PaNOSC WP2 - Data Policy

16 June 2020

Andy Götz, on behalf of PaNOSC WP2

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 823852

Outline

- Background
- Lessons learned
- Updating to FAIR
- Final Result
- Next Steps

PaNdata data policy framework

- Deliverable D2.1 from PaN-Europe project in 2010
- Adopted by 10 Photon and Neutron sources in Europe
- Written 6 years before FAIR and GDPR

PaNOSC lessons learned

- Lessons learned were gathered in an 18 page document

- Experience how Data Policies were applied was documented by institute who already had a data policy e.g. ILL and ESRF

- Lessons learned were re-used

PaN-Data [Comparing Reference Data Policy]	ESRF	ILL	XFEL	CERIC	ESS
1.1. Subjects of data and metadata management policy 1.2. Condition of the award of beamtime 1.3. Users data managing limitations 1.4. Policy infringements 1.5. Data protection legislation	**Differences:** **None** **Additions:** **None** **Deletions:** **None**	**Differences:** **None** **Additions:** ***Point 1.6.*** **Unique identifier defined as Digital Object Identifier (DOI) to identify a data set < proposal number [+ reference to the ILL experiment] >** **Deletions:** **None**	**Differences:** ***Point 3.1.(Pt 1.1. in PaN-Data)*** **+ Management policy pertains also archiving data and metadata** **+ Possibility to use subcontractors to perform Scientific Data Policy obligations** ***Point 3.4.(Pt 1.4. in PaN-Data)*** **+ Punishment in case of repeated non-culpable infringements** ***Point 3.5.(Pt 1.5. in PaN-Data)*** **≠ Application of data protection legislation only if contains personal data** **Additions:** ***Point 3.6.***	**Differences:** **None** **Additions:** **None** **Deletions:** **None**	**Differences:** ***Point 1.1.*** **- Ownership and curation are not mentioned** **- Proprietary access is not considered** **+ The usage of the data pertains to this policy** **Additions:** ***Point 1.2.*** **Revision to this policy may occurred for any reason at any time and with immediate effect. It must be approved by "Council" even retroactively in case it was previously signed by General Director** ***Point 1.2.*** **The revision will be notified and the affected parties will be**

Writing a FAIR data policy framework #1

– Analysed what existing data policies found was missing or needed to be changed. Result was an spreadsheet which gathered all information

– Started with PaNdata framework text and the changed, improved and added new items as found missing during 2 days workshop + 5 online meetings

panosc

Writing a FAIR data policy framework #2

- Compared the new PaNOSC data policy framework with the FAIR Data Maturity Model (FDMM) of the RDA

- Added missing items to ensure all FDMM items were covered

- Work was done together with ExPaNDS WP2 experts

Applying the FAIR Data Maturity Model

FAIR	INDICATOR_ID	INDICATORS	PRIORITY	DP METRIC	DP ITEM
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

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

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.

