Research Methods – The Great Integrator

Rhys Francis
eResearch Futures
rhys_francis@icloud.com

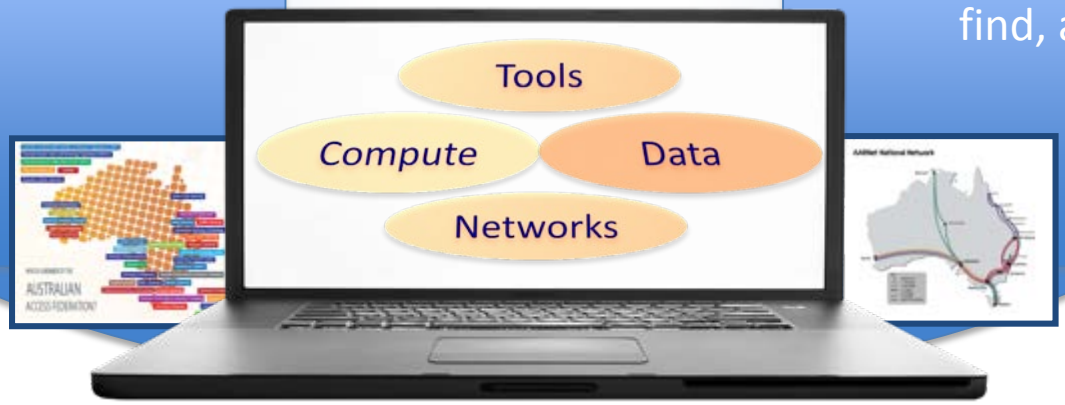
Use new tools, apps, work remotely and collaborate in the cloud

NeCTAR

eResearch Infrastructure

Do computational modeling, complete data analysis, visualize results

NCI
Pawsey



Keep data and observations, describe, collect, share, find, and re-use them

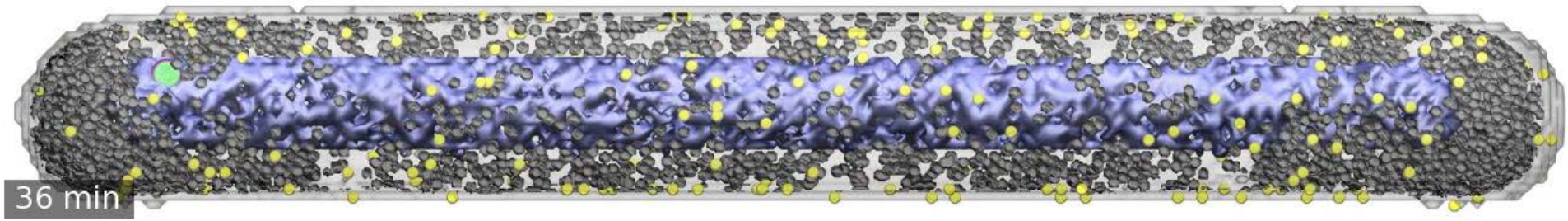
ANDS
RDSI

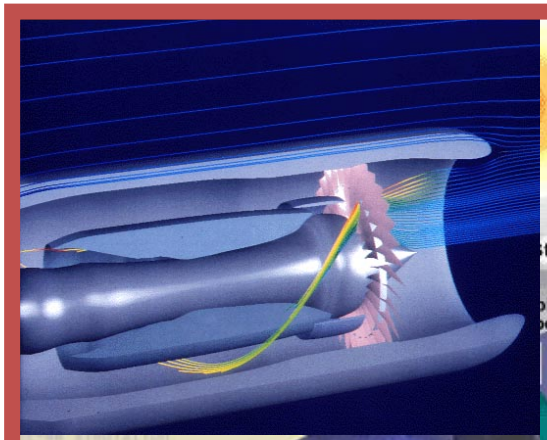
Understand mechanisms impossible to observe or experiment with directly

