Marjan Grootveld

Data Archiving and Networked Services - DANS The Netherlands

"But can I trust your data?"



Swiss Research Data Day 2018, ETH Zürich, 12 June 2018



appenaire eu @MarjanGrootveld **@DANSKNAW**







"No progress without reuse" OpenAIRE

- The potential of existing data
- The importance of documentation
- The support of trustworthy repositories





What is DANS?

Mission: promote and provide permanent access to digital research resources

> Institute of Dutch Academy and Research Funding Organisation (KNAW & NWO) since 2005

European

First predecessor dates back to 1964 (Steinmetz Foundation), Historical Data Archive 1989



		HOME	M.J. GROOT	VELD	MIT DATASETS
E	ASY				
er for the new pe	er reviewed, onli	ne-only o	pen access	Resear	ch Data Journal
For more i	nfo: brill.com/rdj				
archiving of research o	data and access to thou	usands of da	tasets.		
			SEAR	CH .	Search bala
			SEAR		2 Search neip
Browse					
	\bigcirc				
	CORE	, ne	stor		
ds (PDF).	SEAL	2	Seal 2016	WORLD	DATA SYSTEM
	\sim				

EASY: certified long-term **Electronic Archiving System** for self-deposit

by our training sessions, consultancy and information material

CONSULTANCY

DANS assists in developing data management policy and obtaining certification. »»



INFORMATION MATERIAL

Watch the video 'Why share data' or download other information material. »

OpenAIRE OpenAIRE

- Open Access Infrastructure for Research in Europe
- Funded by Horizon2020 to develop and maintain the infrastructure to support OA policy of the EU
- Supports H2020 OA mandates **0100%** OA on scientific publications **Open Research Data Pilot**
- A National Open Access Desk in each country
- 2018 2020: OpenAIRE Advance
- DANS leads taskgroup Research Data Management





SWITZERLAND



ANDRE HOFFMANN







OpenAIRE RDM support (selection!)

- Briefing papers, factsheets, webinars, workshops, FAQs (with example DMPs)
- Information on:
 - Open Research Data Pilot
 - Creating a data management plan
 - Selecting a data repository



OpenAIRE Horizon2020 FactSheets

Open Research Data Pilot in Horizon 2020 How can OpenAIRE help?

What is the Open Research Data Pilot?

Updated on 19 October 20

WHAT IS THE OPEN RESEARCH DATA PILOT?

Open data is data that is free to access, reuse, repurpose, and redistribute. The Open Research D Horizon 2020 projects accessible with as few restrictions as possible, while at the same time protect

If your Horizon 2020 project is part of the pilot, and your data meets certain conditions, your must deposit your data in a res findable and accessible for others. Don't panic - you are not expected to share sensitive data or breach any IPR agreem need to deposit all the data you generate during the project either – only that which underpins published research findings a to supporting your research's integrity, openness has many other benefits. Improved visibility means your research will a impart – for science, society and your own career. Recent studies have shown that citations increase when data is made an papers also have a longer shell-life.

Anonymisation tool Amnesia

https://www.openaire.eu/webinars/ https://www.openaire.eu/what-is-the-open-research-data-pilot https://www.openaire.eu/support https://www.openaire.eu/support/faq https://amnesia.openaire.eu/



OpenAIRE



OpenAIRE Research Data Management Briefing paper

Understanding Research Data Management

April 2017



H2020-EINFRA-2014-1 Topic: e-Infrastructure for Open Access Research & Innovation action Grant Agreement 643410

QUESTION:



OpenAIRE

Who has ever used data that was generated or collected by someone else?

What is the *oldest* data that you have used?





A OpenAIRE

Case study: Viking Lander data

When the US space agency NASA sent two Viking Landers to Mars in 1975 to find out whether life might exist on the red planet, it was assumed that the datasets painstakingly compiled by scientists at the time would be available for future generations of scientists on magnetic tape.

