

eResearch: Building the Scientific Instruments of the 21st Century - 10 Lessons Learned

Prof C. Paul Bonnington

Director, Monash e-Research Centre

Monash University (Sites: Australia, South Africa, Italy, Malaysia, India, China)



About Paul

- A New Zealander and All Black Supporter
- Academic Background: Mathematician
 - *Computational (Combinatorics)*
 - *with, e.g., applications to Bioinformatics*
- Post Docs in Austria and USA
- University of Auckland, NZ 1992-2007
- Monash University, Australia 2008-
- First eResearch Director in NZ, established the New Zealand National eResearch Programme in 2003 (became “BeSTGRID”, and now “NESI”)
 - *Initially a partnership between Massey, Auckland and Canterbury Universities*
- Now Director of Australia’s first and largest University eResearch Centre (45 staff) - Monash eResearch Centre



Monash today



- Over 63,000 students, ~13,000 are international
- Over 8,000 staff (FTE)
- Ten faculties with a metropolitan, regional and international presence
- Six Australian campuses
- Global footprint:
 - Prato Centre, Italy
 - IITB Research Academy (India)
 - South East University-Monash University Joint Graduate School (Suzhou, China)
 - Sunway (Malaysia)
 - South Africa
 - Warwick Alliance

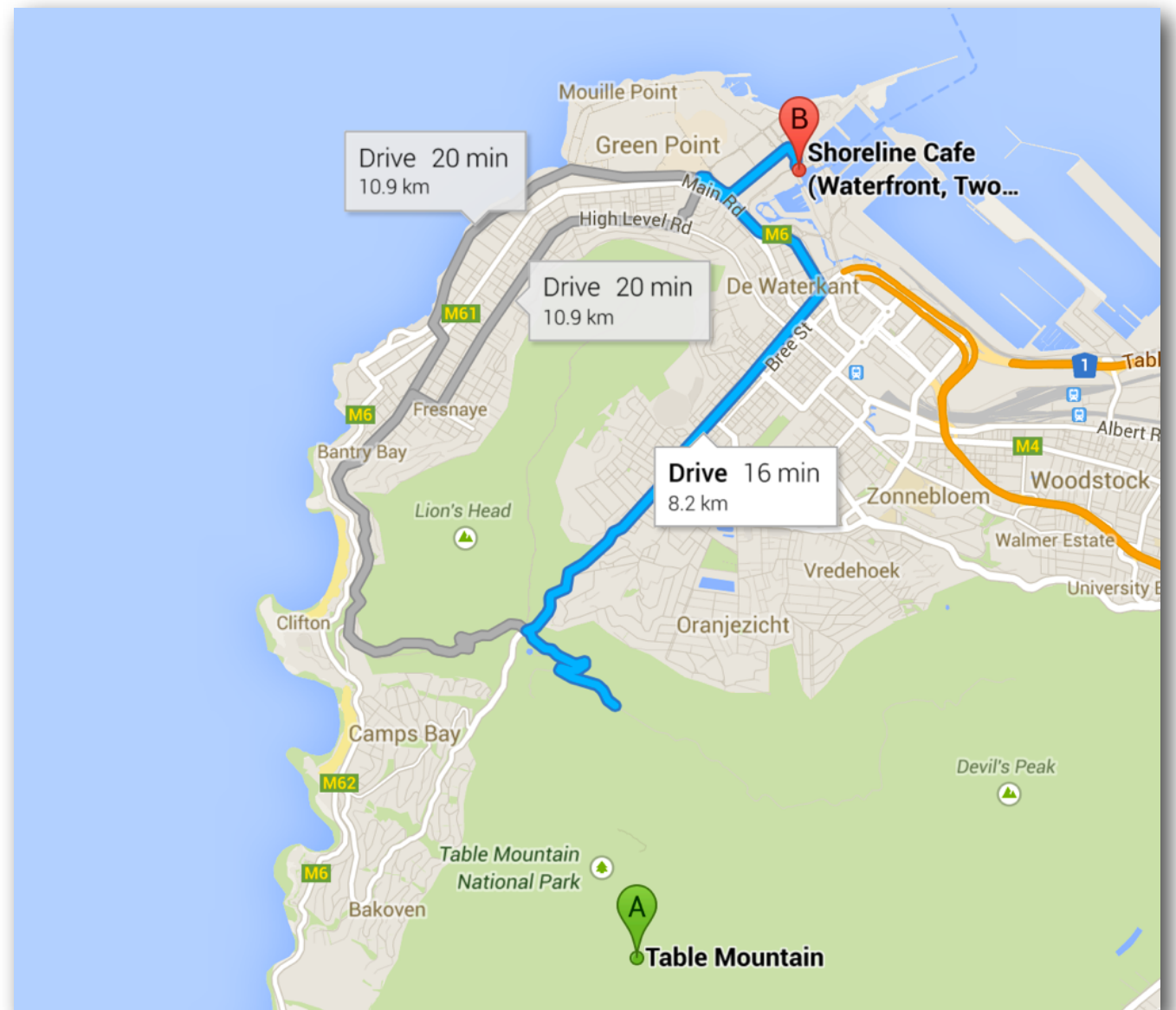
- Ranked in top 1% of universities worldwide, and e.g. a top 5 university globally in Pharmaceutical Sciences....

eResearch: Building the Scientific Instruments of the 21st Century - 10 Lessons Learned

- Personal Reflections on 10 years of building eResearch
Collaborations, Infrastructure, Services and Relevance to Researchers

Lesson 0. eResearch is Exciting!