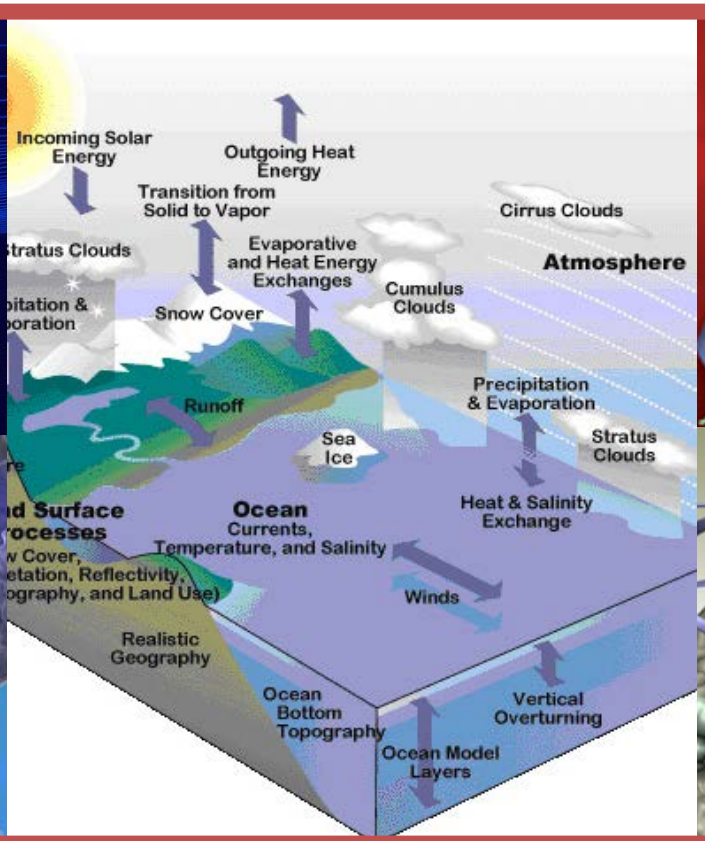
Undertake novel research studies more extensive than ever before

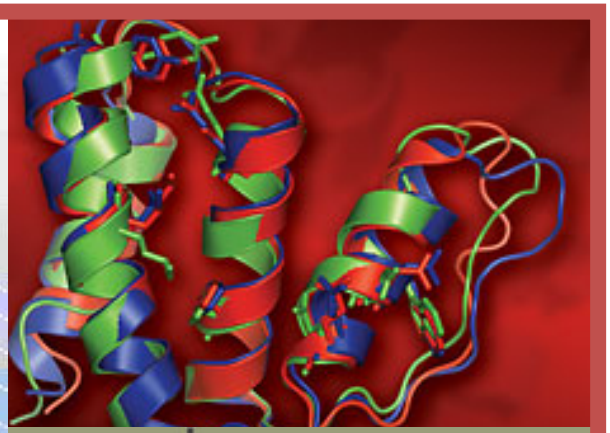
Generate new theories using data at scales previously inconceivable

