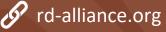


# RDA Global Adoption week

15 - 19 June 2020

**RDA #AdoptionWeek @resdatall** 

11/06/20









# (RDA Objective

 RDA Global Adoption week aims to **demonstrate the wide** variety of RDA adoptable and adopted solutions to data sharing challenges that people in the field encounter in their daily jobs.

# **Purpose of the week:**

- Learn about RDA Outputs
- Converse with speakers from all around the world who have created and implemented them
- Determine how best to integrate those data sharing solutions into your own projects

### **RDA #AdoptionWeek @resdatall**







# Monday 15<sup>th</sup> June

# Data Management

Develop best practices and tools for non-static, machine-readable data management plans which can evolve throughout the research data lifecycle, as well as be machine-readable by collaborators and stored with the data.

- RDA DMP Common Standard for Machine-actionable
   Data Management Plans
   Tomasz Miksa (TU Wien)
- SEAGrid Science Gateway: Adoption of the PID Kernel Information and Data Type Registry Utilizing the E-RPID Testbed toward FAIR Scientific Workflows - Rob Quick (Indiana University)
- <u>23 Things Revisited: Field guides to research data</u> <u>management, an adoption project</u> - Mijke Jetten, Cees Hof (LCRDM task group)

Recommendations & Outputs Catalogue short-link: https://bit.ly/2UhOxyH



# Monday 15<sup>th</sup> June

# Data Management

Develop best practices and tools for non-static, machine-readable data management plans which can evolve throughout the research data lifecycle, as well as be machine-readable by collaborators and stored with the data.

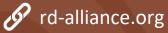
- <u>RDA DMP Common Standard for Machine-actionable</u> <u>Data Management Plans</u> - Peter Neish (University of Melbourne)
- SEAGrid Science Gateway: Adoption of the PID Kernel Information and Data Type Registry Utilizing the E-RPID Testbed toward FAIR Scientific Workflows - Rob Quick (Indiana University)
- <u>23 Things Revisited: Field guides to research data</u> <u>management , an adoption project</u> - Cees Hof, Margriet Miedema (LCRDM task group)

Recommendations & Outputs Catalogue short-link: https://bit.ly/2UhOxyH



# Go to www.slido.com













# Machine-actionable Data Management Plans

Tomasz Miksa RDA Austria tmiksa@sba-research.org







## > Shortcomings of existing DMPs

>manually completed, vague, not updated, considered bureaucracy, completed last minute, ...

	Data Officer	Who is responsible for the data management and the DMP of the project (name/email address)?
I	Data Characteristics	
I.1	Description of the data	What kinds of data/source code will be generated or reused (type, format, volume)? How will the research data be generated and which methods will be used? How will you structure the data and handle versioning? Who is the target audience?
Ш	Documentation and Metadata	a
II.1	Metadata standards	What metadata standards (if any) will be in use and why? (see <u>Digital Curation Centre</u> )
II.2	Documentation of data	What information is needed for the data to be findable, accessible, interoperable and re-usable ( <u>FAIR</u> ) in the future? Is the data machine-readable? How are you planning to document this information?
11.3	Data quality control	What quality assurance processes will you adopt? How will the consistency and quality of data collection be controlled and documented? (This may include processes such as repeat samples or measurements, standardised data capture, peer review of data or representation with controlled vocabularies.)
ш	Data Availability and Storage	
III.1	Data sharing strategy	How and when will the data be shared and made accessible? What repository will you be using? What persistent identifier will be used?
III.2	Data storage strategy	What data are to be preserved for the long-term, and what data will not be stored? How and where will the data be stored and backed up during the research? How and where will the data be stored after the project ends? For how long will the data be stored? Are there any costs that need to be covered for storage? At what point during or after the project will the data be stored? Are there any technical barriers to making the research data fully or partially accessible?



https://www.fwf.ac.at/fileadmin/files/Dokumente/Open\_Access/FWF\_DMPTemplate\_e.pdf









How to discover these tools? Which one do I need to use? Why do I have to provide the same information again? Why haven't they consulted us before? Who is going to pay for this? We don't have enough people for that!

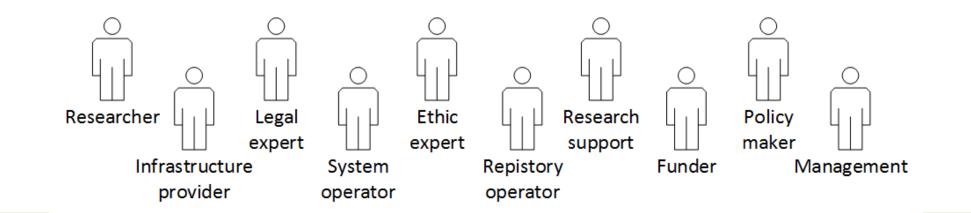






## > Stakeholders involved in research data management

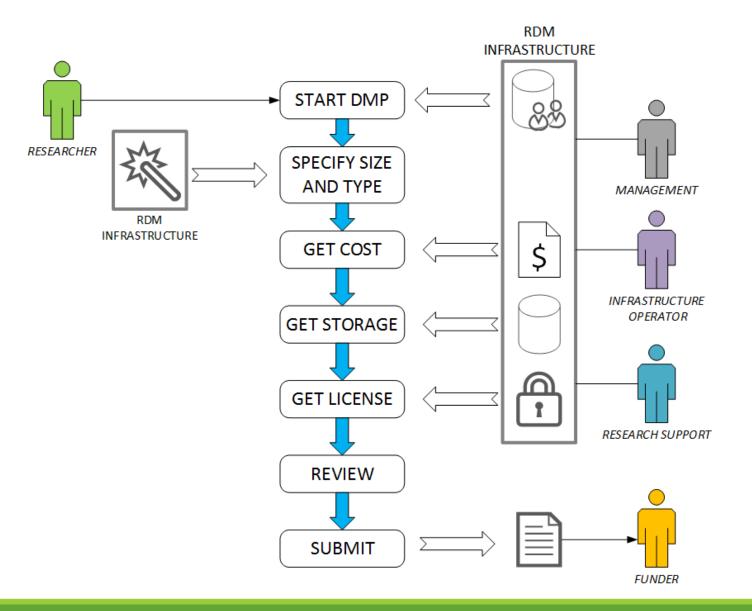
- > require information at certain stages
- > can provide information if requested at a proper stage
- > Many problems can be avoided when
  - timing is right
  - > information flow is ensured







# Automated Data Management Workflow







## >living documents

> automate data management
 > collect information from systems
 > trigger actions in systems
 > facilitate validation

>This requires

- > well-defined RDM workflows
- > data management infrastructure
- <u>common data model</u>





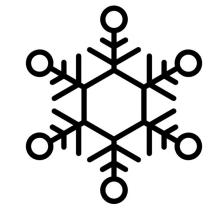
DMP Common Standards - Outputs

## > Common standard for machine-actionable DMPs

- > to model information from standard DMPs
- > NOT a template
- > NOT a questionnaire
- > modular design
  - > core set of elements
  - > domain specific extensions

# > Reference implementations

- ready to use models
  - > JSON, RDF, etc.