- Research and IT together =
- Car chases around the streets of beautiful cities in exotic places





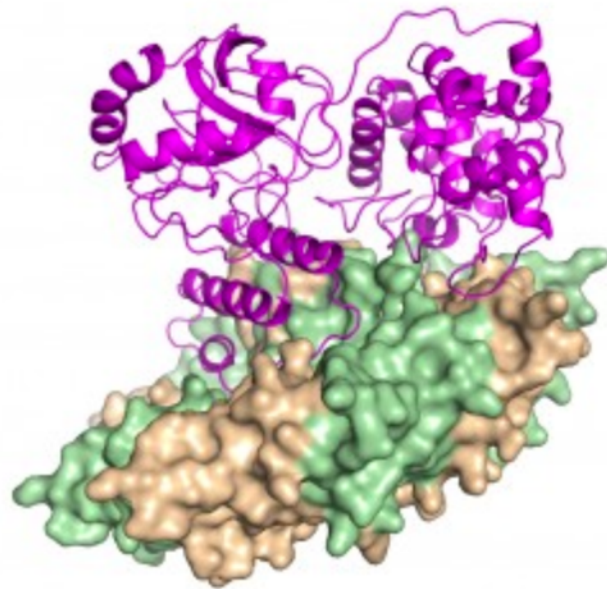
Lesson 1: Research is the “Journey into the Unknown”

A story.....

Proceedings of the National Academy of Sciences of the United States of America

X-ray crystal structure of the streptococcal specific phage lysin PlyC

Sheena McGowan^{a,1}, Ashley M. Buckle^{a,1}, Michael S. Mitchell^{b,1}, James T. Hoopes^b, D. Travis Gallagher^c, Ryan D. Heselpoth^b, Yang Shen^b, Cyril F. Reboul^a, Ruby H. P. Law^a, Vincent A. Fischetti^{d,2,3}, James C. Whisstock^{a,e,2,3}, and Daniel C. Nelson^{b,d,f,2,3}



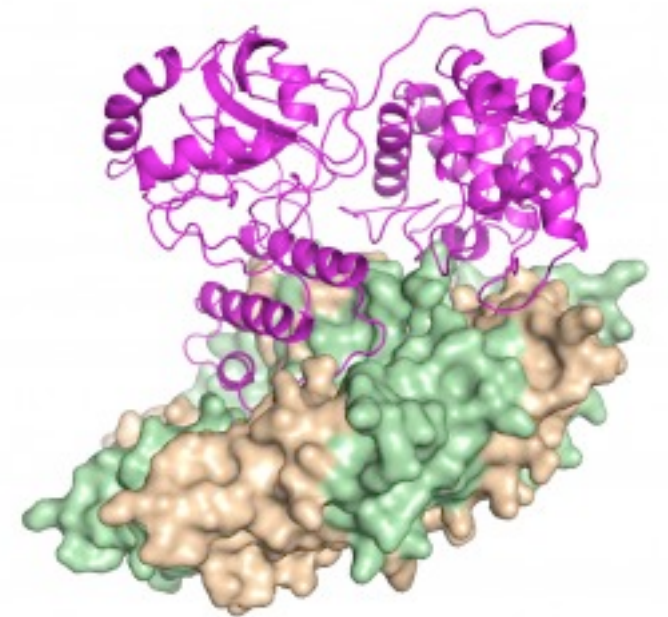
An illustration of the bacteriophage lysin Ply C

^aDepartment of Biochemistry and Molecular Biology, Monash University, Clayton 3800, Australia; ^bInstitute for Bioscience and Biotechnology Research, Rockville, MD 20850; ^cNational Institute of Standards and Technology, Rockville, MD 20850; ^dLaboratory of Bacterial Pathogenesis, New York University, New York, NY 10021; ^eAustralian Research Council Centre of Excellence in Structural and Functional Microbiology, Clayton 3800, Australia; and ^fDepartment of Veterinary Medicine, University of Maryland, College Park, MD 20742



The PlyC Story

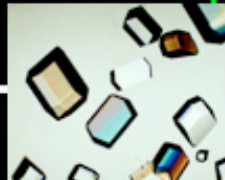
- discovered the structure and operating procedures of a powerful anti-bacterial killing machine that could become an alternative to antibiotics
- The real story.....
- “useless” data capture in 2008
- data sat unused until 2011 and was going to be “thrown away” because no analysis tools existed
- eResearch infrastructure preserved it from point of capture