Yet, just a few decades later, despite the space agency's best efforts to keep the tapes in a climate-controlled environment, time has left them cracking and brittle. Furthermore, when scientists attempted to re-use some of the data in the late 1990s, they found that they could not decode the formats used. In the end they had to track down old printouts and retype everything.¹



http://www.dpconline.org/docman/miscellaneous/advocacy/340-mind-the-gap-assessing-digital-preservation-needs-inthe-uk/file Data now available from https://pds-imaging.jpl.nasa.gov/volumes/viking.html

Viking Lander High Resolution Mosaics, Stereo Images and Range Data			
The follo copying reformation	wing are I data direc ted.	NOT PDS formatted volumes. They were produced by the Science Digital tly off of old, decaying tape media onto more stable CD-WO media. They	
vl_2011		tapes DNM_001-008, DNR_001-143, and DNS_001-016	
vl 2012	H	tapes FNM 001-012 and FNS 001-024	



Sets - SDDPT

Data Preservation Task by have not been otherwise



Climatological database for the world's oceans



Image copied from https://www.knmi.nl/kennis-en-datacentrum/achtergrond/cliwoc Every yellow dot represents a ship report. Project website: http://pendientedemigracion.ucm.es/info/cliwoc/











Research data should be freely accessible to everyone - for scientists as well as for the general public.

The SNSF agrees with this principle. Since October 2017, researchers have to include a data management plan (DMP) in their funding application for most of the funding schemes. At the same time, the SNSF expects that data generated by funded projects are publicly accessible in digital databases provided there are no legal, ethical, copyright or other issues.

4. Data sharing and reuse

4.1 How and where will the data be shared?

- Ouestions you might want to consider - On which repository do you plan to share your data
- How will potential users find out about your data?

4.2 Are there any necessary limitations to prote sitive data?

Questions you might want to consider: Under which conditions will the data be made as (timing of data release, reason for delay if applicable

4.3 I will choose digital repositories that are confe the FAIR Data Principles. [CHECK BOX]

4.4 I will choose digital repositories maintaine non-profit organisation. [RADIO BUTTON yes/no] \rightarrow If the answer is no: "Explain why you cannot shar data on a non-commercial digital repository '



http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf http://www.snf.ch/en/theSNSF/research-policies/open_research_data/Pages/default.aspx





H2020 Programme

Guidelines on

FAIR Data Management

-

1?		C a s	Consider how and on which repository the data will be made available. The methods applied to data sharing will depend on several factors such as the type, size, complexity and sensitivity
		v tl	valued and acknowledged by other researchers. (This relates to he <i>FAIR Data Principles</i> F1, F3, F4, A1, A1.1, A1.2 & A2)
t s			Sala have to be shared as soon as possible, but at the latest at
ail	abl	ti s le ti	he time of publication of the respective scientific output. Re- strictions may be only due to legal, ethical, copyright, confiden- iality or other clauses. Consider whether a non-disclosure
?	ſ	a (*	agreement would give summent protection for confidential data. This relates to the <i>FAIR Data Principles</i> A1 & R1.1)
orn	n t	о Т	The SNSF requires that repositories are conform to the
orn	n t	о Т F	The SNSF requires that repositories are conform to the FAIR Data Principles (Section 5 c
orn	n t	0 T F	The SNSF requires that repositories are conform to the FAIR Data Principles (Section 5 c searchers, SNSF's explanation of the <u>FAIR Data Principles</u>).
orn	n t	o T F <u>S</u> Ii	The SNSF requires that repositories are conform to the FAIR Data Principles (Section 5 c searchers, SNSF's explanation of the <u>FAIR Data Principles</u>). If there are no repositories completed
orn	n t	o T F S Ii	The SNSF requires that repositories are conform to the FAIR Data Principles (Section 5 c earchers, SNSF's explanation of the <u>FAIR Data Principles</u>). If there are no repositories completed on the the sector of th
orn	nt	o T F <u>S</u> Ii n d	The SNSF requires that repositories are conform to the FAIR Data Principles (Section 5 conservation of the FAIR Data Principles). If there are no repositories complete the formation of the form
orn	n t	o T F Ii Ii d	The SNSF requires that repositories are conform to the FAIR Data Principles (Section 5 content of the FAIR Data Principles). If there are no repositories complete the fair of
orn 1 b	nt	o T F Ii d Ii a T	The SNSF requires that repositories are conform to the FAIR Data Principles (Section 5 content of the FAIR Data Principles). If there are no repositories complete the fair of
orn	n t	o T F Ii d Ii a T d	The SNSF requires that repositories are conform to the FAIR Data Principles (Section 5 content of the FAIR Data Principles). If there are no repositories complete the FAIR Data Principles). If there are no repositories complete the second s
i b re y	n t	o T F In d In d In d ur n	The SNSF requires that repositories are conform to the FAIR Data Principles (Section 5 c hearchers, SNSF's explanation of the FAIR Data Principles). If there are no repositories complete ments in your research field, please deposit a copy of your data on a generic platform (see <u>examples</u>). If no data can be shared, this is a statement of principles. The SNSF supports the use of non-commercial repositories for data sharing. Costs related to data upload are only covered for non-commercial repositories.

