user for accessing the resources in the repository</li> <li>2.2 The repository ensures ongoing access to resources for a publicly stated time frame</li> <li>2.3 The repository supports access to its documentation and metadata for persons with disabilities</li> <li>2.4 Device neutrality – no specific device needed for users to access the repository</li> </ul>	<ul style="list-style-type: none"> <li>2.5 The repository provides a mechanism to make very large files available to users outside of the normal user-interface (in cases where the size of the file becomes unwieldy for the user)</li> <li>2.6 In cases where there is restricted access to a resource, the repository facilitates an indirect way to access this resource (e.g. by contacting the author)</li> </ul>

<b>Reuse</b>	<p>3.1 The repository includes licensing information in the metadata record which stipulates reuse conditions</p> <p>3.2 The repository provides citable PIDs<sup>1</sup> (see 1.4)</p> <p>3.3 The repository provides a list of preferred, non-proprietary formats</p> <p>3.4 The landing pages include the metadata about the item including information required for citation in machine and human readable format</p>	<p>3.5 The repository has open API's to support full text harvesting and/or text and data mining</p> <p>3.6 The resources are stored in machine-readable, community standard formats</p>
<b>Integrity and authenticity</b>	<p>4.1 The repository provides documentation or has a policy outlining its practices that prevent unauthorized access/manipulation of resources</p> <p>4.2 The repository keeps a record of all changes to metadata and resources in the repository</p> <p>4.3 The repository supports versioning of metadata and resources after deposit</p>	<p>4.4 The repository provides information about the content provider(s) in the metadata including the name of the person(s) and/or institution(s) responsible for the resource</p>
<b>Quality assurance</b>	<p>5.1 The repository undertakes basic curation of metadata (and resources, if applicable)<sup>2</sup></p> <p>5.2 The repository provides documentation or has a policy outlining what curation processes are applied to the resources and metadata</p>	<p>5.3 The repository supports external annotation, commenting or reviewing of resources and metadata</p>
<b>Privacy of sensitive data (e.g. human subjects, etc.)</b>	<p>6.1 In cases where the repository is collecting sensitive research data, there are mechanisms that allow data owners to limit access to authorized users only</p>	<p>6.2 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</p>

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

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



<b>Preservation</b>	<p>7.1 The repository (or organization that manages repository) has a long term plan for managing and funding the repository</p> <p>7.2 The repository provides documentation or has a policy that defines the duration of time the resources will be managed over the long term and documentation about preservation practices</p>	<p>7.3 Repository has a documented approach to preservation, that adopts widely accepted preservation practice</p> <p>7.4 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</p>
<b>Sustainability and governance</b>	<p>8.1 The repository clearly indicates what organization is responsible for managing the repository</p> <p>8.2 The repository clearly indicates the nature of the governance of the services (or the organization that manages the repository)</p>	
<b>Other</b>	<p>9.1 The repository has a contact point or helpdesk to assist depositors and users</p> <p>9.2 The repository has a public notice stating it will respond to queries within a certain time frame (which is no longer than 14 days).</p> <p>9.3 The repository provides documentation or has a policy that outlines the scope of content accepted into the repository</p> <p>9.4 The repository collects and shares usage information using a standard methodology (e.g. number of views, downloads)</p>	<p>9.5 The repository functions on well-supported operating systems and other core infrastructural software</p> <p>9.6 The submission / deposit system supports both individual creator uploads and bulk uploads of records and resources.</p>