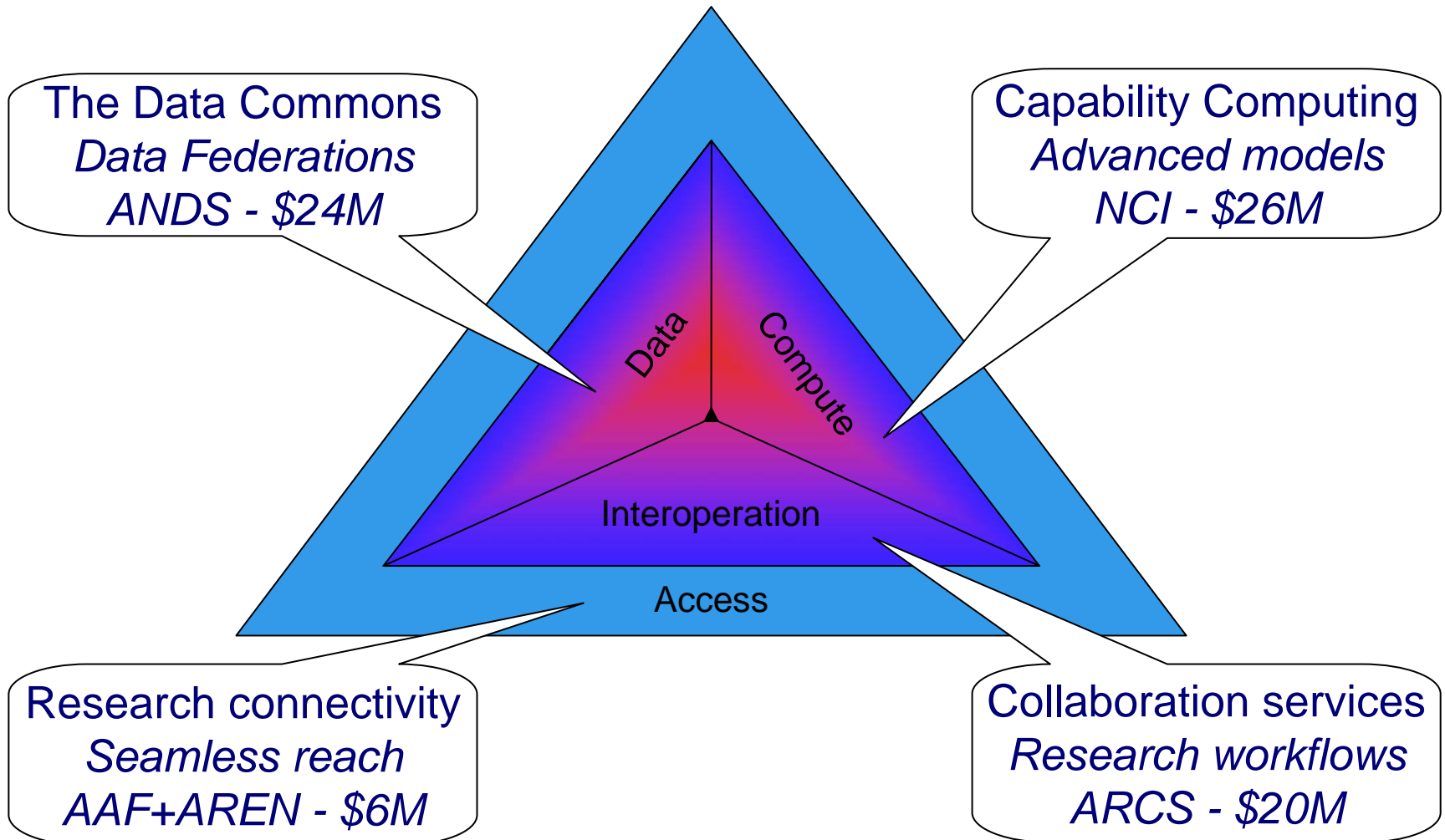


AeRIC and ANDS: OVERVIEW AND UPDATE

Cathrine Harboe-Ree

Platforms for Collaboration: Major Investments 2007-2011

2



Other PfC

3

- ARCS: Tony Williams (reports through VPAC)
- NCI: Lindsay Botten (reports through ANU)
- Both have business plans, with KPIs being resolved
 - Engagement
 - Measurable KPIs
 - Service definition, and managing expectations
- AAF
 - To be hosted by Universities Australia
 - Not enough traction in universities

ANDS Establishment Project

4

- Monash, ANU, CSIRO asked to set up ANDS
- Four main deliverables:
 1. Formal collaboration agreement to deliver ANDS
 2. Funding agreement with DIISR for ANDS
 3. Selection process leading to an acceptable shortlist for the ANDS Executive Director
 4. ANDS Business Plan for FY 2008/2009
- Originally funded to June 2008; extension for deliverables to Aug 29

