eResearch Strategic Positioning for the Future
-The National Collaborative Research Infrastructure Strategy Influence

Dr M.A. Sargent

A Decade of Strategic Investment -1

  - the establishment of the Australian Partnership for Advanced Computing to provide access to high performance computing capability

- 2000: Establishment of Advanced Networks Programme
  - Establish demonstrator advanced networks

- 2002: Higher Education Bandwidth Advisory Committee Report
  - the establishment of Australian Research and Education Network Advisory Committee, and investment in the Australian research and education network

- 2003: Higher Education Information Infrastructure Advisory Committee Report
  - the establishment of ARIIC and funding of projects to improve the access of Australian researchers to information
A Decade of Strategic Investment -2

• **2004: Research Infrastructure Taskforce Report**
  – the establishment of the National Collaborative Research Infrastructure Strategy Committee to implement a program of strategic investment in research infrastructure

• **2006: eResearch Coordinating Committee Report**
  – outlines an integrated program of skills development and of middleware and computer science research


The National Collaborative Research Infrastructure Strategy Principles

• Investments must result in excellent research infrastructure that addresses the national requirements of the relevant capability area described in the NCRIS Roadmap.

• Investments must result in research infrastructure that is accessible by researchers on the basis of merit, at reasonable prices.

• Investments must include a facility ownership and management structure that will result in the efficient and effective operation of the infrastructure.

• Investments must include a business plan that will result in the efficient implementation and effective ongoing financial management of the infrastructure.

A single integrated proposal from the research sector
National Collaborative Research Infrastructure Strategy

Australian Government spend of about $540M over the five years: 2007-2011
• Evolving bio-molecular platforms and informatics
• Integrated biological systems
• Characterisation
• Fabrication
• Biotechnology products
• Networked biosecurity framework
• Optical and radio astronomy
• Integrated marine capability
• Structure and evolution of the Australian continent
• Population health and clinical data linkage
• Terrestrial ecosystem research network
• Platforms for Collaboration (allocated $75M)

Key Outcomes

• Distributed capabilities
• Collaborative governance regimes
• Data focus - create, store, use
• Data intensity
• Linking of physical facilities, operational frameworks and data
• Modelling opportunities

….The eResearch agenda seen from the user end
Platforms for Collaboration

Where are we up to? What is going on?

( APAC + SII Projects ) \rightarrow \begin{array}{l}
\text{AeRIC, NEAT} \\
\text{ANDS, ICI, APAC*} \\
\text{AAF, AREN}
\end{array}

\begin{array}{c}
\ldots 2000 - 2007 \\
\text{Now}
\end{array}
### Platforms for Collaboration

<table>
<thead>
<tr>
<th>NCRIS ROADMAP</th>
<th>RELATED ISSUES</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Data access and discovery, storage and management</td>
<td>• Sharing researcher identities across all systems</td>
</tr>
<tr>
<td>• Grid enabled technologies and infrastructures</td>
<td>• Allowing researchers and resource owners to control who can do what to their resources, directly with the users</td>
</tr>
<tr>
<td>• Technical expertise</td>
<td>• Expanding the e-Research community to “non-builder” users</td>
</tr>
<tr>
<td>• High performance computing</td>
<td>• Engaging the broader NCRIS community</td>
</tr>
<tr>
<td>• High capacity communication networks</td>
<td></td>
</tr>
</tbody>
</table>

### Related Issues

- NCRIS ROADMAP
  - Data access and discovery, storage and management
  - Grid enabled technologies and infrastructures
  - Technical expertise
  - High performance computing
  - High capacity communication networks

- RELATED ISSUES
  - Sharing researcher identities across all systems
  - Allowing researchers and resource owners to control who can do what to their resources, directly with the users
  - Expanding the e-Research community to “non-builder” users
  - Engaging the broader NCRIS community

### Platforms for Collaboration

Consulting and user support

Services (Tools) ((Software))

**Middleware**

**Data Federation**

**Interoperability**

- Sensors
- Compute infrastructure *capability/capacity*
- Instruments
- AAF (trust)
- AREN (access)
The three basic layers

- **Tools and Applications**
- **Resource & Services Infrastructure**
- **Foundation services**