FAIR data principles **OpenAIRE**

- Findable Easy to find by both humans and computer systems and based on 1. mandatory description of the metadata that allow the discovery of interesting datasets;
- 2. Accessible Stored for long term such that they can be easily accessed and/or downloaded with well-defined license and access conditions (Open Access when *possible*), whether at the level of metadata, or at the level of the actual data content;
- 3. Interoperable Ready to be combined with other datasets by humans as well as computer systems;
- 4. Re-usable Ready to be used for future research and to be processed further using computational methods.

http://www.dtls.nl/fair-data/

www.forcell.org/group/fairgroup/fairprinciples http://www.nature.com/articles/sdata201618



SCIENTIFIC DATA

SUBJECT CATEGORIES

OPEN Comment: The FAIR Guiding Research data Principles for scientific data management and stewardship

Mark D. Wilkinson et al."

ccepted: 12 February 2016

There is an urgent need to improve the infrastructure sup representing academia, industry, fu come together to design and jointly endorse a concise and mea to as the FAIR Data Principles. The intent is that these may act as a guideline for those wishing to enhance the reusability of their data holdings. Distinct from peer initiatives that focus on the humi polar, the FAIR Principles out specific emphasis on enhancing the ability of machines to automatical find and use the data, in addition to supporting its reuse by individuals. This Comment is the first formal publication of the FAIR Principles, and includes the rationale behind them, and some exempl tations in the communit

FAIR FOR REUSE







How FAIR are your data?

It should be possible for others to discover your data. Rich metadata should be available online in a searchable resource, and the data should be assigned a persistent identifier.

- A persistent identifier is assigned to your data
- There are rich metadata, describing your data
- □ The metadata are online in a searchable resource e.g. a catalogue or data repository
- The metadata record specifies the persistent identifier

It should be possible for humans and machines to gain access to your data, under specific conditions or restrictions where appropriate. FAIR does not mean that data need to be open! There should be metadata, even if the data aren't accessible.

Following the persistent ID will take you to the data or associated metadata
 The protocol by which data can be retrieved follows recognised standards e.g. http
 The access procedure includes authentication and authorisation steps, if necessary
 Metadata are accessible, wherever possible, even if the data aren't

Data and metadata should conform to recognised formats and standards to allow them to be combined and exchanged.

"Lots of documentation is needed"

eferably open formats

ntologies are used where possible her related data

Lots of documentation is needed to support data interpretation and reuse. The data should conform to community norms and be clearly licensed so others know what kinds of reuse are permitted.