### • Current DMPs – model questionnaires

<administrative\_data>

<question>Who will be the Principle Investigator?</question> <answer>The PI will be John Smith from our university.</answer> </administrative data>

• Machine-actionable DMPs – model information

```
"dc:creator":[ {
	"foaf:name":"John Smith",
	"@id":"orcid.org/0000-1111-2222-3333",
	"foaf:mbox":"mailto:jsmith@tuwien.ac.at",
	"madmp:institution":" AT-Vienna-University-of-Technology"
```

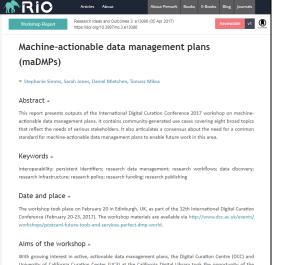
}],





- >Launched in October 2017
- Result of a consultation made by Active DMPs IG
- > Focus on machine-actionable DMPs
- >200 members from all continents
- >DMP tool owners are part of it





With growing interest in active, actionable data management plans, the Digital Curation Centre (DCC) and University of California Curation Center (UC3) at the California Digital Library took the opportunity of the IDEC17 conference to convene a workshop on the topic. The aim of the workshop was to understand different stakeholder requirements and bring together a diverse international group to develop specific use cases for machine-actionable data management plans. The workshop participants included 47 people from 16 different countries, representing funders, developers, librarians, service providers, and the research community, In prestat, barintorming exercises, the groups discussed use cases focused on interoperability with research systems, leveraging persistent identifiers, evaluation and monitoring, and repository and institutional perspectives, and prioritized future work. The DCC and UC3 will use the workshop outputs to implement and pitch use cases in the DMPRodomp platform.

#### https://doi.org/10.3897/rio.3.e13086



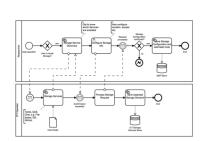


- > 1<sup>st</sup> consultation (user stories) went broad
  - > to define scope of maDMPs
- > 2<sup>nd</sup> consultation went deep
  - > to identify models for specific requirements
- > Proof of concept tools
  - > to demonstrate how model can be used to automate tasks
- > BPMN processes
  - > to identify systems and stakeholders involved
- > Model development



mmon / user-stories	Observe 7 Wonley 8 White
Internet a Strain a UNI L	Inights 0 Settings
Lobel issues and pull requests for no Area, Dirka atticking particular to our area societ with <b>size and atticked</b> of the to be Later.	must ment term
tood Lident Mindones and going Illine, and solo	New York
# 197 Owned Autory 1	abet + Popict + Millionet + Asignet Set
Transper, I used to been slowd IT reasons requirements a resource acquisition processing <b>Descent Solution Descent</b>	ariy in the project lifesyste to
nte rearrages, l'unant la sensare vales and sequenalidities are to one perturne const () me soo	nil debred, stantag state
ntermenager, l'unant he ensure the resourcher has thought the the planning phase. Testanteen the source of 21 day oper	nogh their data requirements 🔅
ets manager. Event the description of data area collected or of Effective to make the descriptions are current and add data mission communications.	used, updated throughout the Col

in K di Dellar	x +			
DATA MPORT	PRESERVATION ON			
Correlati	ng Alcohol Consu	imption and	UFO Sighting	s in the
Authors				
Marc Detrictuter - Dour 4/000 - e01279054	1000-4000-3400 Ballet Carlet AL al			
Documen 25-03-2024	t Version and Date	,		
Gathered	Data			
10.114.00000	-	1400 pare lips		
		Differi spence it par		
Software				
0.00-0-0-0		2018 a present \$ years		
Antonio interio El mediator po Antonio Materio		UNITE Appendix 1 peak		
Documen	tation			
arithmeture.proj		Utilities from		
declara tallibiliti Recription (M	CONTRACTOR OF THE OWNER	this prove if you		
Carlos Million				









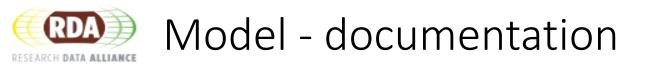




#### RDA DMP Common Standard for Machine-actionable Data Management Plans Recommendations of the RDA DMP Common Standards WG Tomasz Miksa, Paul Walk, Peter Neish Purpose Overview This application profile is meant for exchange Figure 1 presents concepts used within the of machine-actionable DMPs between application profile. Each concept is further systems. It is independent of any internal data broken down into specific fields (not depicted). organisation used by these systems. The The full application profile specification can be application profile does not prescribe how found online. Below we outline main concepts information must be presented to the end user used within the application profile that are and does not enforce any specific logic on how depicted in Figure 1. this information must be collected or used. The DMP - Provides high level information about application profile is an information carrier and the DMP, e.g. its title, modification date, etc. It the full machine-actionability can only be is the root of this application profile. achieved when systems using the application profile implement appropriate logic. Project - Describes the project associated with the DMP, if applicable. It can be used to This application profile is intended to cover a describe any type of project: that is, not only wide range of use cases and does not set any funded projects, but also internal projects, PhD business (e.g. funder specific) requirements. It theses, etc. represents information over the whole DMP lifecycle, that is, it can express planned actions, Funding - For specifying details on funded as well as actions already performed. projects, e.g. NSF of EC funded projects. The application profile is NOT intended to be a Contact - Specifies the party which can provide prescriptive template or a guestionnaire, but to information on the DMP. provide a re-usable way of representing Contributor - For listing all parties involved in machine-actionable information on themes the process of data management described by covered by DMPs. Contact Contributor DMP Cost Funding Project Security and Privacy Technical Resource Metadata Distributio Dataset License Figure 1: Overview of the application profile for the machine-actionable DMPs 1

Miksa, T., Walk, P., & Neish, P. (2019). RDA DMP Common Standard for Machine-actionable Data Management Plans. <u>https://doi.org/10.15497/rda00039</u>





#### Properties in 'contact'

Name	Description	Data Type	Cardinality	Example Value
contact_id	Identifier for a contact person	String	Exactly One	http://orcid.org/0000-0000-0000-0000
mail	E-mail address	String	Exactly One	cc@example.com
name	Name of the contact person	String	Exactly One	Charlie Chaplin

#### Properties in 'cost'

Name	Description	Data Type	Cardinality	Example Value
currency_code	Allowed values defined by ISO 4217.	Term from Controlled Vocabulary	Zero or One	EUR
description	Description	String	Zero or One	Costs for maintaining
title	Title	String	Exactly One	Storage and backup
type	Туре	Term from Controlled Vocabulary	Zero or One	
value	Value	Number	Zero or One	1000

https://github.com/RDA-DMP-Common/RDA-DMP-Common-Standard/blob/master/docs/index.md