From Project to Service

5

- Establishment Project has met all its deliverables
- ANDS 'proper' should come into existence this month
- First Business Plan available at <http://ands.org.au/andsinterimbusinessplan-final.pdf>
 - This will run until June 2009
 - Next Business Plan needs to be complete by March 2009 for consideration and approval
- ANDS will run until July 2011

ANDS governance

6

- Independent Chair (part-time)
- Minimum 4, maximum 8 members
 - 1 each from Monash, ANU, CSIRO
 - Specialists from research disciplines and institutions
- Several reference groups to be established
- 50% ANDS funding to be distributed beyond lead institutions

The ANDS Blueprint

7

- *Towards the Australian Data Commons (TADC)*
- Developed during 2007 by ANDS Technical Working Group
- Mapped out coherent vision of what needs to be done in the data space
- Available at <http://www.pfc.org.au/bin/view/Main/Data>

ANDS Assumptions

8

- ANDS doesn't have enough money to fund storage
 - ▣ And so is predicated on institutionally-supported solutions (Australia doesn't have discipline-specific data centres like the UK)
- ANDS aims to leverage existing activity, and coordinate/fund new activity
- ANDS will only **start** to build the Australian Research Data Commons
- ANDS governance and management arrangements are sized for the current funding

Changing Data, Changing Research

9

- New scientific instruments
 - ▣ Large Hadron Collider at CERN generates 1.5 gigabytes of data per second
 - ▣ the Square Kilometre Array (1 EB/day!)
- New scientific Models
 - ▣ The mapping of the Human Genome: A billion DNA letters in a human sequence
 - ▣ Global climate models
- New teams
 - ▣ 195 scientists mapped the genome of the fruit-fly
- New knowledge from unlocked data
 - ▣ Most research from Hubble telescope data was not “first use”
 - ▣ Common data sets unlocked the power of search technology – TREC

ANDS Delivery Structure

10

- ANDS has therefore been structured as four inter-related and co-ordinated service delivery programs:
 - ▣ Developing Frameworks
 - ▣ Providing Utilities
 - ▣ Seeding the Commons
 - ▣ Building Capabilities
- Plus candidate service development activities funded through National eResearch Architecture Taskforce projects

Developing Frameworks (Monash)

11

- Influencing relevant national policies
- Building common understanding of data management issues and solutions across government, research funding agencies, and research intensive organisations
- Encouraging moves in favour of discipline-acceptable default sharing practices
- Largely centralised, with some specialised outsourcing

Providing Utilities (ANU)

12

- Building and delivering national technical services to support the data commons
- Examples:
 - ▣ Discovery
 - Both “you come to us” and “we come to you” flavours
 - Probably a two-step process for some collections
 - Includes surfacing of ISO2146 entities (next slide) for web harvesting
 - ▣ Persistent identifier minting and management
 - ▣ Collections registry to underpin discovery
- Outsourced delivery and insourced technical framework development
- Providing capability within ANDS for integration of existing systems into Australian Data Commons

Seeding the Commons (Monash)

13

- In targeted areas (because not enough resource to do everything), working to improve:
 - ▣ fabric for data management
 - ▣ amount of content
 - ▣ state of data capture and management
- Plus, opportunistic content recruitment in first year
- Selection process to identify targets
- Placement of ANDS-funded staff, together with co-investment

Building Capabilities (ANU)

14

- Improving level of capability for research data management and research access to data
 - ▣ Train-the-trainer model
- Building community around data management concerns
- Largely distributed

ANDS has its critics

- Why hasn't ANDS started?
- It's not necessary – data is already being well managed
- What's it got to do with me – isn't it just money for Monash ANU and CSIRO?
- ANDS has been captured by librarians

Australian Strategic Roadmap Review

16

- <http://www.innovation.gov.au/ScienceAndResearch/Documents/Strategic%20Roadmap%20Aug%202008.pdf>
- Data Storage (p.21)
 - National data-fabric, based on institutional nodes
- Shared Data (p. 22)
 - More ANDS
- Coordination Component (p. 23)
 - Integration of eResearch activities
- Expertise as an enabling infrastructure (p. 23)

Changing Data, Changing Research

17

- New scientific instruments
 - ▣ Large Hadron Collider at CERN generates 1.5 gigabytes of data per second
 - ▣ the Square Kilometre Array (1 EB/day!)
- New scientific models
 - ▣ The mapping of the human genome: a billion DNA letters in a human sequence
 - ▣ Global climate models
- New teams
 - ▣ 195 scientists mapped the genome of the fruit-fly
- New knowledge from unlocked data
 - ▣ Most research from Hubble telescope data was not “first use”