- The data are accurate and well described with many relevant attributes
- The data have a clear and accessible data usage license
- It is clear how, why and by whom the data have been created and processed
- The data and metadata meet relevant domain standards



'How FAIR are your data?' checklist, CC-BY by Sarah Jones & Marjan Grootveld, <u>EUDAT</u>. <u>https://doi.org/10.5281/zenodo.1065991</u> Image CC-BY-SA by <u>SangvaPundir</u>



- Needed to locate the research data and get a first idea of the content.
- SNSF:
 - intrinsic metadata (e.g. author's name, content of dataset, associated publication, etc.)
 - submitter-defined metadata (e.g. definition of variable names, etc.)
- Use relevant standards to enable interoperability.
- Check which standards the long-term repository supports or expects.



https://fairsharing.org/



http://rd-alliance.github.io/metadata-directory

https://rdamsc.dcc.ac.uk/ Extra: metadata tools: https://rdamsc.dcc.ac.uk/tool-index



http://www.snf.ch/en/theSNSF/research-policies/open research data/Pages/data-management-plan-dmp-guidelines-forresearchers.aspx



Index of metadata tools

- AgriMetamaker
- ANZ-MEST (Metadata Entry and Search Tool)
- AVM Adobe Metadata Panels
- AVM Web Tool
- Bio-Formats
- CF Compliance Checker
- CIF2Cell
- CIM Comparator Tool
- CIM Questionnaire Generator
- CIM Viewer Tool
- CKAN
- CMOR (Climate Model Output Rewriter)
- Converis
- Darwin Core Archive Assistant
- Darwin Core Archive Validator
- Data Package libraries
- Data Package Validator
- Data Package Viewer
- Data Packagist
- DataCite Metadata Store API

OpenAIRE Documentation?

- Code book explaining the variables
- Study design
- Lab journal
- iPython or Jupyter notebook
- Statistical queries
- Software or instruments to understand or to reproduce the data
- Machine configurations
- Informed consent information
- Data usage licence
- . . .

In short: document and preserve everything that is needed to replicate the study – ideally following the standard in your discipline





DATA REPOSITORIES

OpenAIRE Where to find a repository?

2

Use an external data archive or repository already established for your research domain to preserve the data according to recognised standards in your discipline. More information for selecting a data repository.

If available, use an institutional research data repository, or your research group's established data management facilities.

Use a cost-free data repository such as Zenodo. 3



Search for other research data repositories in http://re3data.org/

re3data.org

- More information: https://www.openaire.eu/opendatapilot-repository
- Zenodo: http://www.zenodo.org
- Re3data.org: http://www.re3data.org





OpenAIRE How to select a repository?



For giving (i.e. archiving & sharing) and taking (i.e. reusing) data:

- Certification as a 'Trustworthy Digital Repository' with an explicit ambition to keep the data available for the long term
- Matches your particular data needs: e.g. file formats accepted; mixture of open and restricted access; licences
- The costs for e.g. depositing the data, data documentation, and support
- Gives your submitted dataset a persistent and globally unique identifier for sustainable citations and to link back to particular researchers and grants
- **Provides guidance on how to cite the deposited data**



	Туре	 Preferred format(s)
	Text documents	 PDF/A (.pdf)
OpenAIRE	Plain text	 Unicode text (.txt)
	Markup language	 XML (.xml) HTML (.html) Related files: .css, .xslt, .js, .es
	Spreadsheets	 ODS (.ods) CSV (.csv)
	Databases	 SQL (.sql) SIARD (.siard) DB tables (.csv)
	Statistical data	 SPSS Portable (.por) SPSS (.sav) STATA (.dta) DDI (.xml) data (.csv) + setup (.txt)
	Raster images	 JPEG (.jpg, .jpeg) TIFF (.tif, .tiff) PNG (.png) JPEG 2000 (.jp2)
	Vector images	 SVG (.svg)

