

Policies and machine actionability

Milica Ševkušić

Institute of Technical Sciences of SASA &
University of Belgrade Computer Centre
Serbia



**National Initiatives for
Open Science in Europe**

Why policies matter

- ❑ Policies (for services) inform users (humans and machines) about the principles of operation, rights, roles and responsibilities, uses and reuse, etc.
- ❑ Role in building trust in infrastructure
- ❑ Required in certification frameworks for repositories (e.g. Core Trust Seal)
- ❑ Required in the EOSC onboarding process

<https://v2.sherpa.ac.uk/id/repository/4006>

Open Access Policies

Metadata:	http://dais.sanu.ac.rs/Files/policy-dais-en.html
Data:	http://dais.sanu.ac.rs/Files/policy-dais-en.html
Content:	http://dais.sanu.ac.rs/Files/policy-dais-en.html
Submission:	http://dais.sanu.ac.rs/Files/policy-dais-en.html
Preservation:	http://dais.sanu.ac.rs/Files/policy-dais-en.html

Policy information is not available

Open Access Policies

<https://v2.sherpa.ac.uk/id/repository/4573>

Open Access Policies

Metadata:	http://crco.cssd.ac.uk/policies.html
Data:	http://crco.cssd.ac.uk/policies.html
Content:	http://crco.cssd.ac.uk/policies.html
Submission:	http://crco.cssd.ac.uk/policies.html
Preservation:	http://crco.cssd.ac.uk/policies.html

Terms (optional)

Policies + Add policies

Database access **Type** ▼

Restrictions + Add restrictions

Database licenses + Add databaseLicenses

Data accesses + Add dataAccesses

Data uploads + Add dataUploads

Data upload licenses + Add dataUploadLicenses

FAIRsharing

Access / Retrieve Data

Conditions of Use

Applies to: Data use

Individual licenses based on creative commons are annotated to each dataset
[Creative Commons CC0 1.0 Universal \(CC0 1.0\) Public Domain Dedication](#)

Applies to: Database software

[MIT Licence](#)
[Apache License 2.0](#)
[GNU General Public License \(GPL\) 3.0](#)

Data Release

Data Collection Policy	https://data.4tu.nl/info/fileadmin/user_upload/Documenten/Data_collect ...
Guideline For README-file Creation	https://data.4tu.nl/info/fileadmin/user_upload/Documenten/Guidelines_f ...

Data Curation

Data Preservation Policy	https://data.4tu.nl/info/fileadmin/user_upload/Documenten/4TU.Preserva ...
List Of Preferred File Formats	https://data.4tu.nl/info/fileadmin/user_upload/Documenten/Preferred_Fi ...
Upload Your Data	https://data.4tu.nl/info/en/use/publish-cite/upload-your-data-in-our-d ...

Data Access

Browse By Subject	https://data.4tu.nl/portal
-------------------	---

EOSC Provider Portal - Inclusion Criteria

Key information must be in English due to the limitations of current project resources (thought this may change in future)

- The provider and resource profiles must be in English
- The basic information in the User Interface for the service must be available in English
- Privacy statements, terms of use and Service Level Agreements, Specifications and Descriptions must be available in English. Other documentation may be in the native language only.
- The Helpdesk or support function must be able to answer queries in English at a minimum.

CoreTrustSeal

- ❑ Complex requirements
- ❑ Detailed documentation
- ❑ PolicIES (preservation policy is particularly important)

CoreTrustSeal Standards and Certification Board. (2019). CoreTrustSeal Trustworthy Data Repositories Requirements: Extended Guidance 2020–2022 (v02.00-2020-2022). Zenodo.
<https://doi.org/10.5281/zenodo.3632533>

Problems

- ❑ Many repositories lack policies
- ❑ Aggregators require technical standards to be met, but policies are not mandatory
- ❑ It is difficult for repository owners and managers to envisage all aspects that should be covered

An attempt to standardize repository policies:

Templates and guidelines are provided by OpenDOAR (<https://v2.sherpa.ac.uk/opensoar/policytool/>) but they do not cover all the aspects. In the old version of OpenDOAR policies were more extensively covered.

