TetherlessWorld

Rensselaer

Reflections on the Research Data Alliance and Global Data Infrastructure

Mark A. Parsons
0000-0002-7723-0950

Symposium on International Coordination for Science Data Infrastructure
National Academy of Sciences
Washington, DC
1 November 2017

1st: Save the data. This is hard.

TetherlessWorld

2nd: Share the data. This is harder.

TetherlessWorld

Dynamics of Infrastructure

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

TetherlessWorld

- 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.”

Infrastructure is

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

(that helps data flow and be useful)

Research Data Alliance

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.

Rapid growth and many groups

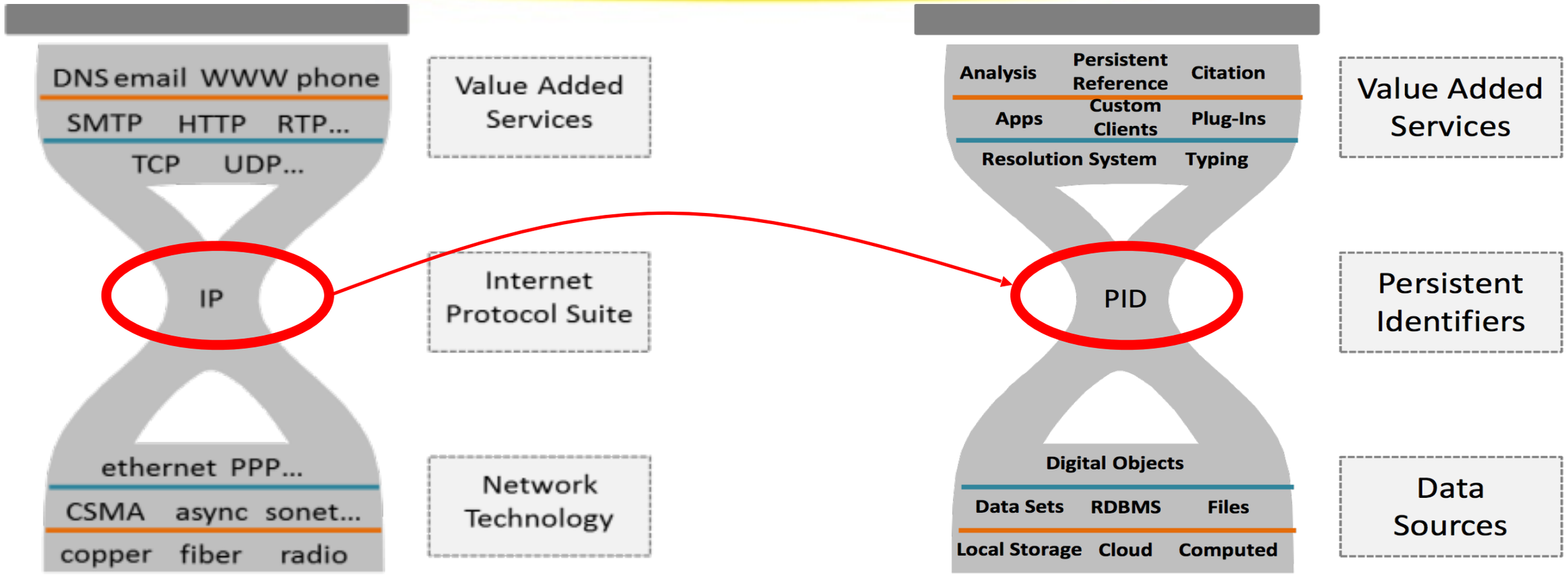
- 6000+ members from 120+ countries
- 2 dozen plus Working Groups
- 4 dozen plus Interest Groups

Some themes amidst the difference (from 2015)

RESEARCH DATA ALLIANCE

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.*

An Area of Convergence and Agreement



Internet Domain
nodes with IP numbers
packages being exchanged
standardized protocols

Data Domain
objects with PID numbers
objects being exchanged
standardized protocols

Slide courtesy P. Wittenberg from L. Lannom from D. Clark

The Five Persistences

- Persistence of object
  – Or mechanism to handle its non-persistence
- Persistence of identifier
- Persistence of binding between identifier and object
- Persistence of service to resolve from identifier to object
- Persistence of service to allow for updating of binding between identifier and object

