

On the Path to a Quality Indicator for Software and Data Publications for the Helmholtz Association

A Workshop Report

Marcel Meistring

Helmholtz-Association, Helmholtz Open Science Office

Spring Symposium 2024 of the Fachgruppe Database Systems (FG DB) "Beyond Silos: Next Steps in Research Data Management"

Jena, 11.-12.03.2024

Agenda

- Helmholtz
- Key Performance Indicators in the Program-Oriented Funding (PoF)
- Quality Indicator Work in Progress
 - Software Publications (Overview)
 - Data Publications (Detail)
- Opportunities, Outlook, Challenges

Helmholtz: Mission and Strategy Research for Grand Challenges

- Systems solutions for grand challenges based on:
 - Scientific excellence
 - Interdisciplinarity and critical mass
 - long term research programs
- Helmholtz provides a highly attractive environment for talents and brilliant brains
- Profound expertise in large scale research infrastructure
- Helmholtz as a prime strategic partner at the local, national and international level
- Transfer of knowledge into economy and society



Helmholtz research centers 18 Centers in 6 Research Fields

- Alfred-Wegener-Institut Helmholtz-Zentrum f
 ür Polarund Meeresforschung (<u>AWI</u>)
- <u>CISPA</u> Helmholtz Center for Information Security
- Deutsches Elektronen-Synchrotron <u>DESY</u>
- Deutsches Krebsforschungszentrum (<u>DKFZ</u>)
- Deutsches Zentrum f
 ür Neurodegenerative Erkrankungen (<u>DZNE</u>)
- German Aerospace Center (<u>DLR</u>)
- Forschungszentrum Jülich (<u>FZJ</u>)
- <u>GEOMAR</u> Helmholtz-Zentrum für Ozeanforschung Kiel
- GSI Helmholtz Center for Heavy Ion Research
- Helmholtz Munich
- Helmholtz-Zentrum Berlin f
 ür Materialien und Energie (HZB)
- Helmholtz Center Dresden Rossendorf (HZDR)



- Helmholtz Center for Infection Research (HZI)
- Helmholtz Center for Environmental Research UFZ
- Helmholtz-Zentrum Hereon
- GEOMAR Helmholtz Center for Ocean Research Kiel
- Helmholtz Center Potsdam –
 German Research Center for Geosciences GFZ
- Karlsruhe Institute of Technology (KIT)
- Max Delbrück Center for Molecular Medicine in the Helmholtz Association (MDC)

Research Fields:

(1) Energy, (2) Earth and Environment
 (3) Health, (4) Information
 (5) Aeronautics, Space and Tansport, (6) Matter

Helmholtz Open Science Office: Core & Focus Topics 2023/2024



Key Performance Indicators in the Program-Oriented Funding (PoF)

Program-Oriented Funding (PoF)

- Direction and structure for research at Helmholtz
- Strategic evaluation und scientific evaluation
 - at the level of the centers and the existing programs
- Central lever

 \rightarrow Among others:

Collection of key performance indicators



Program-Oriented Funding - KPI "Status Quo"

"The scientific evaluation is focused on scientific quality. The results serve to provide a review of the performance of both the Helmholtz Center and the individual, usually cross-center programs. Both aspects are equally important: The centers are the driving force behind the Association, since this is where research is carried out and where new insights are gained..."

With regard to output (in the sense of published research artifacts) so far:

- Number of articles in peer-reviewed journals (indexed in Web of Science and/or Scopus or published by Open Research Europe ORE)
- Share of Open Access

Focus on the "Final Product" of science





Program-Oriented Funding - KPI "Status Quo"

"The scientific evaluation is focused on scientific quality. The results serve to provide a review of the performance of both the Helmholtz Center and the individual, usually cross-center programs. Both aspects are equally important: The centers are the driving force behind the Association, since this is where research is carried out and where new insights are gained..."

Focus on the "Final Product" of science



 give value to all relevant research activities and scientific outputs including high-quality FAIR data and metadata, well-documented and reusable software, protocols and workflows, machine-readable summaries of findings, and teaching, outreach and engagement of societal actors;

Program-Oriented Funding – Quality Indicator

Mandate of the Helmholtz General Assembly (2022):

• Development of a multidimensional quality indicator for data products

Goals:

- Broadening / Improvement of the evaluation of science within Helmholtz
- Improving the quality of published research data
- Promotion of Open Science Practices

Expansion of the mandate to include the aspect of research software



Open Research Data

All Centers will establish detailed procedures for managing research data in publicly available policies,⁶ and will regularly examine and if necessary adapt these procedures.