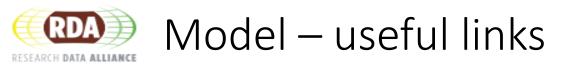




↔ Code ① Issues 1 ⑦ Pull requests 0	Insights 🔅 Settings
Branch: master   RDA-DMP-Common-Standard / docs / FAQ.md	Find file Copy p
R TomMiksa Update FAQ.md	edd9820 21 hours a
1 contributor	
85 lines (54 sloc) 8.34 KB	Raw Blame History 🖵 🖋
Frequently Asked Questions	
Trequentity Asked Questions	
Index:	
index:	
When to use the model?	
When to use the model?	
<ul><li>When to use the model?</li><li>Do I need to populate all fields?</li></ul>	
<ul><li>When to use the model?</li><li>Do I need to populate all fields?</li><li>What is the granularity of a Dataset?</li></ul>	
<ul> <li>When to use the model?</li> <li>Do I need to populate all fields?</li> <li>What is the granularity of a Dataset?</li> <li>What is a difference between Dataset and a Distribution?</li> </ul>	
<ul> <li>When to use the model?</li> <li>Do I need to populate all fields?</li> <li>What is the granularity of a Dataset?</li> <li>What is a difference between Dataset and a Distribution?</li> <li>How versioning works?</li> </ul>	
<ul> <li>When to use the model?</li> <li>Do I need to populate all fields?</li> <li>What is the granularity of a Dataset?</li> <li>What is a difference between Dataset and a Distribution?</li> <li>How versioning works?</li> <li>How to express something is planned?</li> <li>How to indicate actions that were performed?</li> <li>How to model embargoes?</li> </ul>	
<ul> <li>When to use the model?</li> <li>Do I need to populate all fields?</li> <li>What is the granularity of a Dataset?</li> <li>What is a difference between Dataset and a Distribution?</li> <li>How versioning works?</li> <li>How to express something is planned?</li> <li>How to indicate actions that were performed?</li> <li>How to model embargoes?</li> <li>Why Metadata is referenced from a Dataset?</li> </ul>	
<ul> <li>When to use the model?</li> <li>Do I need to populate all fields?</li> <li>What is the granularity of a Dataset?</li> <li>What is a difference between Dataset and a Distribution?</li> <li>How versioning works?</li> <li>How to express something is planned?</li> <li>How to indicate actions that were performed?</li> <li>How to model embargoes?</li> <li>Why Metadata is referenced from a Dataset?</li> <li>Are there any other serialisations planned different than JSON?</li> </ul>	
<ul> <li>When to use the model?</li> <li>Do I need to populate all fields?</li> <li>What is the granularity of a Dataset?</li> <li>What is a difference between Dataset and a Distribution?</li> <li>How versioning works?</li> <li>How to express something is planned?</li> <li>How to indicate actions that were performed?</li> <li>How to model embargoes?</li> <li>Why Metadata is referenced from a Dataset?</li> <li>Are there any other serialisations planned different than JSON?</li> <li>Is there a JSON Schema?</li> </ul>	
<ul> <li>When to use the model?</li> <li>Do I need to populate all fields?</li> <li>What is the granularity of a Dataset?</li> <li>What is a difference between Dataset and a Distribution?</li> <li>How versioning works?</li> <li>How to express something is planned?</li> <li>How to indicate actions that were performed?</li> <li>How to model embargoes?</li> <li>Why Metadata is referenced from a Dataset?</li> <li>Are there any other serialisations planned different than JSON?</li> </ul>	
<ul> <li>When to use the model?</li> <li>Do I need to populate all fields?</li> <li>What is the granularity of a Dataset?</li> <li>What is a difference between Dataset and a Distribution?</li> <li>How versioning works?</li> <li>How to express something is planned?</li> <li>How to indicate actions that were performed?</li> <li>How to model embargoes?</li> <li>Why Metadata is referenced from a Dataset?</li> <li>Are there any other serialisations planned different than JSON?</li> <li>Is there a JSON Schema?</li> </ul>	

#### https://github.com/RDA-DMP-Common/RDA-DMP-Common-Standard/blob/master/docs/FAQ.md

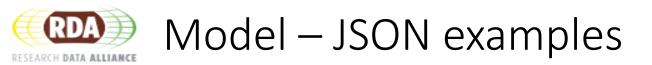




Code 🕛 Is	sues 1	្រ៉ា Pull re	equests 0	III Projects 0	🔳 Wiki	III Insights	🔅 Settin	gs					
anch: master 🔻	RDA-DMF	-Commo	on-Standar	d / docs / links.ı	nd					[	Find file	e Co	P
TomMiksa Upda	te links.md										f846491	L 2 da	ys
contributor													
9 lines (45 sloo	3.84	B						Raw	Blam	e Histo	ory 🖵		
Links													
	ollected h	ere links t	to all impor	tant resources c	reated by th	e RDA DMP C	ommon Sta	andards	s WG	(official	websi	te).	
We have co				tant resources o	-	ie RDA DMP C	ommon Sta	andards	s WG	(official	websi	te).	
We have co	sultati	on - s	coping		-	e RDA DMP C	ommon Sta	andards	s WG	(official	websi	te).	
We have co <b>1st Con</b> Collection o	<b>sultati</b>	<b>on - s</b> ories to ic	<b>coping</b> dentify scop	the maDM	-	ie RDA DMP C	ommon St	andards	s WG	(official	websi	te).	
We have co <b>1st Con</b> Collection of • Description	sultati	<b>on - s</b> pries to ic	<b>coping</b> dentify scop	the maDM	-	ie RDA DMP C	ommon Sta	andards	s WG	(official	websi	te).	
We have co <b>1st Con</b> Collection o • Descrip • User st	sultati	<b>on - s</b> pries to ic ne consul anised or	<b>coping</b> dentify scop	the maDN be of maDMPs.	-	ie RDA DMP C	ommon St	andards	s WG	(official	websi	te).	
We have co <b>1st Con</b> Collection of Descrip User st Interac	sultati	on - s ories to ic ne consul anised or lisation c	coping dentify scop ltation n a project l of user stori	the maDN be of maDMPs.	1Ps	ie RDA DMP C	ommon St	andards	s WG	(official	l websi	te).	
We have co <b>1st Con</b> Collection of Descrip User st Interac Report	sultati of user sto ption of the cories org trive visua from Vie	<b>on - s</b> pries to ic ne consul anised or lisation c nna work	coping dentify scop ltation n a project l of user stori cshop for co	the maDM be of maDMPs. board	1Ps pries	ie RDA DMP C	ommon St	andards	s WG	(official	websi	te).	
We have co <b>1st Con</b> Collection of Descrip User st Interac Report iPres c	sultati	ON - S ories to ic ne consul anised or lisation c nna work e paper su	coping dentify scop ltation n a project l of user stori cshop for co ummarising	the maDM be of maDMPs. board es bollecting user sto	1Ps pries	ie RDA DMP C	ommon St	andards	s WG	(official	l websi	te).	







