Community Approaches to Digital Preservation

Panelists:

- Fran Berman, Rensselaer Polytechnic Institute
- Bradley Daigle, University of Virginia
- Aaron Rubinstein, University of MA Amherst
- Jaime Schumacher, Northern Illinois University

Moderator: Meg Phillips, NARA

Research Data Alliance

Fran Berman and Larry Lannom

research data sharing without barriers
rd-alliance.org

Research Data Alliance Created to Accelerate Development of Research Data Sharing Infrastructure Worldwide

- RDA community focuses on building **social, organizational and technical infrastructure** to

  - reduce barriers to data sharing and exchange

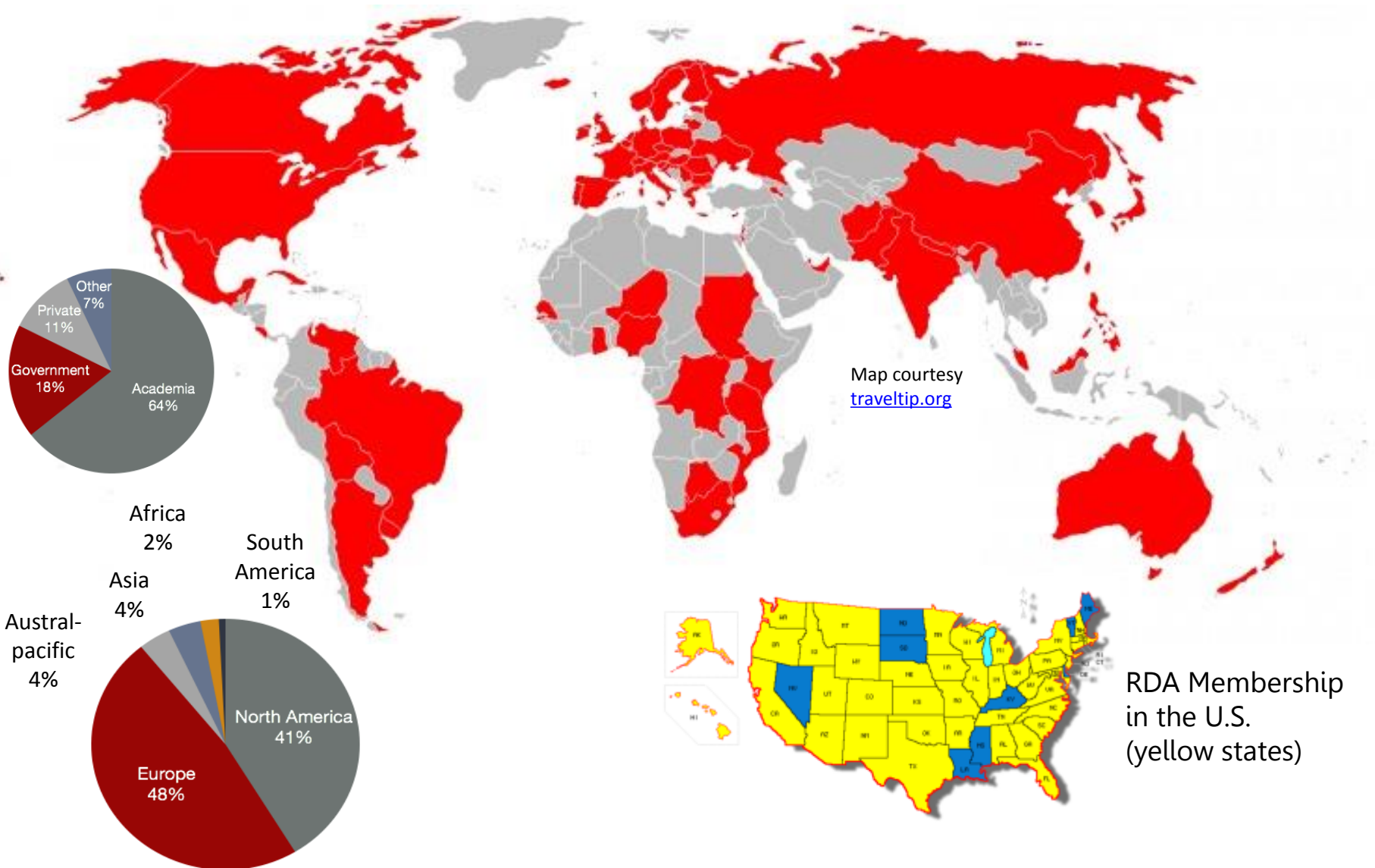 - accelerate the development of coordinated global data infrastructure

Plenary 2
Washington, DC

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 1850 members from 80+ countries (as of 6/14)

Map courtesy traveltip.org

RDA Membership in the U.S. (yellow states)

Pie chart 1:
- Other 7%
- Private 11%
- Government 18%
- Academia 64%

Pie chart 2:
- Africa 2%
- Asia 4%
- South America 1%
- Austral-pacific 4%
- North America 41%
- Europe 48%

RDA Interest (IG) and Working Groups (WG) by Focus Area (as of 6/14)   rd-alliance.org