In 2023, a basic indicator for the presentation of citable research data publications will be established as an incentive within the framework of the PoF.

By 2024, a Helmholtz quality indicator for research data publications will be developed and established, which will be deployed within the framework of the PoF and will replace the aforementioned basic indicator.

Open Research Software

All Centers will aim to establish detailed research software management procedures in publicly available policies by 2025.⁷

A basic indicator for the presentation of citable research software publications will be established in 2023 as an incentive within the framework of the PoF together with research data publications.

By 2024, a **Helmholtz quality indicator for research software publications** will be developed and established, which will be deployed within the framework of the PoF and will replace the aforementioned basic indicator.

Task Group

Helmholtz Quality Indicators for Data and Software Products

- Since March 2022, ongoing
- Inclusive approach: Representatives of all Helmholtz Centers
- Work in 3 groups:
 - 1. Whole group
 - 2. Sub-group research data
 - 3. Sub-group research software
- Relevant products and events
 - Report on Helmholtz Open Science Forum "Indikatoren für Open Science": <u>https://doi.org/10.48440/os.helmholtz.024</u>
 - Report on Helmholtz Open Science Forum "Research Evaluation, Reputation Systems, and Openness" <u>https://doi.org/10.48440/os.helmholtz.065</u>

Basic Indicator, Lessons Learned, Challenges

Reporting year 2022: initially basic indicator (internal test collection)

	Helmholtz Definition											
Zitierbar publizierte Forschungs-	Anzahl von <mark>zitierbar publizierten</mark> Forschungsdaten-Publikationen und Forschungssoftware-Publikationen sind je <mark>separat zu erfassen</mark> . Die zitierbar publizierten Forschungsdaten- und Forschungssoftware-Publikationen											
daten- und Forschungs- software- Publikationen	müssen in einem Repositorium mit Metadaten gespeichert und mit einem persistenten Identifikator (insb. Digital Object Identifier – DOI) versehen sein. Erlaubte Repositorien sind solche, die entweder											
[Anzahl]	 a) unter Beteiligung eines Helmholtz-Zentrums betrieben werden oder b) extern sind und in re3data.org gelistet oder zertifiziert sind. c) Für die Publikation von Forschungssoftware können etablierte Software- Repositorien genutzt werden. 											
	Bei Softwarepublikationen wird jeder Release als Publikation gezählt.											

Challenges for all Centers

- \rightarrow Infrastructures for publication of RD/RSW
- \rightarrow Processes for collecting information on RD/RSW-publications
- \rightarrow Instruments for collection of (records on) RD/RSW-publications

Quality Indicators - Work in Progress Challenges

- Text publications
- ightarrow established mechanisms of quality assurance
- \rightarrow established formats (articles, monographs, proceedings,...)
 - = more or less "comparable units"
- Research Data/ Research Software
- \rightarrow Missing "Quality" framework
- \rightarrow Comparability
- ightarrow Granularity, versions, data types
- ightarrow Heterogeneity of centers, data and software (research areas)
- \rightarrow Quantification



Multidimensional Indicator

ightarrow Mapping several quality dimensions to one value

Questions

- → Which dimensions of quality do we define for data and software publications?
- → Which attributes define the quality dimensions?
- \rightarrow How can the attributes/dimensions be quantified?
- → How to get a single value via the attributes/dimensions?

Approach: Multidimensional Indicators

Iterative and inclusive process with all involved people

- 1. Definition of suitable dimensions for assessing the quality of RD- & RSW-publications
- 2. Collection of specific attributes for each dimension
- 3. Application of a generic maturity model to the attributes and assignment of numerical values for maturity levels in attribute
- 4. Determination of maturity level per dimension based on weighted average values of the attributes of each dimension
- 5. Summarized quality assessment (visualization using radar plot)

Quality Indicators - Work in Progress Approach: Maturity model

3. Application of a generic maturity model to the attributes and assignment of numerical values for maturity levels in attribute

Concept after COBIT Maturity Model

Characteristics of the Maturity levels



Sally Godfrey, Public domain, via Wikimedia Commons (beispielhaft)

Research Data

- (0) Non-existent: no information available or not applied
- (1) Most necessary information provided or measure taken
- (2) Basic information provided or measure taken (sensible level of information/measures)
- (3) Advanced information provided or measure taken, allowing to generally understand and (re)use the published data
- (4) Complete and accurate information provided or measure taken, to an extend that allows maximal understanding and usage of data

