PaNOSC
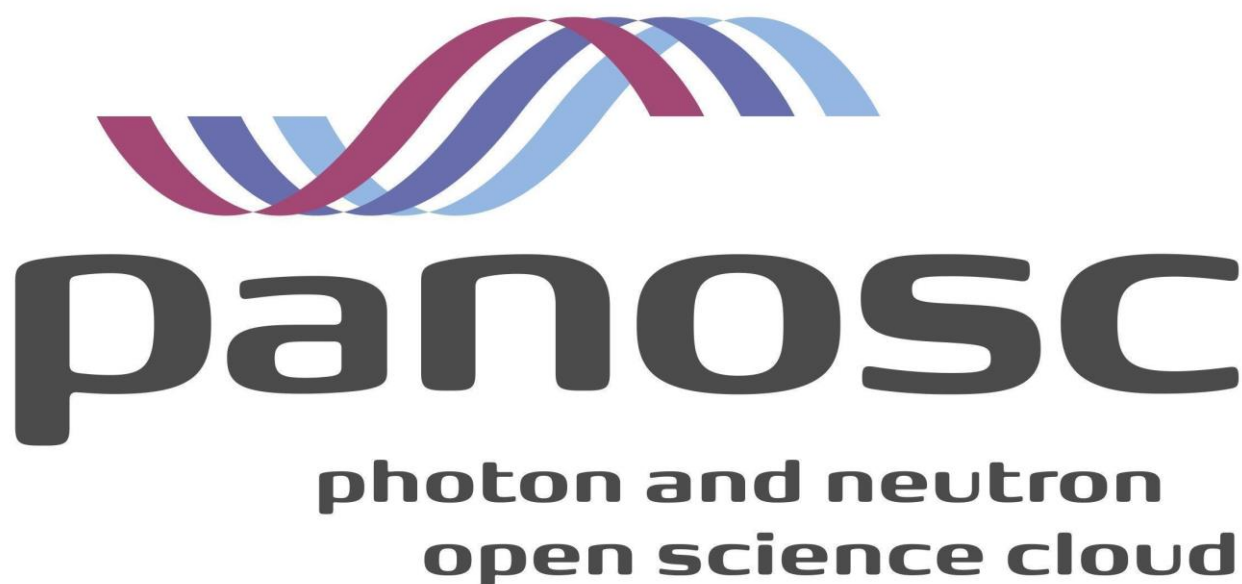# Photon and Neutron Open Science Cloud
H2020-INFRAEOSC-04-2018
Grant Agreement Number: 823852

Deliverable: D2.1 - PaNOSC data policy framework

Project Deliverable Information Sheet

Project Reference No.	823852
Project acronym:	PaNOSC
Project full name:	Photon and Neutron Open Science Cloud
H2020 Call:	INFRAEOSC-04-2018
Project Coordinator	Andy Götz (andy.gotz@esrf.fr)
Coordinating Organization:	ESRF
Project Website:	www.panosc.eu
Deliverable No:	D2.1
Deliverable Type:	Report
Dissemination Level	Public
Contractual Delivery Date:	31 May 2020
Actual Delivery Date:	26 May 2020
EC project Officer:	Rene Martins

Document Control Sheet

Document	Title: PaNOSC Data Policy Framework
	Version: DRAFT FINAL
	Available at: https://doi.org/10.5281/zenodo.3826040
	Files: 1
Date	14 May 2020
Authorship	Written by: A.Götz, J.Taylor, R.Dimper, J-F.Perrin, F.Gliksohn, D.Roccella, K.Wrona
	Contributors: A.Ashton, H. Fangohr, A. Gonzalez-Beltran, A.McBirnie, A.Markvardsen, B.Matthews, R.Pugliese, D.Salvat
	Reviewed by: WP2 PaNOSC, WP2 ExPaNDS
	Approved: J.Bodera Sempere