Non-preferred ferred

- ODT (.odt)
- MS Word (.doc, .d
- RTF (.rtf)
- PDF (.pdf)
- Non-Unicode text
- SGML (.sgml)
- MS Excel (.xls, .xls
- PDF/A (.pdf)
- OOXML (.docx, .d
- MS Access (.mdb, later)
- dBase (.dbf)
- HDF5 (.hdf5, .he5
- SAS (.7dat; .sd2; .
- R (* under exami
- · DICOM (.dcm) (by agreement)
- Illustrator (.ai)
- EPS (.eps)



https://dans.knaw.nl/en/deposit/information-about-depositing-data/before-depositing/file-formats?set_language=en

ormat(s)	
docx)	
t (.txt)	
sx)	
locm) , .accdb) (v. 2000 or	
5, .h5)	
.tpt) ination)	
y mutual	

Licensing research data and software **OpenAIRE**

EUDAT licensing wizard helps you pick licences for data & softwar

Choose a License	×
Answer the questions or use the search to find the license you want	
O Start again ← →	
What do you want to deposit?	
Software Data	
Search for a license	
Public Domain Mark (PD)	
The work identified as being free of known restrictions under copyright law, including all related and neighboring rights.	
Publicly Available	
Public Domain Dedication (CC Zero)	
CC Zero enables scientists, educators, artists and other creators and owners of copyright- or database-protected content to waive those interests in their works and thereby place them as completely as possible in the public domain, so that others may freely build upon, enhance and reuse the works for any purposes without restriction under copyright or database law.	
Creative Commons Attribution (CC-BY)	
This is the standard creative commons license that gives others maximum freedom to do what they want with your work.	

Creative Commons Attribution-ShareAlike (CC-BY-SA)

This creative commons license is very similar to the regular Attribution license, but requires you to release al

- Type of data
- Original licenses
- rights

(EUDAT services are provided through EOSC-hub.)



http://ufal.github.io/public-license-selector/ and part of https://b2share.eudat.eu/



Choose a public license by answering some questions regarding access to your dataset. **Suggestions depend on several factors:**

Data consumer access and distribution

Short facts about Zenodo OpenAIRE

- Catch-all repository for EU-funded research
- Up to 50 GB per upload
- Data stored in the CERN Data Center
- Persistent identifiers (DOIs) for every upload, with DOI versioning
- Includes article-level metrics
- Free for the long tail of science
- Open to all research outputs from all disciplines
- GitHub integration
- Easily add EC funding information and report via OpenAIRE



Zenodo: https://zenodo.org/



OpenAIRE Standards of trust





Part of the requirements

R2. The repository maintains all applicable licenses covering data access and use and monitors compliance.

R3. The repository has a continuity plan to ensure ongoing access to and preservation of its holdings.

R4. The repository ensures, to the extent possible, that data are created, curated, accessed, and used in compliance with disciplinary and ethical norms.

R7. The repository guarantees the integrity and authenticity of the data.

R8. The repository accepts data and metadata based on defined criteria to ensure relevance and understandability for data users.

R10. The repository assumes responsibility for long-term preservation and manages this function in a planned and documented way.

R11. The repository has appropriate expertise to address technical data and metadata quality and ensures that sufficient information is available for end users to make quality-related evaluations.

R13. The repository enables users to discover the data and refer to them in a persistent way through proper citation.

R14. The repository enables reuse of the data over time, ensuring that appropriate metadata are available to support the understanding and use of the data.



