Research Data Driving Solutions to Complex Scientific and Societal Challenges

Who is most at risk to contract asthma?

How can we increase wheat yields?

How accurate is the Standard Model of Physics?

How can we best address energy needs and sustain the environment?
Data Infrastructure Needed to Support Data-Driven Research

- Research Dissemination and Reproducibility
- Data Use and Re-use
- Data Access (now) and Preservation (later)
- Data Discovery and Data Sharing

- Data Access and Distribution Policy
- Institutional Data Sharing Practice
- Data Discovery Tools
- Data Citation Standards

- Digital Object Identifiers
- Common Metadata Standards
- Data Scientists and Expert Support

- Data Preservation Practice
- Data Analytics Algorithms
- Curation Practice and Policy
- Sustainable Economic Models

- Auditing, Certification and Reporting Practice
World-wide Efforts Focusing on Infrastructure to Support Research Data Sharing, Access, Use

Science, Humanities, Arts Communities

E-Infrastructure professionals, data analysts, data center staff, ...

Libraries, Archives, Repositories, Museums

National Data Sharing and Accessibility Policy-2012 (NDSAP-2012)

Data Scientists

Australian National Data Service

Our Vision: More Australian researchers reusing research data more often

ANDS is enabling the transformation of:

Data that are: to: Structured Collections that are:

Unmanaged Managed
Disconnected Connected
Invisible Findable

E-Infrastructure professionals, data analysts, data center staff, ...

Department of Science & Technology
Ministry of Science & Technology
Government of India

A Europe-Japan-United States GNSS data-sharing pilot project
for the Geohazard Supersites and Natural Laboratories

Falk Anselmi, University of Miami, USA (GEO task lead)
Craig Dobson, NASA and Committee of Earth Observation Satellites (CEOS)
Rui Fernandez, EROS and EUREF <rmanuel@di.ubi.pt>
RDA community focuses on building **social, organizational and technical infrastructure** to

- reduce barriers to data sharing and exchange
- accelerate the development of coordinated global data infrastructure

CREATE → ADOPT → USE

RDA Working Group Infrastructure Deliverables are:

- **Focused pieces of adopted code, policy, infrastructure, standards, or best practices** that enable data to be shared and exchanged
- **“Harvestable” efforts** for which 12-18 months of work can eliminate a roadblock for a substantial community
- **Efforts that have substantive applicability** to “chunks” of the data community, but may not apply to everyone
- **Efforts for which working scientists and researchers can start today** while more long-term or far-reaching solutions are appropriately discussed in other venues
The RDA Community Today: Over 2300 members from 96 countries (as of 9/14)

Map courtesy traveltip.org
RDA Interest (IG) and Working Groups (WG) by Focus 1 (as of 9/14)

**Domain Science - focused**
- Toxicogenomics Interoperability IG
- Structural Biology IG
- Biodiversity Data Integration IG
- Agricultural Data Interoperability IG
- Wheat Data Interoperability WG
- Digital Practices in History and Ethnography IG
- Geospatial IG
- Marine Data Harmonization IG
- Metabolomics IG
- RDA/CODATA Materials Data Infrastructure and Interoperability IG
- Research Data Needs of the Photon and Neutron Science Community IG
- Defining Urban Data Exchange for Science IG*
- The BioSharing Registry: Connecting data policies, standards and databases in the life sciences WG*
- Urban Quality of Life Indicators WG*

**Community Needs - focused**
- Community Capability Model IG
- Engagement IG
- RDA / CODATA Summer Schools in Data Science and Cloud Computing in the Developing World WG*
- Development of Cloud Computing Capacity and Education in Developing World Research IG
- Data for Development IG
- Education and Training on handling of research data IG
# RDA Interest (IG) and Working Groups (WG) by Focus 2 (as of 9/14)

**Reference and Sharing - focused**
- Data Citation WG
- Standardization of Data Categories and Codes WG
- RDA/CODATA Legal Interoperability IG
- Reproducibility IG*
- Data Description Registry Interoperability Working Group
- RDA / WDS Publishing Data Bibliometrics WG

**Data Stewardship and Services - focused**
- Research Data Provenance IG
- Preservation e-infrastructure IG
- RDA / WDS Publishing Data Services WG
- RDA / WDS Publishing Data Workflows WG
- Long-tail of Research Data IG
- RDA/WDS Publishing Data IG
- RDA/WDS Repository Audit and Certification WG
- Domain Repositories Interest Group
- Brokering Interest Group
- ELIXIR Bridging Force IG*
- Libraries for Research Data IG*
- RDA / WDS Certification of Digital Repositories IG
- RDA / WDS Publishing Data Cost Recovery for Data Centres IG

**Base Infrastructure - focused**
- Data Foundation and Terminology WG
- Metadata Standards Directory WG
- Practical Policy WG
- PID Information Types WG
- Data Type Registries WG
- Data in Context IG
- Big Data Analytics IG
- Data Brokering WG*
- Federated Identity Management IG
- Metadata IG
- PID Interest Group
- Service Management IG
- Data Fabric IG
Precipitous Growth

First RDA organizational telecon: August 2012

Global Data Planning Meeting: October 2012

RDA Launch / First Plenary

First Working Groups and Interest Groups

240 participants

First “neutral space” community meeting (Data Citation Summit)

First Organizational Partner Meet-up

First BOFs

380 participants from 22 countries

RDA Second Plenary

RDA Third Plenary

RDA Fourth Plenary

First Organizational Assembly

6 co-located events

14 BOF, 12 Working Groups, 22 Interest Groups

497 participants

RDA Plenary 1 / Launch
Gothenburg, Sweden

RDA Plenary 2
Washington, D.C.