Structural Biology Pipeline

Cloning Purification
Expression Crystallisation

DNA

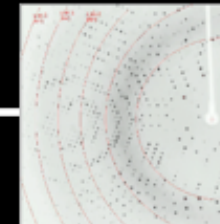


High throughput robotics
and technologies

X-ray diffraction



Australian
synchrotron

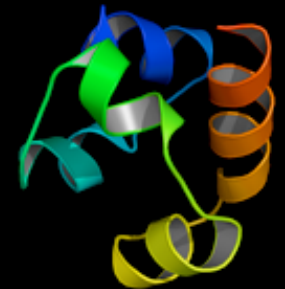


Data
Grid

Determine structure



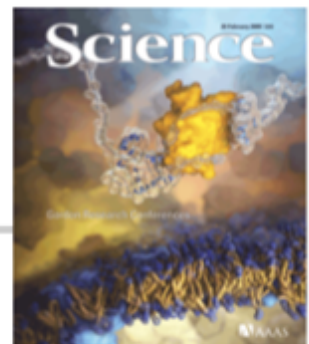
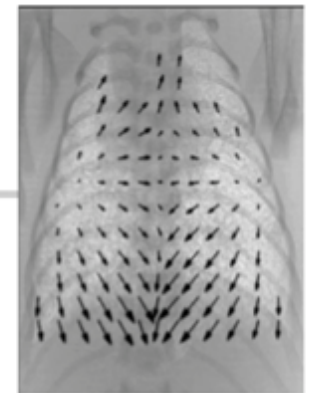
