



RESEARCH DATA ALLIANCE EUROPE

Project Acronym	RDA Europe
Project Title	Research Data Alliance - Europe 3
Project Full Title	RDA Europe 3 - the European plug-in to the global Research Data Alliance
Project Number	653194
Deliverable Title	Planning and Activities Report
Deliverable No.	D4.6
Delivery Date	18/06/2016
Author	Costantino Thanos, Franco Zoppi

ABSTRACT:

This WP4 Planning and Activities Report describes the actions that are intended within WP4 in the coming 12 months to engage with the practitioners and to organize uptake. It is the second report of this type describing what has been done and what is being planned.

DOCUMENT INFORMATION

PROJECT		
Project Acronym	RDA Europe	
Project Title	Research Data Alliance – Europe 3	
Project Full Title	RDA Europe 3 – the European plug-in to the global Research Data Alliance	
Project Start	1 st September 2015	
Project Duration	30 months	
Funding	H2020-EINFRA-2014-2	
Grant Agreement No.	653194	
DOCUMENT		
Deliverable No.	D4.6	
Deliverable Title	Planning and Activities Report	
Contractual Delivery Date	31 March 2016	
Actual Delivery Date	June 2016	
Author(s)	Costantino Thanos, Franco Zoppi (CNR-ISTI)	
Editor(s)	Costantino Thanos, Franco Zoppi (CNR-ISTI)	
Reviewer(s)	Peter Wittenburg (Athena-RC), Raphael Ritz (MPG)	
Contributor(s)	CNR-ISTI, MPG, Athena-RC, BSC	
Work Package No. & Title	WP 4: Practitioner Level Engagement and Uptake	
Work Package Leader	Raphael Ritz (MPG)	
Work Package Participants	MPG, CSC, TRUST-IT, CNR-ISTI, Athena, STFC, ACU, CNRS, RIA, NLI	
Distribution	Public	
Nature	Report	
Version / Revision	1.0	
Draft / Final	Final	
Total No. Pages (including cover)	12	
Keywords	Technology, Computer Science	

DISCLAIMER



Communications Networks, Content and Technology
European Commission Directorate General

DG CONNECT

RDA Europe (653194) is a Research Infrastructures Coordination and Support Action (CSA) co-funded by the European Commission under the Research and Innovation Framework Programme, Framework Programme Horizon 2020 (H2020).

This document contains information on RDA Europe (*Research Data Alliance Europe*) core activities, findings and outcomes and it may also contain contributions from distinguished experts who contribute as RDA Europe Forum members. Any reference to content in this document should clearly cite the authors, source, organisation, and publication date.

The document has been produced with the funding of the European Commission. The content of this publication is the sole responsibility of the RDA Europe Consortium and its experts, and it cannot be considered to reflect the views of the European Commission. The authors of this document have taken any available measure in order for its content to be accurate, consistent and lawful. However, neither the project consortium as a whole nor the individual partners that implicitly or explicitly participated in the creation and publication of this document hold any sort of responsibility that might occur as a result of using its content.

The European Union (EU) was established in accordance with the Treaty on the European Union (Maastricht). There are currently 28 member states of the European Union. It is based on the European Communities and the member states' cooperation in the fields of Common Foreign and Security Policy and Justice and Home Affairs. The five main institutions of the European Union are the European Parliament, the Council of Ministers, the European Commission, the Court of Justice, and the Court of Auditors (http://europa.eu/index_en.htm).

Copyright © The RDAEurope Consortium 2014.

For more information on the project, its partners and contributors please see <https://europe.rd-alliance.org/>. You are permitted to copy and distribute verbatim copies of this document containing this copyright notice, but modifying this document is not allowed. You are permitted to copy this document in whole or in part into other documents if you attach the following reference to the copied elements: "Copyright © The RDA Europe Consortium 2014."

The information contained in this document represents the views of the RDA Europe Consortium as of the date they are published. The RDA Europe Consortium does not guarantee that any information contained herein is error-free, or up to date. THE RDA Europe CONSORTIUM MAKES NO WARRANTIES, EXPRESS, IMPLIED, OR STATUTORY, BY PUBLISHING THIS DOCUMENT.