List of participants

Participant No.	Participant organisation name	Country
1	European Synchrotron Radiation Facility (ESRF)	France
2	Institut Laue-Langevin (ILL)	France
3	European XFEL (XFEL.EU)	Germany
4	The European Spallation Source (ESS)	Sweden
5	Extreme Light Infrastructure Delivery Consortium (ELI-DC)	Belgium
6	Central European Research Infrastructure Consortium (CERIC-ERIC)	Italy
7	EGI Foundation (EGI.eu)	The Netherlands

Table of Content

1.Introduction

The increasingly high profile of issues surrounding preservation and access of research data makes it important to work towards harmonisation of data policies across the research base. PaNOSC[1] and its sister project ExPaNDS[2] bring together a significant number of major world-class European photon and neutron research facilities to lay the foundation for a fully integrated, pan-European, data platform supporting the complete scientific cycle from experiment proposal to publication.

This document[3] describes the common framework for management of scientific data at photon and neutron facilities. The new framework is an evolution of the original PaNdata data policy framework[4] from 2010 and integrates practical experience gathered since the original inception of this data policy.

The participating facilities are used by researchers from universities, publicly-funded research entities, and industry. A common outcome of all these facilities includes the generation of raw data from each experiment, which is then analysed by the research team. The results of publicly-funded research are expected to be published in peer-reviewed scientific journals and made publicly available. In case of proprietary research, beam time is purchased by the experimental team and the results usually remain confidential.

The original PaNdata framework has been strongly influenced by the OECD "Principles and guidelines for access to research data from public funding"[5]. The current framework builds on the PaNdata framework by including the recommendations of the "Turning FAIR data into reality" report by the European Commission's FAIR Expert Group[6]. The report outlines how to interpret the FAIR principles[7], a set of guiding principles to make data Findable, Accessible, Interoperable, and Reusable. This framework, like the previous one, strives for a careful balance between competition and collaboration in science.

By definition, to be FINDABLE, any data object should be uniquely and persistently identifiable; a data object is ACCESSIBLE by machines and humans under the conditions explained in this policy; data use a formal, accessible, shared, and broadly applicable language for knowledge representation in order to be INTEROPERABLE; the data object has a plurality of accurate and relevant attributes (usage license,

---

[1] PaNOSC - Photon and Neutron Open Science Cloud (https://panosc.eu)
[2] ExPaNDS - European Open Science Cloud (EOSC) Photon and Neutron Data Service (https://expands.eu)
[3] This document, the PaNOSC data policy framework, should be cited using this DOI: https://doi.org/10.5281/zenodo.3826040
[4] PaNdata Common policy framework on scientific data: https://doi.org/10.5281/zenodo.3738497
[5] OECD (2007), OECD Principles and Guidelines for Access to Research Data from Public Funding, OECD Publishing, Paris, https://doi.org/10.1787/9789264034020-en-fr.
[6] "Turning FAIR data into reality" , (https://doi.org/doi:10.2777/54599)
[7] Wilkinson, M., Dumontier, M., Aalbersberg, I. et al. « The FAIR Guiding Principles for scientific data management and stewardship". Sci Data **3,** 160018 (2016). https://doi.org/10.1038/sdata.2016.18

provenance, community standards) to be REUSABLE.

The policy framework presented in sections 3-5 has been evaluated according to the FAIR Data Maturity Model published by the RDA[8]. The results of the evaluation are presented in Appendix 2.

Having an open access data policy with data in well-defined formats has many benefits:

- It makes previously measured data available for further analysis without the necessity to repeat the experiment.
- It promotes data use, cross-disciplinary research and machine learning.
- Raw data becomes open to scrutiny by other researchers, which ensures scientific integrity and reproducibility of experiments.
- Scientists can mine data in previously unknown ways or reapply new methods to existing data.

The data format is an essential part of making data inter-operable and machine readable. Fortunately the photon and neutron community has a standard data format (Nexus/HDF5[9]) that has been adopted by a majority of photon and neutron sources, and is supported by some detector suppliers and more and more data analysis software. The data format recommended therefore by PaNOSC for the raw data is NeXUS/HDF5, which in addition to the detector data includes sample, instrument and scientific metadata. The full strength of this digital approach will be reached when all data from the detector to the final publication are included in a digital object which is machine readable, giving full advantage to the experimental team and the scientific community.