The Five Persistences

- Persistence of object
  - Or mechanism to handle its non-persistence
- Persistence of identifier
- Persistence of binding between identifier and object
- Persistence of service to resolve from identifier to object
- Persistence of service to allow for updating of binding between identifier and object

Some themes amidst the difference (from 2015)

RESEARCH DATA ALLIANCE

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 themes amidst the difference (from 2015)

RESEARCH DATA ALLIANCE

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.*

Trust

Some amateur thoughts on trust and sharing and infrastructure

- When or do we need to certify trust? Do we?
- We must preserve the freedom to tinker.
- Build in decentralization where possible. Any centralization must be community governed.
- Trust is built through
  - (routine) shared experience— e.g., RDA Plenaries,
  - shared perspectives — RDA is a forum for engagement and constructive disagreement
  - actual reuse and adoption — in RDA consensus or “standardization” is defined through use.
  - sustained performance — RDA seeks to build a broad coalition of international support

Social and Technical Trust

TetherlessWorld

Social

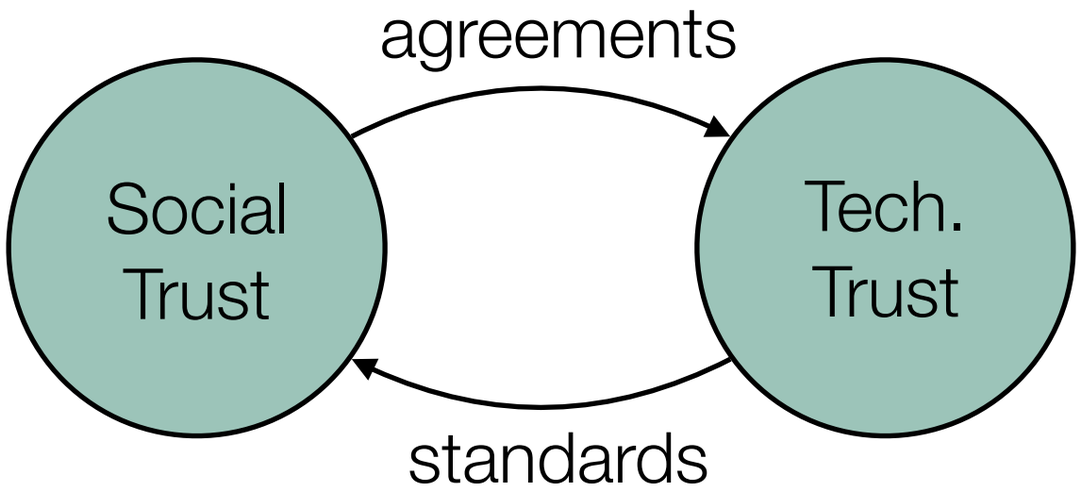
- The *authority* question — Do I believe you? (Who are you to say…?)
    - This bumps into the so-called social contract of science (social knowledge in exchange for funding) which requires requires monitoring and incentives.

Technical

- The *authenticity* question — Do I believe the object? (content, description, bit verification, location, etc.)
    - This must include the “Do I believe the binding?“ per Treloar.
    - So there are issues of malware, neglect, disaster, …,
    - but also the social trust in a negotiated binding.

An Ongoing Interplay between Social & Technical Trust

TetherlessWorld



- The act of creating ‘standards’ (i.e. consistency), when coupled with implementation or adoption, forces you through this cycle
  - to develop an equilibrium of conflicting community interests
  - that is negotiated between individuals
- Must be done in a neutral place *and* space (folks must feel included)
  - must consciously recognize and work through the friction
  - must be done glocally.

Rensselaer

Effective trust requires glocal work

TetherlessWorld

Glocalization "means the simultaneity—the co-presence—of both universalizing and and particularizing tendencies."

Roland Robertson.