Research Software

- (0) Non-existent: no information available
- (1) Initial: initial information available being obtained in an ad-hoc, unorganized manner
- (2) Repeatable: the information is complete, being produced in a repeatable, yet intuitive manner
- (3) Defined: a process is established guaranteeing the complete compilation of the required Information

(4) Managed: the process being established is managed, i.e. monitoring/measuring is included(5) Optimized: practices are put in place optimizing the way the process is operated, leading to improved quality over time

Approach: In general applying FAIR-Principles

- 1. Definition of suitable dimensions for assessing the quality of RD- & RSW-publications
- 2. Collection of specific attributes for each dimension



Wilkinson, M. et al. (2016). https://doi.org/10.1038/sdata.2016.18 FAIR Data Maturity Model Specification and Guidelines 2020

RESEARCH DATA ALLIANC



RDA FAIR Data Maturity Mod el Working Group (2020). https://doi.org/10.15497/rda00050 FAIR Principles for Research Software (FAIR4RS Principles)

Authors

Neil P. Chue Hong*, Daniel S. Katz*, Michelle Barker*; Anna-Lena Lamprecht Carlos Martinez, Edis E. Psomonoulos, Jen Harrow, Levia Jael Castro Morane Gruenpeter, Paula Andrea Martinez, Tom Honeyman; Alexander Struck, Allen Lee, Axel Loewe, Ben van Werkhoven, Catherine Jones, Daniel Gariio Esther Plomp, Francoise Genova, Hugh Shanahan, Joanna Leng, Maggie Hellström, Malin Sandström, Manodeep Sinha, Mateusz Kuzak, Patricia Herterich, Qian Zhang, Sharif Islam Susanna-Assunta Sansone, Tom Pollard, Udavanto Dwi Atmoio Alan Williams, Andreas Czerniak, Anna Niehues, Anne Claire Fouilloux, Bala Desinghu, Carole Goble, Céline Richard, Charles Gray, Chris Erdmann, Daniel Nüst, Daniele Tartarini, Elena Ranguelova Hartwig Anzt Ilian Todorov James McNally Javier Moldon Jessica Burnett Julián Garrido-Sánchez, Khalid Belhaijame, Laurents Sesink, Lorraine Hwang, Marcos Roberto, Tovani-Palone, Mark D. Wilkinson, Mathieu Servillat, Matthias Liffers, Merc Fox, Nadica Miliković, Nick Lynch, Paula Martinez Lavanchy, Sandra Gesing, Sarah Stevens, Sergio Martinez Cuesta, Silvio Peroni, Stian Soiland-Reves, Tom Bakker, Tovo Rabemanantsoa, Vanessa Sochat, Yo Yehudi; and the FAIR4RS WG

Chue Hong, N. P. et al. (2021). FAIR Principles for Research Software (FAIR4RS Principles). Research Data Alliance. https://doi.org/10.15497/RDA00065

Approach: Software - Dimensions (Overview)

FAIR-ST Framework

Findable - Accessible - Interoperable - Reusable

Scientific basis

Research software is an integral part of the research process. It therefore has to also follow community/organization specified common standards in performing research. While the contribution of software to a certain scientific achievement must be evaluated by dedicated experts in the field, some aspects of scientific quality can also be considered in wider generality.

Technical basis

Quality of software also reflects general aspects of software engineering. At the end, software has to guarantee to actually produce what it has been aimed to do. Professional software development aims at producing software following state-of-theart software engineering concepts. Thus, aspects of professional software engineering must also be considered as part of a quality assessment of research software development.

FAIR Principles for Research Software (FAIR4RS Principles)