RDA-DMP-Common / RDA-DMP-Common-Standard	O Unwatch ▼	3 🖈 Star	0 ¥	Fork 1	
Code (!) Issues 0 (!) Pull requests 0 [!] Projects 0	Insights	🔅 Settings			
Branch: master   RDA-DMP-Common-Standard / examples / JSON /		Create new file	Upload files	Find file	History
RomMiksa missing ,			Latest commit	ca8c7e6 12	days ago
ex1-header-fundedProject.json	missing ,			12 (	days ago
ex2-dataset-planned.json	JSON examples			12 (	days ago
ex3-dataset-finished.json	JSON examples			12 (	days ago
ex4-dataset-embargo.json	JSON examples			12 (	days ago
ex5-dataset-planned-host.json	JSON examples			12 (	days ago
ex6-dataset-closed.json	JSON examples			12 (	days ago
ex7-dataset-many.json	JSON examples			12 (	days ago

https://github.com/RDA-DMP-Common/RDA-DMP-Common-Standard/tree/master/examples/JSON







Minimal maDMP

32 li	ines (32 sloc) 679 Bytes	
1	{	
2	"dmp": {	
3	"title": "Minimal DMP",	
4	"contact": {	
5	<pre>"contact_id": {</pre>	
6	"identifier": "http://orcid.org/0000-0000-	3000-0000",
7	"type": "orcid"	
8	},	
9	"mbox": "cc@example.com",	
10	"name": "Charlie Chaplin"	
11	},	
12	"created": "2018-07-23T10:10:23.6",	
13	"dmp_id": {	
14	"identifier": "https://doi.org/10.0000/00.0.1234",	
15	"type": "doi"	
16	},	
17	"dataset": [	
18	{	
19	"dataset_id": {	
20	"identifier": "https://doi.org/10.	3000/00.0.5678",
21	"type": "doi"	
22	},	
23	"title": "Placeholder dataset",	
24	"personal_data": "unknown",	
25	"sensitive_data": "unknown"	
26	}	
27	],	
28	"ethical_issues_exist": "unknown",	
29	"language": "eng",	
30	"modified": "2019-02-06T15:30:42.1"	
31	}	
32	}	





- > **DMP Online** by Digital Curation Centre (DCC) in the UK
- > **DMP Tool** by California Digital Library (CDL) in the US
- > **DMP OPIDoR** by Centre national de la recherche scientifique (CNRS) in France
- > RDMO by Leibniz-Institut für Astrophysik Potsdam in Germany
- > Data Stewardship Wizzard by Elixir research infrastructure in the EU
- > Argos OpenDMP by OpenAIRE and EUDAT research infrastructures in the EU
- > F1000Research open research publisher in the UK
- > NSD DMP in Norway
- > Haplo repository in the UK
- > TU Wien, TU Graz, Uni Wien via FAIR Data Austria project

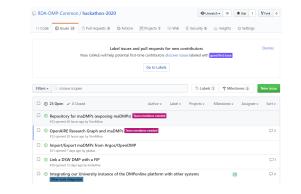
>....





- > 71 participants, 12 teams, 21 countries
- >Integrations, mappings, etc.
- > <u>Results</u>
  - > Reports, slides
  - Grand finale recording





https://github.com/RDA-DMP-Common/hackathon-2020







#### Results of the hackathon

The **recording of the Grand Finale** session can be found here. Each of the 12 teams participating in the hackathon has summarized the results. The recording also includes live demonstrations.

Table below presents a list of teams, topics they dealt with and links to the work completed during the hackathon

Team	Торіс	Links
The Unaturals	New version of the DCSO (DMP Common Standard Ontology)	DCSO Repository, Issue Document, slides
maDMP Link	Integrating a Converis CRIS/RIMS with DMPRoadmap and Export/Import maDMP from Figshare	results
DSW	Export/import of maDMPs from/to Data Stewardship Wizard	report, slides
The Datatypists	Add data types and design suppor for datasets to make EasyDMP more compatible with RDA DMP Common Standard	report, json to test imports
InsTmaDMP	Aligning institution RDM to maDMP Common Standard	results
TigTag	Mapping of maDMP standard to funder templates	report
DMP Exchange	Exchange DMPs between DMP Tools using the RDA Common Standard	report
Something	maDMP integration with project data management workflow	slides
DMP InvenioRDM	maDMP integration into InvenioRDM	results
RDMO	maDMP export from RDMO	slides
DMP Ninjas	Import/Export maDMPs from Argos/OpenDMP	results
Fancycatmeme	maDMP integration with Research Data Connectome Data Pipeline	article, slides

#### https://github.com/RDA-DMP-Common/hackathon-2020/blob/master/results.md





### > Active Data Management Plans IG

> Umbrella group

> Place to bring your new ideas!

## Exposing Data Management Plans WG

> Opening DMPs for public

> Use cases for sharing DMPs

### > <u>DMP Common Standards WG</u>

Standard specific discussion

>Adoptions of the standard





Building the social and technical brid	lges to enable open sharing and re-use of data	RDA EU RI	DA US CONTACT US LOGIN R	REGISTRATION	≫ 🕞 🛱 🛄 🦉
RDA RESEARCH DATA ALLIANCE	O&A Members 61 Active Organisational & Affiliate members	MEMBERSHIP Becoming a member of RI open to both individuals a Register now		RDA Groups Discover what RDA Workl Groups and all other Gro out how to join them. Exp	ups are up to and find
	ED • GROUPS • RECOMMENDATIONS Active Data Management Plans IG ta Management Plan		SCIPLINES - PLENARIES	& EVENTS Y NEWS & N	NEDIA - C
Posts		utputs Charter	Plenaries Member	Cleate new	v content
Group Status: 🔗 IG Establ	ished			<b>&amp;</b>	Join Group

#### 6

IG

NEW! February 2020: In order to make it easier for you to collaborate with your teams, we have improved the user experience of your Groups' online space. A series of icons and labels now guide your activity and help you post messages to the group members, create and organise wiki pages, send events' announcements, publish and organise the outputs and case statements resulting from your group's activity and browse all the members of your Group. One new area also collects the Plenary sessions your group contributed to.

We hope that you'll find this useful! Please do not hesitate to send your comments and suggestions to the RDA Secretariat here.

#### Group details

Status: Recognised & Endorsed Chair (s): David Giaretta, Kevin Ashley, Sarah Jones, Tomasz Miksa, John Chodacki TAB Liaison: Jane Wyngaard Case Statement: Download

IG Established

The proposed activity of this group is to act as a nucleus for discussing requirements for and identifying developments needed to support active (i.e. able to evolve and be monitored) data management planning. Working groups will be proposed to carry out work on specific areas of interest. Currently research data management plans (DMP), created at the proposal stage of a project, do not evolve and cannot be monitored in any detail. The DMP should begin at the planning stage for any dataset evolving through its particular to the source for any data is appropriately.