High level goals

<table>
<thead>
<tr>
<th>Foundations</th>
<th>Resources</th>
<th>Tools</th>
</tr>
</thead>
</table>
| National networking  
  - From desktop to national and global facilities  
  - All desktops to all facilities  
  - High speed  
  - Not volume charged  
National access control  
  - Home institution based  
  - Researcher controlled authorisation  
  - Single sign on  
  - Community wide roles and rights | Collaborative computing  
  - Relatively tightly integrated  
  - Able to include shared and dedicated facilities  
Federated data  
  - National reach (location independence)  
  - Inter-operating authentication/authorisation  
Seamless integration  
  - Compute and Data  
  - Large instruments and sensor networks  
  - Able to expand over time | Current e-Researchers …  
  Users = Builders  
For builders we need:  
  - Operational services  
  - Development tools  
  - Standard components  
  - Supported software  
For users we need:  
  - Supported applications  
  - Simple tools & interfaces  
In expertise we need:  
  - Training for the builders  
  - Consulting and help desk support for the users |
Main PfC Activities & Investments

<table>
<thead>
<tr>
<th>Resource &amp; Services Infrastructure</th>
<th>NCRIS</th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data Management: ANDS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Federation services, describing, finding, accessing</td>
<td>21</td>
<td>45</td>
</tr>
<tr>
<td>Stewardship for specific collections</td>
<td>7</td>
<td>17</td>
</tr>
<tr>
<td>Outreach for researchers and institutions</td>
<td>7</td>
<td>14</td>
</tr>
<tr>
<td>Interoperation: ICI</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Workflow, portals, collaboration services</td>
<td>20</td>
<td>32</td>
</tr>
<tr>
<td>Data and data movement services</td>
<td>8</td>
<td>13</td>
</tr>
<tr>
<td>Job submission, VMs, AAA services</td>
<td>4</td>
<td>6</td>
</tr>
<tr>
<td>Computing: NCI (APAC)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>National facility capability system</td>
<td>26</td>
<td>58</td>
</tr>
<tr>
<td>Regional application capability systems</td>
<td>18</td>
<td>38</td>
</tr>
<tr>
<td>Techniques and outreach for researchers</td>
<td>3</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>6</td>
</tr>
</tbody>
</table>

Overall summary in threes...

<table>
<thead>
<tr>
<th>New elements</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAF – the Australian Access Federation and the deployment and use of its services</td>
</tr>
<tr>
<td>ANDS – to assist teams and institutions develop and implement strategies for research data</td>
</tr>
<tr>
<td>NEAT - an e-Research architecture process that can assist align research communities, middleware and operational services</td>
</tr>
</tbody>
</table>
### Overall summary in threes...

<table>
<thead>
<tr>
<th>New elements</th>
<th>Changed functions</th>
<th>Enhanced policy goals</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAF – the Australian Access Federation and the deployment and use of its services</td>
<td>ICI - repurposing the grid to researchers across the spectrum (data and web services)</td>
<td>Harmonisation of campus, regional and national network and authorisation infrastructures</td>
</tr>
<tr>
<td>ANDS – to assist teams and institutions develop and implement strategies for research data</td>
<td>RNSPs – new role for regional service providers to include linking major research resources</td>
<td>Policies and standards that enhance e-Research and enable easier collaboration</td>
</tr>
<tr>
<td>NEAT - an e-Research architecture process that can assist align research communities, middleware and operational services</td>
<td>APAC – adding a tier of compute infrastructure to the national facility focussed on capability or discipline needs</td>
<td>Support for the growth, enhancement and sharing of e-Research expertise; (rewards, incentives and career recognition)</td>
</tr>
</tbody>
</table>
Proposed arrangement

Researchers and research communities
Resource owners and operators
Users and User-Builders

New tools, services & expertise

Data Tools & Discipline Services
Interoperation & Collaboration Infrastructure
Compute Tools & Discipline Services

Foundation Services: AREN and AAF

NEAT
National e-Research Architecture Taskforce
More services in everyday operation

Researchers and research communities
Resource owners and operators
Users and User-Builders

New tools, services & expertise

Data Tools & Discipline Services
Interoperation & Collaboration Infrastructure
Compute Tools & Discipline Services