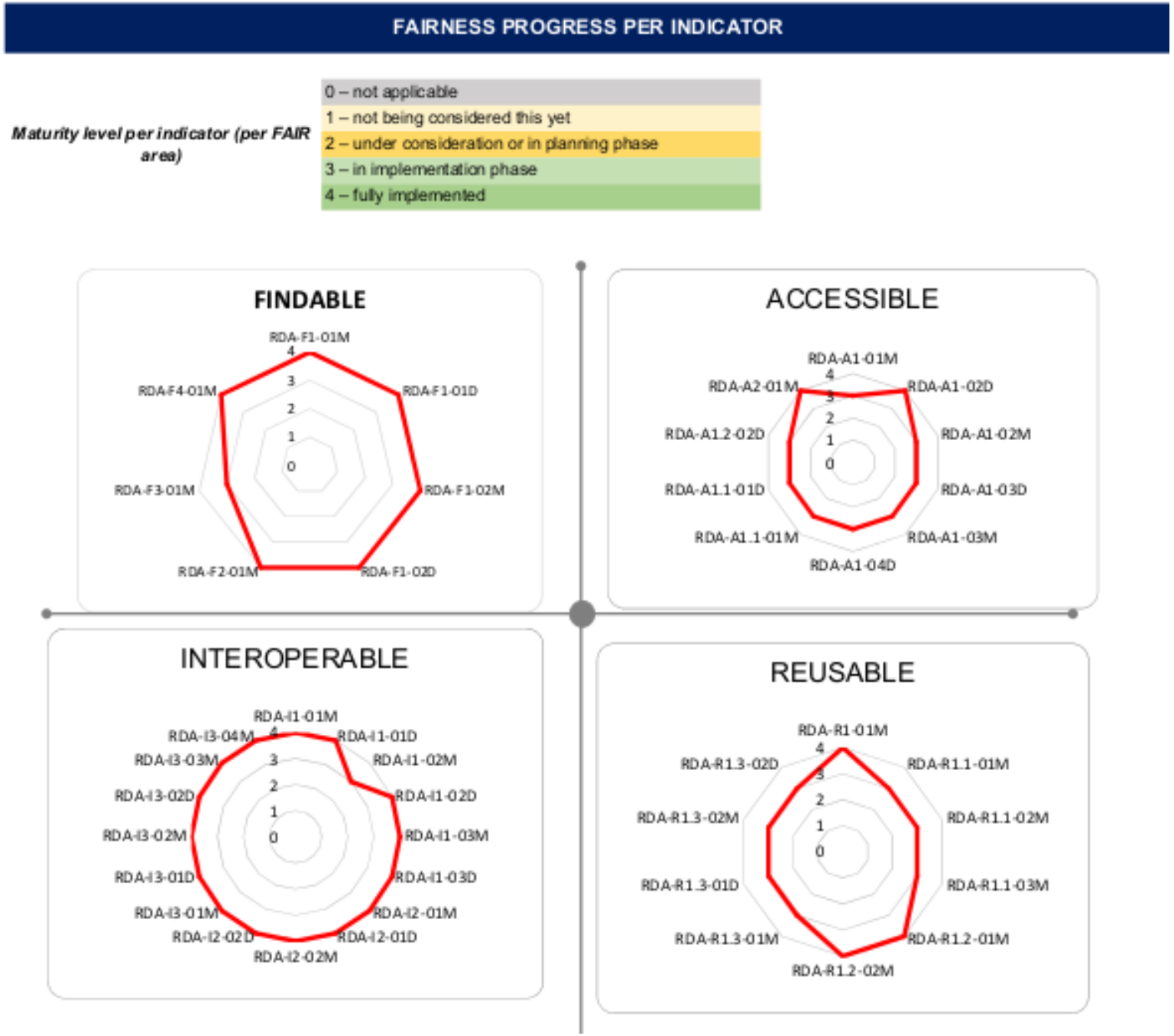
panosc

Conformity of

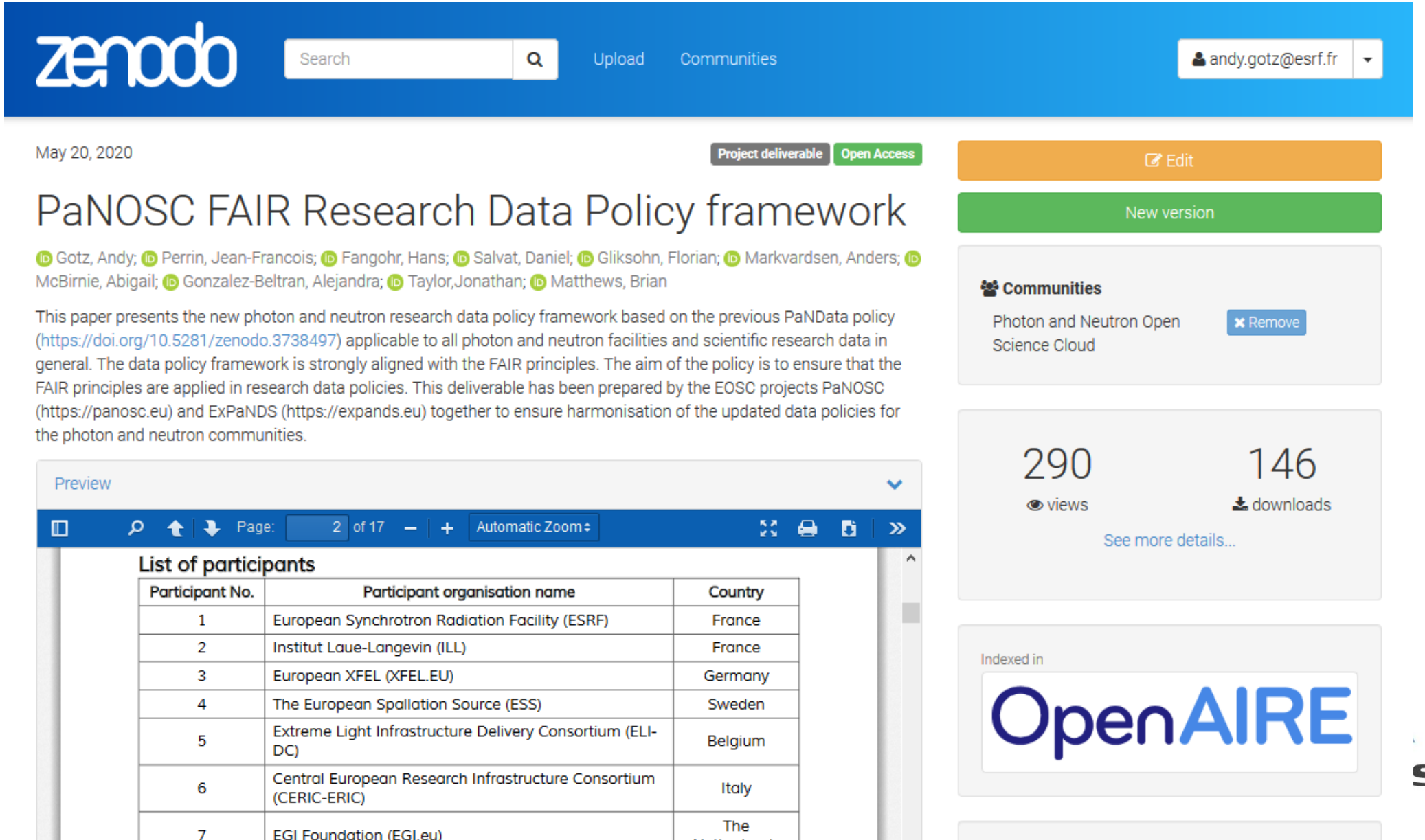
PaNOSC Data Policy Framework

with the

FAIR Data Maturity Model



FAIR Data Policy Framework



https://doi.org/10.5281/zenodo.3826039

PaNOSC KPIs

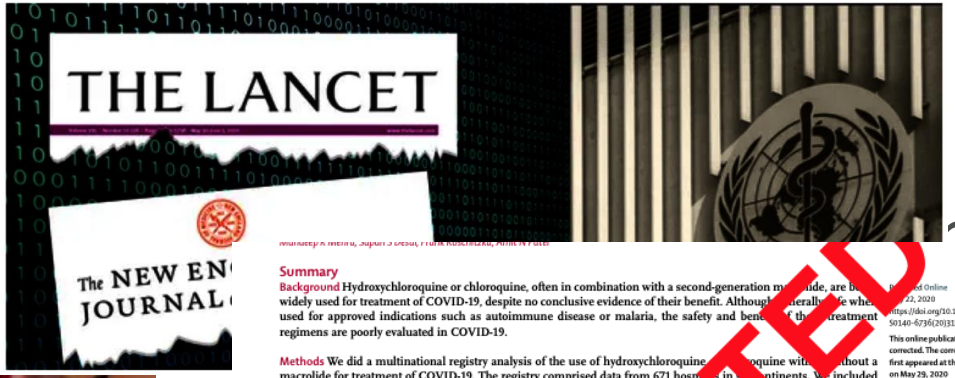
	ILL	ESRF	CERIC	XFEL	ELI	ESS
Data/year 2018	0.2 PB	8 PB	1 PB	3PB	< 1 PB	0
Data/year 2023	0.6 PB	50 PB	15 PB	100 PB	10 PB	< 1 PB
Data Policy 2018	2011	2016	2014(3/8)	2017	in prog	2017
Data Policy 2023	Updated	Updated	2021	Updated	2021	Updated
Metadata catalogue 2018	Local	Icat	Local	myMdC	No	SciCat
Metadata catalogue 2023	Local	Icat	Icat	myMdC	[TBD]	SciCat
Metadata definition 2018	Nexus	Nexus	custom	myMdC	?	Nexus
Metadata definition 2023	Nexus	Nexus	Nexus	Nexus	[Nexus]	Nexus
DOI 2018	yes	yes	no	yes	no	yes
DOI 2023	yes	yes	yes	yes	yes	yes

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 823852

PaNOSC Certifying Data Repositories

Surgisphere: governments and WHO changed Covid-19 policy based on suspect data from tiny US company

Surgisphere, whose employees appear to include a sci-fi writer and adult content model, provided database behind Lancet and New England Journal of Medicine hydroxychloroquine studies



RETRACTED

Summary

Background Hydroxychloroquine or chloroquine, often in combination with a second-generation m ide, are b widely used for treatment of COVID-19, despite no conclusive evidence of their benefit. Although generally safe whe used for approved indications such as autoimmune disease or malaria, the safety and ben of th treatment regimens are poorly evaluated in COVID-19.

Methods We did a multinational registry analysis of the use of hydroxychloroquine oquine wit out a macrolide for treatment of COVID-19. The registry comprised data from 671 hosp s in continents. We included patients hospitalised between Dec 20, 2019, and April 14, 2020, with a positive laboratory for SARS-CoV-2. Patients who received one of the treatments of interest within 48 h of diagn included in of four treatment groups (chloroquine alone, chloroquine with a macrolide, hydroxychl ne alone, or hydroxychloroquine with a macrolide), and patients who received none of these treatments formed control gr Patients for whom one of the treatments of interest was initiated more than 48 h after diagnosis or le they w on mechanical ventilation, as well as patients who received remdesivir, were excluded. The main outc of in t were in-hospital mortality and the occurrence of de-novo ventricular arrhythmias (ained or ed ventricular tachycardia or ventricular fibrillation).