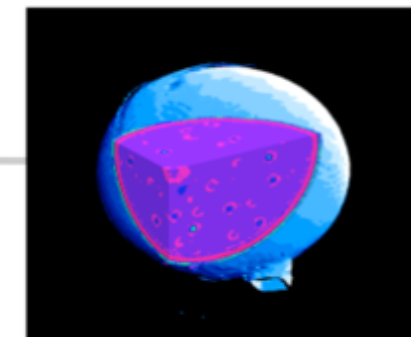
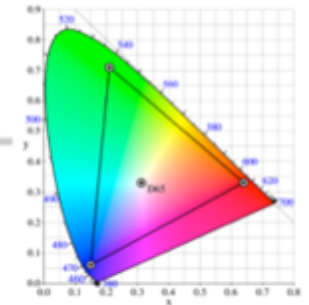
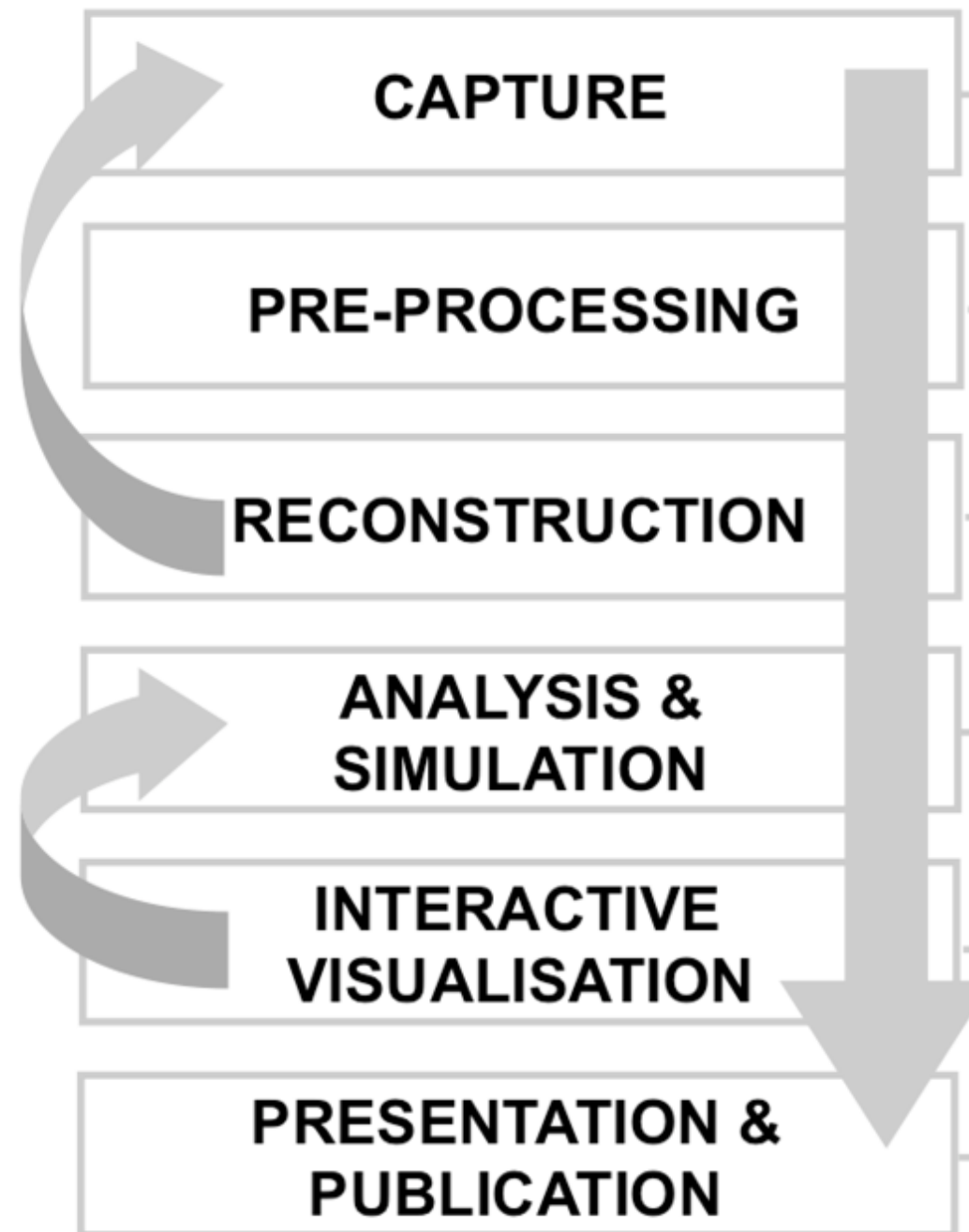
Data processing & structure
determination



HARDWARE

SOFTWARE

EXPERTISE



[\[Register\]](#) [\[Log In\]](#)

X-ray crystal structure of the streptococcal specific phage lysin PlyC

Description

Datasets (2)

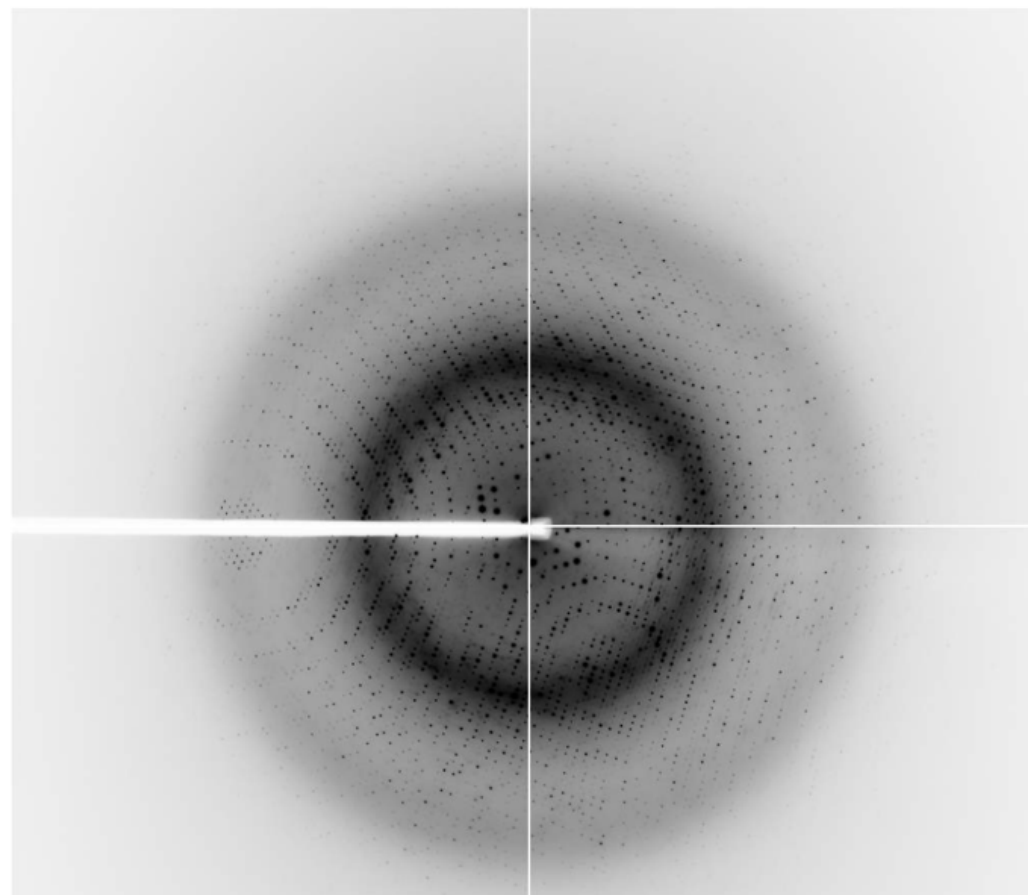
Authors:

Sheena McGowan, Ashley Buckle, James Whisstock

Abstract:

DOI for this data: **10.5072/03/50171CC7EEC1D**

Sample diffraction image (PlyCB: plycb0001.img):



[Browse by Subject Area](#)[Advanced Search](#)[All](#)[Collections](#)[Parties](#)[Activities](#)[Services](#)

1 results (0.003 seconds)

Page: 1/1 | [First](#) [1](#) [Last](#)



X-ray crystal structure of the streptococcal specific phage lysin PlyC

Bacteriophages deploy lysins that degrade the bacterial cell wall and facilitate virus egress from the host. When applied exogenously, these enzymes destroy susceptible microbes and,...

1 results (0.003 seconds)

Page: 1/1 | [First](#) [1](#) [Last](#)

Search for Research Data



Browse by Subject Area



Browse by Map Coverage



Advanced Search

Home / Monash University / Collection

X-ray crystal structure of the streptococcal specific phage lysin PlyC

Bacteriophages deploy lysins that degrade the bacterial cell wall and facilitate virus egress from the host. When applied exogenously, these enzymes destroy susceptible microbes and, accordingly, have potential as therapeutic agents. The most potent lysin identified to date is PlyC, an enzyme assembled from two components (PlyCA and PlyCB) that is specific for streptococcal species. Here the structure of the PlyC holoenzyme reveals that a single PlyCA moiety is tethered to a ring-shaped assembly of eight PlyCB molecules. Structure-guided mutagenesis reveals that the bacterial cell wall binding is achieved through a cleft on PlyCB. Unexpectedly, our structural data reveal that PlyCA contains a glycoside hydrolase domain in addition to the previously recognized cysteine, histidine-dependent amidohydrolases/peptidases catalytic domain. The presence of eight cell wall-binding domains together with two catalytic domains may explain the extraordinary potency of the PlyC holoenzyme toward target bacteria.

Identifiers

Local: experiment/view/794



Share

Access

<https://mytardis.its.monash.e...>

Access rights

This experiment was closed access at time of registration.

Rights statement

Creative Commons Attribution-Non-Commercial 2.5 Australia

Contacts

Monash University Clayton Campus
Building 75 Clayton 3800 Victoria

Connections

Researchers

Prof James Whisstock

Dr Ashley Buckle

Dr Sheena McGowan

[\[Register\]](#) [\[Log In\]](#)

X-ray crystal structure of the streptococcal specific phage lysin PlyC

Description

Datasets (2)

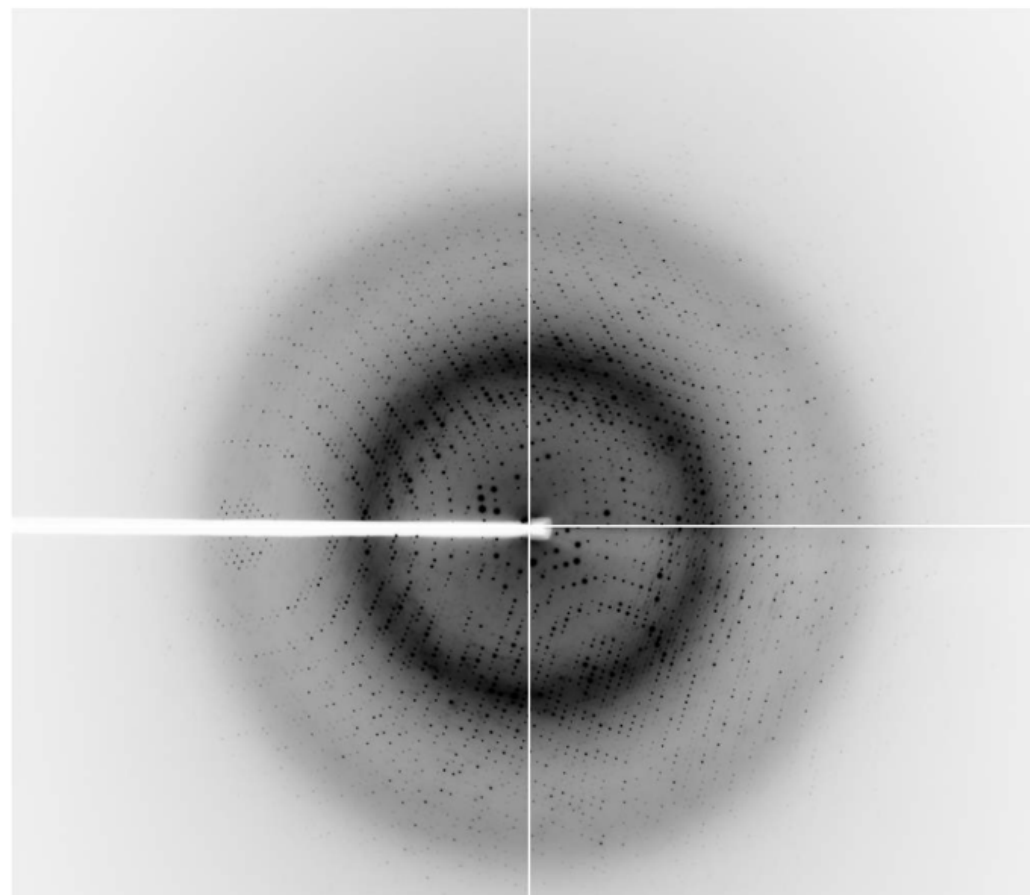
Authors:

Sheena McGowan, Ashley Buckle, James Whisstock

Abstract:

DOI for this data: **10.5072/03/50171CC7EEC1D**

Sample diffraction image (PlyCB: plycb0001.img):



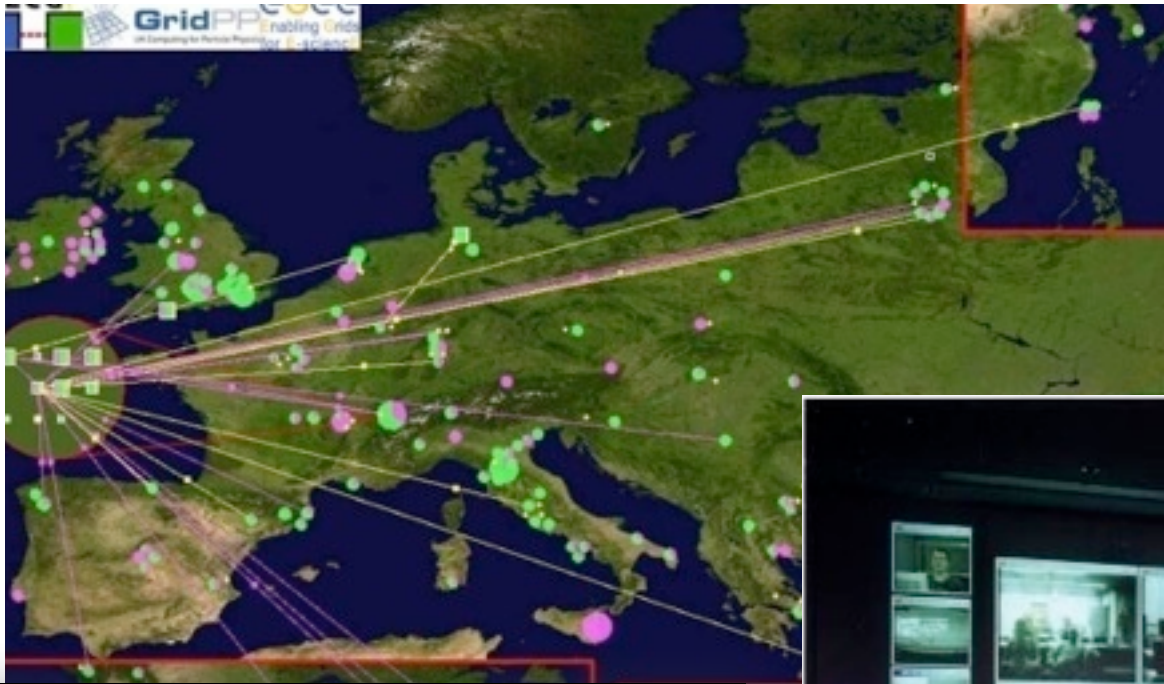
Lesson 1: Research is the “Journey into the Unknown”

- Investment in the unknown future
- Not *losing* data, and *reusing* data
- Challenge: how do you build and sustain capabilities and support research where
 - *there is sometimes no “business” case (why do it?)*
 - *there is sometimes no business model (who pays?)*
 - *ambiguities abound*
- the answer is coming.....

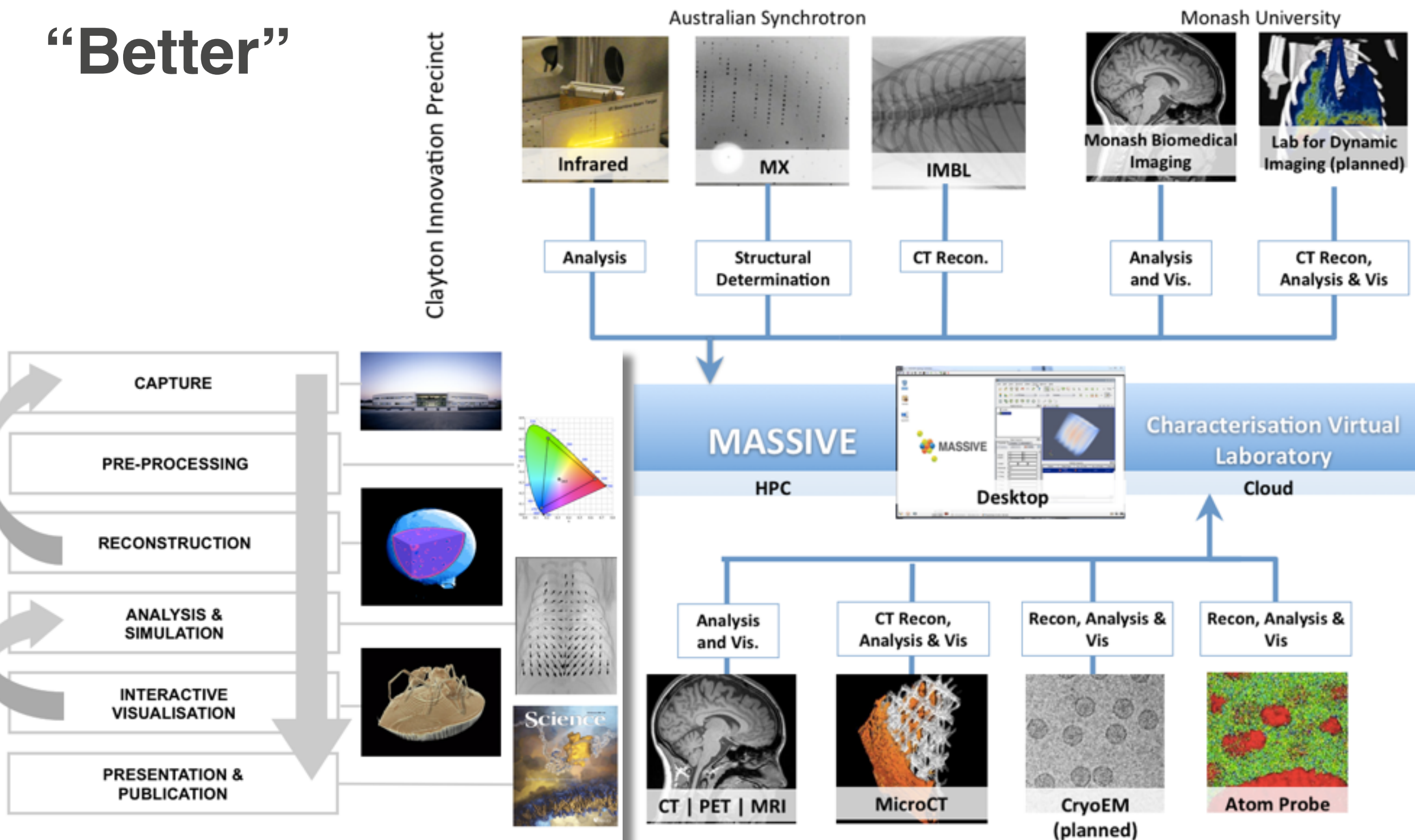
Lesson 2: The institution (university) is key