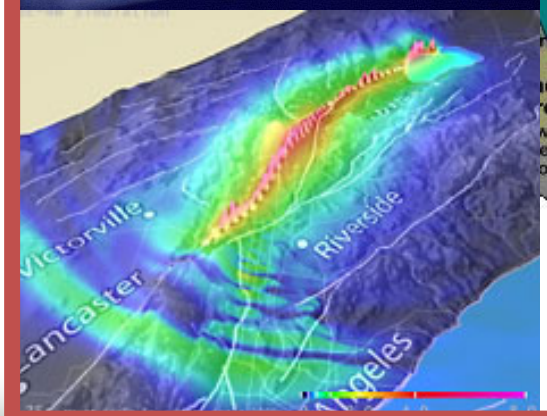
Computation and modelling

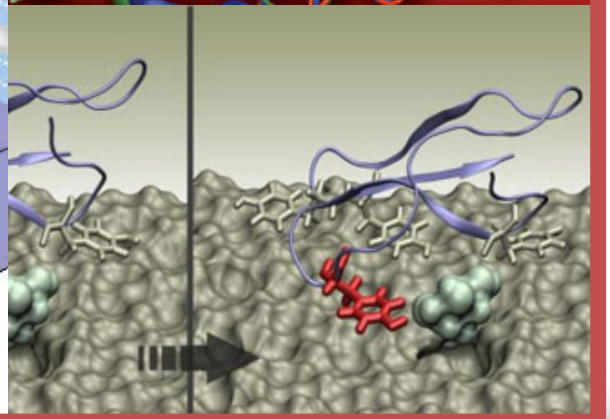




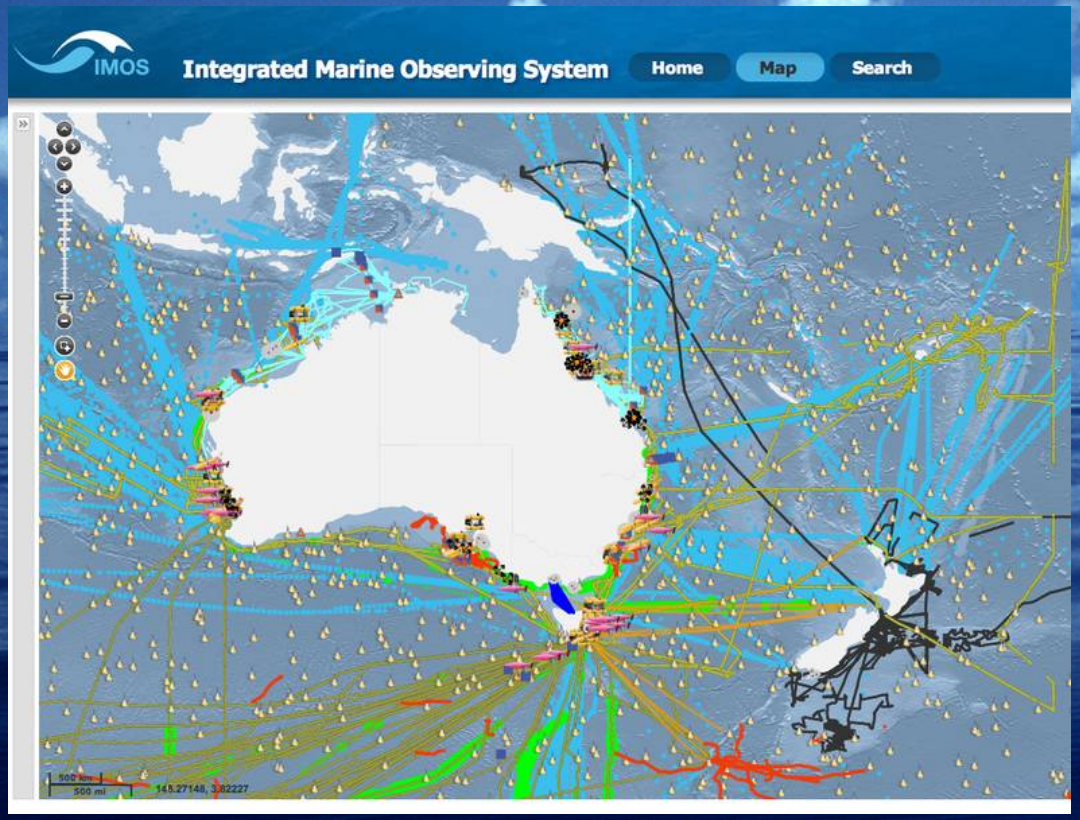


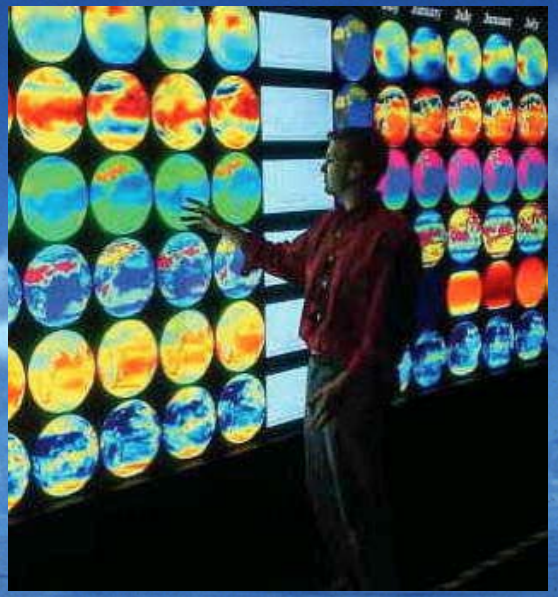






Data is king

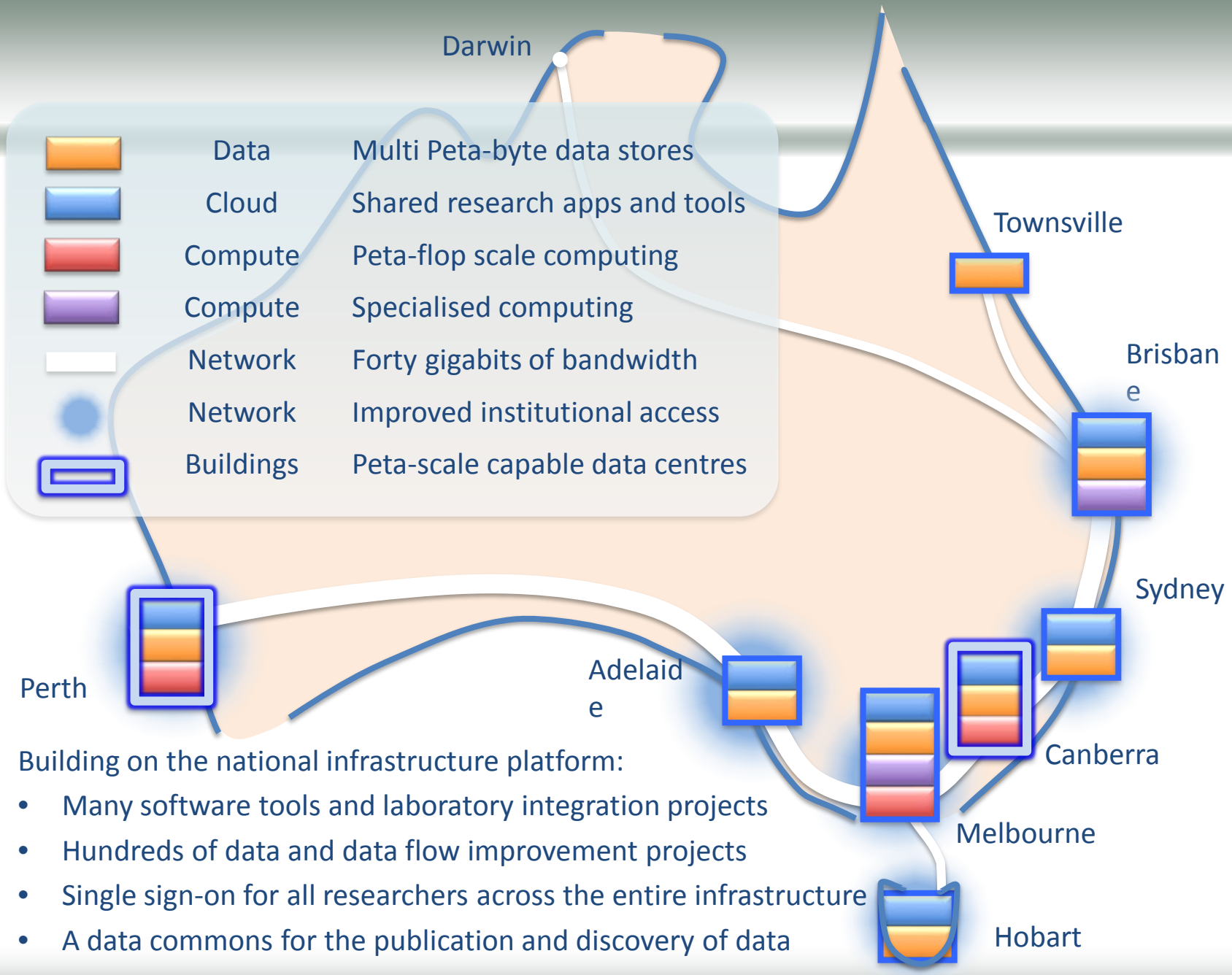




Research teams– biosecurity access

Scale Of National Investment