Foundation Services: AREN and AAF
### AeRIC

**The Australian eResearch Infrastructure Council**

<table>
<thead>
<tr>
<th>Composition</th>
<th>Role</th>
</tr>
</thead>
</table>
| • An independent Chair  
• An Executive Director  
• A DEST Principal Adviser  
• 8 members drawn from providers and users of e Research capabilities  
• The executive director will have a small secretariat to assist programme management | • Determine strategic directions and variations to components  
• Convene an e Research community forum to provide advice  
• Review the annual business plans prepared by the Executive Director and PfC components  
• Make recommendations to DEST and NCRIS |

---

### Overall Development Trajectory

- NCRIS PfC process  
=> AeRIC and NEAT
- APAC National Facility  
=> National Computational Infrastructure
- SII projects + APAC grid  
=> Interoperation and Collaboration Infrastructure
- SII + 3rd party investments  
=> Australian National Data Services  
=> Australian Access Federation
- AARNet + Grangenet + Centie + ...  
=> AREN
=> Advanced Collaboration Network
Contractual Structure

AeRIC
Executive Director
NEAT
Secretariat

Computational Infrastructure
Agent

Interoperation & Collaboration Infrastructure
Agent

Data Management Infrastructure
Agent*

Contract between DEST and agent

AAF*
AARNet

Rough development path

<table>
<thead>
<tr>
<th></th>
<th>Jul</th>
<th>Aug</th>
<th>Sep</th>
<th>Oct</th>
<th>Nov</th>
<th>Dec</th>
<th>Jan</th>
<th>Feb</th>
<th>Mar</th>
<th>Apr</th>
<th>May</th>
<th>Jun</th>
</tr>
</thead>
<tbody>
<tr>
<td>AeRIC</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Membership established, first meeting 23 July, meet quarterly</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NEAT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Membership established, first meeting August</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ICI</td>
<td>NCRIS contract</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Partner JV</td>
<td>Revised plan</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NCI</td>
<td>NCRIS contract + NF</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Specialist #1</td>
<td>Revised plan</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ANDS</td>
<td>Definition</td>
<td>Members</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>NCRIS contract</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AAF</td>
<td>Establishment</td>
<td></td>
<td></td>
<td>Implementation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AREN</td>
<td>Adv AREN</td>
<td>Non-volume charging</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
ICI – Current Infrastructure

Portal Tools:
- GridSphere

Workflow Tools:
- Kepler*

Security:
- APAC CA
- MyProxy
- VOMRS

Systems:
- Gateways
- Partners’ Facilities

A virtual system of computing, data storage and visualisation facilities

ICI Partners
- sign JV

ICI Affiliates
- sign affiliate agreement

Goal is greater reach
Interoperation and Collaboration Infrastructure

ICIs J/V

Needs

Development Project

Business planning Reviews

Executive Committee

Director

User Support
Applications, Portals
Collaboration Services

NCRIS Agreement

Operation support
Data and Data Transport
Job submission, VM, AAA

NEAT
AeRIC
DEST

Initial ICI activities

Operational aspect-
- Information, Portals & Collaboration Services
- Data management and Data Transport Services
- Job submission, Gateways
- Virtualisation
- AAA
- Help desk

Development aspect-
- Next generation tools and services
- NEAT determined
Interoperation and Collaboration Infrastructure

• Member classes
  – Partners, operate resources and services provided by ICI
  – Affiliates, integrate resources with core infrastructure

• Agreements
  – NCRIS agreement with VPAC
  – Partners sign unincorporated J/V to implement ICI Program
  – Affiliate membership executed with any Partner

• Executive Committee
  – Chair appointed by Executive Committee
  – A nominee of every Partner
  – A nominee elected by the Affiliates
  – The ICI Director (appointed by the Executive Committee)
  – The ‘other’ PfC Directors

National Computational Infrastructure

Peak system

<table>
<thead>
<tr>
<th>A strategic requirement</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Demands expert users</td>
</tr>
<tr>
<td>• Capability use limits users</td>
</tr>
</tbody>
</table>

Decision points

| Need $4M per annum to buy capability |
| Double that to operate, add outreach… |
| So $10M pa – but can share costs |
| ANU, CSIRO, … will buy shares |
| Expertise is rare |
| Current facility works, has international standing and reputation |
| Relocation would be damaging |
| ANU have agreed to continue |
### National Computational Infrastructure