GLOSSARY

ABBREVIATION	DEFINITION
BoD	Board of Directors
CB	Consortium Board
DoA	Description of the Action
EC	European Commission
EU	European Union
IG	Interest Group in the global RDA initiative
PM	Person Months
PMO	Project Management Office
RDA EU 3	Research Data Alliance - Europe 3
SyA	Synchronisation Assembly
TWM	Technical Working Meetings
WG	Working Group in the global RDA initiative
WP	Work Package in the RDA EU project
WPL	Work Package Leader

TABLE OF CONTENTS

Content

1	Executive Summary	6
2	Activity Report	7
2.1	Topic.....	7
2.2	Questions Addressed	8
2.3	Agenda	8
2.4	Participants	9
2.5	Time Plan for Reporting and Actions	10
3	Activities Planned	11
3.1	Activities Planned for the second year	11
3.2	Activities Planned for the third year	11
4	Appendix.....	12

1 Executive Summary

This Deliverable focuses, essentially, on the activities carried out by the Task 4.3. The main planned activity is the organization of a Technical Working Meeting (TWM) per year and to follow up on its impacts.

The first report of this task 4.3 included:

- the main objective and the activities to be carried out
- the concept of the Technical Working Meetings¹
- the focus on bringing experts from various backgrounds together, in particular to strengthen the computer science perspective in the RDA work
- the plans for a first TWM in the RDA EU3 project

The first TWM was organised under the RDA EU2 project and had as topic an analysis of the results of the WG on Dynamic Data Citation. The TWM was organised in the end phase of the work of the Working Group to give feedback to the WG. Basically it resulted in a confirmation and partly a refinement of the rules that were presented by the WG. Therefore, we can claim that this workshop had very positive consequences for the RDA work. One of the WG chairs who participated in the TWM was extremely satisfied about the results, indicating the correctness of the approach and the found rules.

This report presents

- the results of the second TWM being organised which is the first within the RDA EU3 project and
- a timeline for further actions

¹ By TWM we mean a meeting attended by max 20 attendees, where participation is only by invitation and there is a balanced audience of technical experts/computer science researchers and RDA group leaders.

2 Activity Report

While the first TWM was meant to review the results that had been worked out by the RDA Working Group on Dynamic Data Citation based on computer science knowledge, the second TWM was meant to work on a topic which has a high relevance, but where up till now no activity in RDA could be observed, i.e. the question implicitly addressed it whether a TWM can stimulate specific activities in RDA to overcome hurdles that had been addressed during the working meeting.

This second TWM took place at the Island of Santorini at 21-22 April 2016. CNR-ISTI was responsible of the content-related organisation and ATHENA as Greek member takes care of the local organisation.

- **Organizing Team**

- Costantino Thanos organizer
- Catherine Bosio secretariat support
- Katerina Papadaki administrative support
- Eleni Sotiropoulou local arrangements

- **Editorial Board**

- Leonardo Candela
- Umberto Straccia

2.1 Topic

The chosen theme for this second TWM was

Discoverable Data Service/Tool Registry.

Service/tool discoverability is an increasingly important topic given the increasing amount of usable algorithms that are being developed globally to carry out in particular data analytics. In particular Open Science will require to not just make data much more accessible, but to also make methods and tools available to other scientists and even beyond.

Already now only few have an overview of what has been developed and could be used even within research fields. Finding an appropriate tool therefore is most often the result of personal networking instead of systematic search. Given the Open Science trend this situation will become worse. Researchers, scientists, data scientists, and practitioners in all kinds of scientific fields are creative in developing new data services and tools including data mining, data visualization, data analysis tools, etc.

The main question that was addressed during the workshop therefore was how to build appropriate registries and thus how to describe a data service/tool on the different levels, which modelling approaches are there or can be thought of and which kind of discovery mechanisms can make sense. It will influence efficiency and competitiveness how in particular researchers will be able to locate data services/tools in a globally networked scientific world that fulfil their research needs. Efficiency of discovering appropriate data services/tools and possibly composing them to build complex scientific workflows is a requirement of modern science. A crucial aspect in making this happen is the ability to semantically describe the

functionality of a service/tool. It becomes increasingly necessary to create registries that maintain semantically enriched descriptions (metadata/profiles) of data services/tools. It is also necessary to develop discovery mechanisms that match the service/tool description against the description of a user need.

While there are many detailed activities in the realm of metadata taking place in RDA at the moment covering a broad range of topics, at present, semantic descriptions (metadata) of data service/tools are missing.

References to related work were sent to the participants and are listed in the appendix.

2.2 Questions Addressed

The organiser raised a number of questions to also open a discussion about the possibility of creating an RDA WG on "Research Data Service Discoverability". [See <https://rd-alliance.org/working-and-interest-groups/goals-and-outcomes.html>].