Period	Australian Government Investments	Overall	eResearch
2001 - 2005	MNRF and Strategic Infrastructure Initiative	$400M	$180 M
2005 - 2011	National Collaborative Research Infrastructure Strategy	$542M	$80 M
2009 - 2014	Super Science Initiative	$1,100M	$312 M
2012 - 2016	Continuity (CRIS, NCRIS2013, NCRIS2015)	~$390M	~$75 M
2016 - 2017	NCRIS 2016	$150M	?
	… and beyond		



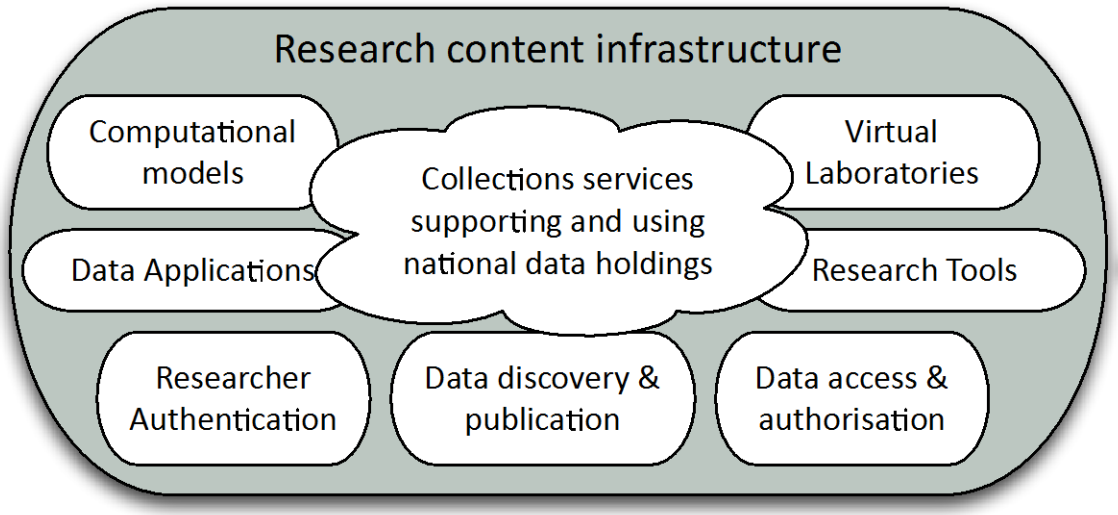
Building on the national infrastructure platform:

- Many software tools and laboratory integration projects
- Hundreds of data and data flow improvement projects
- Single sign-on for all researchers across the entire infrastructure
- A data commons for the publication and discovery of data