Authors

Neil P. Chue Hong*, Daniel S. Katz*, Michelle Barker*; Anna-Lena Lamprecht, Carlos Martinez, Fotis E. Psomopoulos, Jen Harrow, Leyla Jael Castro, Morane Gruenpeter, Paula Andrea Martinez, Tom Honevman; Alexander Struck, Allen Lee, Axel Loewe, Ben van Werkhoven, Catherine Jones, Daniel Garijo, Esther Plomp, Francoise Genova, Hugh Shanahan, Joanna Leng, Maggie Hellström, Malin Sandström, Manodeep Sinha, Mateusz Kuzak, Patricia Herterich, Qian Zhang, Sharif Islam, Susanna-Assunta Sansone. Tom Pollard. Udavanto Dwi Atmoio: Alan Williams, Andreas Czerniak, Anna Niehues, Anne Claire Fouilloux, Bala Desinghu, Carole Goble, Céline Richard, Charles Gray, Chris Erdmann, Daniel Nüst, Daniele Tartarini, Elena Ranguelova, Hartwig Anzt, Ilian Todorov, James McNally, Javier Moldon, Jessica Burnett, Julián Garrido-Sánchez, Khalid Belhajjame, Laurents Sesink, Lorraine Hwang, Marcos Roberto Tovani-Palone, Mark D. Wilkinson, Mathieu Servillat, Matthias Liffers, Merc Fox, Nadica Miljković, Nick Lynch, Paula Martinez Lavanchy, Sandra Gesing, Sarah Stevens, Sergio Martinez Cuesta, Silvio Peroni, Stian Soiland-Reyes, Tom Bakker, Tovo Rabemanantsoa, Vanessa Sochat, Yo Yehudi; and the FAIR4RS WG

Approach: Software - Dimensions (Overview)

FAIR-ST Framework

Findable - Accessible - Interoperable - Reusable

Attributes Findable:

- Open Publication Repository
- Versioning
- Persistent Identifier (PID)
- Rich Metadata

Approach: Software - Dimensions (Overview)

FAIR-ST Framework

Findable - Accessible - Interoperable - Reusable

Attributes Findable:

- Open Publication Repository
- Versioning
- Persistent Identifier (PID)
- Rich Metadata

Attributes Accessible:

- Access Conditions (organizational)
- Access Conditions (process)
- Technical Accessibility (run/start)

Approach: Software - Dimensions (Overview)

FAIR-ST Framework

Findable - Accessible - Interoperable - Reusable

Attributes Findable:

- Open Publication Repository
- Versioning
- Persistent Identifier (PID)
- Rich Metadata

Attributes Interoperable:

- Input/Output Formats
- Adaptability/Flexibility of Use

Attributes Accessible:

- Access Conditions (organizational)
- Access Conditions (process)
- Technical Accessibility (run/start)

Approach: Software - Dimensions (Overview)

FAIR-ST Framework

Findable - Accessible - Interoperable - Reusable

Attributes Findable:

- Open Publication Repository
- Versioning
- Persistent Identifier (PID)
- Rich Metadata

Attributes Interoperable:

- Input/Output Formats
- Adaptability/Flexibility of Use

Attributes Accessible:

- Access Conditions (organizational)
- Access Conditions (process)
- Technical Accessibility (run/start)

Attributes Reusable:

Reusability Conditions

Approach: Software - Dimensions (Overview)

FAIR-ST Framework

Scientific basis - Technical basis

Attributes Scientific basis:

- Community Standards
- Team Expertise
- Scientific Embedding

Approach: Software - Dimensions (Overview)

FAIR-ST Framework

Scientific basis - Technical basis

Attributes Scientific basis:

- Community Standards
- Team Expertise
- Scientific Embedding

Attributes Technical basis:

- Project Management
- Repository Structure
- Code Structure
- Reproducibility (Code)
- Code Change Process
- Security

POCME-Framework

Findable, Accessible, Interoperable, Reusable?

 \rightarrow Regrouped and extended to

P-ublication

O-penness

C-uration

M-etadata

E-xternal View





FAIR Data Maturity Model

https://os.helmholtz.de 27

POCME-Framework

Publication - Openness - Curation - Metadata - External View

• Published with Identifier

(0) No identifier (resource may only be found via personal communication)

(1) Basic Uniform Resource Identifier

(2) Dataset is identifiable via internal handle (does not resolve globally, generally no metadata)(3) Dataset is basically identifiable via formalized, standardized, persistent identifier (resolves globally, general metadata provided)

(4) Dataset is identifiable via globally unique, formalized, standardized, persistent identifier supported by general metadata (e.g. DOI).