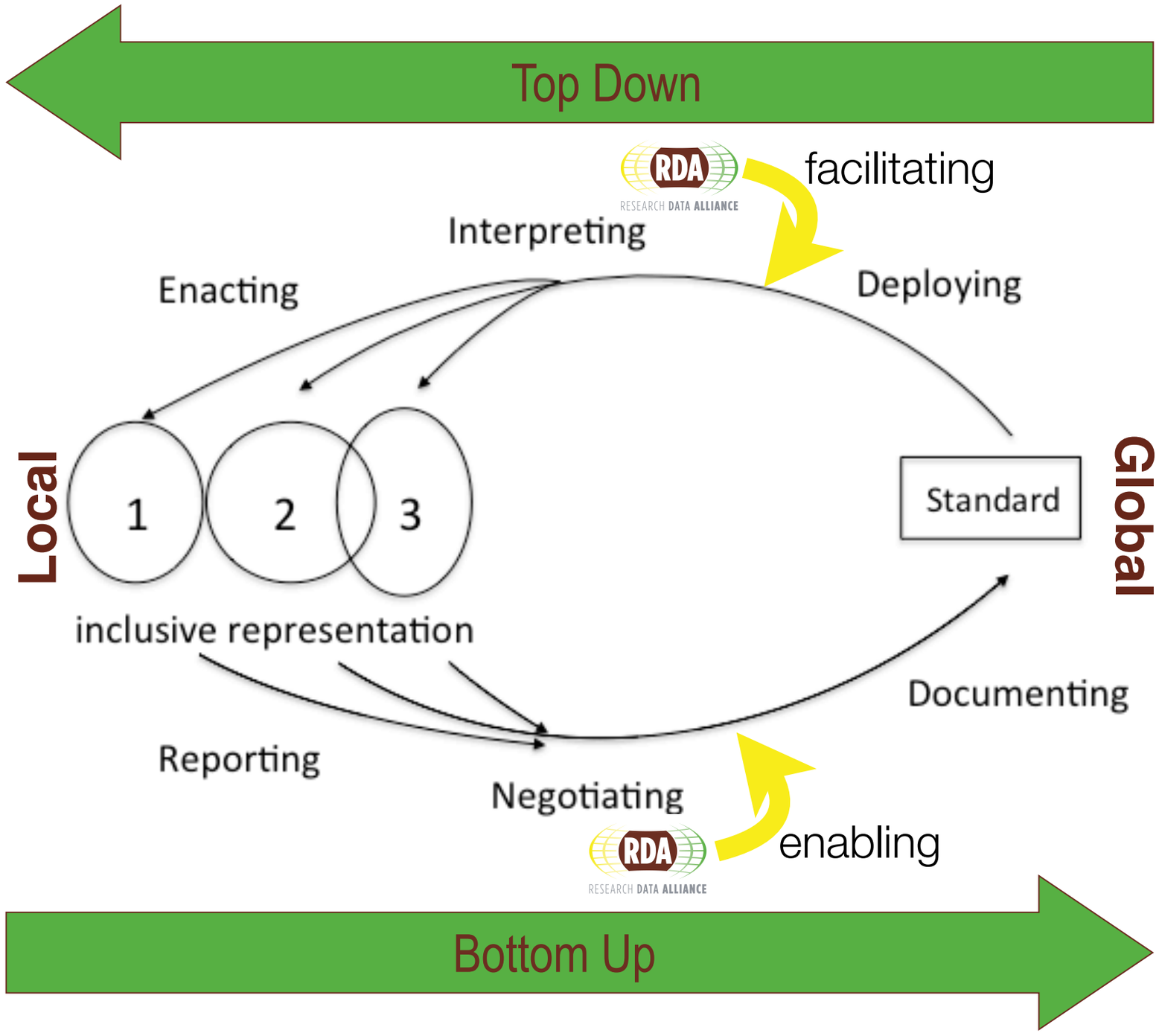


Figure adapted from Yarmey and Baker (2013) http://dx.doi.org/10.2218/ijdc.v8i1.252

Rensselaer

Solving the problem must include adopters in the process

Public Radio International

The Inquirer

bigthink.com

Open problem solving is key.

What else seems to work?

TetherlessWorld

1. Keep working timelines short (12-18 months) to focus effort,

   But keep discussion/interest timelines open.

2. Seed funding can be a huge help for adoption and deployment,

   But not for coordination,

   And one must find the right balance and separation of concerns between local and central funding.

3. Foster discussion fora and neutrality,

   But central facilitation needs community buy-in,

   And one must work to create both a neutral space and place.

4. Solid principles guide difficult decisions.

5. Openness makes for more durable decisions.

6. Friction and disagreement are necessary and productive,

   But disagreement rooted in power dynamics is destructive (see above).

Thank You

Mark A. Parsons
parsom3@rpi.edu
@chutneyboy
0000-0002-7723-0950