Inside the eResearch Infrastructure Activities

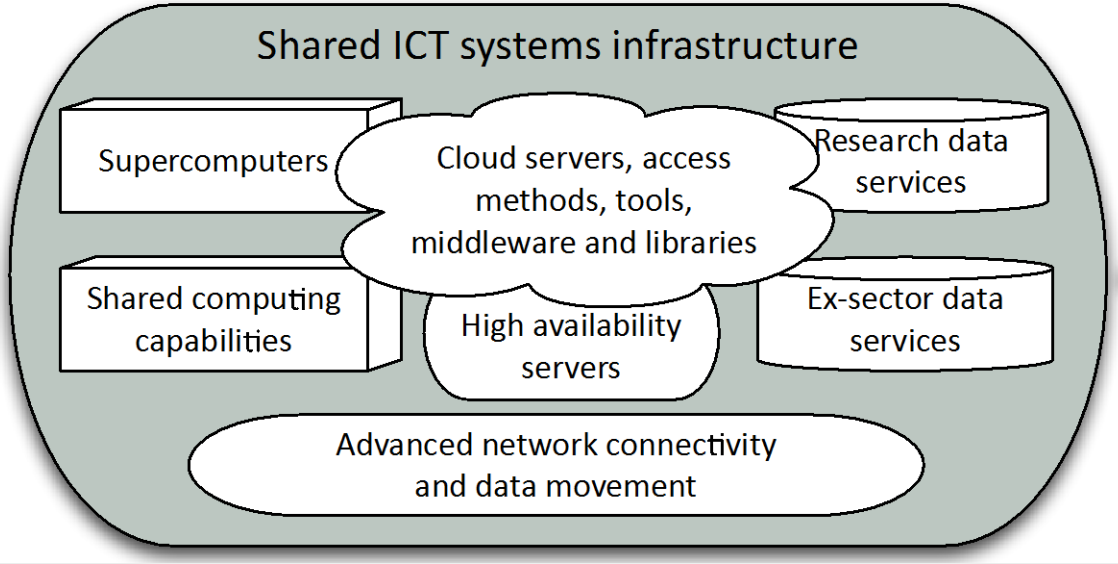
Tools	Provide a dedicated research cloud computing resource • associated with data intense centres Create a library of re-usable research Apps in the cloud Assist research communities create virtual laboratories in the cloud
Data	Create a national corpus of published research data assets Sustain a sector wide commitment to improved research collections Create a network of high capacity, scalable, data storage centers
Computation	Construct two buildings for next generation supercomputers • fit for purpose over next 10 years Install new systems for research use • Each providing peta-scale plus computation resources
Networks	Expand connectivity to the research network backbone • improve regional and metropolitan access Extend the fibre backbone - connect east and west coasts • connect through to the ASKAP and SKA sites

The next generation to come…



Content Capabilities

- problem or discipline oriented capabilities
- part of differentiated national and global regimes
- comprising general and specialised data and tools
- supported by multiple institutions
- mapped to approrpiate systems infrastructure