OpenDOAR Policy Support (new version)

<https://v2.sherpa.ac.uk/opensoar/policytool/>

- Metadata Policy (for information describing items in the repository)
- Data Policy (for full-text and other full data items)
- Content Policy (for document types and datasets)
- Submission Policy (concerning depositors, quality and copyright)
- Preservation Policy

▼ Open access repository policy (minimum recommended)

▼ Open access repository policy (optimum recommended)

OpenDOAR Policies Tool (old version)

Policies Tool

[Introduction](#) | [Purpose of this Tool](#) | [Non-Aims](#) | [Policies Covered](#) | [Customisation](#) | [EPrints Software](#) | [Demonstration](#)

Repository name *	<input type="text"/>
Repository URL *	<input type="text" value="http://"/>
OAI Base URL *	<input type="text" value="http://"/>
* Optional but recommended information.	
<input type="button" value="Start"/> <input type="button" value="Clear"/>	

Introduction

In a survey for OpenDOAR in early 2006, Peter Millington discovered that about two thirds of Open Access repositories did not have publicly stated policies for the permitted re-use of deposited items or for such things as submission of items, long term preservation, etc. This complicates matters for organisations wishing to provide search services, which in turn reduces the visibility and impact of these repositories.

Reference

Peter Millington (2006) [Moving Forward with the OpenDOAR Directory](#), 8th International Conference on Current Research Information Systems, Bergen, 11th-13th May 2006

Purpose of this Tool

To improve the situation, OpenDOAR has created this simple tool to help repository administrators to formulate and/or present their repository's policies. It provides a series of check boxes and pick lists for all the key policy options, which can be very quickly selected.

We have even provided recommended options for minimum compliance with the aims of the Open Access movement, and for optimising usage of your repository. For example, the minimum policy recommends allowing re-use of metadata for not-for-profit purposes, but prohibits commercial re-use. On the other hand, the optimum policy also allows free commercial re-use because the extra exposure this gives to your material probably outweighs any disadvantages.

Several output formats are provided, including: HTML (for inclusion in your website), plain text, and source code for EPrints software configuration files ([see below](#)). Other formats may follow. We welcome suggestions.

Non-Aims

This tool is not intended to generate policy statements for legal purposes. The emphasis is on clear plain language for repository users. If your lawyers wish to be more specific, we recommend you put legal statements on separate web pages. This tool lets you include links to such pages, if you wish.

Policies Covered

- **Metadata Policy**- for information describing items in the repository.

Access to metadata; Re-use of metadata

- **Data Policy** - for full-text and other full data items.

Access to full items; Re-use of full items

- **Content Policy** - for types of document and dataset held.

Repository type; Type of material held; Principal languages

- **Submission Policy** - concerning depositors, quality and copyright.

Eligible depositors; Deposition rules; Moderation; Content quality control; Publishers' and funders' embargos; Copyright policy

- **Preservation Policy**

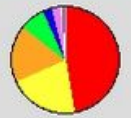
Retention period; Functional preservation; File preservation; Withdrawal policy; Withdrawn items; Version control; Closure policy

<https://web.archive.org/web/20120511075630/http://opendoar.org/tools/en/policies.php>

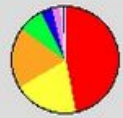
OpenDOAR (old version) - Statistics

OpenDOAR Charts - Worldwide

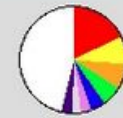
Click on a thumbnail image to display the full-size chart and its key.



