RDA-pt: a Portuguese community in the Research Data Alliance

Cristina Ribeiro, Irene Rodrigues

The Research Data Alliance

- Community-driven organisation

  - Starts 2013, EU+ NSF+ NIST+ Austrália

- ~9.000 members, 137 countries (November 2019)

  - Free participation for individuals, organizations contribute with fee

- Organisation

  - Interest Groups (67): focused on specific area, launch WG for tasks

  - Working Groups (35): 18 months, solution and its dissemination

THE RESEARCH DATA ALLIANCE
www.rd-alliance.org

building the social and technical bridges that enable open sharing of data

32 FLAGSHIP OUTPUTS

of which 4 ICT Technical Specifications

75 ADOPTION CASES

across multiple disciplines, organisations & countries

104 GROUPS WORKING ON GLOBAL DATA INTEROPERABILITY CHALLENGES

of which 35 WORKING GROUPS & 67 INTEREST GROUPS

8,642 INDIVIDUAL MEMBERS FROM 137 COUNTRIES

67,9% Academia & Research
14,2% Public Administration
12,7% Enterprise & Industry

50 ORGANISATIONAL MEMBERS & 8 AFFILIATE MEMBERS

Vision
Researchers and innovators openly share data across technologies, disciplines, and countries to address the grand challenges of society.

Mission
RDA builds the social and technical bridges that **enable open sharing** of data.

What is RDA?

RDA is an international **member based organization** focused on the development of infrastructure and community activities that reduce barriers to data sharing and exchange, and the acceleration of data driven innovation worldwide.

With more than 8,600 members globally representing 137 countries, RDA includes **researchers, scientists and data science professionals** working in multiple disciplines, domains and thematic fields and from different types of organisations across the globe.

RDA is building the social and technical bridges that enable open sharing of data to achieve its vision of researchers and innovators openly sharing data across technologies, disciplines, and countries to address the grand challenges of society.

RDA-Europe: European plugin to the RDA

- Pioneer nodes in RDA-4
  - **Greece** Athena Research Center
  - **UK** STFC (Science and Technology Facilities Council)
  - **Spain** BSC (Barcelona Supercomputing Centre)
  - **Italy** CNR-ISTI (National Research Council)
  - **Netherlands** DANS (Data Archiving and Networked Services)
  - **Finland** CSC (Tieteen Tietotekniikan Keskus)
  - **Germany** Max Planck Institute
  - **France** CNRS (Centre National de la Recherche Scientifique)
  - **Ireland** National Library of Ireland

RDA-Europe: Build a network of nodes

- 1st call August 2018, 4 new nodes
  - **Austria** RDA Austria

  - **Denmark** DM Forum (National Forum for Research Data Management)

  - **Portugal** INESC TEC (leading a group of 8 institutions)

  - **Slovenia** ADP (Social Science Data Archives)

RDA-Europe: Build a network of nodes

- 2nd call April 2019, 6 new nodes
  - **Croatia** SRCE (University of Zagreb Computing Center)
  - **Czech Republic** Plan4all
  - **Hungary** HRDA
  - **Lithuania** *Group of institutions*
  - **Norway** NSD (leading a group of 7 institutions)
  - **Sweden** SND (Swedish National Data)

RDA-Europe: Build a network of nodes

- 3rd call September 2019
    - **Bulgaria** Sofia University St. Kliment Ohridski
    - **Estonia** University of Tartu
    - **Romania** Exec. Agency for Higher Ed., R & D & I Funding

- Accomplished goal of RDA-4: network of **22 nodes**
    - 9 pioneer, 4 first call, 6 second call, 3 third call

Final
22 nodes

The RDA Europe National Node Network

RDA-pt   RDA in Portugal

- **Vision**
  - Researchers, institutions and FCT define data management practices, choose technologies and use a network infrastructure for data access, preservation and sharing

- **Mission**
  - RDA-pt builds a community with interest in open data; disseminates, proposes technical solutions and surveys case studies in multiple domains

RDA activity in Portugal before RDA-pt

- Participation in Interest Groups/ Working Groups
  - 58 people registered
  - Contributions to documents and presentations
- Responsibility in Groups
  - *Repository Platforms for Research Data*: João Rocha is a co-chair
  - *RDA in Portugal*: Cristina Ribeiro, Irene Rodrigues coordinators
- RDA plenaries: 19 participations
  - Including various contributions: special sessions in projects, session organization, presentations in sessions, posters