Systems Capabilities

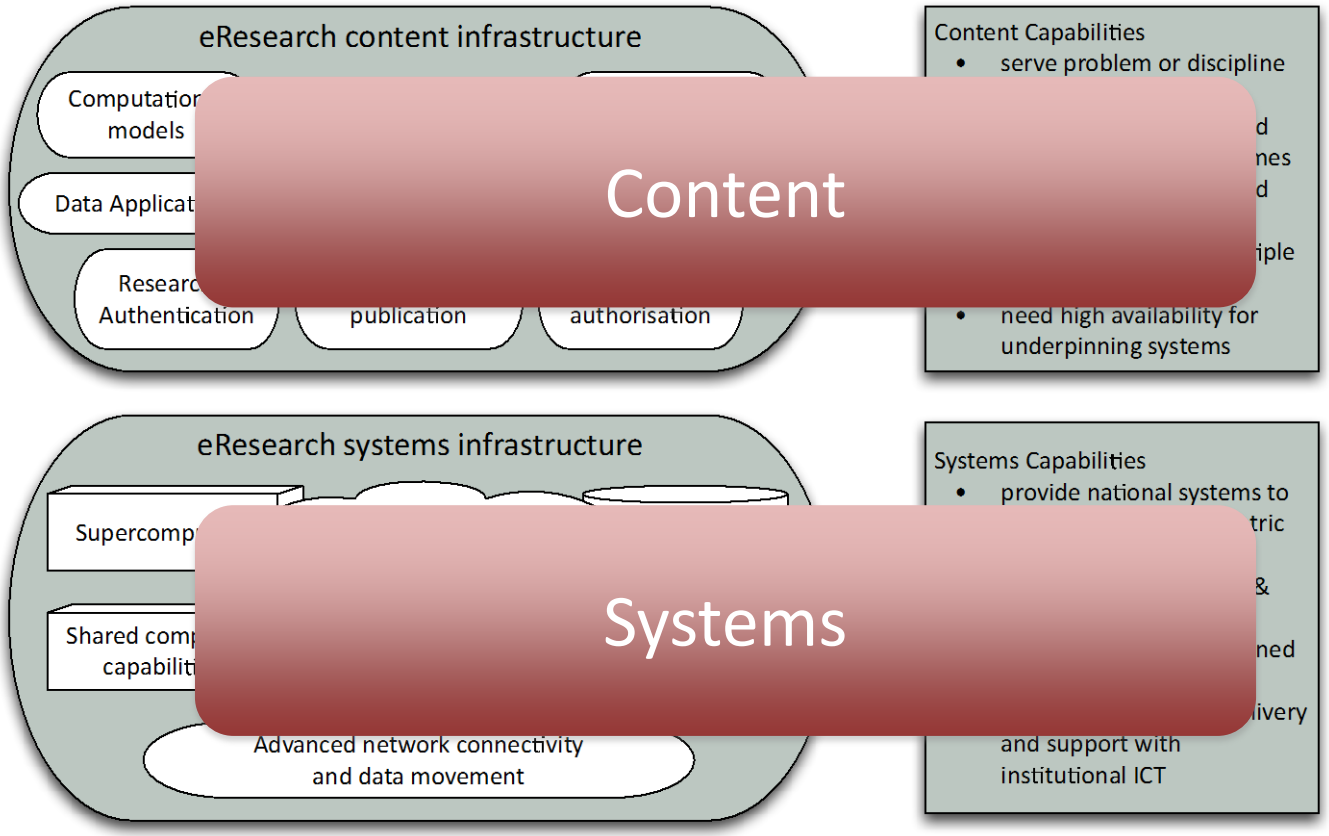
- provision of national systems that support informatics centric research activities
- coordinated & interoperable
- owned and/or governed by sector participants
- integrated for delivery and support with institutional ICT

