



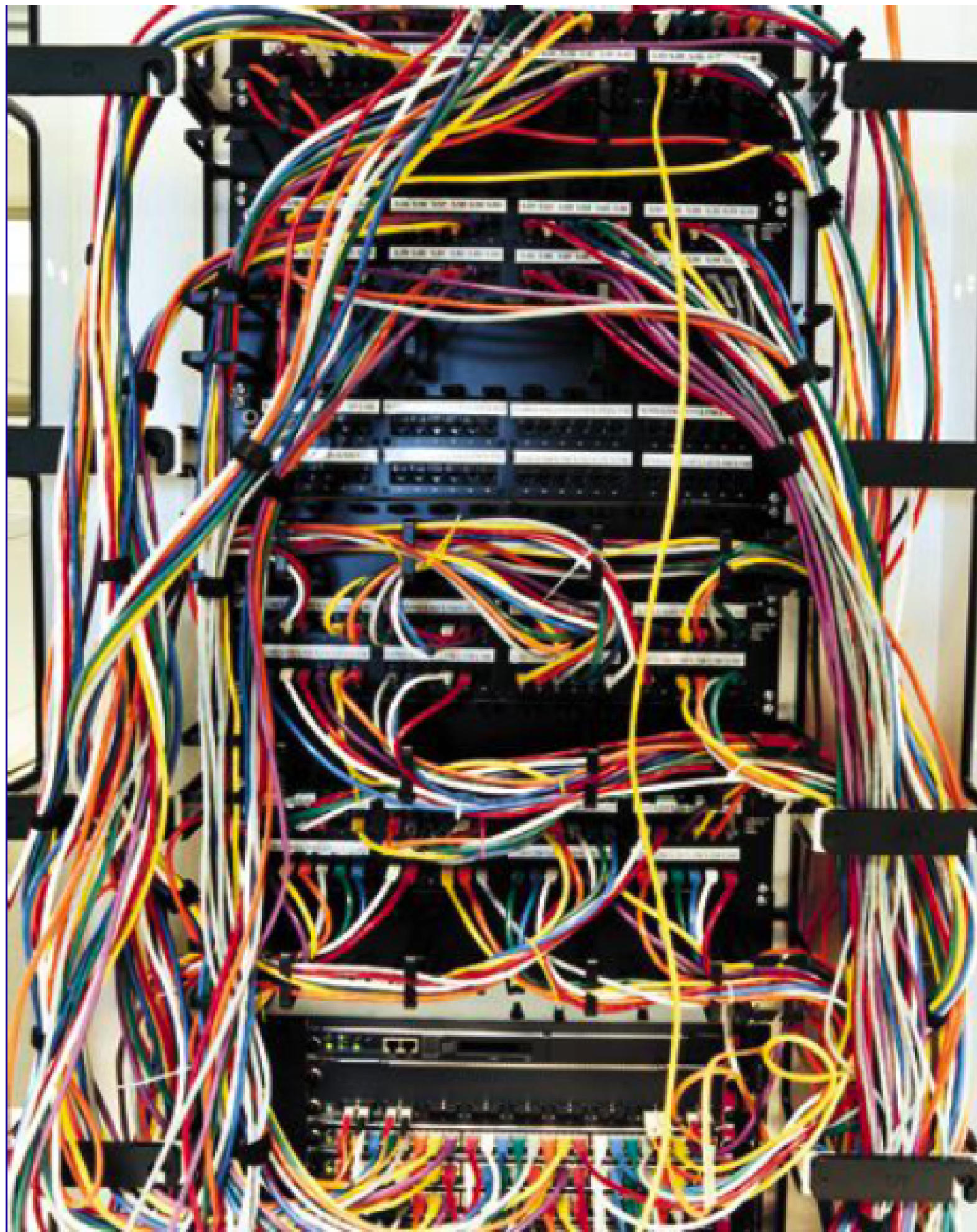
The Research Data Alliance. Because open data is not enough.