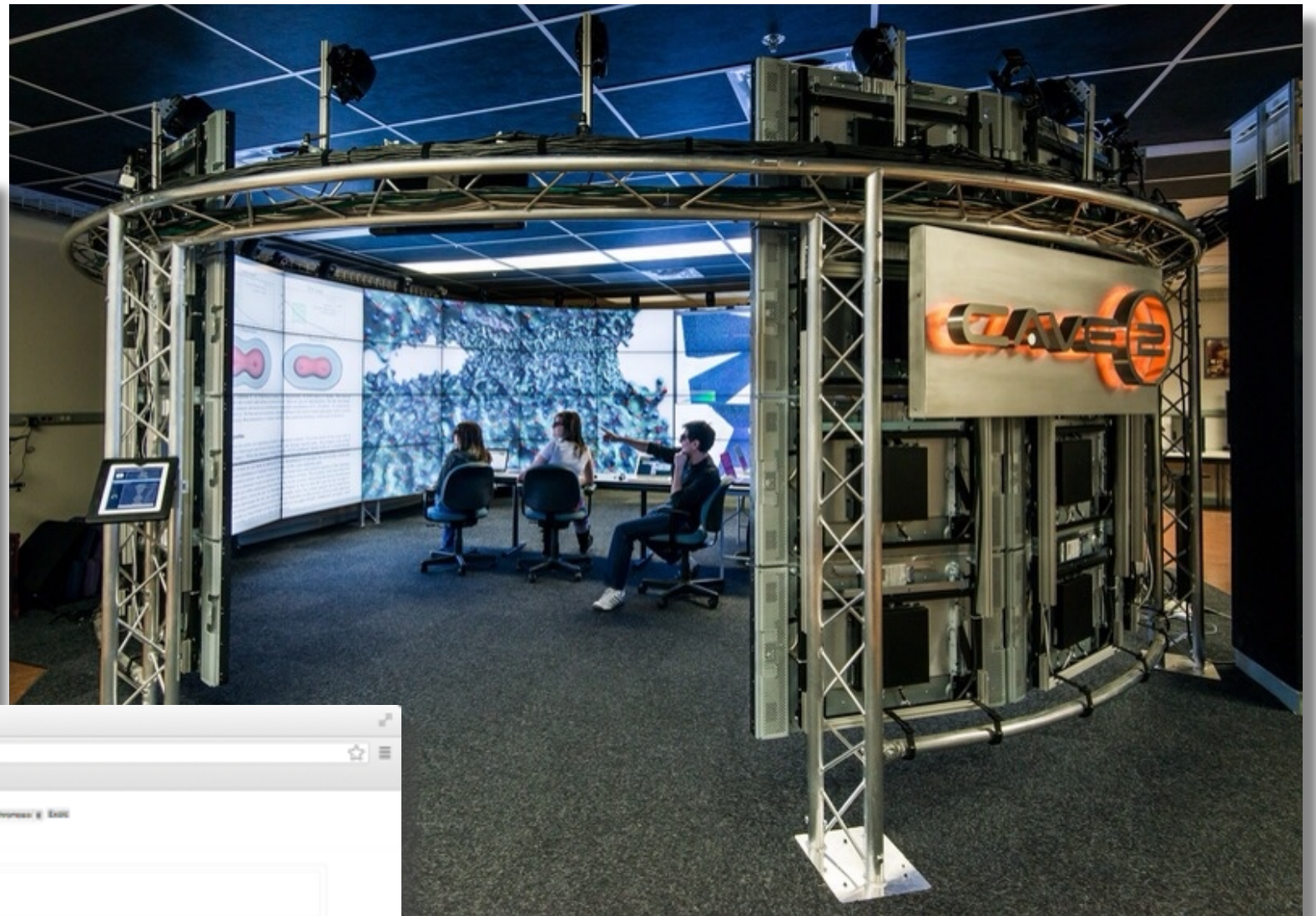
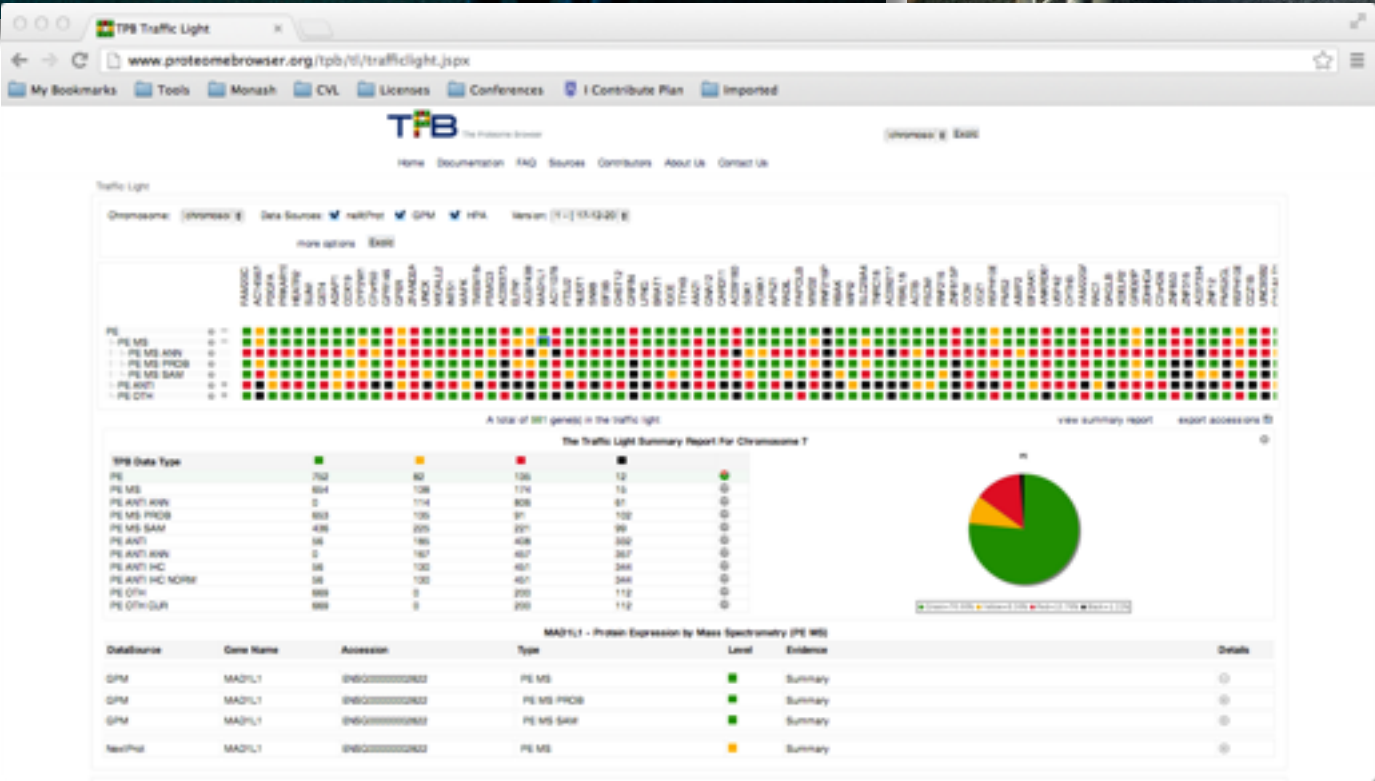
- Universities need to take responsibility, transform, and build capability
- Universities are **enduring** - capabilities (human, data etc) need an enduring **home**
- Responsibility: “Light bulbs” and “data/compute on demand”
- The history of email
- Universities have invested into the “unknown future” for centuries
 - *Universities invest in themselves, and take strategic “gambles”*
 - *Started with the creation of university libraries*
 - *Requires constant transformation,*
 - *and cannot be entirely outsourced*
- Step 1: Getting senior management onside: *The eResearch **Programme***
 - ***Dont talk about technology -- talk about the research***

