

Developing infrastructure for data sharing around the world

March 10 2015, by Aaron Dubrow



RDA's Wheat Data Interoperability working group is comprised of members from the French National Institute for Agricultural Research, the International Maize and Wheat Improvement Center and other agriculture-related organizations. The group's objective is to build an integrated Wheat information system for the international community of wheat researchers, growers and breeders. With approximately three-quarters of all U.S. grain products made from wheat flour, advancing and sustaining wheat-related science is critical and finding ways to improve the sharing of data is an important first step. Credit: Thinkstock

How can we support agricultural productivity around the world? How can we develop public health models that leverage social data, health data and environmental data? What are best practices to ensure the stewardship of research data today and tomorrow?

Solutions to these and other critical challenges are being advanced through the sharing and exchange of [research data](#). To increase data sharing and overcome the critical challenges associated with making data accessible, an international group of leaders in the data community joined together in 2013 to form the Research Data Alliance (RDA).

With support from the U.S. National Science Foundation (NSF), the European Commission and the Australian government, RDA has grown in just two years from a core group of committed agencies to a community that now comprises more than 2,600 members from more than 90 countries, all dedicated to pragmatically removing the barriers to data sharing and raising awareness of those challenges among regions, disciplines, and professions.

NSF supports U.S. participation in RDA as part of a grant to promote coordination and develop infrastructure for data sharing.

Twice a year, RDA members meet face to face at plenary meetings held in various locations worldwide to coordinate activities and advance their efforts. At these meetings, researchers, policymakers and representatives from funding agencies speak on pressing data issues, and members meet to collaborate on projects through interest and working groups.

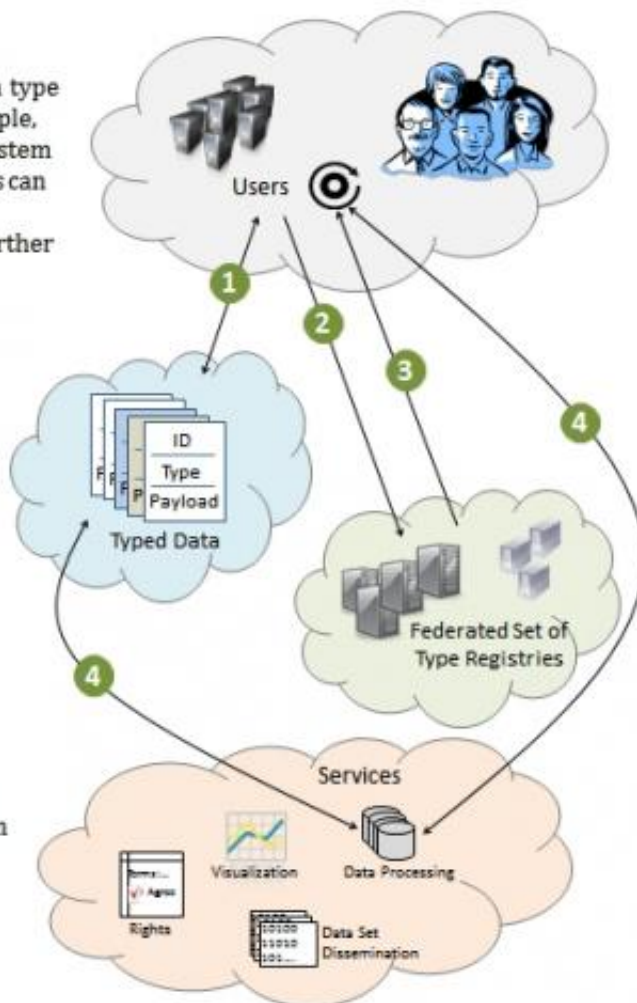
RDA holds its 5th Plenary through March 11 in San Diego, hosted by the U.S. members of RDA. The event featured an "Adoption Day", focusing on the use of new RDA-developed products and guidelines in various domains.

1 Users (processors or people) encounter data of an unknown type as part of an id/type/value triple, e.g., an identifier resolution system or a data repository. Registries can explicate those types and/or provide structure to enable further processing of the data.