• Published with Information on Access to the Data

(0) No metadata available

(1) Metadata available, but no data access-information available in the metadata

(2) Metadata available, data access-information available only in human-readable form

(3) Metadata available, data access-information available only in human readable form, including general license information

(4) Metadata available, data access-information available in human-readable and machinereadable form, including license information

• Published via a Repository or Collection,

that is indexed in a Meta-Repository

(0) No information available, the data is not published via a repository/collection

(1) The data is published in a repository/ collection which is not listed in an eligible metarepository

(2) The repository/collection is listed in an eligible meta-repository, basic no. of quality indicators assigned by the meta-repository are achieved

(3) The repository/collection is listed in an eligible meta-repository, medium no. of quality indicators assigned by the meta-repository are achieved

(4) The repository/collection is listed in an eligible meta-repository, high no. of quality indicators

assigned by the meta-repository are achieved

POCME-Framework

Publication - Openness - Curation - Metadata - External View

General Degree of Openness

(0) No information on open accessibility/availability of the data at all

(1) Information available: no open accessibility/availability of the data. No justification, no information on possible contact or restrictions

(2) Like (1) + information on possible contact, restrictions or potential use cases on request available

(3) Like (2) + with justification AND/OR date of moratorium

(4) Open accessibility with corresponding license (no login or contact needed or otherwise with justification)

• Primary Data Formats

(0) No primary data available in digital form

(1) Primary data generally available

(2) Primary data stored in common proprietary data formats

(3) Primary data stored in open formats

(4) Primary data makes use of common, domain specific terminologies (e.g., codelists)

POCME-Framework

Publication - Openness - Curation - Metadata - External View

Level of Curation

(0) Data is published in raw form without any curation or documentation (e.g. raw long-tail data)

(1) Data is published in raw form without curation but according to standard with basic documentation like readme (e.g. automatic generated sensor data, long-tail data following a basic scheme)

(2) Data is published in cleaned form with some curation (e.g. brief checking, documentation according to standard)

(3) Data is published in cleaned form with enhanced curation and/ or reprocessing (e.g. conversion to new formats, enhancement of documentation)
 (4) Data is published after undergoing extensive curation and/or reprocessing according to discipline specific standards in order to enhance to max.
 quality (like (3) + additional editing of deposited data for accuracy)

Community Standards for Data

(0) Data is prepared/ recorded in a non-standardized form, no documentation

(1) Data is prepared/ recorded in a non-standardized form, but there is documentation of the used format (e.g. readme, codebook)

(2) Data is prepared/recorded following an institutional Standard Operation Procedures (SOP) with traceable reference to the used SOP-scheme

(3) Data is prepared/recorded following a Standard Operation Procedures (SOP) or Community standard reflecting a general level of domain specific standards; with traceable reference to the used standard or scheme

(4) Data is prepared/recorded following an acknowledged and openly accessible Standard Operation Procedure (SOP) or Community standard reflecting a high level of domain specific standards; with traceable reference to the used standard or scheme

POCME-Framework

Publication - Openness - Curation - Metadata - External View

Metadata to find/retrieve a Resource / Formal Metadata

(0) No metadata available

(1) Metadata available for/with the data publication that is not structured according to a commonly accepted scheme (i.e. no scheme applied)

(2) Metadata provided with the data publication that is structured in a basic way according to a commonly accepted scheme (e.g. completed DataCite mandatory-properties/discovery ; Dublin Core, etc.)

(3) Metadata provided with the data publication that is structured in an advanced way, according to a commonly accepted scheme (e.g., completed Datacite mandatory- and recommended-properties for discovery + discovery-supporting basic content metadata according to DataCite scheme)

(4) Full Metadata provided with the data publication (complete DataCite mandatory- and recommended- and optional-properties for discovery + comprehensive discovery-supporting content metadata according to DataCite scheme)

Content related Metadata

(0) No content related metadata available

(1) Some content related metadata available, following a (generic) scheme (e.g. DataCite)

(2) Complete content related metadata available following a (generic) scheme (e.g. DataCite)

(3) Some content related metadata available following standardized form or domain specific scheme

(4) Complete and curated content related metadata available following a standardized form and domain specific scheme (see 3)

POCME-Framework

Publication - Openness - Curation - Metadata - External View

- Score from Domain Specific Fair Assessment Tool
- (0) 0-20% Score reached
 (1) 21-40% Score reached
 (2) 41-60% Score reached
 (3) 61-80% Score reached
 (4) 81-100% Score reached

Quality Indicators - Work in Progress Approach: Maturity model

Publication - Openness - Curation - Metadata - External View



Approach: Visualization and Aggregation



Aggregation (Considerations):

- Number of RD/RSW publications that fulfill a minimum value in each dimension?
- Number of RD/RSW with a minimum maturity (weighted across all dimensions)?

• ...?