1. What is a "Research Data Service"?
(our definition at an abstract level: *a research data service is a rule of correspondence between two sets*).
2. Can a "rule of correspondence" be described "algorithmically"?
3. What kind of metadata semantics are needed for the description of the "input set" (domain) and the "output set" (co-domain)?
4. Do we need to specify the syntax and semantics of the elements of the domain and co-domain?
5. Which are the differences between a Research Data Service and a Web Service?
6. Are Research Data Services "stateless" or "state-based" services?
7. What should be included in a Research Data Service profile in order to appropriately describe its functionality?
8. Are the scientific workflows appropriate for describing the process model of a Research Data Service?
9. Is the OWL-S framework appropriate for the description of Research Data Services?
10. Do we need discipline-specific classification of data services (classes of data services) supported by discipline-specific ontologies?
11. Which are the main characteristics and capabilities of a knowledge representation language appropriate for the description of the functionality of a data service as well as for effectively supporting reasoning in the matchmaking process?
12. Which is the role and architecture of Registries/Directories/Catalogues of Services?
13. How to classify Research Data Services? "stateless/state-based", "type of input data: discrete data/vectors/functions/streaming data", "types of output data", etc.
14. Is "scalability" a functional characteristic of a research data service?
15. Is "citation" instrumental in making research data service discoverable?

2.3 Agenda

21 April 2016

9.00 – 9.15	Introduction
9.15 – 9.30	Presentation of the Participants
9.30 – 11.10	Current Practices (Elixir, OpenAIRE, GEOSS, BioDiversity, Eudat) (20 min each presentation)

11.10 – 11.40	A Case Study (CNR)
11.40 – 12.00	Coffee Break
12.00 – 13.00	Presentation of Participants' Views (20 min each presentation)
13.00 – 14.00	Lunch
14.00 – 16.00	Presentations of Participants' Views (cont.)
16.00 – 16.15	Coffee Break
16.15 – 17.00	Presentations of Participants' Views (cont.)
17.00 – 17.30	Discussion
17.30	End of the first day

22 April 2016

9.00 – 11.00	Discussion & Identification of the main Research Issues
11.00 – 11.30	Coffee Break
11.30 – 13.00	Definition of a Research Agenda
13.00 – 14.30	Lunch
14.30 – 16.00	Discussion about the possibility of creating an RDA WG to address the issues identified by the Workshop
16.00 – 16.30	Conclusions

2.4 Participants

Computer scientists

1. **Buneman Peter** (University of Edinburgh)
2. **Corcho Oscar** (Universidad Politecnica de Madrid)
3. **D'Aquin Mathieu** (The Open University UK)
4. **Fensel Anna** (University of Innsbruck)
5. **Guarino Nicola** (ISTC-CNR)
6. **Klan Friederike** (University Jena)
7. **Kritikos Kyriakos** (ICS-FORTH)
8. **Mannens Eric** (University of Ghent)
9. **Toma Ioan** (Semantic Technologies Institute)
10. **Velegrakis Yannis** (University of Trento)

Research Data Infrastructures

- | | |
|--------------------------|---------------------------|
| 1. GEOSS | Nativi Stefano |
| 2. ELIXIR | Kristoffer Rapacki |
| 3. BiodiversityCatalogue | Niall Beard |
| 4. OpenAIRE | Manghi Paolo |

RDA Working Groups

- | | |
|-------------------------------------|----------------|
| 1. Metadata Standards Catalogue | Ball Alex |
| 2. Brokering Governance | Nativi Stefano |
| 3. Data Type Registries | Broeder Daan |
| 4. Metadata Standards Directory | Ball Alex |
| 5. RDA/WDS Publishing Data Services | Manghi Paolo |

Service Providers

- | | |
|----------|--------------|
| 1. EUDAT | Broeder Daan |
|----------|--------------|

2.5 Time Plan for Reporting and Actions

The plan is that the editors produce a report (titled “White Paper on Research Data Service Discoverability”) until **30 June 2016**. It will describe the research problems identified during the live discussions that need to be addressed and directions to be followed in order to solve them. The content of the report will be organized according to the following index:

1. Introduction
2. Definitions
3. Research Data Service in Context
4. The Research Data Service Discovery Process
5. Research Data Service Discovery: Enabling Technologies
 - 5.1. Knowledge Representation Languages
 - 5.2. Reasoning/Matchmaking Support
 - 5.3. Domain Specific Ontologies
 - 5.4. Research Data Service Metadata
 - 5.5. Mediation Support
 - 5.6. Digital Service Identifier
 - 5.7. Research Data Service Catalogues/Registries
6. Recommendations
7. References

Most of the reports have been completed, whilst work is in progress on sub-sections 5.1., 5.2., and 5.7. Section 6 will be finalized once all the other sections will be completed. Contributions and suggestions made by some of the participants are being integrated as well. Finally the report will be widely distributed via RDA channels to collect additional comments by the community.

An analysis and assessment of the TWM results will be performed by **30 September 2016**.