The PanOSC Data Policy framework has been prepared in three phases: (1) a first draft was prepared based on the PaNdata data policy framework during the breakout sessions of WP2 at the first Annual Meeting of PaNOSC, and then (2) a series of ten review meetings was conducted with the experts from PaNOSC and ExPaNDS to review and enhance the contents; and finally (3) the framework was evaluated according to the FAIR Data Maturity Model (Appendix 2), which led to further improvements of the data policy.

---

[8] "FAIR Data Maturity Model: specification and guidelines" (http://dx.doi.org/10.15497/RDA00045)
[9] Nexus Data Format - https://www.nexusformat.org/

2. Recommendations

The PaNOSC photon and neutron community strongly recommends that research facilities adopt and adapt the data policy framework described in this document. Having an identical approach to scientific data management will ease the life of scientists using more than one facility and add to the overall transparency of the scientific process. It is suggested that the approved data policy at each facility be versioned and citable using a persistent identifier. It is furthermore recommended that facilities adopting the data policy framework foresee that the data policy can be updated when the need arises to reflect changes in the scientific data management either locally or internationally.

The document describes a data policy framework and is not intended to be adopted exactly as is. The minimum that the facilities adopting the framework need to do is remove this section (Recommendations), add their institute's name and the list of instruments the data policy applies to, add links to their data portal, add a suitable introduction, modify any items in the General Data Policy section, and update the Implementation Notes[10]. The framework should be adapted to the needs of the facility and local legal constraints. Nonetheless, the main objectives of the framework, namely to make scientific data produced at research facilities FAIR, must be kept.

To help ensure facilities fully implement the FAIR principles, the data policy framework has been evaluated according to the RDA FAIR Data Maturity Model and the results presented in Appendix 2. Facilities must be aware that modifying or removing items from the general framework will have an impact on the FAIRness of the resulting data policy as identified in Appendix 2. For example removing item 3.3.1 will remove compliance with all F1 criteria and therefore make the resulting data policy not FAIR compliant anymore. Facilities adopting this data policy framework are encouraged to re-evaluate their resulting data policy according to the FAIR Data Maturity Model.

---

[10] PaNOSC will provide a template In the future as part of Deliverable 2.3 to facilitate the adaptation of the Data Policy framework to new sites

3. Generic scientific data management policy

3.1. Definitions

3.1.1. For the purposes of this policy the term **facility** refers to a large-scale research centre.

3.1.2. The term **raw data** refers to data collected from experiments performed on facility instruments. [see **Implementation Note 1**]

3.1.3. The term **metadata** describes information referring to data collected from instruments, including (but not limited to) the context of the experiment, the experimental team, experimental conditions, electronic logbooks generated during the experiment and other logistical information.

3.1.4. The term **auxiliary data** refers to data that provide contextual information regarding the experiment and its datasets but which are collected outside the context of the experiment conducted at the research facility, such as information about the sample images, provenance and preparation, data processing scripts, processing environment information such as software tools and versions used, etc.

3.1.5. The term **principal investigator** (PI) refers to the main proposer identified on the experiment proposal as the main person interacting with the research facility on behalf of the experimental team. For experiments outside of the facilities proposal system, the principal investigator can be considered to be the person initiating or performing the experiment.

3.1.6. The term **experimental team** includes the PI and any other persons to whom the PI assigns the right to access resultant raw data and associated metadata.

3.1.7. The term **beam time** refers to the period of time when the experimental team has access to the facility resources to conduct an experiment.

3.1.8. The term **users** refers to the members of experimental teams, which have obtained access to beam time.

3.1.9. The term **embargo period** refers to the period during which the users have exclusive access to the data.

3.1.10. The term **public research** refers to publicly funded research which has been allocated access to the facility resources through a peer-review process and which is intended to lead to publication(s).

3.1.11. The term **proprietary research** refers to research done through

purchased (commercial) access to the research facility.

3.1.12. The term **on-line catalogue** refers to a database of metadata containing links to raw data files, that can be accessed by a variety of methods, including (but not limited to) web-based browsers on desktop and mobile devices.