<table>
<thead>
<tr>
<th>Peak system</th>
<th>Specialist systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>A strategic requirement</td>
<td>An effectiveness investment</td>
</tr>
<tr>
<td>• Demands expert users</td>
<td>• Support applications and tools needed by a broader community of less expert users</td>
</tr>
<tr>
<td>• Capability use limits users</td>
<td>• Locate near investments and expertise in applications and related research</td>
</tr>
<tr>
<td><strong>Decision points</strong></td>
<td>• Gain higher leverage (States, Institutes)</td>
</tr>
<tr>
<td>• Need $4M per annum to buy capability</td>
<td>• Contenders: Bio, Nano, Physics, Imaging, Chemistry, Materials, Engineering</td>
</tr>
<tr>
<td>• Double that to operate, add outreach…</td>
<td><strong>Issues</strong></td>
</tr>
<tr>
<td>• So $10M pa – but can share costs</td>
<td>• Need applications+tools+communities</td>
</tr>
<tr>
<td>• ANU, CSIRO, … will buy shares</td>
<td>• Need to strengthen regional providers</td>
</tr>
<tr>
<td>• Expertise is rare</td>
<td>• Target at 25% of peak, $2M pa each</td>
</tr>
<tr>
<td>• Current facility works, has international standing and reputation</td>
<td>• Could see 8-10 but do 3 (and therefore evaluate community readiness)</td>
</tr>
<tr>
<td>• Relocation would be damaging</td>
<td>• ANU have agreed to continue</td>
</tr>
</tbody>
</table>

### Some potentials…

- Geo
- Bio
- Generic
- Social
- Climate
- Engineering
Some potentials… but everybody wants bio…

National Computational Infrastructure
National Computational Infrastructure

- **Member classes**
  - Partners, operate or co-fund NCI facilities
  - Affiliates, have an interest in HPC

- **Agreements**
  - NCRIS agreement with ANU
  - Partners have contracts with ANU related to NCI functions
  - Affiliate membership executed with ANU

- **Executive Committee**
  - An independent Chair
  - A second independent member
  - A nominee of each Partner, a nominee of the Affiliates
  - The NCI Director & the chair of the merit allocation committee

The Australian National Data Service

- **A** Community agreed data investment; intended for sharing
- **B** CRCs, MNRFs, Centres; Data may be used, shared or published
- **C** Collections; Data may be shared, local repositories
- **D** Independent research; Data used and retained
- **E** Services: Find, Access, Authorise
- **F** Services: stores, repositories, tools, expertise

NUMBER

SIZE
### Responsibilities

<table>
<thead>
<tr>
<th>Institutions</th>
<th></th>
<th>Researchers</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Develop and implement policy on data ownership</strong></td>
<td></td>
<td><strong>Determine what data to keep, considering research community practice and any project or legal requirements</strong></td>
</tr>
<tr>
<td><strong>Provide guidelines to researchers on ownership and their responsibilities</strong></td>
<td></td>
<td><strong>Ensure research data is retained (5 years from publication of results) using institutionally provided mechanisms</strong></td>
</tr>
<tr>
<td><strong>Maintain durable records on data held and ensure that it is under the control of the institution where the work was performed</strong></td>
<td></td>
<td><strong>Ensure that data retention passes to the institution</strong></td>
</tr>
<tr>
<td><strong>Provide secure systems for holding data and for granting access to that data</strong></td>
<td></td>
<td><strong>Keep confidentiality where it exists</strong></td>
</tr>
</tbody>
</table>
### Responsibilities

<table>
<thead>
<tr>
<th>Institutions</th>
<th>Researchers</th>
<th>ANDS</th>
</tr>
</thead>
</table>
| • Develop and implement policy on data ownership  
• Provide guidelines to researchers on ownership and their responsibilities  
• Maintain durable records on data held and ensure that it is under the control of the institution where the work was performed  
• Provide secure systems for holding data and for granting access to that data | • Determine what data to keep, considering research community practice and any project or legal requirements  
• Ensure research data is retained (5 years from publication of results) using institutionally provided mechanisms  
• Ensure that data retention passes to the institution  
• Keep confidentiality where it exists | • Provide implementations of the required federated services (E and F)  
• Broker solutions for collections to fast track e-Research development where appropriate (C)  
• Build expertise and provide outreach services that can assist others, ‘training the trainers’  
• Ensure promulgation and use of relevant (simple) legal frameworks and access templates |