The main findings of the TWM will be introduced in the Atlas of Knowledge by **31 October 2016**.

Based on the agreed findings the participants will discuss whether they will start an RDA Working Group, an Interest Group or first organise a BoF session at the plenary in March 2017 in Barcelona.

3 Activities Planned

3.1 Activities Planned for the second year

A second TWM will be organized during the second year of the project.

The foreseen timing is the following:

- | | |
|--|-------------------|
| 1. Identification of the theme: | 30 September 2016 |
| 2. Submission of the proposal to the BoD | 31 October 2016 |
| 3. List of Experts | 31 January 2017 |
| 4. Date | April/May 2017 |

3.2 Activities Planned for the third year

A third TWM will be organized during the third year of the Project.

It is too early for producing a timing related to this event.

A final consideration concerns the topics to be addressed by the second and third TWMs. The choice of these topics will very much depend on the progress of the work carried out by the RDA WGs and IGs as the goal of the TWMs is to critically analyse the results obtained and help to optimize them.

In addition, the lessons learned from the organization of the first TWM will guide us in organizing the successive TWMs more efficiently and effectively.

4 Appendix

Automatic Location of Services

Uwe Keller, Ruben Lara, Holger Lausen, Axel Polleres, and Dieter Fensel: Digital Enterprise Research Institute (DERI) Innsbruck, Austria. Tecnologa, Informacion y Finanzas, Madrid, Spain.

Web Services Description Language (WSDL) 1.1

E. Christensen, F. Curbera, G. Meredith, and S. Weerawarana. <http://www.w3.org/TR/wsdl>, March 2001

The Web Service Modeling Framework WSMF.

D. Fensel and C. Bussler. *Electronic Commerce Research and Applications*, 1(2), 2002.

Description logics for matchmaking of services.

J. Gonzalez-Castillo, D. Trastour, and C. Bartolini. In *KI-2001 Workshop on Applications of Description Logics*, September 2001.

A Logical Framework for Web Service Discovery.

M. Kifer, R. Lara, A. Polleres, C. Zhao, U. Keller, H. Lausen, and D. Fensel. In *Semantic Web Services Workshop at ISWC, 2004*.

Semantics for Web Service Discovery and Composition

R. Lara, W. Binder, I. Constantinescu, D. Fensel, U. Keller, J. Pan, M. Pistore, A. Polleres, I. Toma, P. Traverso, and M. Zaremba. Technical report, *Knowledge Web*, December 2004

Web Service Modeling Ontology (WSMO)

H. Lausen, D. Roman, and U. Keller (editors). Working draft, DERI, March 2004.
<http://www.wsmo.org/2004/d2/v0.2/>

A Software Framework for Matchmaking Based on Semantic Web Technology

Lei Li and I. Horrocks. In *WWW'03*, Budapest, Hungary, May 2003.

The Description Logic Handbook

D. Nardi, F. Baader, D. Calvanese, D. L. McGuinness, and P. F. Patel-Schneider (eds.). Cambridge, Jan 2003.

Trust Negotiation for Semantic Web Services

D. Olmedilla, R. Lara, A. Polleres, and H. Lausen. In *SWSWPC Workshop at ICWS 2004*, July 2004.

Semantic Matching of Web Service Capabilities

M. Paolucci, T. Kawamura, T. Payne, and K. Sycara. In *ISWC*, pages 333–347. Springer Verlag, 2002.

A Conceptual Architecture for Semantic Web Services

Chris Preist. In *Proceedings of the International Semantic Web Conference 2004 (ISWC 2004)*, November 2004.

METEOR-S WSDI: A Scalable P2P Infrastructure of Registries for Semantic Publication and Discovery of Web Services.

K. Verma, K. Sivashanmugam, A. Sheth, and A. Patil. *Journal of Information Technology and Management*, 2004.

INSPIRE REGISTRY

Modeling Data Services (ORACLE Documentation), Book Chapter, *Datenbanken und Informationssysteme, Festschrift zum 60. Geburtstag von Gunter Schlageter*, Publication Hagen, October 2003-09-26

Metadata and Semantics for Web Services and Processes

Kaarthik Sivashanmugam, Amit Sheth, John Miller, Kunal Verma, Rohit Aggarwal, Preeda Rajasekaran. Large Scale Distributed Information Systems (LSDIS) Lab, University of Georgia, USA.

Web Services Metadata Exchange (WS-Metadata Exchange)

Discovery of Web Services in a Federated Registry Environment

UDDIe: An Extended Registry for Web Services

Ali ShaikhAli, Omer F. Rana, RashidAl-Ali, David W. Walker. Department of Computer Science Cardiff University, UK