RDA-pt key personnel

Cristina Ribeiro (INESC TEC, University of Porto)

Irene Rodrigues (University of Évora)

Maria Manuel Borges (University of Coimbra)

José Borbinha (Instituto Superior Técnico, University of Lisbon)

João Miranda (Polytechnic Institute of Portalegre)

Pedro Moura Ferreira (ICS, University of Lisbon)

Eloy Rodrigues (University of Minho)

Francisco Barbedo (DGLAB, Portuguese National Archives)

RDA-pt   SWOT analysis

- Strengths
  - Large experience in Open Access (RCAAP)
  - Strong international connections
  - Motivated community
- Weaknesses
  - Few people
  - Ambiguity in funding mechanisms
  - Little visibility of the national policies in this area

RDA-pt   SWOT analysis

- Opportunities
  - European mandates (strong) and FCT mandates
  - European funding
  - Growing awareness in disciplinary groups
- Threats
  - Poor alignment of incentives and penalties
  - "Band-aid" solutions compromise more solid ones
  - "Buzzwords" create saturation

RDA-pt planned activities

	pre-contract				Part 1								Part 2							
	2018			2019									2020							
MONTH	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun	Jul	Ago	Set	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May
EVENT		E1						E2	E3				E4	E5				E6		
PLENARY		P12					P13						P14						P15	

E1	RDA-pt presentation in Forum GDI, national RDM forum
E2	(in disciplinary event) RDA-pt co-located event
E3	(in disciplinary event) RDA-pt co-located event
E4	Organisation of co-located RDA-pt event in ConfOA (Brazil)
E5	(in disciplinary event) RDA-pt co-located event
E6	(in disciplinary event) RDA-pt co-located event
(P12)	Plenary 12, no plan by RDA-pt
P13	Plenary 13, Philadelphia, RDA-pt participation
P14	Plenary 14, Helsinki, RDA-pt participation
P15	Plenary 15, ??, RDA-pt participation

RDA-pt ongoing activities

	pre-contract			Part 1											Part 2				
	2018			2019											2020				
MONTH	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun	Jul	Ago	Set	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May
EVENT		E1						E2	E3	E4		E5	E6	E7				E8		
PLENARY		P12					P13							P14					P15	

E1	RDA-pt presentation in Forum GDI, national RDM forum		
E2a, E2b, E2c	Participation in disciplinary events	Porto4Ageing	C4G Coordination Council	WS Comp. Data Analysis Num. Methods
E3, E4	RDA-pt participation in disciplinary event+ National Science	Biodata.PT- ELIXIR	Ciência 2019
E5	Organisation of co-located RDA-pt event in OSF	Open Science Fair: RDA-pt The Portuguese Node of the Research Data Alliance
E6	Organisation of co-located RDA-pt event in ConfOA (Brazil)	WS at the Portuguese-Brazilian Conference on Open Access
E7	(in disciplinary event) RDA-pt co-located event	Session at 5th Forum GDI (national RDM meeting)
E8	(in disciplinary event) RDA-pt co-located event	

P13	Plenary 13, Philadelphia, RDA-pt participation	João Rocha da Silva, RDA-pt participation	Joana Rodrigues, RDA Early Career Grant
P14	Plenary 14, Helsinki, RDA-pt participation	Cristina Ribeiro, Irene Rodrigues, RDA node	4 RDA-pt grants + 2 RDA Early Career Grants
P15	Plenary 15, Melbourne, RDA-pt participation		

The RDA Plenaries

- Meetings where IG and WG present work

    - IG and WG propose sessions

    - New groups can test their proposal and recruit people

    - Twice a year, see previous and upcoming

- P14: 10 Portuguese participants

    - 2 Early Career grants: Yulia Karimova, João Cardoso
    - 4 RDA-pt Grants: André Pacheco, Angela Lomba, Helena Amaral, Susana Barbosa
    - Cristina Ribeiro, Irene Rodrigues, Pedro Príncipe, Mário J. Silva

RDA-pt@P14: Yulia Karimova (ECR)

- Researcher at INESC TEC, PhD student at FEUP

  - Research data management, Data Management Plan (creation and monitoring)

- DMP Common Standards WG

  - "The collaborative method between curators and researchers in the preparation of a Data Management Plan (DMP) and Privacy Impact Assessment (PIA)"

- IG Repository Platforms for Research Data, IG RDA/WDS Certification of Digital Repositories

  - rdm.inesctec.pt

RDA-pt@P14: João Cardoso (ECR)

- PhD Thesis on **Machine-Actionable DMPs**
  - **Semantic** representation and value

- DMP Common Standards WG
  - Long time collaborator

- DCS Ontology
  - An ontology to **represent** the **DMP Common Standard Model**.