3.1.13. The term **results** pertains to data, and other outcomes arising from the analysis of raw data, e.g. algorithms, workflows etc. This does not include publications, which are handled by journals.

3.1.14. The term **open access** means available freely for use by anyone without fees, copyright or patent restrictions.

3.1.15. The term **DMP** pertains to a Data Management Plan, which is a defined strategy that covers the data produced, volumes, metadata requirements, data retention periods, data disposal, processing and analysis requirements and tools. The DMP shall enable the clarification of all aspects of data management between the facility and the users before the experiment takes place.

3.1.16. The term **processed data** pertains to the data obtained by processing raw data in an automated manner usually done at the facility.

3.1.17. The term **data object** is a sequence of bytes with a persistent identifier that refers to the collection of metadata, data, files, and (possibly) software describing a data collection. In the context of this document, the data collection is the output from one or multiple experimental sessions.

3.2. General principles

3.2.1. This data policy governs the curation of and access to scientific data and metadata collected and/or stored at the facility. This includes raw, processed and auxiliary data.

3.2.2. Acceptance of this policy is a condition for the award of beam time.

3.2.3. Users shall not attempt to access, exploit or distribute raw data or metadata unless they are entitled to do so under the terms of this policy.

3.2.4. Deliberate infringements of the policy may lead to denial of access to raw data or metadata and/or denial of future beam time requests at the facility.

3.2.5. Users shall ensure raw data and processed data are collected with accurate metadata such that raw and processed data fulfill the FAIR principles. The facility will define a minimum subset of metadata as an appendix to this policy.

3.2.6. Users shall endeavour to include auxiliary data to augment the experimental data.

3.2.7. Users are required to follow any recommendation provided by the facility on what constitutes good data management and any guidelines for completing DMPs

3.2.8. GDPR[11] compliance of data and metadata at the facility is ensured by that facility. [see **Implementation Note 2**]

3.2.9. The facility will at its own discretion apply all reasonable efforts to ensure an accurate storing and curation of data as well as an uninterrupted access to data. However, failures caused by technical or human mistakes cannot be ruled out. The facility cannot warrant an absolutely accurate storing and curating. Access to data might be temporarily limited or impossible, especially due to necessary maintenance, service updates or failure of third-party service providers. [see **Implementation Note 3**]

3.2.10. The facility cannot be made liable in case of unavailability or loss of data or data analysis software.

3.2.11. Access to raw data, facility processed data, auxiliary data, results (if uploaded), and the associated metadata is restricted to the experimental team during the embargo period. Thereafter, they will become openly accessible.

3.2.12. The embargo period begins at the end of the experiment session.

3.2.13. Raw data, facility processed data, auxiliary data, and results (if uploaded) will be stored by the facility for a minimum duration of 10 years. Metadata will be stored forever. [see **Implementation Note 4**]

3.2.14. The PI can request an extension of the embargo period by following the facility defined procedure.

3.2.15. Data can always be made openly accessible earlier on request of the PI.

3.2.16. Access to raw data, processed data, auxiliary data, results (if uploaded), and the associated metadata in the facility is via a remotely searchable and indexed on-line catalogue using an open protocol. [see **Implementation Note 5**]

3.2.17. Facility support staff (e.g. instrument scientists, computing staff) have access to all data or metadata curated by the facility in order to provide support to users. The facility reserves the right to use data still under embargo to improve facility processes and performance.

3.2.18. The research facility will release open data under an appropriate license. [see **Implementation Note 6**]

---

[11] General Data Protection Regulation

3.3. Persistent identifiers

3.3.1. Persistent identifiers, for example DOIs, shall be generated for raw data and metadata. [see **Implementation Note 7**]

3.3.2. Persistent identifiers shall be generated for processed data that is generated by facility-maintained automated systems.

3.3.3. The experimental team shall be able to create a DOI for one or more specific datasets to be cited in a publication.

3.3.4. Users shall cite the persistent identifier in any publication that refers to the data (or to a subset of the data).