Quality Indicators - Work in Progress Opportunities

- more than just a simple number reflects the quality of RD/RSW processes
- Different requirements and starting positions at institutional level can be mapped (radar plot)
- Radar plot delivers indications of potential for improvement at the level of the individual publication
- Incentive to improve internal processes

Research Fields: (1) Energy, (2) Earth and Environment (3) Health, (4) Information (5) Aeronautics, Space and Tansport, (6) Matter



Quality Indicators - Work in Progress Outlook and next steps

 Concept has been published as white paper: <u>https://doi.org/10.48550/arXiv.2401.08804</u>

	Cornell University
arXi	V > cs > arXiv 2401.08804
Compute	er Science > Digital Libraries
(Submitted Towar	on 18 Jan 2024 (r/), ikat revised 28 Jan 2024 (this version, v2) rds a Quality Indicator for Research Data publications and Research Software publications
Wolfgang Bertelma	r zu Castell, Doris Dransch, Guido Juckeland, Marcel Meistring, Bernadette Fritzsch, Rönny Gey, Britta Höpfner, Martin Köhler, Christian Meeß nn
Resea publica for res quality current	ch data and software are widely accepted as an outcome of scientific work However in comparison to tend-based publications, there is not yet an established polytom. The paper presents an attempt to thit the gap instance for the Working Group Core Discrete of the Heinhold Association her Task Group Heinhold Qualitation data and research damape polytomics to be und within the Association. The regot Lammatize the Heinhold Qualitation of the Task Group Heinhold Qualitation and end research damape polytomics to such a the association and end research damape polytomics the source of the polytomic and the core of the foreign and the foreign and the core of the foreign and the core of the foreign and the core of the foreign and the foreign and the core of the foreign and the forei
Comments Subjects Cite as:	21 pages. Tigure Depart Environis (s. K.), Computers and Society (s.C. CY) and/2-011 00014-(s. K.), C. K.), for the version ingression up 914 0001402 (s. K.), for the version ingression up 914 0001402, 2011 0001 (
Submiss From: Guid Ir1) Tue, 18	lon history & Jockind (ine ensil) J. Ja 2024 26 00 27 UTO (134 KB)

- Paper and Pencil-exercise with actual software and data publications
 - by groups from different research areas