#### **Adoption Stories**

Home » Recommendations & Outputs » Adoption Stories

The Research Data Alliance (RDA) currently hosts over 60 Interest Groups and more than 30 Working Groups consisting of experts who are working on various topics related to (open) research data and innovation. These working groups produce the RDA outputs: the technical and social infrastructure solutions enabling data sharing, exchange, and interoperability.

For you, to see how to implement the RDA outputs to improve the sharing, exchange and interoperability of your own data, we've asked RDA members who have already adopted RDA outputs, to share their experience and lessons learned in a story. Find below a series of RDA adoption stories by individuals, organisations and projects. NEW! Recommendations & outputs catalogue Adoption Use Cases Adoption Stories RDA Europe Adoption Grants Interest in RDA Recommendations Standards



Next Event

#### DMPs in Sweden: Sharing Good Practice. 13.00-15.00 (CET)

We are pleased to announce a half-day online workshop on Data Management Plans, organised by the Swedish RDA node. During the workshop we will take a look at the DMP related work within RDA and hear from universities that are working activly with DMP's today.

SUBMIT YOUR STORY HERE

https://www.rd-alliance.org/recommendations-outputs/adoption-stories







- Tomasz Miksa, Stephanie Simms, Daniel Mietchen, Sarah Jones (2019) Ten principles for machine-actionable data management plans. PLOS Computational Biology 15(3): e1006750.
- Tomasz Miksa, Peter Neish, Paul Walk, Andreas Rauber: Defining requirements for machineactionable Data Management Plans. iPres 2018
- Tomasz Miksa, João Cardoso, José Luis Borbinha: Framing the scope of the common data model for machine-actionable Data Management Plans. BigData 2018: 2733-2742
- Asztrik Bakos, Tomasz Miksa, Andreas Rauber: Research Data Preservation Using Process Engines and Machine-Actionable Data Management Plans. TPDL 2018: 69-80





# Standard specification

https://github.com/RDA-DMP-Common/RDA-DMP-Common-Standard

- > Participate in model adoption!
- Contact group chairs
  - Questions
  - Ideas
  - Success stories



Tomasz Miksa



Paul Walk



Peter Neish









# **Questions?**





# Toward FAIR Data Workflows with SEAGrid and RPID

Rob Quick, Yu Luo, and Guangchen Ruan Cyberinfrastructure Integration Research Center Pervasive Technology Institute Indiana University rquick@iu.edu



Researcher uses a Science Gateway to do research

- They are concerned the digital objects they use and produce are aligned with **FAIR** principles
- They want the research to be reproducible beyond the Science Gateway environment
- Share the steps used (not just the data) to calculate the results simply and easily within a publication or with collaborators
- Reuse a workflow with 'tweaks' without having to recreate the entire computational workflow











### FAIR PRINCIPLES

### Findable

F1. (meta)data are assigned a globally unique and persistent identifier;

F2. data are described with rich metadata

F3. metadata clearly and explicitly include the identifier of the data it describes;

F4. (meta)data are registered or indexed in a searchable resource;

### Accessible

A1. (meta)data are retrievable by their identifier using a standardised communications protocol;

A1.1. the protocol is open, free and universally implementable;

A1.2. the protocol allows for an authentication and authorisation procedure, where necessary;

A2. Metadata are accessible, even when the data are no longer available;

### Interoperable

11. (meta)data use a formal, accessible, shared and broadly applicable language for knowledge representation;

12. (meta)data use vocabularies that follow FAIR principles;

I3. (meta)data include qualified references to other (meta)data;

### Reusable

R1. (meta)data are richly described with a plurality of accurate and relevant attributes;

R1.1. (meta)data are released with a clear and accessible data usage license;

R1.2. (meta)data are associated with detailed provenance;

R1.3. (meta)data meet domain-relevant community standards;



Slide provided by Luiz Bonino



DOI: 10.1038/sdata.2016.18



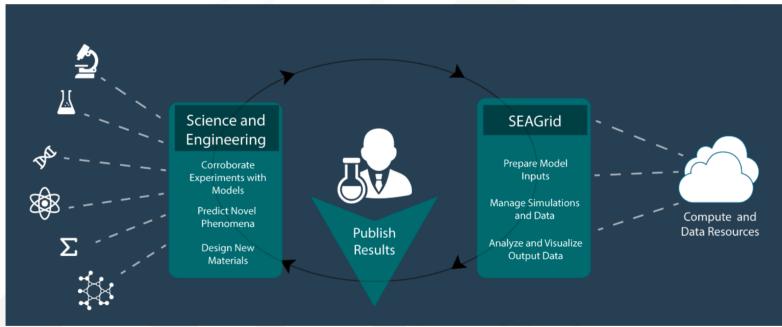




# Science and Engineering Application Grid

Science Gateway built with Apache Airavata Middleware Framework

This adoption centered on small molecules and fluorescent properties



15/06/20





# RDA What is the RPID Testbed?

- Data cyberinfrastructure for minting PIDs, resolution of PID metadata, data type registry, and protocol for operations on digital objects
- Basically the cyberinfrastructure to implement technical components of FAIR principles
- Services leverage several RDA Outputs and Recommendations
  - PID Kernel Information Strawman Profile
  - Data Type Registry
  - $\odot$  Heavily leverages work done in the Data Fabric WG
- NSF Funded Testbed (Grant No. 1839013)
- More at <u>https://rpidproject.github.io/rpid/</u>



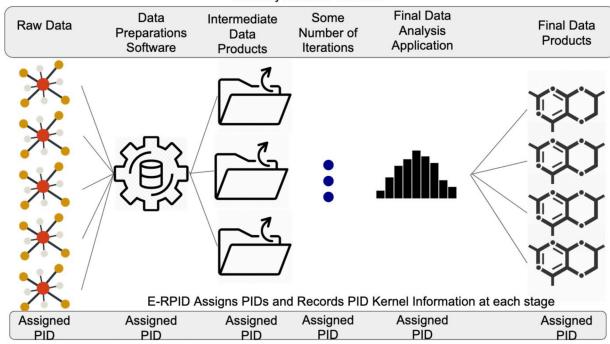




# RDA The RDA Adoption Project

# 6-Month project funded by RDA-US

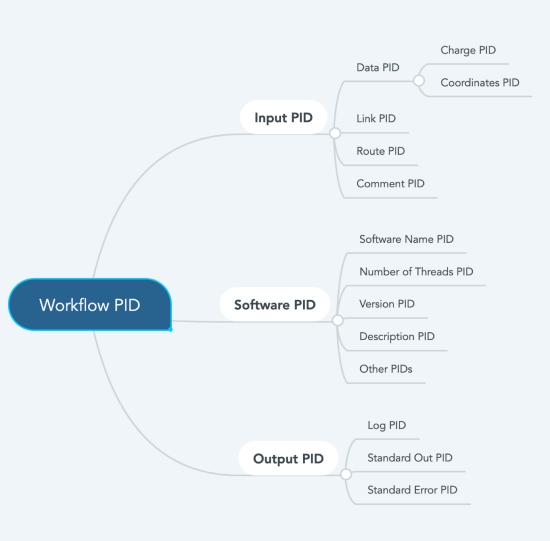
### Integrates SEAGrid with ERPID Services to provide a FAIR Science Gateways Gateway Enabled Workflow













#### Grd-alliance.org









https://rpid.seagrid.org if time allows.