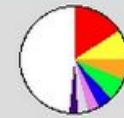
Repositories by Continent



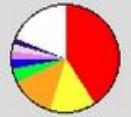
Repository Organisations by Continent



Repositories by Country



Repository Organisations by Country



Usage of Open Access Repository Software



Open Access Repository Types



Repository Operational Statuses



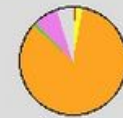
Most Frequent Content Types



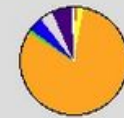
Most Frequent Languages



OpenDOAR Subjects



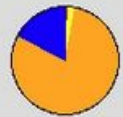
Metadata Re-use Policy Grades



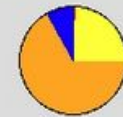
[Full Text] Data Re-use Policy Grades



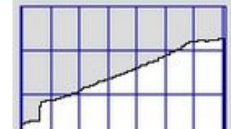
Content Policy Grades



Submission Policy Grades



Preservation Policy Grades



Growth of the OpenDOAR Database

© 2006-2013, [University of Nottingham, UK](http://www.nottingham.ac.uk). Last updated: 14-Jan-2013



- ❑ Can be upgraded to generate various policy types (currently, templates for repository policies and privacy policies)
- ❑ Modular structure
- ❑ An attempt to standardize policy content
- ❑ XML export supported
- ❑ If saved (as XML), policy may be imported back in and edited
- ❑ Designed with machine actionability in mind

<https://repol.ni4os.eu/>

Policy elements

OpenDOAR	RePol
	Preamble + Designated community (adopted from CTS)
Metadata policy (access rights and permissions for information describing items in the repository, and the minimum metadata requirements)	Metadata policy
Data policy (access rights and (re-)use permissions for full-text and other full data items)	Data policy
Content policy (tating the types and versions of documents and datasets held)	Content policy
Submission policy (eligible depositors, quality control and copyright statements)	Submission policy
	Privacy policy (also mentioning third-party applications)
	Curation policy (adopted from CTS)
Preservation policy (long-term retention, migration, and withdrawal protocols)	Preservation policy
	Harvesting policy (suggested by CORE)

Repository Policy - Repo

Date (yyyy-mm-dd): 2021-02-16

About Repo

TODO: TO REVIEW AND EDIT BY POLICY WRITER, (SEARCH FOR PARAGRAPHS WITH "TODO:")

TODO: Organization has established the institutional repository Repo). It is the digital repository of ALL (OR THOSE QUALIFIED/CONSTRAINED WITH, E.G., “research”) MEMBERS OR DEPARTMENTS OR DEPOSITING/ASSOCIATED INSTITUTIONS of Organization.

TODO: The aim of the repository is to provide open access to DIGITAL COLLECTIONS OR RESEARCH OUTPUTS OR SCHOLARLY OUTPUTS of Organization IN SOME SPECIFIC FIELD (IN LINE WITH THE REPOSITORY TYPE (institutional) AND SUBJECTS (social sciences, multidisciplinary)). The main aim of this repository is to increase the visibility of its research output / OR of the MEMBERS OR DEPARTMENTS OR ALLOWED DEPOSITORS (librarians, researchers, staff) of Organization and, at the same time, provide free access, in compliance with the LINK TO Open Access/Science Policy (ALSO PRINTABLE URL) of Organization passed by the GOVERNING BODY (DATE OF POLICY, E.G., 1 April 2020).

TODO: The aim of this policy is to ensure that the contents of Repo reflect and strengthen the Organization’s tradition of research excellence, by increasing the visibility of Organization’s research, and providing open access to its research outputs. This policy also establishes the framework within which Repo is managed.

TODO: DESCRIPTION OF REPOSITORY PLATFORM AND VENDOR; EXAMPLE: The repository uses a [DSpace](#)-based software platform developed and maintained by the MAINTAINER/SUPPORTER (WITH LINK AND PRINTABLE URL). The software platform is compliant with the [OpenAIRE Guidelines for Literature Repositories v3](#).