Mark A. Parsons
0000-0002-7723-0950
Secretary General

CREMLIN WP2 Workshop on Big data Management
Moscow, Russia
15 February 2017



Infrastructure is hard to conceive and describe because when it works, it's transparent, ubiquitous, and embedded in our daily work.











Dynamics of Infrastructure

Edwards, et al. 2007 Understanding Infrastructure: Dynamics, Tensions, and Design.



- Infrastructures become “ubiquitous, accessible, reliable, and transparent” as they mature.
- Systems  Networks  Inter-networks
 - “system-building, characterized by the deliberate and successful design of technology-based services.”
 - “technology transfer across domains and locations results in variations on the original design, as well as the emergence of competing systems.”
 - Finally, “a process of **consolidation characterized by gateways** that allow dissimilar systems to be linked into **networks**.”

Not what, but

When is infrastructure?

Not what, but

When and

Who is infrastructure?

Bridges and Gateways

Gateways are often wrongly understood as “technologies,” i.e. hardware or software alone. A more accurate approach conceives them as combining **a technical solution with a social choice**, i.e. a standard, both of which must be integrated into existing users’ communities of practice. Because of this, gateways rarely perform perfectly.

— Edwards et al. 2007



Infrastructure is

Relationships, interactions, and connections
between people, technologies, and institutions

(that helps data flow and be useful)

THE RESEARCH DATA ALLIANCE

www.rd-alliance.org

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

17 FLAGSHIP OUTPUTS

of which 4 ICT
Technical
Specifications

75 ADOPTION CASES

across multiple
disciplines,
organisations &
countries

85 GROUPS WORKING ON GLOBAL DATA INTEROPERABILITY CHALLENGES

of which 35 WORKING GROUPS
& 50 INTEREST GROUPS

4,908 INDIVIDUAL MEMBERS FROM 118 COUNTRIES

66% Academia & Research
15% Public Administration
11% Enterprise & Industry

46 ORGANISATIONAL MEMBERS & 6 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.

“Create - Adopt - Use”

(in 12-18 months)



Adopted Policy



Systems
Interoperability



Common Types,
Standards, Metadata



Sustainable Economics



Adopted Community
Practice



Training, Education,
Workforce

*Traffic Image:
Mike Gonzalez*

RDA: Accelerate Data Sharing and Interoperability Across Cultures, Communities, Scales, Technologies

■ **Technical parts of the data engine:**

- Data type registries reference model
- Wheat data interoperability framework

■ **Rules of the road:**

- Common agreement on data citation
- Common practice for data repositories
- Principles of legal interoperability

■ **Better drivers**

- Summer schools in data science and cloud computing in the developing world (with CODATA)
- Active data management plan development and monitoring



Systems
Interoperability



Common Types,
Standards, Metadata



Policy and Practice



Sustainable
Economics



Training, Education,
Workforce

Solving the problem must include adopters in the process.