3.4. Raw data and associated metadata

3.4.1. All raw data and the associated metadata obtained as a result of publicly-funded access to the research facility will be open access after the embargo period, with the research facility acting as the custodian.

3.4.2. All raw data and the associated metadata obtained as a result of proprietary research will be accessible exclusively by the experimental team. Proprietary research users must agree with the facility management how they wish their raw data and metadata to be managed before the start of any experiment.

3.4.3. It is the responsibility of the PI to ensure that the metadata collected meets the minimum requirements by the facility and domain standards [see **Implementation Note 8**].

3.4.4. All raw data will be curated in well-defined formats, for which the means of reading the data will be made available by the facility. [see **Implementation Note 8**]

3.4.5. Metadata that are automatically captured by instruments will be curated in a catalogue or similar repository which links the metadata to the raw data they are describing.

3.4.6. It is recommended to add additional rich metadata, as relevant for the domain, specifically considering functionality such as data discovery.

3.4.7. Data taken on or metadata from user supplied equipment must be provided to the facility for curation.

3.4.8. Raw data and metadata will be read-only for the duration of its lifetime,

3.4.9. Data will be migrated or copied to archival facilities for curation.

3.4.10. Open data are machine downloadable via an open protocol from the remote data catalogue.

3.4.11. Each dataset will have a unique identifier. Anybody publishing results based on open access data must quote the unique identifier (and related publications if available and appropriate).

3.4.12. The on-line catalogue will enable the linking of experimental data to experimental proposals and reports. Access to proposals will only be

provided to the experimental team, reviewers and appropriate facility staff, unless otherwise authorized by the PI. Access to experiment reports is open to all.

3.4.13. The PI has the right to transfer or grant parts or all of their rights to another registered person at any time.

3.4.14. The PI has the right to create and distribute copies of the raw data at any time.

3.5. Processed data and associated metadata

3.5.1. All processed data and metadata that are generated by facility-maintained automated systems during publicly-funded experiments will be made open access after the embargo period with the facility as custodian.

3.5.2. All processed data and metadata generated by facility-maintained automated systems during proprietary research will be accessible exclusively by the client who obtained the access. Proprietary users must agree with the facility management how they wish their processed data and metadata to be managed before the start of any experiment.

3.5.3. Processed data generated by facility-maintained systems shall be curated in well-defined formats. [see **Implementation Note 8**]

3.5.4. The facility does not guarantee readability for user-generated processed data in the case the processed data are stored in a non-standard format.

3.5.5. Users must include appropriate metadata describing the provenance of the processing carried out.

3.5.6. The metadata for processed data should be interpretable across domains and communities.

3.6. Auxiliary data

3.6.1. Auxiliary data stored together with raw data for publicly-funded experiments will be made open access after the embargo period with the facility acting as custodian.

3.6.2. Auxiliary data shall be curated in the original format.

3.6.3. The facility does not guarantee the readability of auxiliary data.

3.7. Results

3.7.1. The intellectual property rights for results derived from the analysis of the raw data are determined by the contractual obligations of the

person(s) performing the analysis.

3.7.2. The facility will provide a means for users to upload results and associated metadata to the facility and enable them to associate these results with raw data collected from the facility.

3.7.3. Users must include the appropriate metadata, which describe the provenance of the results.

3.7.4. These results might be stored long-term by the research facility. It will not be the responsibility of the facility to curate these data, e.g. to ensure that software to read / process these data is available

3.7.5. Access to the results of analyses performed on raw data and metadata is restricted to the person or persons performing the analyses during the embargo period, unless otherwise decided by the PI.

4. Good practices

4.1. The experimental team is encouraged to ensure that experiments' metadata are as complete as possible, as this will enhance the possibilities for the experimental data to search for, retrieve and interpret their own data in the future.

4.2. The experimental team is strongly encouraged to provide a complete log of the protocol carried out and what happened during the experiment. The logs must be entered in the electronic logbook if the facility provides one. In the absence of a facility electronic logbook, the experimental team must use other means (electronic if possible) and link the logbook to the data.

4.3. The facility undertakes to provide means for the capture of such metadata items that are not automatically captured by an instrument, in order to facilitate recording the fullest possible description of the raw data.