If there's not time for a demo go to http://hdl.handle.net/11723/SEAGrid.96f51339-8bfe-4b69-a11e-597f21f31ddb and explore.



38

15/06/20





# (RDA)) Conclusion

SEAgrid has been integrated with the RPID testbed services to implement FAIR science workflows

- We are currently exercising the integrated system and considering publishing run results to a public fluorophores data repository
- Workflow PIDs describe the components of computational processes used during a scientific workflow
- With the PID you could recreate the entire computation workflow
  - $\,\circ\,$  Though in the real world you would probably not recreate it in its entirety
  - A realistic scenario would be to mix new or updated data, software, or parameters into a previously used workflow method
  - No access to SEAGrid is necessary to get the metadata necessary for this scenario
- Directly impacts the FAI principles, not as much for the R
- Is extendable in both metadata and to other Airavata based science gateways
- This project leverages outputs and resources made available by the RDA Community







# **Questions?**





# Monday 15<sup>th</sup> June

### Data Management

Develop best practices and tools for non-static, machine-readable data management plans which can evolve throughout the research data lifecycle, as well as be machine-readable by collaborators and stored with the data.

23 Things Revisited: Field guides to research data management, an adoption project

#### Mijke Jetten, Cees Hof & Margriet Miedema

Representing the 23 Things task group of the (Dutch) National Coordination Point Research Data Management (LCRDM)

**RESEARCH DATA ALLIANCE** 



# Adopting the 23 Things by the Dutch community

What? Creating versions of the 23 Things for different audiences, by the Dutch community

Why? 23 Things as a shared reference tool for knowledge on RDM. Improving national cooperation, and a common understanding of RDM among diverse practitioners and supporters

**How?** For this RDA adoption grant, a LCRDM task group generates four deliverables between June '19 and July '20

#### **Today's session**

- 1. Summary of the project
- 2. Showcase: 23 Things tool
- 3. How can you contribute, as a potential future adopter



#### 23 Things: Libraries for Research Data

An overview of practical, free, online resources and tools that you can begin using today to incorporate research data management into your practice of librarianship.

#### **Research Data Sharing Without Barriers**

#### Learning Resources

Librarians are learning how to apply the principles of library science to solve problems and to provide new services related to research data.

- 1. A "top ten" list of recommendations for libraries to get started with research data management from LIBER, http://bit.ly/RDAthing1
- 2. Relevant concepts are presented and mapped in the e-Science Thesaurus, http://bit.ly/RDAthing2
  - 3. Understanding the life of research data with the DCC Curation
  - Lifecycle Model,

10. Questions about data answered by experts on the DataQ forum, http://bit.ly/RDAthinq10

#### **Data Management Plans**

Librarians are becoming familiar with funder requirements and consulting with researchers to help them write and implement effective data management plans.

11. One example is the DMPTool that lists funder requirements in the United States and builds a plan by asking the researcher to answer a series of questions. Other countries such as the U.K. and Canada have similar tools. http://bit.ly/RDAthing11

The original 23 Things was created by the RDA Libraries for Research Data IG

Learning Resources

Data Reference and Outreach

Data Management Plans

Data Literacy

**Citing Data** 

Data Licensing and Privacy

**Digital Preservation** 

**Data Repositories** 

and a Community of Practice

..to help librarians engage in research data management!

Data Reference & Outreach

Librarians are answering questions

conducting outreach to assess the

data needs of their researchers and

about data from patrons and







# Creating a Dutch nationwide commitment

The project's implementation plan is shared in the LCRDM pool of experts and other Dutch stakeholder groups on data management.

Deliverable: doi.org/10.5281/zenodo.3337870

https://doi.org/10.5281/zenodo.3337870







## Adjusting the 23 Things

We started creating local versions of the 23 Things for the Dutch community. Via joint sprint sessions, the original resource is being updated and adapted to the Dutch community, and different audiences.

Deliverable: doi.org/10.5281/zenodo.3465895

https://doi.org/10.5281/zenodo.3465895







### Audience-specific versions

Join the community effort in creating an overview of RDM resources. Help us finalize audience-specific versions of the 23 Things for researchers & PhD candidates, Bachelor & Master students, data & subject librarians, data stewards, IT support staff & IT specialists, research software engineers, and policy makers.

https://doi.org/10.5281/zenodo.3773663







### Getting the 23 Things adopted

Particularly in training, the 23 Things are expected to be useful, as it may help creating a consensus on the content of RDM courses in the Netherlands. One way of doing this is by developing an online tool for browsing the various audience-versions of the 23 Things.

*Deliverable*: Recommendations for use of the 23 Things







# Dissemination of experiences and final versions

The project will be promoted via blogs, webinars, conferences and an article in an international data journal. All materials and outputs (including sharing our experiences to allow others to follow in our footsteps) will be published on Zenodo.

*Deliverable*: Sustainability plan



23 Things for Data Stewards

Policy Development Data Management Plans Compliance

Data Reference & Outreac

research data managemen

management policy,

Community of Practice Metadata Digital Preservation & Repositorie

data management into your data stewardship practices

for the institute, department or project, including services.

edu.nl/br77i & edu.nl/ag6v4 & edu.nl/43pgf

data stewards working with social science data,

edu.nl/ajy6k & edu.nl/ug3w9

ement Plans

edu.nl/p4nke & edu.nl/i8cr6

institute, department or project.

Data Manas

An overview of practical resources and tools that you can begin using today to incorporate

... to help data stewards engage with policy, research and infrastructure oriented stakehol

Data stewards develop, implement and monitor the research data management policy and strategy

1. Use the Learn RDM toolkit to check review the relevant elements of a successful research da

management services; also review LCRDM Put into Practice to ensure that appropriate local

Health-RI facilitates the optimal use of knowledge, tools, facilities, health data and samples in
personalised medicine and health: while the CESSDA Data Management Expert Guide might heli

Data stewards help researchers write and implement effective data management plans in the

4 Understand the life of research data with examples of data life cycles at USC and RDNI

skills on research data mana

2. Use the resources of DCC and RDNL to learn how to develop and deliver effective dat

infrastructures and services are in place to support researchers.