RDA-pt@P14: Angela Lomba (RDA-pt grant)

- Researcher Interface between Biological and Agricultural Sciences

  - Social-ecological systems underlying agricultural ecosystems

  - Impacts of environmental change on agricultural landscapes

- IG Meeting: Agricultural Data: Open Session

  - Overview of the aims and activities of the group

  - Discussion on key activities related to agricultural data

RDA-pt@P14: Angela Lomba (RDA-pt grant)

- IG Meeting: ESIP/RDA Earth, Space and Environmental Sciences

    - ESES data infrastructures & data cubes (EarthCube);

RDA-pt@P14: André Pacheco (RDA-pt grant)

- PhD Thesis on **Information Science @ Univ. Coimbra**

  - **Authenticity** in archival description standards

- IG: RDA/WDS Certification of Digital Repositories

  - And many metadata/standards/vocabularies others...

RDA-pt@P14: Helena Amaral (RDA-pt grant)

- Technical Director of C4G Collaboratorium for Geosciences Research Infrastructure

  - Responsible for the implementation of Services and Products

- Get acquainted with the several IG that relate to C4G's topics: Geospatial, Global Water, Earth, Space, and Environ. Sci.

- Relate to other IG that cover DMPs, Data Policy, Standardization and Data Repositories

RDA-pt@P14: Susana Barbosa (RDA-pt grant)

- Researcher, Earth system science

  - Climate / Data Science / Space-earth interactions

  - Environmental monitoring (marine / atmospheric data)

- Meeting BoF: Curation for FAIR and Reproducible Data and Code

  - Reproducible scientific code (not only data)

RDA-pt@P14: Susana Barbosa (RDA-pt grant)

- Meeting BoF: Management of Computational Notebooks
  - Citation of notebooks
  - Long-term preservation
  - Integration with data sources
- Meeting Interest Group "ESIP/RDA Earth, Space, and Environmental Sciences"

RDA-pt goal: Build a community

- More participation in RDA global

  - Number of members

  - Contributions to Interest Groups/ Working Groups

- Co-organization of events

  - RDA presentations in events: disciplinary and open science

- Open calls for plenary participation

RDA-pt goal: Liaise with infrastructures

- Alignment with international infrastructures

  - ELIXIR (U. Lisboa)

  - EPOS (U. Évora, INESC TEC, P. Portalegre)

  - CESSDA (ICS)

  - CLARIN (U. Évora)

  - LifeWatch (INESC TEC)

- Entry point: infrastructures of the Portuguese roadmap

RDA-pt goal: collaborate Europe + Brazil

- European nodes
  - Take advantage of the network and activities in RDA-4
  - Learn from extensive experience of pioneer nodes
- Brazil
  - RDA in Brazil starting
  - Joint RDA-pt and RDA in Brazil workshop on October 4, 2019, part of CONFOA, the Portuguese-Brazilian Conference on Open Access

RDA Recommendations and outputs

RDA Outputs are the technical & social infrastructure solutions **developed by RDA Working Groups or Interest Groups** that enable data sharing, exchange, and interoperability.

➜ Solving data sharing problems
➜ Adoption in infrastructure environments by individuals, projects and organisations
➜ Improve global data sharing, exchange and interoperability

https://www.rd-alliance.org/recommendations-and-outputs/all-recommendations-and-outputs

RDA-pt goal: Adopt recommendations

RDA works on raising awareness and expanding the adoption of these outputs across all regions of the world.

RDA-pt goal: Adopt recommendations
RDA works on raising awareness and expanding the adoption of these outputs across all regions of the world.

>>> action plan for PP >>> think, check, implement

RECOMMENDATION	GROUP	IMPACT	FORESEEN ADOPTION	USE CASE	IMPLEMENTATION ACTIONS

Scalable dynamic-data Citation Methodology

The Challenge:
Supporting accurate citation of data subjected to change, for the efficient processing of data and linking from publications.

What is the solution?
The Dynamic Data Citation Working Group developed a simple, scalable mechanism that allows the precise, machine-actionable identification of arbitrary sub selections of data at a given point in time irrespective of any subsequent addition, deletion or modification.