Driving the need for new skills…

Long term value resides in

- Data
- Tools and
- Content Skills

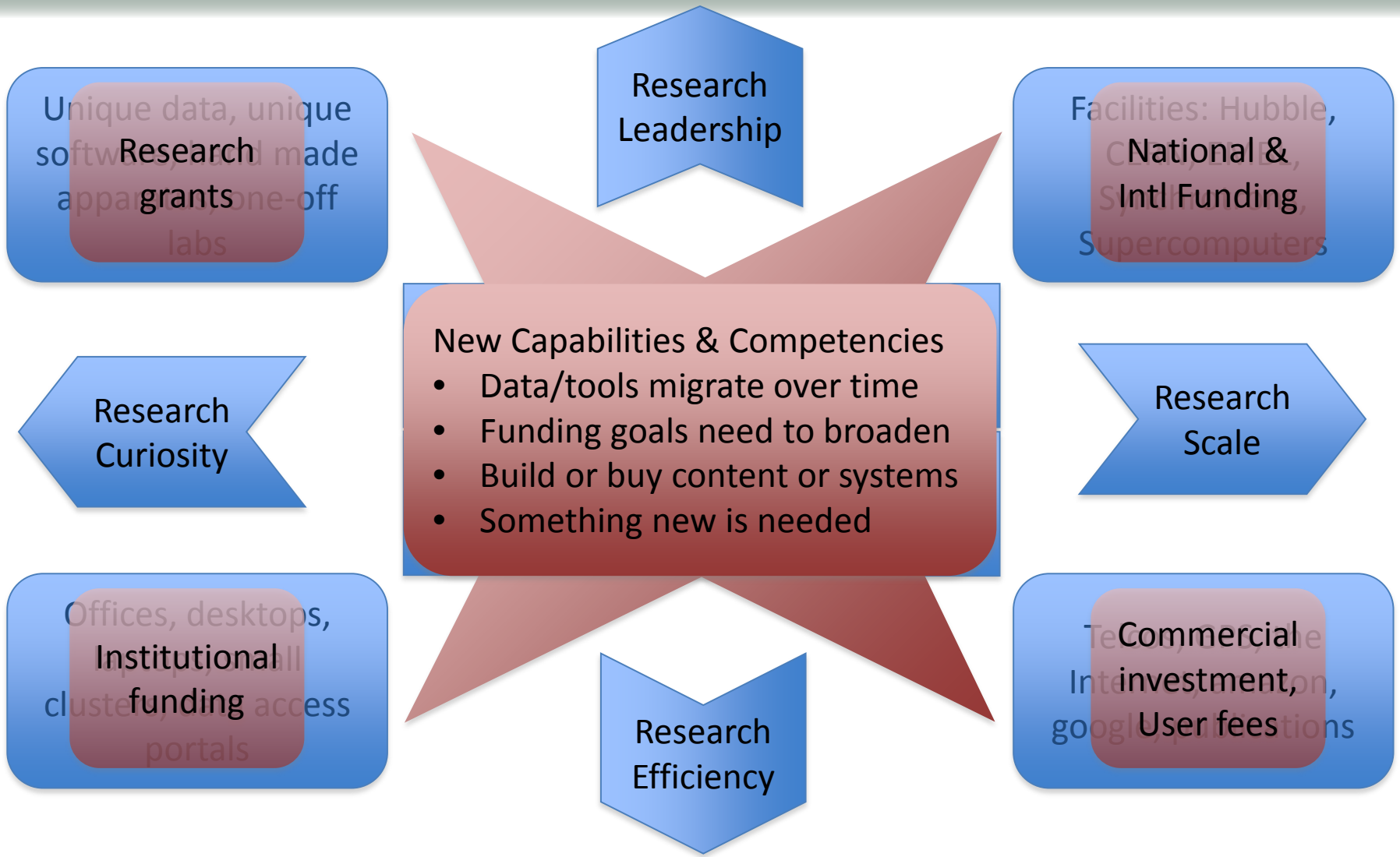
ICT systems complexity, especially at scale, can easily dominate e-infrastructures.



Therefore we need to invest more directly in long lived content

Data & Tools & Skills

The Elephants in the Room



Why these categories endure

Methods

Data

Systems

Networks

- Different maturity
  => different approaches
- Different natural owners
  ⇒Research communities
  ⇒Institutions
  ⇒Facilities
  ⇒Preferred supplier
- Investment value
  => Inverted to maturity

What have we learned?

2002 Networks Compute

2007 Networks Compute Data (Tools)

We energised the “value“ of data separately

2012 Networks Systems Data Tools

We have to bring compute and data “systems” together

And the maturity of tools is a limiting factor

2020 Networks Systems Data Methods

Future science depends on properly qualified Methods

Methods = Data+Tools on suitable Systems+Networks

The re-use value of data leaves it as a unique asset

Where to from here……

Obvious questions

- What e-infrastructure does research need?
- How much of what is optimum?
- How do we organise and provide and support its use
- Where consumable or limited in capacity, how do we ration or prioritise its use?
- How do we continue to know the answer to these questions?

More interesting Questions?

- How joined up should all this be? How do we get frictionless infrastructure?
- Do more people care? if so, what do people new to caring about e-infrastructures for research care about?
- Where does the data researchers use come from and where does data they produce go?
- When researchers use data where do they do that?
- How do we share the knowledge encoded in our data and tools with government, health, industry… etc?
- Whose money does what? What does NCRIS money do and not do?
- How can research e-infrastructure benefit from commercial spending?

General outline of the planning activity

- Terms of Reference by end of July
- Project Committee during August
- Invite your inputs on framework topics
- Develop Discussion Paper by October
- Your written responses this year
- Draft framework by March
- Your written commentary by April
- You leverage your international linkages
- You drive the ideas
- You participate in review meetings