OpenAIRE

https://www.coretrustseal.org/



DATA IN REPOSITORIES



GO-FAIR Initiative: FAIR Metrics Group **OpenAIRE**

- Aim: to define metrics enabling both qualitative and quantitative assessment of the degree to which online resources comply with the 15 Principles of FAIR Data
- Founding Members:
 - Mark Wilkinson, Universidad Politécnica de Madrid
 - Susanna Sansone, University of Oxford
 - Michel Dumontier, Maastricht University
 - Peter Doorn, DANS
 - Luiz Olavo Bonino, VU/DTL
 - **Erik Schultes, DTL**



https://www.dtls.nl/fair-data/fair-metrics-group/ and http://fairmetrics.org/



FAIR "light" assessment

Findable (defined by metadata (PID included) and documentation)

- I. No PID nor metadata/documentation
- 2. PID without or with insufficient metadata
- 3. Sufficient/limited metadata without PID
- 4. PID with sufficient metadata
- 5. Extensive metadata and rich additional documentation available

Accessible (defined by presence of user license)

- I. Metadata nor data are accessible
- 2. Metadata are accessible but data is not accessible (no clear terms of reuse in license)
- 3. User restrictions apply (i.e. privacy, commercial interests, embargo period)
- 4. Public access (after registration)
- 5. Open access unrestricted

Interoperable (defined by data format)

- I. Proprietary (privately owned), non-open format data
- 2. Proprietary format, accepted by Certified Trustworthy Data Repository
- 3. Non-proprietary, open format = 'preferred format'
- 4. As well as in the preferred format, data is standardised using a standard vocabulary format (for the research field to which the data pertain)
- 5. Data additionally linked to other data to provide context





FAIR badge scheme





2 User Reviews **1** Archivist Assessment 24 Downloads

- FAIR as proxy for data "quality" or "fitness for (re-)use"
- We want to create a badge system using the FAIR principles to assess data sets in a Trustworthy Digital Repository
- **Developing the data assessment tool: FAIRdat**
- Score each FAIR dimension on a 5-point scale
- **Operationalise the original principles to ensure no** interactions among dimensions to ease scoring
- **Consider Reusability as the resultant of the other three:**
 - the average FAIRness as an indicator of data quality
 - (F+A+I)/3=R Assessment tool based on questionnaire to evaluate any dataset in any (trustworthy) repository by depositors, data specialists and users
- Manual and automatic scoring
- **Prototype is being tested**

DANS FAIRdat prototype: https://www.surveymonkey.com/r/fairdat





Display FAIR badges in any repository (Zenodo, Dataverse, Mendeley Data Figshare R2SAFE