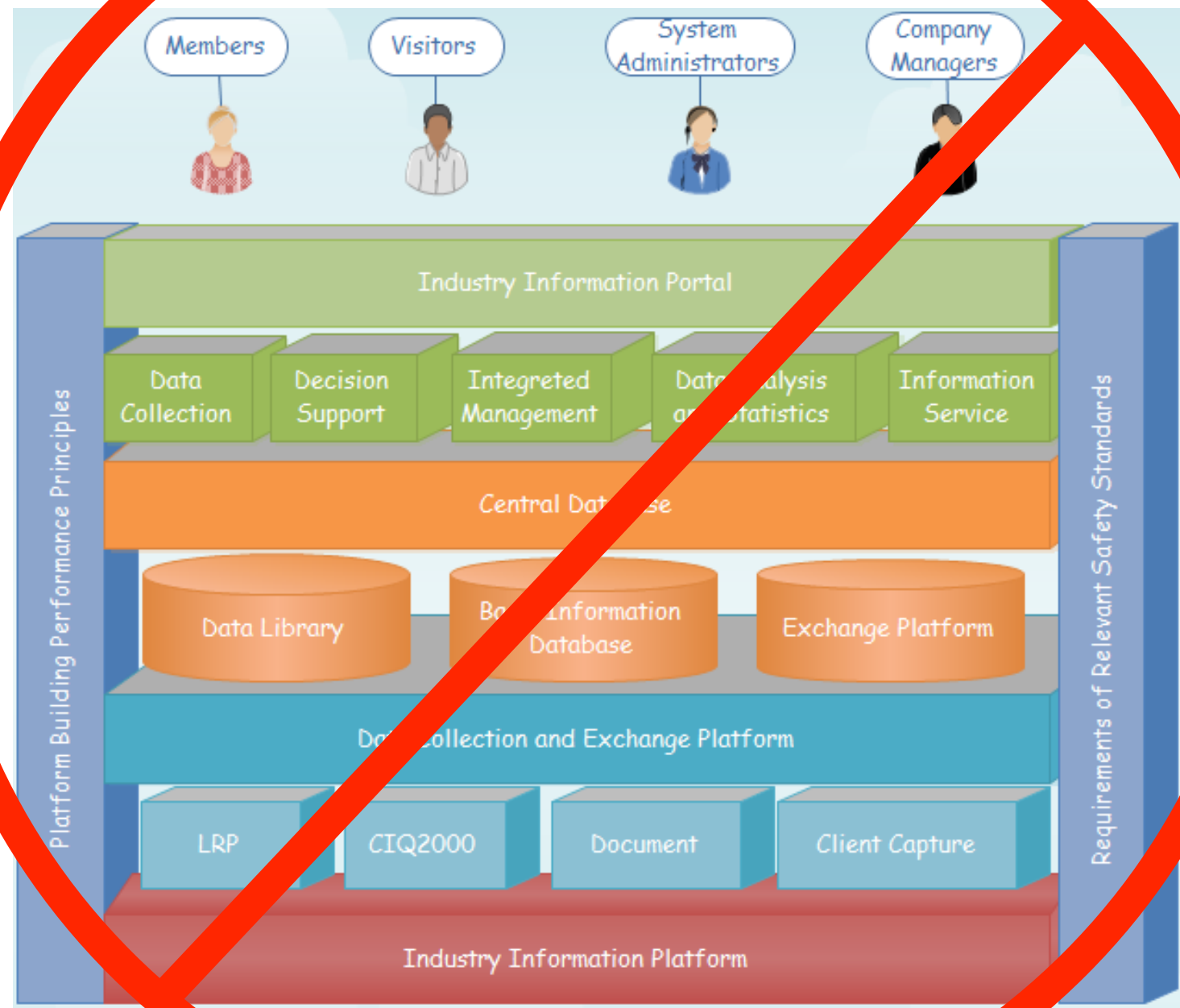
Open problem solving is key.