4.4. Researchers who aim to carry out analyses of open data and metadata are encouraged to contact the original PI to inform them and suggest a collaboration, if appropriate. Researchers must acknowledge the source of the data and cite its unique identifier and any publications linked to the same raw data.

4.5. PIs and researchers who carry out analyses of raw data and metadata are encouraged to link the software used to obtain the results of these analyses with the raw data / metadata using the mechanisms provided by the on-line catalogue. Furthermore, they are encouraged to make such software and results openly accessible.

4.6. Researchers are strongly encouraged to follow best practices adopted by many journals concerning citing the software used and developed for the data analysis.

4.7. For each publication using facility data, authors are strongly encouraged to make available the analysis procedure description, scripts, software and software environments that completely describe the process of data analysis from the raw and metadata to the published results, and which allow others to reproduce that analysis.

4.8. Authors are encouraged to deposit these files at the facility as auxiliary data associated with the dataset at the time of the submission of the manuscript, and to make them available as open access after the publication date.

4.9. Where a software tool cannot be made available, for example for licensing reasons, the analysis procedure description should explain which tool and version has been used, and how the analysis could be repeated if that tool was available.

5. Termination of custodianship

If the facility decides to stop acting as a custodian and/or to maintainer and provider of the metadata catalogue, the facility will inform the PIs concerned in a timely manner allowing them to make a copy of the data, metadata, and results that were generated by their proposal(s), provided the facility is aware of the e-mail address of the PI.

APPENDIX 1 - Implementation Notes (IN)

This section is reserved for supplementary information about the implementation of the Data Policy. It should be adapted according to the implementation at each facility.

IN-1. Raw data refers to the experiment data generated at the facility which are persisted and is implementation specific. Raw data do not necessarily only refer to the output generated directly by the detector but may refer to data produced further down the processing pipeline.

IN-2. Some descriptive metadata (e.g. authors, date, facility, instrument, and abstract of experiment) are considered necessary for the publication of data and therefore covered by Article 85 of GDPR[12] which exempts them from the right to be removed unless they are incorrect.

IN-3. All data management and processing services are provided on a 'best effort' basis.

IN-4. The facility will strive to curate data for 10 years. However local restrictions on capacity might restrict the amount of data curated. It is preferable to keep the same curation period amongst facilities and reduce the volume of data stored rather than decrease the curation period.

IN-5. Access to data and metadata will be guaranteed via a searchable online catalogue via a web interface and open protocols like OAI-PMH or the PaNOSC search API for harvesting metadata, and https or gridftp for data.

IN-6. It is recommended to adopt one of the licences from the Creative Commons family of licences. Most commonly used ones are CC-BY, CC-BY-NC and CC0.

IN-7. The metadata should include the persistent identifier to the raw data. The recommended persistent identifier for raw data is the Digital Object Identifier (DOI) system[13]. Where possible include other persistent identifiers in use at the facility e.g. Orcid for authors.

IN-8. The facility should keep a list of supported data formats and provide documentation and tools on how to read and write data in these formats. The most widely accepted domain standard for photon and neutrons is the Nexus format[14] conventions stored in HDF5 file format. Sites are strongly encouraged to use Nexus/HDF5 as much as possible and propose extensions where Nexus is lacking in definitions.

IN-9. The local implementation of the metadata catalogue for the PaNOSC Data Policy needs to ensure the following: (1) metadata includes a link to the data persistent identifier, (2) the metadata identifiers resolve to the corresponding metadata records, and (3) the data identifier resolves to a data object.

---

[12] https://gdpr-info.eu/art-85-gdpr/
[13] https://www.doi.org/
[14] https://nexusformat.org

APPENDIX 2 - FAIR Data Maturity Model analysis

The PaNOSC Data Policy Framework have been analysed according to the FAIR indicators in the RDA FAIR Data Maturity Model[15]. The results are presented in the table and in the plots below.