2 Users query Type Registries with the unknown types.

3 Type Registries response (model under discussion) includes type definitions, relationships, properties and, potentially, pointers to relevant services or software that can be used to interpret or process the typed data or used to request an external service.

4 Optionally, the typed data or a reference to the typed data can be sent to a type-appropriate service or application to be rendered or processed as needed.



A good example of the infrastructure RDA is developing is the Data Type Registries, an effort to make it easier to create machine-readable and researcher-accessible data by creating an archive of common data structures that researchers can turn to when deciding how to organize their data. The creation of such a registry will support the accurate use of data to reproduce experiments, confirm findings and interoperate among data sets. Formed at the first RDA Plenary in early 2013, the Data Type Registries working group has collaborated during the last year to develop and test its new system. The infrastructure products of this group are already being adopted by European Data Infrastructure, the National Institute of Standards and Technology in the United States and other groups who are applying it to their own research activities. Credit: Daan Broeder, Max Planck Institute for Psycholinguistics

"RDA continues to experience tremendous growth in response to global interest," said Bob Chaddock, a program officer at the National Science Foundation. "RDA-developed tools will have a tremendous impact throughout science, and the plenary provides a place where interested communities from around our world get an opportunity to test-drive the tools."

RDA members collaborate to develop, coordinate and adopt data sharing infrastructure, addressing a broad spectrum of challenges. RDA's working and interest groups design and implement specific tools, recommendations or products within a 12 to 18 month time frame, and these products are adopted and used by other organizations and communities within the alliance. Leveraging diverse perspectives, these groups tackle data sharing challenges pertaining to interoperability, stewardship, sustainability, policy and use.

"Impact is a primary focus for RDA," said Fran Berman, chair of RDA/U.S. "In only two years, RDA has begun fulfilling its mission to build the social and technical bridges that enable the open sharing of data. It's exciting to see the start of a pipeline of adopted infrastructure efforts that will accelerate data sharing and data-driven innovation."

A good example of such infrastructure is RDA's Data Type Registries. The registries make it easier to create machine-readable and researcher-accessible data by designing an archive of common data structures that researchers can turn to when deciding how to organize their data. The creation of such a registry will support the accurate use of data to reproduce experiments, confirm findings and interoperate among data sets.

Formed at the first RDA Plenary in early 2013, the Data Type Registries working group has collaborated during the last year to develop and test its new system. The infrastructure products of this group are already

being adopted by European Data Infrastructure (EUDAT), the National Institute of Standards and Technology in the U.S., and other groups who are applying it to their own research activities.

Another effort underway, which is still in its early stages, is RDA's Wheat Data Interoperability working group. Comprised of members from the French National Institute for Agricultural Research, the International Maize and Wheat Improvement Center, and other agriculture-related organizations, the group's objective is to build an integrated wheat information system for the international community of wheat researchers, growers and breeders. With approximately three-quarters of all U.S. grain products made from wheat flour, advancing and sustaining wheat-related science is critical. Finding ways to improve the sharing of data is an important first step.

As RDA enters its second year, its community of data researchers continues to grow. The organization is working closely with countries, communities and agencies to expand the alliance to include new participants. These include partners in Japan, Brazil, Canada and South Africa, and U.S. projects and organizations such as CENDI (Commerce, Energy, NASA, Defense Information Managers Group), the National Data Service, EarthCube and the Sustaining Digital Repositories group.

Throughout its expansion, the alliance's focus will remain on the development of products that promote data sharing and exchange and the establishment of diverse collaborations.

"With its tremendous success in the first two years, its growing reputation as a gathering place for the global research [data](#) community and its targeted focus on impact and infrastructure, RDA is capitalizing on its momentum to reach a broader community, and fulfill its goal of research sharing without barriers," Berman said.

Provided by National Science Foundation

Citation: Developing infrastructure for data sharing around the world (2015, March 10) retrieved 20 April 2024 from <https://phys.org/news/2015-03-infrastructure-world.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.