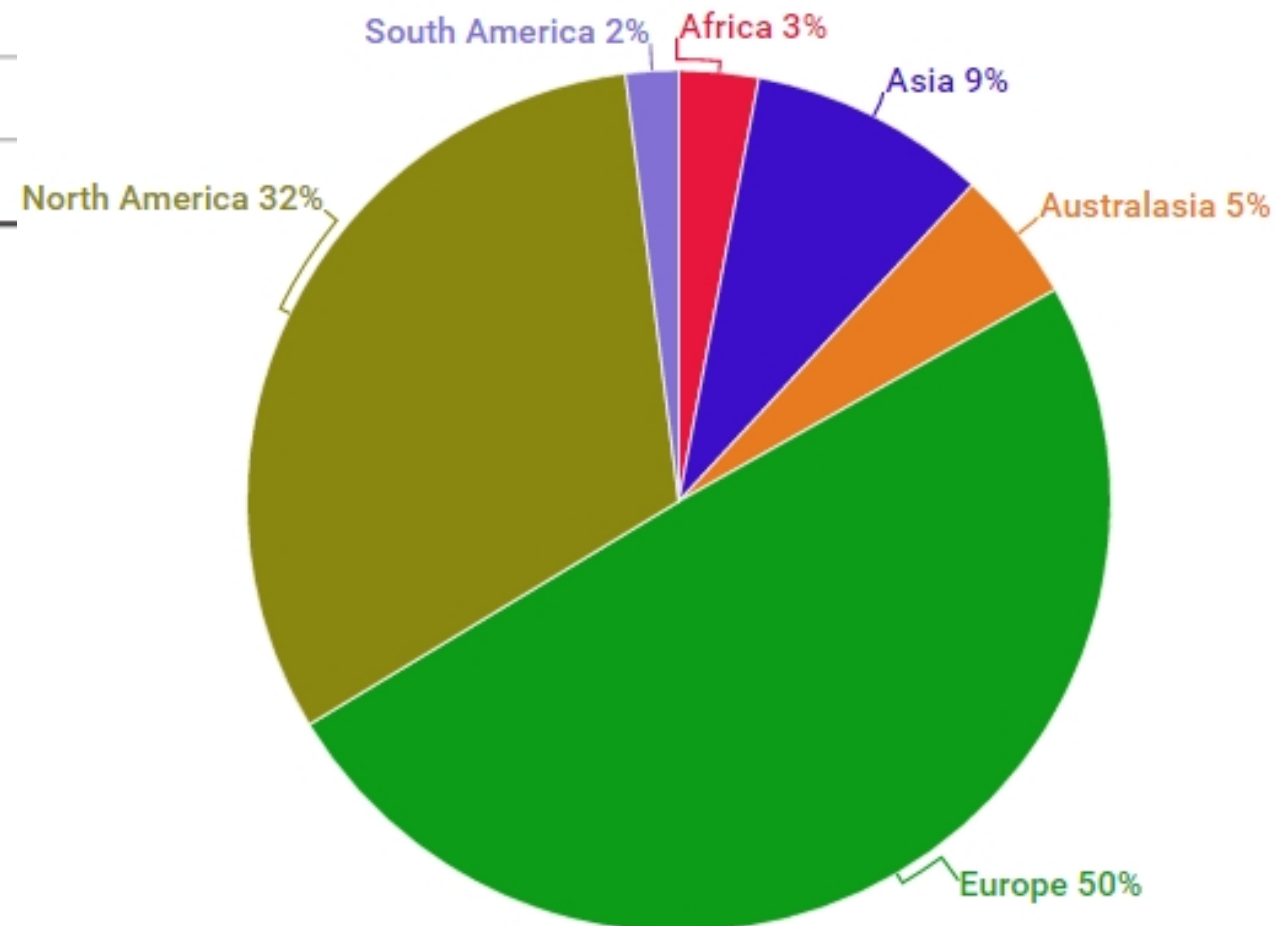