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
- Defining Urban Data Exchange for Science IG
- Geospatial IG
- Marine Data Harmonization IG
- RDA/CODATA Materials Data Infrastructure and Interoperability IG
- Research Data Needs of the Photon and Neutron Science Community IG

Community Needs - focused

- Community Capability Model IG
- Engagement IG
- Development of Cloud Computing Capacity and Education in Developing World Research IG
- Ethics and Social Aspects of Data IG

Reference and Sharing - focused

- Data Citation WG
- Standardization of Data Categories and Codes WG
- RDA/CODATA Legal Interoperability IG
- Data Description Registry Interoperability Working Group

Data Stewardship - focused

- Research Data Provenance IG
- RDA/WDS Certification of Digital Repositories IG
- Preservation e-infrastructure IG
- Long-tail of Research Data IG
- RDA/WDS Publishing Data IG
- RDA/WDS Repository Audit and Certification Working Group
- Domain Repositories Interest Group

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 IG
- Federated Identity Management IG
- Metadata IG
- PID Interest Group
- Service Management IG

Academic Preservation Trust aptrust.org

Presentation at Digital Preservation 2014

July 22, 2014

Bradley Daigle (University of Virginia)

Academic Preservation Trust

Academic Preservation Trust, a consortium of 17 institutions, is taking a community approach in building and managing a repository infrastructure that will provide long-term preservation of the scholarly record. APTrust will also be a DPN first node.

APTrust Institutions

- Columbia University
- Johns Hopkins University
- Indiana University
- North Carolina State University
- Penn State University
- Stanford University
- Syracuse University
- University of Chicago
- University of Cincinnati
- University of Connecticut
- University of Maryland
- University of Miami
- University of Michigan
- University of North Carolina
- University of Notre Dame
- University of Virginia
- Virginia Tech

www.aptrust.org

APTrust is hosted by the University of Virginia, which fully supports 5 ½ staff, including space and equipment.

- Program Director
- Lead Engineer
- Junior Engineer
- Systems Engineer
- Content Lead

readiness and action

digital preservation at the five colleges

Amherst College

Hampshire College

Mt Holyoke College

Smith College

UMass Amherst

Aaron Rubinstein
University and Digital Archivist
UMass Amherst
@rubinsztajn

phase I: assessment

define scope

understand domains of practice

assessment of organizational, technological, and resource readiness
    AIDA self-assessment survey

talk to other consortia
    Florida Digital Archive
    Colorado Alliance of Research Libraries
    Orbis Cascade
    OhioLink
    MetaArchive

phase II: three-pronged plan

education

   Nancy McGovern workshop
   POWERR workshop
   digital preservation guide
   outreach opportunities
   affinity groups

best practices/standardization

   stakeholders and decision making

experimentation

   archivematica pilot
   collaboration
   hands-on experience

the task force

Thea Atwood (Hampshire/UMass)
Meghan Bergin (UMass)
Leslie Fields (MHC)
Elisa Lanzi (Smith)
Aaron Rubinstein, Chair (UMass)
Shaun Trujilo (MHC)
Sarah Walden (Amherst)

and Kelcy Shepherd, who chaired the group during phase I

~90 Seconds To Enlightenment~

POWRR

Preserving (Digital) **O**bjects **W**ith **R**estricted **R**esources

Our Background

Our Mission

What We Discovered

What We Did About It

What Still Needs To Be Done

Preserving (digital) Objects With Restricted Resources

Sponsored By:

POWRR
Preserving (Digital) Objects With Restri... ources
UA!

Tool Grid

Digital POWRR Tool Evaluation Grid	Ingest					Processing						Access		Storage					Maintenance			Other		
	Copy	Fixity Check	Virus Scan	File Dedupe	Auto Unique ID	Auto Metadata Creation	Auto Metadata Harvest	Manual Metadata	Rights Management	Package Metadata	Auto SIP Creation	Public Interface	Auto DIP Creation	Auto AIP Creation	Reliable, Long-Term Bit Preservation	Redundancy	Geographically Dispersed Data Storage Model	Exit Strategy	Migration	Monitoring	Auto Recovery	Open Source	Clear Documentation	Cost
ACE (Audit Control Environment)		x																		x		x		Free
AFF Open Source Computer Forensics Software							x															x		Free
Amazon Cloud	x																	x					x	Varies
Archive-It	x	x	x	x	x	x	x	x	x	x	x	x	x	x		x	x		x	x		x	x	Varies
Archivematica		x	x		x	x	x	x	x	x	x	x	x	x	x					x		x	x	Free
BagIt Library		x			x	x				x				x					x					
BagIt Transfer Utilities																								Free
BWF MetaEdit		x						x	x													x		free
Carbonite	x	x										x			x	x			x	x	x			$149 – $599
Chronopolis	x	x		x	x	x	x	x	x	x	x			x	x	x	x	x	x	x	x		x	$1,500 – $2,200

Moving Forward….How?

ILLINOIS STATE UNIVERSITY
Illinois' first public university

ILLINOIS WESLEYAN UNIVERSITY

NIU
Northern Illinois University

CHICAGO STATE UNIVERSITY

WESTERN ILLINOIS UNIVERSITY

Sponsored By: INSTITUTE of Museum and Library SERVICES