RDA Plenary 3
Dublin, Ireland

RDA Plenary 4
Amsterdam

First Working Group exchange meeting
Who Attended Plenary 3?
March 26-28, 2014
Dublin, Ireland

<table>
<thead>
<tr>
<th>Professional Title</th>
<th>Total</th>
<th>% (rounded)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Researcher</td>
<td>87</td>
<td>18%</td>
</tr>
<tr>
<td>Professor</td>
<td>42</td>
<td>8%</td>
</tr>
<tr>
<td>Student</td>
<td>38</td>
<td>8%</td>
</tr>
<tr>
<td>IT Specialist / IT Architect</td>
<td>53</td>
<td>11%</td>
</tr>
<tr>
<td>CTO / IT Director</td>
<td>20</td>
<td>4%</td>
</tr>
<tr>
<td>Librarian</td>
<td>27</td>
<td>5%</td>
</tr>
<tr>
<td>Programme Manager / Project Manager</td>
<td>62</td>
<td>12%</td>
</tr>
<tr>
<td>CEO / Managing Director / Chief Executive</td>
<td>35</td>
<td>7%</td>
</tr>
<tr>
<td>Advisor/Consultant</td>
<td>22</td>
<td>4%</td>
</tr>
<tr>
<td>Policy Development Manager / Policy Consultant</td>
<td>12</td>
<td>2%</td>
</tr>
<tr>
<td>Journalist / Editor / Copywriter</td>
<td>6</td>
<td>1%</td>
</tr>
<tr>
<td>Other</td>
<td>93</td>
<td>19%</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>497</td>
<td><strong>100%</strong></td>
</tr>
</tbody>
</table>
Organizations Committed to Joining RDA

### Organizational Members:
- Alliance for Permanent Access
- American University Library
- Australian National Data Service
- Barcelona Supercomputing Center - Centro Nacional de Supercomputación
- Columbia University Library
- CNRI
- CSC
- Digital Curation Center
- EIROForum IT Working Group
- eResearch Services and Scholarly Application Development Division of Information Services, Griffith University
- European Data Infrastructure (EUDAT)
- National Institute of Advanced Industrial Science and Technology (AIST), Japan
- International Association of STM Publishers
- Internet2
- Microsoft Research
- NZ eScience Infrastructure
- Purdue University Libraries
- Research Data Canada
- Scholarly Publishing and Academic Resources Coalition (SPARC)
- Washington University in St. Louis Libraries
- Science and Technology Facilities Council

### Affiliates
- CODATA
- ICSU World Data System
- ORCID
- DataCite
- Global Alliance for Genomics and Health
- CASRAI
RDA/US Goals:

- Contribute to RDA “international” efforts and leadership
- Bring US efforts to broader RDA community
- Build the RDA community within the US
- Leverage and implement RDA deliverables in the US to amplify impact
- Collaborate closely with other RDA “regions” on key programs and initiatives

NSF-supported RDA/US initiatives:
- Outreach (RDA → RDA/US)
- RDA Deliverables Amplification
- Student / Early Career Engagement

RDA/US Steering Committee
- Fran Berman, RPI
- Kathy Fontaine, RPI
- Larry Lannom, CNRI
- Beth Plale, IU
RDA/US Community Activities

- **FY14/15:**
  - Individual outreach activities with Earth Science Information Partners, SC Conference, National Data Stewardship Alliance, American Geophysical Union, CENDI, EarthCube, NIST, etc.
  - Joint workshops with CAMP-4-DATA, National Data Service, ICPSR Repository Group
  - Over a dozen RDA/US Student Interns and Fellows working with RDA Interest and Working Groups

- **March 9-11, 2015: RDA/US hosting RDA Plenary 5 in San Diego**
  - Working Meeting of the RDA with co-located community meetings
  - SDSC hosting RDA Adoption Day on March 8
Next Steps for the RDA

More Infrastructure
- Continuing pipeline of infrastructure deliverables adopted and used to accelerate data sharing
- Increasing coordination of infrastructure

Effective Community
- Increasing cross-boundary collaborations between domains, sectors, organizations

Synergistic Programs
- International and regional programs focusing on workforce, outreach, expansion of infrastructure impact

Partnership with Industry
- New partners in the Organizational Assembly
- Focused strategy to support development of industry infrastructure for data sharing
Thank you!

research data sharing without barriers
rd-alliance.org