TODO: DESCRIPTION OF FRONT-END APP SUPPORTING DEPOSITORS OR OTHER USERS, EXAMPLE: The repository is supplied with the external application FRONT END APP NAME (LINK AND PRINTABLE URL), developed by the SOLUTION DEVELOPER (LINK AND PRINTABLE URL). In FRONT-END APP, authors and funding information are presented as normalized entities. FRONT-END APP also enables metadata export and displays [Altmetric scores](#) and [Dimensions](#) citations.

Related documents:

- [Creative Commons Licences Explainer \(below\)](#)
- **TODO:** OPTIONAL LINK TO User Manual (ALSO PRINTABLE URL)
- **TODO:** LINK TO Open Access/Science Policy (ALSO PRINTABLE URL)

TODO: Designated Community

Describe briefly the groups of current and potential users and communities for whom the content of the repository is relevant. The designated community may include various sub-communities and it may be relevant to distinguish between primary and secondary communities or internal and external users. For example:

The Designated Community of Repo includes the following user groups:

- **Internal users:** MEMBERS, STAFF OR FACULTY of Organization. They can deposit their research outputs in the repository, browse and search the metadata, and, in most cases, download content from all repository collections.
- **External users:** local and international professional researchers, citizen scientists, teachers, students, informed citizens, the general audience, etc. They can browse and search the metadata, and download the Open Access content from the repository.
- **Associates:** local and international professional researchers who have joint projects with internal users. They can browse and search the metadata, download Open Access content from the repository, access the collection dedicated to their project.

Metadata Policy

- Access to some metadata is restricted.
- The metadata of the items deposited in Repo are made available under the Creative Commons [CC0 Public Domain Dedication option](#) (“no rights reserved”, <https://creativecommons.org/publicdomain/zero/1.0/>). This dedication applies exclusively to the metadata associated with the items deposited and under no circumstances is a substitute for or affects their individual licences or permits.
- Anyone can access, copy, share, modify, harvest and re-use the metadata for any purpose, free of charge and without requiring permission from Organization.
- Organization is dedicated to providing high-quality and machine-readable item-level metadata.
- Mention of Repo is appreciated but not mandatory.

Machine readability/actionability

“While all of the policies we reviewed are machine readable in as much as they are available in HTML and/or PDF formats from the web, we did not find any evidence that any of the policies were described **using a structured data markup schema**. To enable machine-actionable workflows to be carried out in a FAIR ecosystem, it will be crucial that the status of policies governing data sharing can be interpreted unambiguously by both humans and machines.”

Davidson, Joy, Engelhardt, Claudia, Proudman, Vanessa, Stoy, Lennart, & Whyte, Angus. (2019). D3.1 FAIR Policy Landscape Analysis (v1.0_draft). FAIRsFAIR.
<https://doi.org/10.5281/zenodo.3558173>

The basis for machine actionability

Preservation

Label / ID / Description	Type / Mandatory	Displayed - "value" (default: green)
Retention period in years retention_period_years From the date of depositing. Blank or 0 for the unlimited period.	INTEGER	
If the repository is performing backups regular_backups Check if the repository is regularly backed up.	BOOLEAN	√true
Migration to new formats file_format_migration Files will be migrated to new formats to ensure continuity and accessibility, if possible.	BOOLEAN	√true
Withdrawn records are preserved records_withdrawal Check if records removed from public view remain preserved in the repository.	BOOLEAN	√true
Links to withdrawn items links_to_withdrawn_items  Check if withdrawn items' identifiers/URLs are retained indefinitely.	BOOLEAN	√false
Long-term preservation archive committed_to_preserve_repository_database Check if the database is going to be archived for long-term preservation, especially if the repository is closed down.	BOOLEAN	√false

Challenges

- ❑ Reaching consensus regarding the structure of policies and adopting common standards
- ❑ Responsibility: Who will be responsible for what?
- ❑ Sustainability: Who will maintain and update the services?
- ❑ Duplication of effort

Questions?



e-mail: biblioteka@itn.sanu.ac.rs

Twitter: @lessormore4