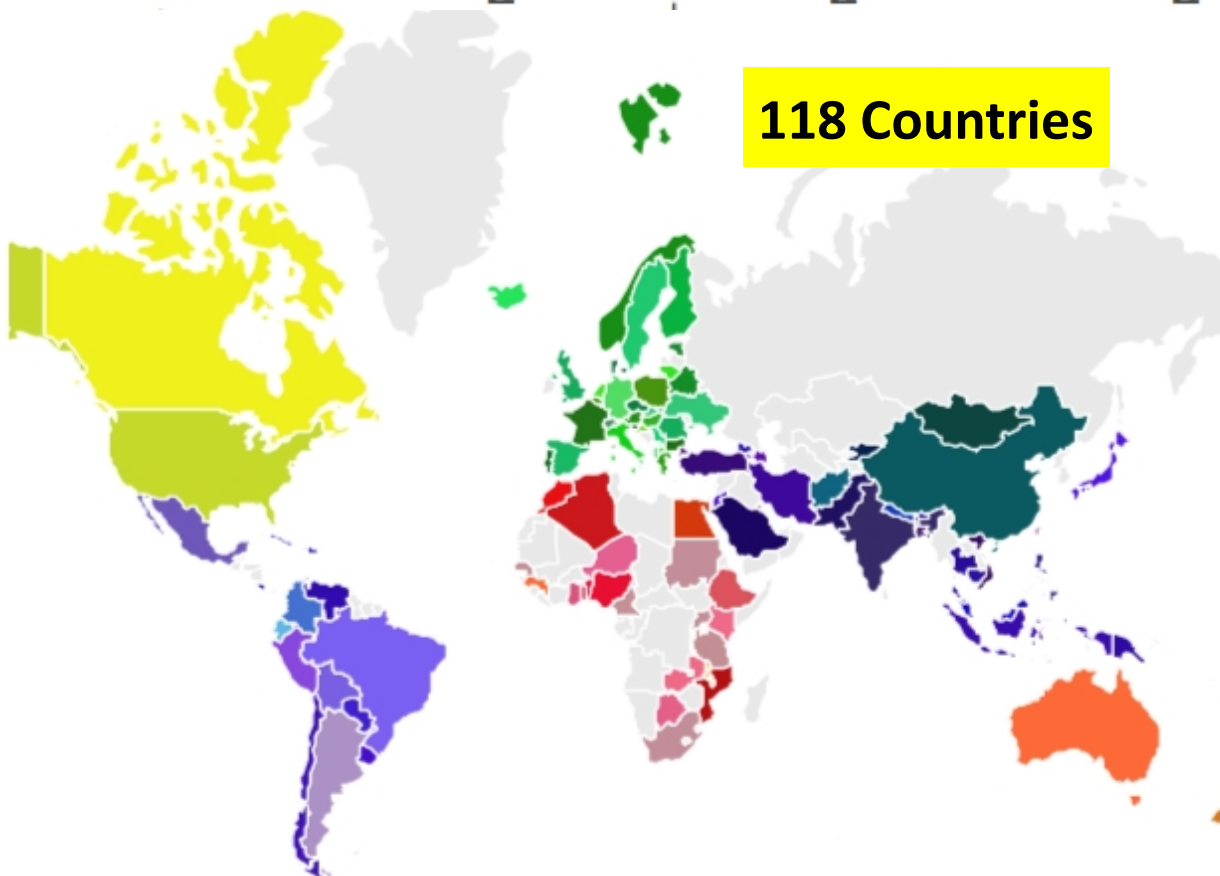