Findings 96 032 patients (mean age 53·8 years, 46·3% women, OVID-19 were hospitalised during the study period and met the inclusion criteria. Of the patie were in the treatment groups (1868 received chloroquine, 3783 received chloroquine wit macr e, 3016 eived hydroxychloroquine, and 6221 received hydroxychloroquine with a macrolide) and pati in the control group. 10 698 (11·1%) patients died in hospital. After controlling for multiple fou ctors, sex, race or ethnicity, body-mass index, underlying cardiovascular disease and its risk fac diabete derlying lung disease, smoking, immunosuppressed condition, and baseline disease severity), w mpared wi ortality in the control group (9·3%), hydroxychloroquine (18·0%; hazard ratio 1·335, 95% 1·2 457), hydr ychloroquine with a macrolide (23·8%; 1·447, 1·368–1·531), chloroquine (16·4%; 1·365, 1·18–1·531), hloroquine with a macrolide (22·2%; 1·368, 1·273–1·469) were each independently associated an increased of in-hospital mortality. Compared with the control group (0·3%), hydroxychloroquine (, 2·36 935–2·900), hydroxychloroquine with a macrolide (8·1%; 5·106, 4·106–5·983), chloroquine (4·3%; 1, 2 0–4·596), and chloroquine with a macrolide (6·5%; 4·011, 3·344–4·812) were independently associate an inc ed risk of de-novo ventricular arrhythmia during hospitalisation.

Interpreta We w unable firm a benefit of hydroxychloroquine or chloroquine, when used alone or with a macr on in spital outcomes for COVID-19. Each of these drug regimens was associated with decreased in-hospit reased frequency of ventricular arrhythmias when used for treatment of COVID-19.

Published Online
22, 2020
https://doi.org/10.1016/S0140-6736(20)31180-6
This online publication has been corrected. The corrected version first appeared at thelancet.com on May 29, 2020
See Online/Comment
https://doi.org/10.1016/S0140-6736(20)31174-0

Brigham and Women's Hospital Heart and Vascular Center and Harvard Medical School, Boston, MA, USA (Prof M R Mehra MD); Surgisphere Corporation, Chicago, IL, USA (S S Desai MD); University Heart Center, University Hospital Zurich, Zurich, Switzerland (Prof F Ruschitzka MD); Department of Biomedical Engineering, University of Utah, Salt Lake City, UT, USA (A N Patel MD); and HCA Research Institute, Nashville, TN, USA (A N Patel)
Correspondence to: Prof Mandeep R Mehra, Brigham and Women's Hospital Heart and Vascular Center and Harvard Medical School, Boston, MA 02115, USA
mmehra@bwh.harvard.edu

PaNOSC WP2 KPIs

KPI	Description	Before	30/11/2019
2.1.1	number of instruments on which Data Policy is implemented	**53**	**62**
2.1.2	number of techniques / instruments for which metadata are defined	**39**	**41**
2.1.25	number of instruments available	**135**	**139**
2.1.26	Percentage of techniques for which metadata is defined	**113%**	**114%**
2.1.3	number of metadata parameters defined	**1310**	**1,880**
2.1.4	number of raw data DOIs	**3591**	**4,612**
2.1.5	number of user defined DOIs	**5**	**12**
2.1.6	number of downloads	**95**	**194**
2.1.7	number of citations of DOIs	**242**	**383**
2.1.8	number of datasets	**1110**	**1,584**
2.1.9	volume of data archived	**5**	**10**
2.1.10	number of datasets cited in publications	**237**	**375**

Next Steps

- Adopt – adopt (CERIC-ERIC, ELI) or update existing data policies (ILL, ESRF, EuXFEL) to be compliant with FAIR
- Common PaN data policy – continue working closely with ExPaNDS to have a common FAIR data policy for the PaN community
- Data Management Plans –working closely with ExPaNDS to implement DMPs for the PaN community
- Share – share the FAIR data policy framework with the FAIRsFAIR community and https://fairsharing.org

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 823852

Conclusion

- FAIR – writing a FAIR data policy which takes into account all the FAIR principles sufficiently e.g. Complying with FDMM, requires paying attention to a lot of details

- FDMM – this output of the RDA has proved to be very useful.

- FAIR WG – we will integrate the recommendations from the EOSC FAIR working group concerning data policies and PIDs

- Future – next steps are adopting the PaNOSC FAIR data policy framework at all partner sites and agreeing on PaN policy with ExPaNDS

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 823852

Thank you

andy.gotz@esrf.fr

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 823852