Produced by: Data Citation (WGDC) WG
https://rd-alliance.org/groups/data-citation-wg.html

What is the impact?
The main impact is on the reproducibility of work: it allows databases to be dynamically updated as needed, while users can still retrieve the data as it was at a specific point in time, providing valuable semantic information on the way the specific dataset was constructed.
As data gets migrated to new representations, the queries to extract the specific cited subset can be migrated accordingly, ensuring stability across technological changes enhancing data interoperability, sharing and re-use.

Find out more about the Data Citation WG Recommendation

Persistent Identifier Type Registry

The Challenge:
Defining standard core PID information types to enable simplified verification of data identity and integrity

What is the solution?
The PID Information Types Working Group identified the essential types of information associated with persistent identifiers, defined and registered a number of core persistent identifiers (PIDs) information types. The working group developed a conceptual model for structuring the typed information, an application programming interface to access to typed information and a demonstrator implementing the interface.

Produced by: PID Information Types (PIT) WG
https://rd-alliance.org/groups/pid-information-types-wg.html

What is the impact?
In complex data domains, unique and persistent identifiers (PIDs) are at the core of proper data management and access, providing a unified access method to all PID service users and the automatic processes to deal with heterogeneous data allowing them to fully exploit the huge amount of data available. Registered PID Information Types are one of the fundamental building blocks to enable semantic interoperability of data.

Developers are able use a simple API to deal with PID related information, such as checksum verification, to assess data identity and integrity, thus enabling software development to become more efficient.

Repository Audit and Certification Catalogues

The Challenge:
Creating harmonized Common Procedures for certification of repositories

What is the solution?
Drawing from the procedures already put in place by the Data Seal of Approval (DSA) and the ICSU World Data System (ICSU-WDS), the RDA Repository Audit and Certification DSA-WDS Partnership Workin Group has produced a convergent DSA-WDS certification standard aiming to eliminate duplication of effort, increase certification procedure coherence and compatibility thus benefitting researchers, data managers, librarians and scientific communities.

Produced by: The Repository Audit and Certification DSA-WDS Partnership WG
https://rd-alliance.org/groups/rdawds-certification-digital-repositories-ig.html

What is the impact?
The convergent DSA-WDS certification standard marks a step towards having more coherent, increasingly stringent and compatible standards for repository certification.
The harmonization of repository certification criteria will increase adoption and the number of certified repositories building trust for data generators, data consumers and funding bodies.

Find out more about the PIT WG Recommendation

Find out more about the Repository Audit and Certification DSA-WDS Partnership WG Recommendation

RDA-pt goal: Adopt recommendations

Repository Audit and Certification Catalogues

The Challenge:

Creating harmonized Common Procedures for certification of repositories

What is the solution?

Drawing from the procedures already put in place by the Data Seal of Approval (DSA) and the ICSU World Data System (ICSU-WDS), the RDA Repository Audit and Certification DSA-WDS Partnership Workin Group has produced a convergent DSA-WDS certification standard aiming to eliminate duplication of effort, increase certification procedure coherence and compatibility thus benefitting researchers, data managers, librarians and scientific communities.

Produced by: The Repository Audit and Certification DSA-WDS Partnership WG

https://rd-alliance.org/groups/
rdawds-certification-digital-repositories-ig.html

What is the impact?

The convergent DSA-WDS certification standard marks a step towards having more coherent, increasingly stringent and compatible standards for repository certification.
The harmonization of repository certification criteria will increase adoption and the number of certified repositories building trust for data generators, data consumers and funding bodies.

Find out more about the Repository Audit and Certification DSA-WDS Partnership WG Recommendation

RDA-pt goal: Adopt recommendations

Persistent Identifier Type Registry

The Challenge:
Defining standard core PID information types to enable simplified verification of data identity and integrity

What is the solution?
The PID Information Types Working Group identified the essential types of information associated with persistent identifiers, defined and registered a number of core persistent identifiers (PIDs) information types. The working group developed a conceptual model for structuring the typed information, an application programming interface to access to typed information and a demonstrator implementing the interface.

Produced by: PID Information Types (PIT) WG

https://rd-alliance.org/groups/
pid-information-types-wg.html

What is the impact?

In complex data domains, unique and persistent identifiers (PIDs) are at the core of proper data management and access, providing a unified access method to all PID service users and the automatic processes to deal with heterogeneous data allowing them to fully exploit the huge amount of data available. Registered PID Information Types are one of the fundamental building blocks to enable semantic interoperability of data.