has	V 2 - Published with Hamilton V 2 - R E - No Mercifier (Resource may only in found at 2 - No	Published via a Neparalamy or Call V 2. Published with Information on Books to the Orienta Information authors, the data's 2. No Matabase website	an Fallahing W-2 - Minatata In Endown on a Ressaurus J Ramat Min. M-2 - Canton C - No Minatata analatis	ni ndana Minakan ni ndané manjaté	6-1 - Leviel of Curation 5-Occile published in ran family athlaud any curation 5- No information or open a													
	Consult is therefore as international place 2. The Consult's barbally derived as international, 3. The	e name and a second of a second se	1 Scenet of Presidency provider with the free Participation that \$ 101 Comparison in Scenet on Presidency provider with the free Participation that \$ 101 Section pre- table president with the free Participation that \$ 101 Section pre- table president with the free Participation that \$ 101 Section pre- table president with the free Participation that \$ 101 Section pre- table president with the free Participation that \$ 101 Section pre- table pre- table pre- table pre- sident pre- table pre- table pre- sident pre- sident pre- table	energia da la canta estado de la cança e prese (la cinera de gli bancas) en estado cantas estado Minera energia de la cinera e de las sente activas			Dimensio	ion Reusable Community (Rendards	Team Deportise	Scientific Embedding	Dimension Scien Project	Management	Repository Structure	Code Structure			
Andref, K., Kargen, C. et al. (2022) Albame, Saliyano Publicaturan																		
and, Malyar, Ribbergan, Carolen, Scheeder, G. Acetholica, de			4		4 A													
ton, N.B. Hawk, G.Froce, Moore, Theodore C.(H. Hottiddi.do		4 4		4 4	4 4								Sec. 199					
utilab, Tylke, Neuhaue, Thefan, Centhella, Allan, 1 Denote		4 4		4 4	4							1+10	wormation on project					
mans, Rubh, Mathiot, Pierre, Jourdain, Nocolas, Denoto	4	4 4		4 4				1 - No inform	store given	1 - No information given.	1 - No information given.	210M	d	1 - No information given.	1 - No information given.			
Carto System re, Sargery, Flankov, Thumas, Solomerko, Denitry, Grait Assessio		4 A	e e	4. A. A.	4. A								Was full hier runter series under Punkt 3	was nicht tiet?				
Earlis Igalee Ini, Elica, Marei, Daniela, Pradkan, Himanos Kus, Orio Rabeato MERC	() () () () () () () () () ()	4 4	4	4														
COTE AND POLICIPAN ALBIA Phys. 2010-02 - Gela Service MEMI			· · ·					1.000		2 - Is it clear that at least expertise from a				all files provided in some				
arve Marin, Velote, Nalut, Hoppe, Clara JM, Po. Units Service	A I	C.	D Headandar	1 August 1 August 1		Q.	H after	105,050 aneround, or similar own.	ection to known (scientific)	single, relevant domain is being part	2 -Is there at least one scientific use case	2-165	ere some kind	ed/unstructured way inside th	he 2 - Is every developer free to			
National .	Clation Software(Keleace) Report	Gory Open Publication Reportary	Versioning	Personal Meridian	Roh Metalata	Darwassion Funda Access Cand	ices	nie including that license	en?	of the software development team?	being documented?	being	sed?	(n)	own style of coding?			
attophet		1 - There is no information available on where in first the set	ware 1 - No software versioning and of	1 - No. PEte store	1 - No metadata dam	L - Net see 1	depe	endencies are at least being checked										
(NEM)							740	nually?						VICTOR				
Curtiers for					2 - Is there some metadata information provided	2 - Is there as	tact given which to inq			3 - Over the orthogon development term	A - to there a basedor origentic context			- Innerhalh won Reno	sitorian scheint e	s winklich sehr wenig Va	riation in der Bewertung	
Burranar atomatic		2 - Is there an anline repository which contains the software	2 - Is there some kind of sersioning for the software	2 - Is there a handle/URL given to identify the softw	are? with the software?	about the right	a use the solouare?		and following standards of	has access to espective to several relevant	hairs documented inclution researd	2.00	ers a service (Service services has	- Innernato von nepo	situl ien scheint e	s wirklich sein neuig ve	racion in der bener cong	
M.A.C., Hypena, Main, Jores, Elisbeth M(2) (NCR)									sertific community?	domains?	examples?	uned?		die bewertung dann t	atsachlich auf die	ser coene vorzunenmen, v	as auch die operationalis	2
Ormetric an		3 - is there some kind of description molithis sides forther	3 - Down the address one structured is a second city	 Is the handled (B) remained with a defined more 	tera 1 - la chamaractera la formación folimaine a silam	As in there all	na dan Amerikan da							sollte trivial sein)	 evtl. könnte man 	n dann sogar sich gemein	sam auf Bewertungen für F	1
Advertised		information on the antheora in this remains to a meeting to	all service of an and an	schara?	manufata schame complete?	100	the providence of the							-Bei einigen Skalen :	scheinen mir Zusam	menhänge zu existieren -	z.B. kann ich die Curati	i
Include .																		
Cartlers to												4-115	ere a version control system in	******				
information										4 - Does the software development team	4 - Is the software development at least	place I	ring part of a code project	0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	e 11 14 1 1 1		c	
ax, Hen, J. V., Fairbook, Osehard(201) Date: chemis		4 - Is there a structured meta data description (e.g. following			4 - Is there a metadata curstion process to reflect	4 - Does the I	tse allow far open use	4 - Does the s	oftware comply with	has access to expertise in all relevant.	loosely connected to some scientific	manag	ement platform Je.g. Gi04ub,	Prinzipieli solite a	uf die Konsistenz (ser skala geachtet werde	n: Es muss 0-5 sein. Dar	3
Annon A		DataCite) given for software in this repository?	4 - Is a description of the versioning scheme available	4 - Is a persistent identifier given?	changes/updates?	sofoware (e.g.	a licenses)?	reevant scien	cinc standards of the held?	donaes?	and active?	OCL4	and an associated ticket system	sorgfältig mit dem z	ugrundeliegenden M	aturity Model gegengepri	ft werden!	
2017		5 - It the providence fished in some compactices and proposition	estant. Such these advancementation on enforcements for the	 S - K - KPA sizes attacked for accounted barrenting 	s - Carl al mecana intervalue rangeire de	See Bulleton and	te des altres come la											
hap.Dick.Schilter.Reiter2091.Madiataria: data Mar		Relivibulia Research Software Directory (RSDI)	salware?	metadata information?	havested?	second in us	the solution?							Ansonsten können die	"Usability" und d;	ie Formulierungen der Fr	agen verbessert werden.	
India -								5 - IL ENERGY	indication on now turther	5 - Is there a fixed, etcablished,	A to the ordering development and of a	5-115	ere a transpareet process for so	Es muss klan hervore	ehen, dass es sich	in einer Snalte um Stud	en handelt und nicht um i	ł
Repairs								evolution of a	ommunicy standards will be	interdisciplinary ceam, working on the	5 - Is the software development part of a	1000	g, cool review by cover develop	and the second second second	enen, aass es sien	in einer operice on ordi	en nandeze und nizene dan i	
Tunnet						G - Is there as	ununity, providing the				all second second		Providence and					
		6 - Is the meta-repositiony performing quality checks (e.g. reli	Eata) 6 - Daws the versioning scheme allow for automatic	6 - Is the PID part of an established community	6 - Is there an external quality assessment of the	opportantly	upport and exchange is	e, are reparted bugs taken into the						Oft genug tritt auch	der Fall aut, das:	s die Stuten 4, 5 oder b	erreicht werden aber Stu	1
-		for the cord publication reportany? What if an active charger, a circule selectic to sublick the	Lagging by LICD processes?	Candand 7	metalata!	Concerning a	cit of using the software? In	Development cycle?						Hier sollte das Leve	1 in 2 oder 3 so n:	iederschwellig		
		software? Still 17												wie möglich mit etwa	s Zusatzarbeit erre	eicht werden können, dam	it man dann auch direkt o	1
		- Edan's get the worth of 67 Why should the meta repo perfe	en la		- 5 and 6 is possible without fulfilling 4, no?			software service is a clearly defined to	(TT)									
		quality checks?			 what is the difference between 4 and 6, 6 is 			ind not used in the sense it is used he	re									
		-4.5.6 is not about the software publication itself but the rep	to it is - 6 is possible without fulfilling 4 and 5, no?		basically metadata curation or vice versa 4 is also			S - Laminot sure, if Lunderstand										
	Commercia	published in	 is 4 really helpful? 		elwaya 6		6	correctly, more description needed?	_									
	Auroborgor, J. (2022), Implys - Impge																	
	Enalysis (0.1) Zenada.	4	3	0	5	4,4	5	4										
	https://doi.org/30.5281/temode.811 Izmode																	
	Scherer, U., rain, B., unerschief,																	
	2023). System for automated	5	3	6	5	4.6	5	4										
	Quality Control - SoOC (2.5.0)				-		-	-										
	Ernede. Zenede																	
		1				0												
														h	ttne·//od	p holmho		
															1103.//03	5.110111110		