FAIR	INDICATOR_ID	INDICATORS	PRIORITY	DP METRIC	DP ITEM
F1	RDA-F1-01M	Metadata is identified by a persistent identifier	Essential	4	3.3.1
F1	RDA-F1-01D	Data is identified by a persistent identifier	Essential	4	3.3.1
F1	RDA-F1-02M	Metadata is identified by a universally unique identifier	Essential	4	3.3.1
F1	RDA-F1-02D	Data is identified by a universally unique identifier	Essential	4	3.3.1
F2	RDA-F2-01M	Sufficient metadata is provided to allow discovery, following domain/discipline-specific metadata standard	Important	4	3.2.5
F3	RDA-F3-01M	Metadata includes the identifier for the data	Important	3	IN-9
F4	RDA-F4-01M	Metadata is offered/published/exposed in such a way that it can be harvested and indexed	Essential	4	3.2.16
A1	RDA-A1-01M	Metadata contains information to enable the user to get access to the data	Important	3	IN-5
A1	RDA-A1-02M	Metadata can be accessed manually (i.e. with human intervention)	Essential	3	IN-5
A1	RDA-A1-02D	Data can be accessed automatically (i.e. by a computer program)	Essential	4	3.4.10
A1	RDA-A1-03M	Metadata identifier resolves to a metadata record	Essential	3	IN-9
A1	RDA-A1-03D	Data identifier resolves to a digital object	Essential	3	IN-9
A1	RDA-A1-04M	Metadata is accessed through standardised protocol	Essential	3	IN-5
A1	RDA-A1-04D	Data is accessible through standardised protocol	Essential	3	IN-5
A1	RDA-A1-05D	Data can be accessed automatically (i.e. by a computer program)	Important	3	IN-5
A1.1	RDA-A1.1-01M	Metadata is accessible through a free access protocol	Essential	3	IN-5
A1.1	RDA-A1.1-01D	Data is accessible through a free access protocol	Important	3	IN-5
A1.2	RDA-A1.2-02D	Data is accessible through an access protocol that supports authentication and authorisation	Useful	3	IN-5
A2	RDA-A2-01M	Metadata is guaranteed to remain available after data is no longer available	Essential	4	3.2.13
I1	RDA-I1-01M	Metadata uses knowledge representation expressed in standardised format	Essential	4	3.4.4 IN-8

I1	RDA-I1-01D	Data uses knowledge representation expressed in standardised format	Important	4	3.4.4 IN-8
I1	RDA-I1-02M	Metadata uses machine-understandable knowledge representation	Essential	3	IN-8
I1	RDA-I1-02D	Data uses machine-understandable knowledge representation	Important	4	3.4.4 IN-8
I2	RDA-I2-01M	Metadata uses FAIR-compliant vocabularies	Important	3	IN-8
I2	RDA-I2-01D	Data uses FAIR-compliant vocabularies	Important	3	IN-8
I3	RDA-I3-01M	Metadata includes references to other metadata	Important	4	4
I3	RDA-I3-01D	Data includes references to other data	Useful	4	4
I3	RDA-I3-02M	Metadata includes references to other data	Useful	4	4
I3	RDA-I3-02D	Data includes sufficiently qualified references to other data	Useful	4	4
I3	RDA-I3-03M	Metadata includes qualified references to other metadata	Important	4	4
I3	RDA-I3-04M	Metadata include qualified references to other data	Useful	4	4
R1	RDA-R1-01M	Plurality of accurate and relevant attributes are provided to allow reuse	Essential	4	3.5.6 IN-8
R1.1	RDA-R1.1-01M	Metadata includes information about the licence under which the data can be reused	Essential	3	IN-6
R1.1	RDA-R1.1-02M	Metadata refers to a standard reuse licence	Important	3	IN-6
R1.1	RDA-R1.1-03M	Metadata refers to a machine-understandable reuse licence	Important	3	IN-6
R1.2	RDA-R1.2-01M	Metadata includes provenance information according to community-specific standards	Important	4	3.5.5
R1.2	RDA-R1.2-02M	Metadata includes provenance information according to a cross-community language	Useful	4	3.5.5
R1.3	RDA-R1.3-01M	Metadata complies with a community standard	Essential	3	IN-8
R1.3	RDA-R1.3-01D	Data complies with a community standard	Essential	3	IN-8
R1.3	RDA-R1.3-02M	Metadata is expressed in compliance with a machine-understandable community standard	Essential	3	IN-8
R1.3	RDA-R1.3-02D	Data is expressed in compliance with a machine-understandable community standard	Important	3	IN-8