RDA

Ten Tins

tion and

# The output....

- Things for researchers & PhD candidates
- Things for Bachelor & Master students
- Things for data & subject librarians
- Things for data stewards
- Things for IT support staff & IT specialists
- Things for research software engineers
- Things for policy makers
- Things for Research Data

https://doi.org/10.5281/zenodo.3773663

National Coordination Point Research Data Management

	Search the Site	Q			
23 RDA Things					
🕈 Partners About					
National Coordination Point Research Data Management (The Netherlands)					
¢)(Type	Data life cycle				
Basic programming skills	Best practices in software archiving	Budget for managing and stewarding data			
Learn advanced programming skills	lake a look at best practices in software archiving	Spend 5% of total research expenditure on properly managing and 'stewarding' data			
Learning Resources	Metadata	Infrastructure			
Budget for managing and stewarding data	Check career paths	Check for available external resources			
Spend 5% of total research expenditure on properly managing and 'stewarding' data	Check the different global goals and career paths with employment conditions from the association of universities in the Netherlands,	Check for external certified available resources to scale out local systems			
Infrastructure	Support for the Data Life Cycle	Support for the Data Life Cycle			
Check for European Services	Check important players RDM Netherlands	Check important players RDM Netherlands Get acousing with the providers			
	nal Coordination Po rch Data Managemu rlands) a (Type Basic programming skills Learn advanced programming skills Learn advanced programming skills Check for European Services Learn about the different services	mal Coordination Point rch Data Management         trans         trans         e       Type         e			

RESEARCH DATA ALLIANCE

# The output....

Interactive site & underlying database!

Filter by:

Audience (7 options) Theme (16 options) Type (13 options) Data life cycle (11 options)

https://23things.sites.uu.nl (beta)

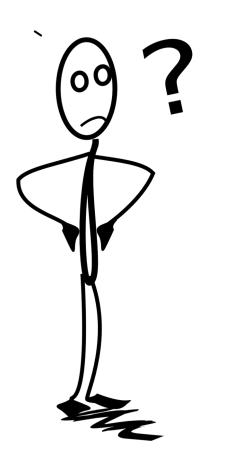






Using the database!

### A user story....



#### **Boris:**

- 1) Is a postdoctoral researcher at the TU Delft
- 2) Has shared his code with one of his PhD students
- 3) She asked him how she should cite his code/software in her thesis?





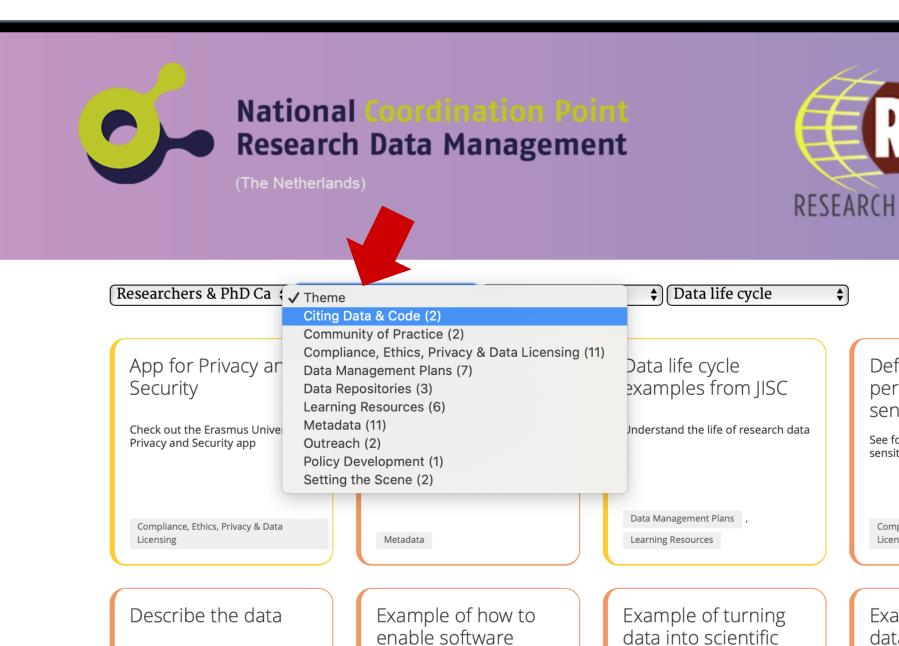
Boris selects his audience...

	al Coordination Po ch Data Managemo	ent Ę	EARCH
<ul> <li>Policy Makers (45)</li> <li>Research Software Engineers (36)</li> <li>Researchers &amp; PhD Candidates (45)</li> </ul>	•) <b>\$</b> (Type	Data life cycle	\$
App for Privacy and Security Check out the Erasmus University's Privacy and Security app	Checklist for assessment of FAIR Use this checklist for assessment of FAIRness of your data	Data life cycle examples from JISC Understand the life of research data	Def per ser See fo sensit
Compliance, Ethics, Privacy & Data Licensing	Metadata	Data Management Plans , Learning Resources	Com Licer
Describe the data	Example of how to enable software	Example of turning data into scientific	Exa dat





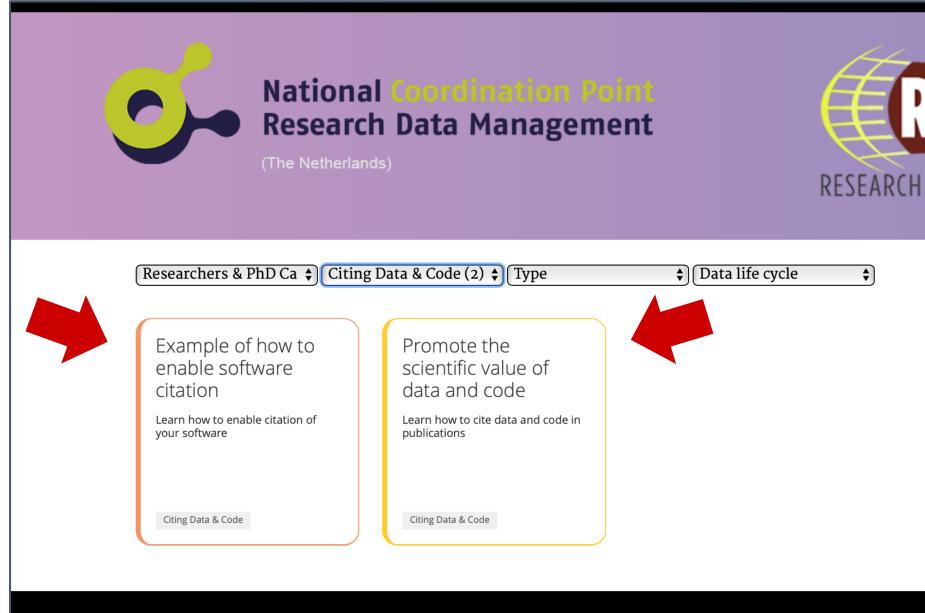
Boris selects his theme...







Boris gets two possible solutions...