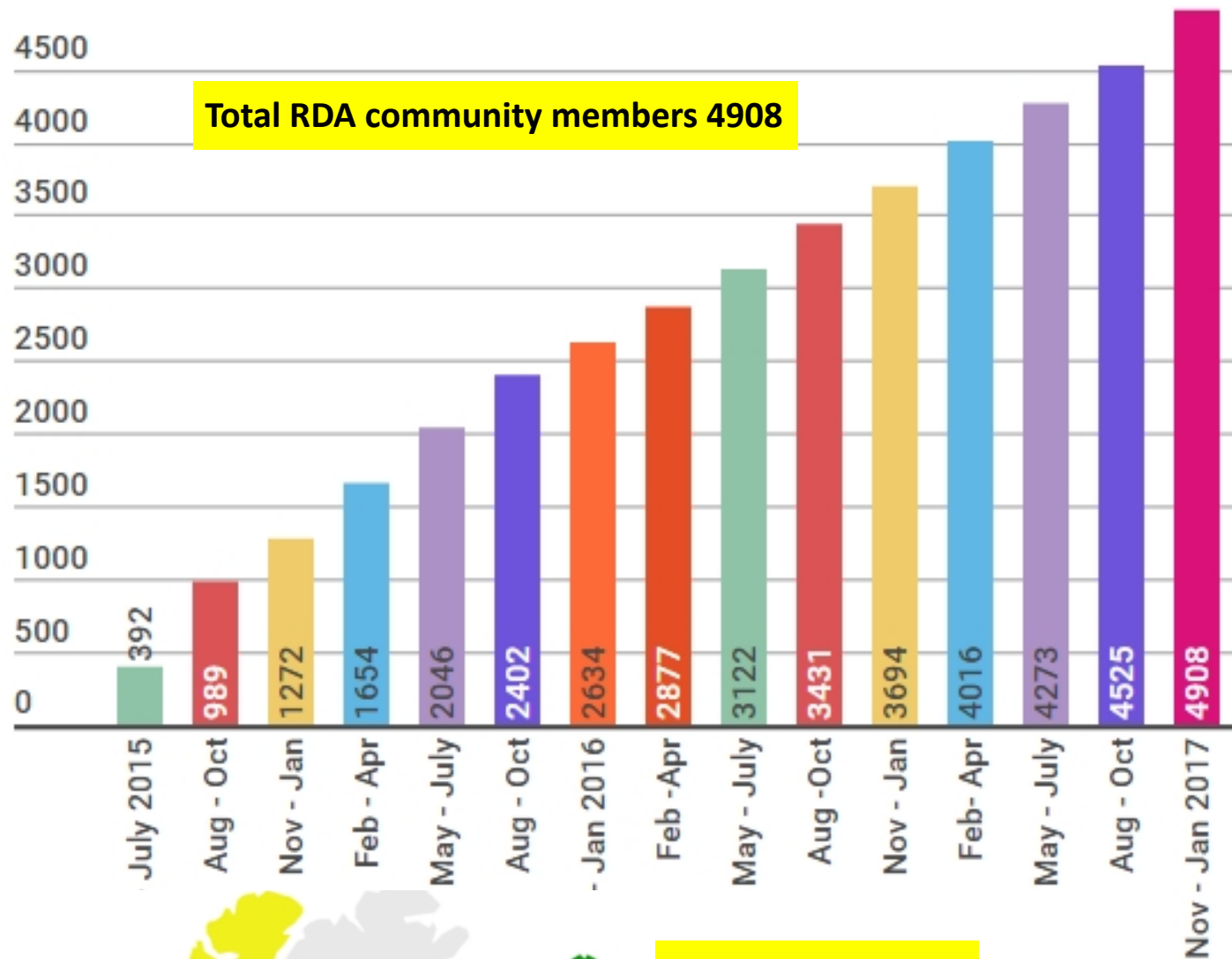
No defined architecture.