Outlook and next steps

- On site workshop of the Task Group (20.&21.03.2024)
 - Evaluation / Analysis results Paper and Pencil-exercise
 - Consensus building on the feasibility of the attributes, Color coding (immediate, medium-term, long-term)
 - Tool market to identify a possible Helmholtz-Indicator-Toolbox Identify cooperations in automation
- Transparent communication:
 - a) to decision-makers what is currently possible/realistic to measureb) to the people who are supposed to collect the figures (create acceptance, addressing concerns



Quality Indicators - Work in Progress Challenges I

- Current status represents conceptual framework (what/how to measure in an ideal world)
- Difference between the centers in terms of the data/software generated/used
- Different starting positions and speeds of the centers
- Automation is challenging

Manual effort currently unavoidable

Quality Indicators - Work in Progress Challenges II

Manual effort currently unavoidable - Discussions with regards to data publications

- Extension of the HMC FAIR Data Dashboard < HMC > HELMHOLTZ Metadata Collaboration
- Creation of a Repository "White List" for Helmholtz Repositories; assessment process required based on defined Dimensions/ Attributes
- Data "in the wild" (aka in Repositories outside Helmholtz)?
- \rightarrow re3data (~100 Repos with Helmholtz connection; scheme >40 attributes)



→ RDA Common Descriptive Attributes of Research Data Repositories <u>https://doi.org/10.15497/RDA00103</u>



нмс .

FAIR DATA

HELMHOLTZ Open Science

Thank you for your attention!

- E-Mail: <u>open-science@helmholtz.de</u>
- Website: <u>https://os.helmholtz.de</u>
- Mailing list for employees of Helmholtz: <u>Helmholtz Open Science Professionals</u>
- Helmholtz Open Science Newsletter
- Social Media: LinkedIn | Mastodon

HELMHOLTZ Open Science

Thank you for your attention!

Marcel Meistring

- marcel.meistring@os.helmholtz.de
- b https://orcid.org/0000-0001-6347-9926
- <u>@marcelmson@troet.cafe</u>



All texts of this presentation, except quotations and images of third parties, are licensed under a Attribution 4.0 International License Agreement: <u>https://creativecommons.org/licenses/by/4.0/deed.de</u>