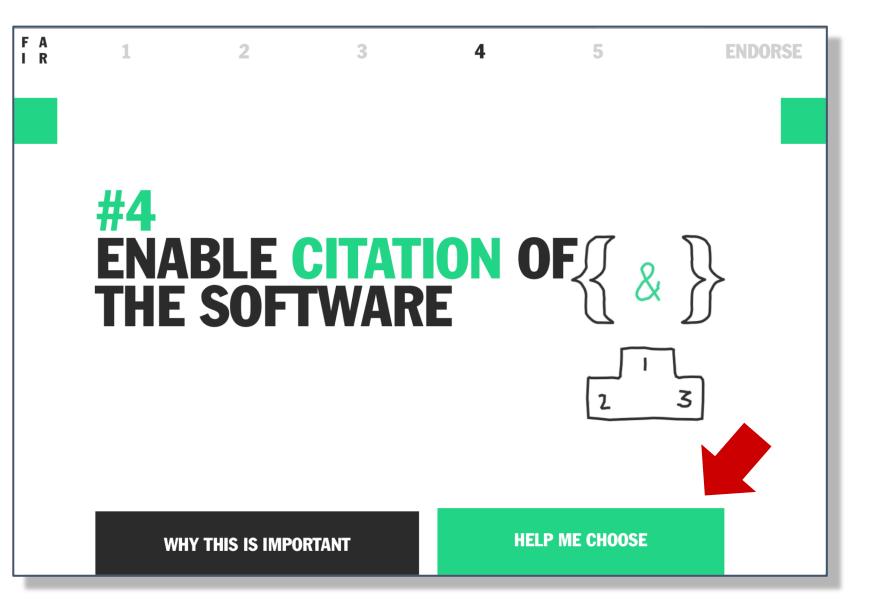




National Coordination Point
 Research Data Management

Boris gets where he wants to be...

FAIR Software Route <u>https://fair-software.eu</u>

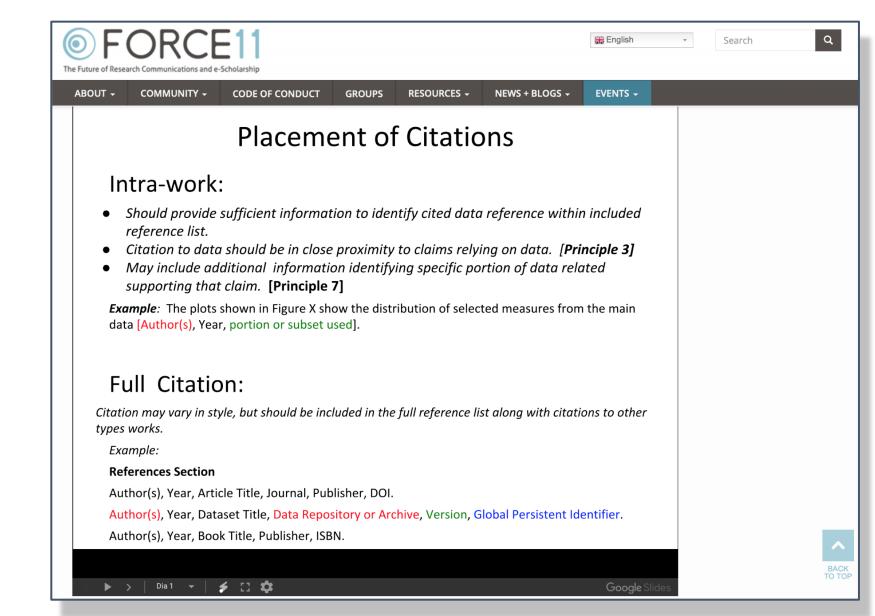




Boris gets where he wants to be...

National Coordination Point Research Data Management

Force11 Slide Deck on citation <u>https://www.force11.org</u>







# Design your own learning path!

RESEARC

Partners About

**National Coordination Point Research Data Management** 

(The Netherlands



Data Stewards (4)

**•** 

Data Management Pla 
RDM tool (4)

planning (4)

\$

Resources for reviewing and assessing DMPs

Review and/or assess the data management plans written by your researchers

Data Management Plans

Resources for reviewing and assessing DMPs

Review and/or assess the data management plans written by your researchers

Data Management Plans

Resources for reviewing and assessing DMPs

Review and/or assess the data management plans written by your researchers

Data Management Plans

#### Resources for reviewing and assessing DMPs

Review and/or assess the data management plans written by your researchers

Data Management Plans





# Design your own learning path!





# Monday 15<sup>th</sup> June

### Data Management

Develop best practices and tools for non-static, machine-readable data management plans which can evolve throughout the research data lifecycle, as well as be machine-readable by collaborators and stored with the data.

23 Things Revisited: Field guides to research data management, an adoption project

Remarks, suggestions, additions? Want to add your own Things to the tool? Please mail: info@lcrdm.nl



# **TIME FOR QUESTIONS!**



# **Questions?**



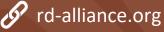
# **Recommendations and outputs catalogue**

- RDA Outputs are classified as RDA Recommendations (official, endorsed results of RDA Groups), Supporting Outputs (useful solutions from our RDA Working and Interest Groups) or other Outputs
- They can be searched according to their status, Data Life Cycle topics or scientific domain



rd-alliance.org/recommendations-and-outputs/catalogue

(RDA







#### (RDA) **Tell your adoption story!**

- Are you an adopter? RDA is actively seeking new adoption stories to inspire the further uptake of RDA outputs.
- Submit your story here: https://www.rdalliance.org/tell-your-rdaadoption-story

#### **RDA ADOPTION STORIES**

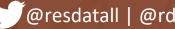
Adopters of RDA outputs share their experiences and lessons learned to inspire further uptake of **RDA** outputs

oro Read the current adoption stories



Submit your story through the webform







#### (RDA **Call for papers: CODATA Data Science Journal**

• CODATA Data Science Journal: <a href="https://datascience.codata.org/">https://datascience.codata.org/</a>

#### RDA special collection:

- Results produced by an IG or WG;
- Description of an Adoption Case outlining how a specific recommendation or output has been implemented;
- Other types of work related to RDA activities.
- https://datascience.codata.org/collections/special/research-data-alliance-results/
- RDA Europe 4.0 has funds available for the publication of articles in DSJ
- Open to all interested applicants regardless of their geographical provenance.
- Deadline 17 July 2020









Day & Topic	Sessions
Monday, 15th June 2020 - Data Management	14:00 UTC + 23:00 UTC
Tuesday, 16th June 2020 - Data Description	06:00 UTC + 14:00 UTC
Wednesday, 17th June 2020 - Identify, Store and Preserve	07:00 UTC + 14:00 UTC
Thursday, 18th June 2020 - Disseminate, Link and Find	07:00 UTC + 12:00 UTC
Friday, 19th June 2020 - Policy, Legal Compliance and Capacity	05:00 UTC + 13:00 UTC









#### WWW.RD-ALLIANCE.ORG/ **@RESDATALL**



#### **RDA Global**

Email - enquiries@rd-alliance.org Web - www.rd-alliance.org **Twitter - @resdatall** LinkedIn - www.linkedin.com/in/ResearchDataAlliance Slideshare - http://www.slideshare.net/ResearchDataAlliance

**RDA Europe Twitter - @RDA\_Europe** 

**RDA US Twitter** - **@RDA\_US** 