RDA Principles

Openness
Consensus
Balance
Harmonization
Community Driven
Non-profit



RDA worldwide growth



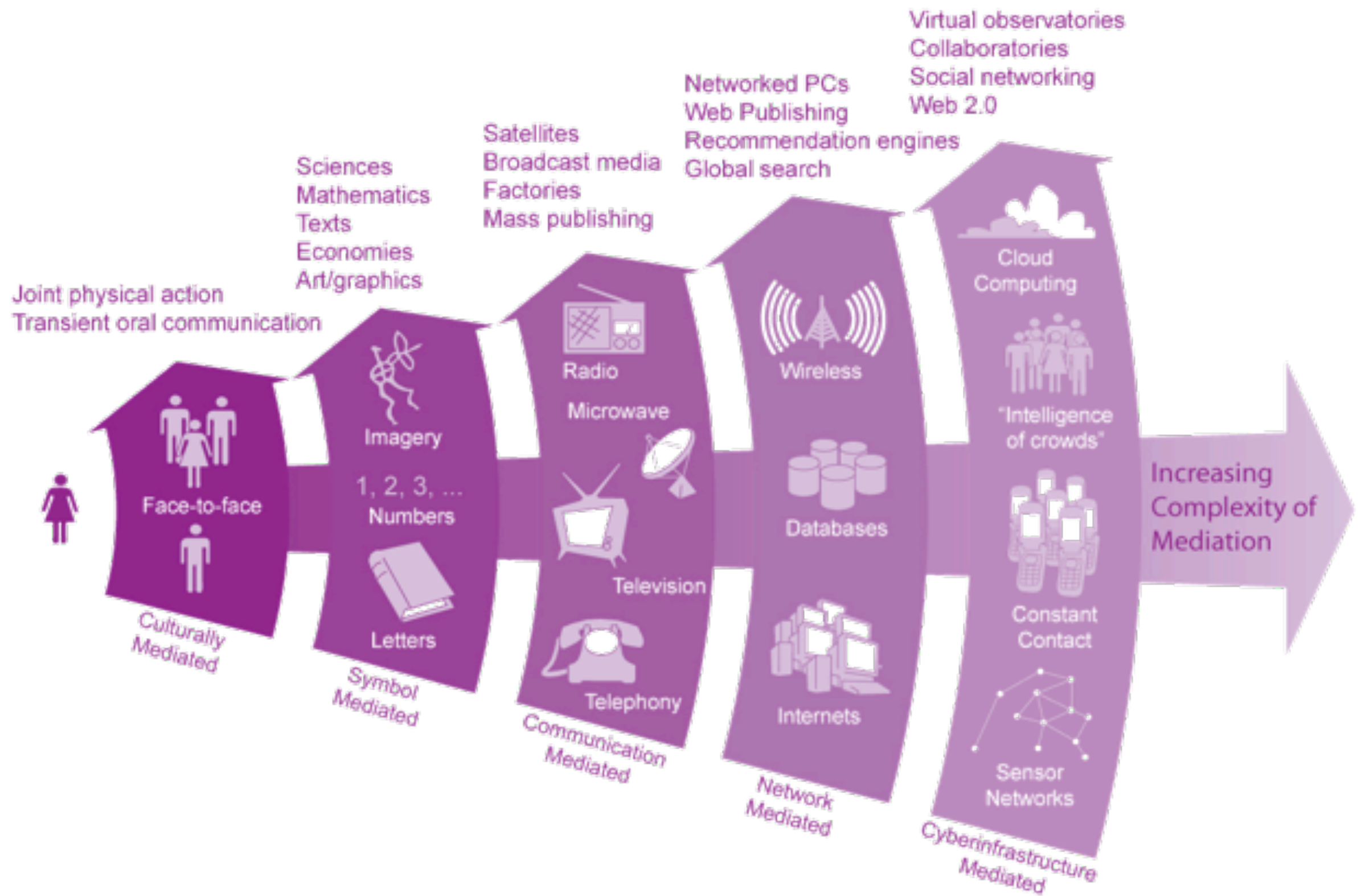
Glocality — Bridging across scales



Glocalization “means the simultaneity—the co-presence—
of both universalizing and and particularizing tendencies.”
— Roland Robertson

Some themes amidst the difference

1. **Persistent Identifiers** for data, documents, people, organisations, instruments — Everything!
2. **Certifying Trust** in assertions, evidence, organisations, processes...
3. The value of **Conversations, Relationships, and Mediation** — an agile network effect.



Increasing Complexity of Mediation

From: C. Borgman, 2008, NSF Cyberlearning Report

Some themes amidst the difference

1. **Persistent Identifiers** for data, documents, people, organisations, instruments — Everything!
2. **Certifying Trust** in assertions, evidence, organisations, processes
3. The value of **Conversations, Relationships, and Mediation** — an agile network effect.

Some reflections on trust and sharing and infrastructure



- When or do we need to certify trust? Do we?
- We must preserve the freedom to tinker. That is challenging in large-scale data facilities.
- Build in decentralization where possible. Any centralization must be community governed.
- Trust is built through
 - shared experience— e.g., RDA Plenaries
 - shared perspectives — RDA is a forum for engagement and constructive disagreement
 - actual reuse and adoption — in RDA consensus is defined through use.
 - sustained performance — an area where we all must find common purpose



RDA 9th Plenary Meeting

Data Infrastructures for Open Science

**5-7 April 2017, Barcelo Sants Hotel,
Barcelona, Spain**

Organised by Barcelona Supercomputing Center (BSC) with the support of RDA Europe