	Data, 1 Igshare, D20711	└,)	
ZEOCO Search	Q Upload Communities Dog in Sign up		
Recent uploads December 31, 2010 Figure Open Access FIGURE 5 in Molecular and bioacoustic differentiat occidentalis with description of a new treefrog from Madagascar	View View Welcome to the improved Zenodo. See what's new and known issues.	B Dataverse	Q About Guides - Support Sign Up Log In verse A controls libe and Harvard Library, Harvard University IT, and IQSS
ences, Miguel; Andreone, Franco; Glos, Julian; Glaw, Frank IGURE 5. Photographs of Boophis occidentalis from Isalo Nation pecimen ZSM 2314 / 2007 in dorsolateral and ventral view, and (osition in a small cavity in a rock above a stream, photographed o ZFMK ploaded on December 8, 2016.	Get exposure and credit for your data: write a data paper for the new peer reviewed, online-only o For more info: brill.com/rdj	Open access Research Data Journal (published by Brill)	ntre - Population Services International (PSI) Dataverse
December 6, 2016 Dataset Open Access Revisiting the phylogeny of phylum Ctenophora: a poerspective	EASY offers sustainable archiving of research data and access to thousands of da Search > Advanced search > Browse	atasets.	Q Find Advanced Search
rteaga-Figueroa, Luis A.; Sánchez-Bermúdez, Valentina; Franco-S aw data used in 'Revisiting the phylogeny of phylum Ctenophora ploaded on December 8, 2016.	32,530 RESULTS IN PUBLISHED DATASETS Image: Interpretation Image: Interpretation Archeologisch booronderzoek verdubbeling N381 Donkerbroek Oosterwolde, gemeente Ooststellingwerf (FR) Date: 2019-06-09 Creators: Krol-Karsten, T.N. (MUG Ingenieursbureau) Audience: Archaeology Submitted: 2016	Sort by: Choose One REFINE Search SEARCH	1 to 10 of 65,724 Results It Sort- Internet Banking Espousal in Bangladesh: A Probing Study Image: Constraint of the search Archive Dataverse Image: Constraint of the search Archive Dataverse Image: Constraint of the search Archive Dataverse Image: Constraint of the search Archive Dataverse Image: Constraint of the search Archive Dataverse, V1 Internet banking (IB) is a distinctive banking improvement with the intention of potentially can convert the monetary services scenery in budding antions such as Bangladesh. Nevertheless, due to the connected near to the ground acceptance rate, its full potential in deepening and Archival Data for Consider the Redirect: A Missing Dimension of Wikipedia Research
	Thematic Collection: Children of Immigrants Longitudinal Survey in the Netherlands (CILSNL) Date: 2017-12-31 Creators: Jaspers, dr. E. (Universiteit Utrecht); Tubergen, prof. dr. F. van (Universiteit Utrecht) Audience: Social sciences Social sciences Social sciences Output Social sciences Van (Universiteit Utrecht) Behang Van (Universiteit Utrecht) Migration, ethnic relations and multiculturalism Access: Restricted (request permission)	Advanced search Audience Behavioural and educational sciences Conomics and Business Administration Humanities (31830) Interdisciplinary sciences (148)	Pre Dec 10, 2016 - Community Data Science Collective Dataverse Hill, Benjamin Mako; Aaron Shaw, 2016, "Archival Data for Consider the Redirect: A Missing Dimension of Wikipedia Research", doi:10.7910/DVN/NQSHQD, Harvard Dataverse, V1 This contains data and software for the following paper: Hill, Benjamin Mako and Aaron Shaw. "Consider the Redirect: A Missing Dimension of Wikipedia Research." In Proceedings of the 10th International Symposium on Open Collaboration (OpenSym 2014). ACM Press, 2014. This is an



DANS FAIRdat prototype: https://www.surveymonkey.com/r/fairdat



https://www.slideshare.net/sjDCC/fair-data-interim-report-and-action-plan

Turning FAIR Data into Reality Interim Report and Action Plan

EOSC Summit 2018 **European Commission Expert Group on FAIR Data**

Simon Hodson, Chair CODATA simon@codata.org @simonhodson99

Sarah Jones, Rapporteur **Digital Curation Centre** sarah.jones@glasgow.ac.uk @sjDCC

Repositories need to be encouraged and supported to achieve CoreTrustSeal certification. The development of rival repository accreditation schemes, based solely on the FAIR principles, should be discouraged.

Rec. 29: Implement FAIR metrics

(...) Repositories should publish assessments of the FAIRness of datasets, where practical, based on community review and the judgement of data stewards.



FAIR data HLEG Recommendations and Action Plan – Published June 11, 2018

Rec. 10: Trusted Digital Repositories



Let's learn to trust

- End of the year you have all downloaded and explored at least one dataset - from a repository outside your organisation
- Credit researchers and others who seek value in and add value to existing data
- Support FAIR and Open data in trustworthy repositories, for instance by
 - sticking to standards for data documentation and file formats ("replication packages")
 - pushing your favourite repository to get certified: make processes and policies transparant
 - educating early-career researchers
 - learn the trade by replicating a study
 - manage data to make them FAIR and as open as

possible



OpenAIRE/EOSC-hub webinar: https://www.openaire.eu/how-to-manage-your-data-to-make-them-open-and-fair Photo by Yuri Catalano - CC0 - https://www.pexels.com/photo/city-landscape-sky-people-127420/



Thank you!

- www.openaire.eu
- @openaire_eu
- facebook.com/groups/openaire
- linkedin.com/groups/OpenAIRE-3893548 in



marjan.grootveld@dans.knaw.nl