---

### The Australian National Data Service

![Diagram of the Australian National Data Service](Diagram.png)

- **NEAT**  
- **AeRIC**  
- **DEST**  
- **ANDS**  
- **Business planning Reviews**  
- **Executive Committee**  
- **Agent(s)**  
- **Stewardship**  
- **Outreach**

**Needs**  
**Technical Services**  
**Mng**  
**Development**  
**Operation**  
**Service agreement**  
**NCRIS Agreement**

---

19
The Australian National Data Service

- **Member classes**
  - ANDS Agent(s), operate technical services, develop community
  - Partners, operate stewardship facilities for ANDS
  - Affiliates, provide outreach services for ANDS

- **Agreements**
  - NCRIS agreement with ANDS Agent(s)
  - Partners sign service agreement with ANDS Agent(s)
  - Affiliates sign service agreement with ANDS Agent(s)

- **Executive Committee**
  - Chair elected from Executive Committee
  - Nominees of the Partners
  - Two nominees elected by the Affiliates
  - The ANDS Director (appointed by the Executive Committee)

---

Co-operative funding

- Researchers and research communities
- Resource owners and operators
- Users and User Builders

- Foundation Services: AREN and AAF

NCRIS funds national components, and co-funds other components, noting that foundation services are member funded

- New tools, services & expertise
- Data Tools & Discipline Services
- Interoperation & Collaboration Infrastructure
- Compute Tools & Discipline Services

- Research supported activities
- Focussed development
- Jointly supported activities
- PfC activities
Platforms for Collaboration

Back to...

... the big picture
Towards the Future - It ain’t what it used to be

- National Collaborative Research Infrastructure Strategy, Platforms for Collaboration:
  - access to leading edge research infrastructure nationally and internationally
  - access to distributed and large datasets
  - advanced modelling & analysis opportunities
    - integration of datasets and computational resources
    - transdisciplinary and interdisciplinary research
  - innovation from data rather than data from innovation
    - shorter innovation cycles
    - new forms of ‘publication’
    - new forms of peer review
  - new forms of collaboration
    - data rather than models or discipline

The Future - eResearch Strategic Directions

- Continuing Need for a Focus
  - through a nationally coordinated approach
- Human Capabilities
  - people with skills and understanding
- Linkage of eResearch Resources
  - seamless access to resources
- Access to Data
  - adopt best practice data management and curation
- Structural and Cultural Change
  - evolution of organisational structures and cultures
- Awareness and Support
  - develop researchers’ ability to adopt eResearch methods
The Future - Physical Platforms

- **Networks:**
  - AREN + Institutional + Commercial + Wide-area wireless
- **Computation:**
  - National facilities + Institutional + grid services
- **Data stores:**
  - ‘Central’ + Institutional
- ‘**Instruments**’
  - International, national, institutional

The Future - eResearch R&DD Issues

- Resource scheduling, marshalling, management, monitoring
- Intelligent and flexible search for resources
- Data management, access, integration
- Security and integrity
- Collaborative tools
- Simulation, modelling, visualisation
- Management of complex and intelligent systems
- Remote operation
The Future - The Hard Issues

- Cultural re-engineering - the institutional factor
- Capability development - the human factor

The Future - Cultural Re-engineering

- Competition vs collaboration vs sharing
  - individual
  - institutional
  - sectoral
  - funding processes
- Discipline traditions vs discipline future
- Universities cf government and industry research performers
- University IT as an enabler cf administrative service
The Future - Role of the Research Institution

• *Host and/or Owner* of intellectual capital
  – critical mass created in a trans-institution manner

• *Repository of knowledge (eg publications) or data*
  – platforms of the past or for the future

• *Competitor, ally or promiscuous collaborator*
  – control of personal and virtual collaboration?
  – individual pre-eminence vs collaborative pre-eminence

• *Provision of resources*
  – physical, access, human support

• *Integrity management*
  – ethics, identity

The Future - Capability Development

• *Researcher confidence & proficiency*
  – underlying assumptions and work arounds

• *Researcher support*
  – domain-ICT linking professionals
  – career structures and positioning

• *Platform evolution*
  – ICT expertise
  – international linkage