Host a **collocated or associated** event on April 3rd, 4th or 7th (PM) – Applications open until **27 Nov 2016 17:00 CET** <https://www.rd-alliance.org/plenaries/ninth-plenary/rda-9th-plenary-collocated-associated-events>

<https://www.rd-alliance.org/plenaries/rda-ninth-plenary-meeting-barcelona>



Call for Poster Session

<https://www.rd-alliance.org/plenaries/rda-ninth-plenary-meeting-barcelona/rda-9th-plenary-poster-session>

APPLICATIONS are due: 12th March 2017 at 23:00 UTC

RDA 10th ème



Montréal, Canada
19-21 September 2017
19 au 21 septembre 2017

Plenary Meeting
Conférence plénière

<https://www.rd-alliance.org/plenaries/rda-tenth-plenary-meeting-montreal-canada>

HOST THE INTERNATIONAL RESEARCH DATA ALLIANCE PLENARY 11 MEETING



CALL FOR PROPOSALS TO HOST RDA PLENARY 11 RDA PLENARY 11,
EARLY 2018

<https://www.rd-alliance.org/host-international-research-data-alliance-community's-meeting-early-2018>



Info:

enquiries@rd-alliance.org

[@resdatall](#)

**Seeking new
Secretary
General**

research data sharing without barriers
rd-alliance.org

RDA Recommendations & Outputs

THE RESEARCH DATA ALLIANCE RECOMMENDATIONS & OUTPUTS



Data Foundation & Terminology: a model for data in the registered domain.

PID Information Types: a common protocol for providers and users of persistent ID services worldwide.

Data Type Registries: allowing humans and machines to act on unknown, but registered, data types.

Practical Policy: defining best practices of how to deal with data automatically and in a documented way with computer actionable policy.

Metadata standards directory: Community curated standards catalogue for metadata interoperability

RDA Recommendations & Outputs

Data Citation: defining mechanisms to reliably cite dynamic data

Data Description Registry Interoperability solutions enabling cross platform discovery based on existing open protocols and standards

Wheat Data Interoperability impacting the discoverability, reusability and interoperability of wheat data by building a common framework for describing, representing linking and publishing wheat data

Brokering Governance WG: Sustainable Business Models for Brokering Middleware to support Research Interoperability

RDA/CODATA Summer Schools in Data Science and Cloud Computing in the Developing World WG: A framework to run a series of Summer Schools in Data Science and data sharing in low and middle income countries (LMICs)

THE RESEARCH
DATA ALLIANCE
RECOMMENDATIONS
& OUTPUTS



RDA Recommendations & Outputs



Repository Audit and Certification DSA–WDS: A convergent DSA-WDS certification standard to help eliminate duplication of effort, increase certification procedure coherence and compatibility thus benefitting researchers, data managers, librarians and scientific communities.

RDA/WDS Publishing Data Bibliometrics: improved research data metrics and corresponding services, with the final goal of increasing the overall availability and quality of citations and research data itself.

RDA/WDS Publishing Data Services: A universal interlinking service between data and the scientific literature.

RDA/WDS Publishing Data Workflows: enhance the possibilities for greater discoverability and a more efficient and reliable reuse of research data benefitting other stakeholders like publishers, libraries and data centres.

RDA Recommendations & Outputs

THE RESEARCH DATA ALLIANCE RECOMMENDATIONS & OUTPUTS



23 Things: Libraries For Research Data An overview of practical, free, online resources and tools that users can immediately take advantage of to incorporate research data management into the practice of librarianship.

Legal Interoperability of Research Data Principles and Implementation

Guidelines: a set of principles and practical implementation guidelines offered as high-level guidance to all members of the research community — the funders, managers of data centers, librarians, archivists, publishers, policymakers, university administrators, individual researchers, and their legal counsel.

The Scholix initiative a high level interoperability framework for exchanging information about the links between scholarly literature and data. It aims to build an open information ecosystem to understand systematically what data underpins literature and what literature references data.