Developers are able use a simple API to deal with PID related information, such as checksum verification, to assess data identity and integrity, thus enabling software development to become more efficient.

Find out more about the PIT WG Recommendation

Metadata Standards Directory

The Challenge:
Enabling discovery of metadata standards for documenting research data, regardless of academic discipline, and addresses issues related to coverage, ease of maintenance and sustainability.

What is the solution?
The Metadata Standards Directory Working Group adopted, enriched, & expanded the Disciplinary Metadata Standards Catalogue set up by the UK Digital Curation Centre (DCC), and developed a functional prototype directory based around the GitHub infrastructure, that places the information from the DCC directory into an environment where it can be maintained transparently with full version control.

Produced by: Metadata Standards Directory (MSD) WG

https://rd-alliance.org/groups/metadata
-standards-directory-working-group.html

What is the impact?
The Metadata Standards Catalogue was built to guide researchers towards the metadata standards and tools relevant for their discipline. The directory drives up adoption of those standards, improving the chances of future researchers finding, accessing, and reusing the associated data.
By raising awareness of existing standards, the directory reduces the proliferation of ad hoc metadata formats and helps direct future standards development efforts towards those areas that most need it.

Find out more about the MSD WG Recommendation

RDA-pt goal: Adopt recommendations

RDA-pt goal: Adopt recommendations

Group	Recommendation - Short description & Impact	Recommendation details	Adopters
Repository Audit and Certification / DSA–WDS Partnership	**Repository Audit and Certification Catalogues** Creates harmonized Common Procedures for certification of repositories at the basic level, drawing from the procedures already put in place by the Data Seal of Approval (DSA) and the ICSU World Data System (ICSUWDS)	Recommendation page: https://www.rd-alliance.org/group/repository-audit-and-certification-dsa-wds-partnership-wg/outcomes/dsa-wds-partnership	• CLARIN • International Oceanographic Data and Information Exchange (IODE) programme • CESSDA SaW Project
Research Data Repository Interoperability WG	**Research Data Repository Interoperability WG Final Recommendations** Provides recommendations with respect to an interoperable packaging and exchange format for digital content.	Recommendation page: https://www.rd-alliance.org/group/research-data-repository-interoperabil... DOI: http://dx.doi.org/10.15497/RDA00025	• Fedora Commons • DARIAH-DE Repository • ICat Project • KIT Data Manager

rd-alliance.org/groups/rda-portugal

RDA-pt goal: Adopt recommendations

Title	Group	Impact
Persistent Identifier Type Registry	PID Information Types Working Group	Defines standard core PID information types to enable simplified verification of data identity and integrity
Scalable Dynamic-data Citation Methodology	Data Citation Working Group	Supporting accurate citation of data subjected to change, for the efficient processing of data and linking from publications.

https://www.rd-alliance.org/recommendations-and-outputs/all-recommendations-and-outputs

RDA-pt goal: Adopt recommendations

- Adoption stories
  - [Check them all](#)
- Example: Agriculture
  - Data standards for Agriculture: [INRA adoption of RDA Outputs](#)
    - French National Institute for Agricultural Research
  - Adopted recommendations
    - Wheat Data Interoperability Guidelines
    - The FAIRsharing Registry and Recommendations
    - Data Citation of Evolving data
- RDA actively searching for adoption stories

RDA-pt goal: contribute to sustainability

- RDA foundation seeks sustainability model

- Regional structure

  - RDA Regional Engagement Group

    - Ongoing discussion

    - A basis for a model of contributions (PIB, members, region dimensions)

- Sustainability in RDA-pt

  - Alignment with Open Science goals in FCT

How can you participate in RDA?

- Join as a member
  - No cost, access to all groups and results
  - Sign up for Interest Groups and Working Groups
- Join RDA in Portugal
  - Highlight the presence of the Portuguese community in RDA
  - 51 of the 150 Portuguese members of RDA have signed up for the node

How can you participate in RDA-pt?

- Promote activities and suggest actions

  - Related to open science and research data

  - Disseminate actions in disciplinary events

- Connect to groups/ institutions in charge of data

  - Disseminate node activities

  - Look up and recommend Interest Groups and Working Groups

Opportunities & Grants

RDA Europe grants

- ○ Participate in plenaries: Early Career Programme (for individuals)

- ○ Participate in plenaries: Experts (for individuals)

- ○ Adoption Grants (for organisations adopting recommendations)

- ○ Ambassadors (for individuals)

- ○ Check out the the open calls: grants.rd-alliance.org