The metrics for the indicators are 0 = not applicable, 1 = not considered, 2 = under consideration, 3 = in implementation, 4 = fully implemented. In the case of the PaNOSC Data Policy framework "fully implemented" means the indicator is included as an item in the data policy framework items, while "in implementation" means the indicator is mentioned in the Implementation Notes and is implementation specific.
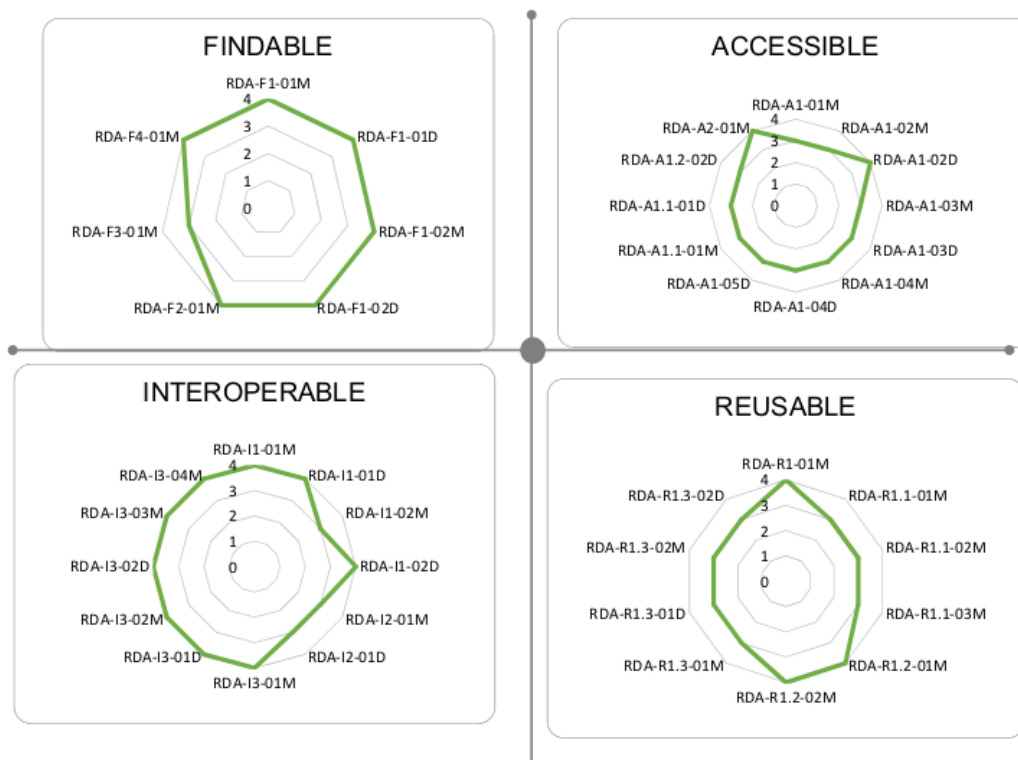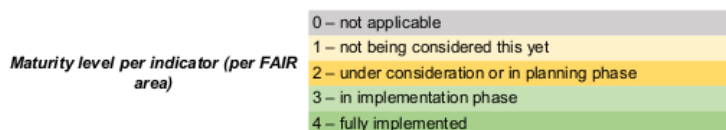
Figure 1: Plots of the FAIRness levels according to the RDA FAIR Data Maturity Model criteria.

The source code (Excel spreadsheet) used to produce the plots can be found on the PaNOSC github repository[16].

---

[16] https://github.com/panosc-eu/panosc/raw/master/Work%20Packages/WP2%20Data%20Policy%20and%20stewardship/Resources/PaNOSC_Data_Policy_FAIR_evaluation_levels_V0.01.xlsx