“Faster”

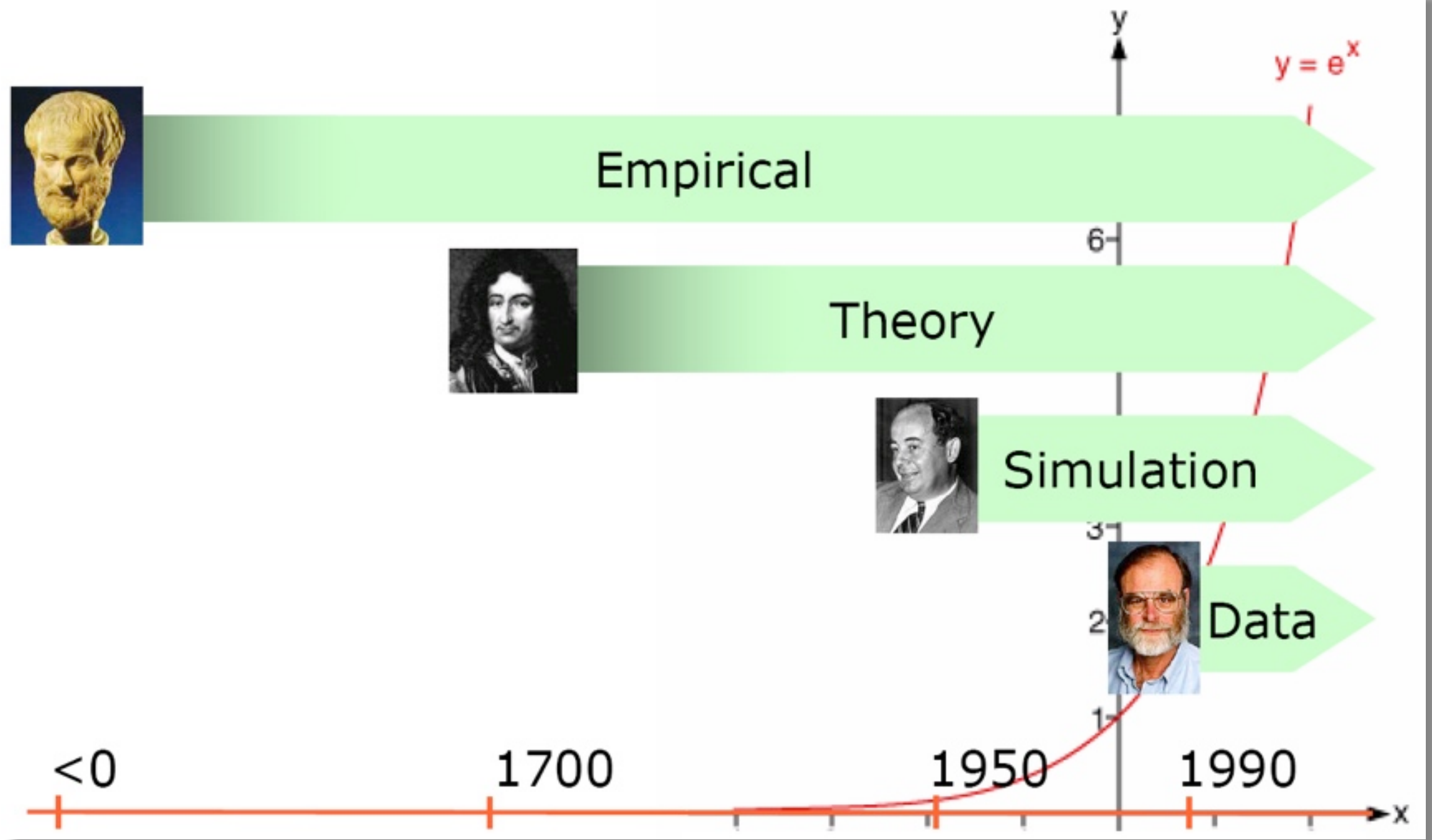


“Better”





Emergence of Research Methodologies





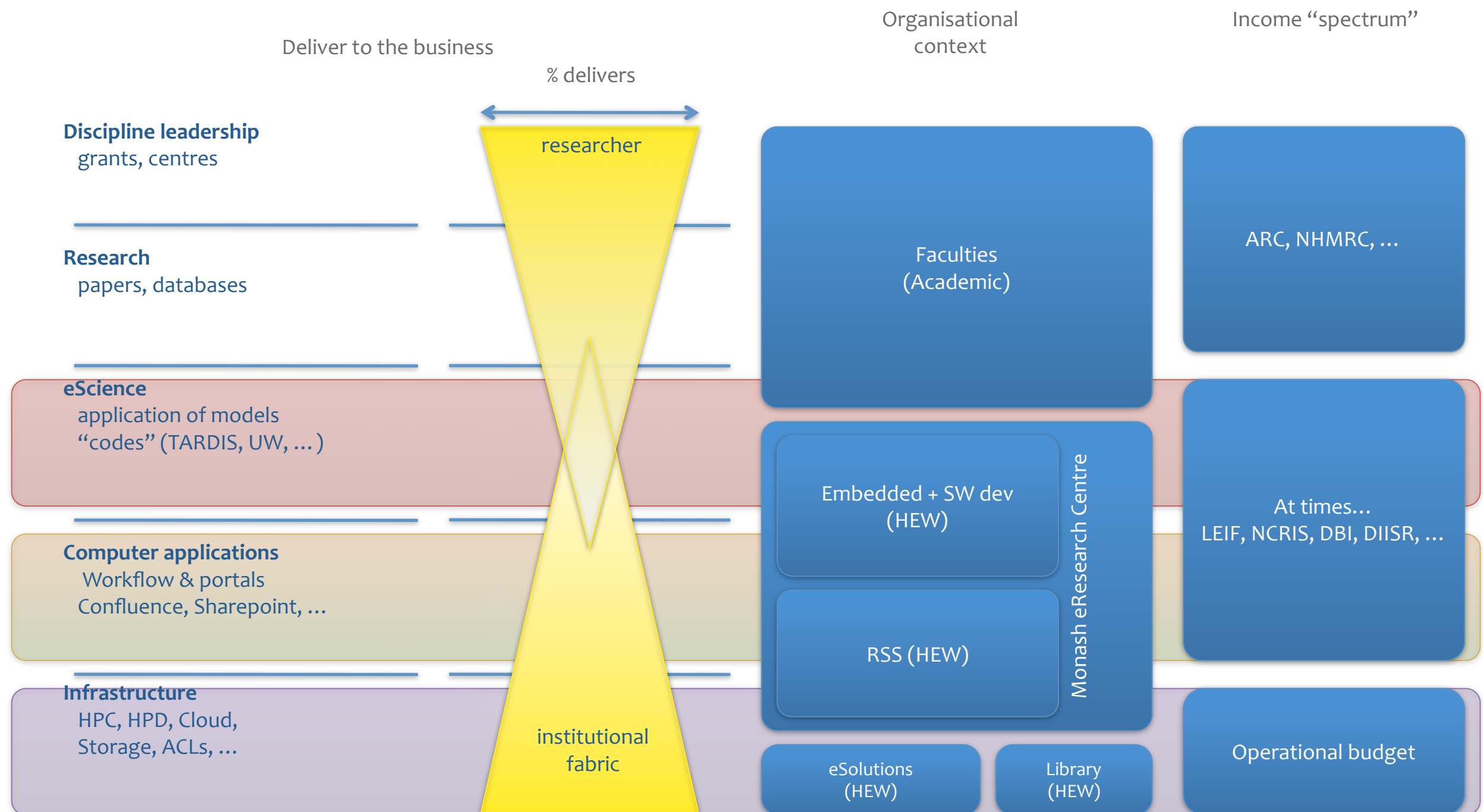
The Institutional Challenges of Cyberinfrastructure and E-Research

cholarly practices across an astounding wide range of disciplines have become profoundly and irrevocably changed by the application of advanced information technology. This

By Clifford Lynch

- “Scholarly practices across an astounding wide range of disciplines has become profoundly and irrevocably changed by the application of advanced information technology. This collection of new and emergent scholarly practices was first widely recognized in the science and engineering disciplines....humanities continue to showcase some of the most creative and transformative examples of the use of information technology to create new scholarship.”

The “3rd” Career Path: eResearch Specialist



Monash Research

Fundamental Research

- Outputs: papers, ...

High profile areas

- Outputs: branding/ reputation, ...

Platforms

- Outputs: utilised infrastructure

Faculty & Institutes

Fabric

Research
capability

IT
capability

eResearch Capability

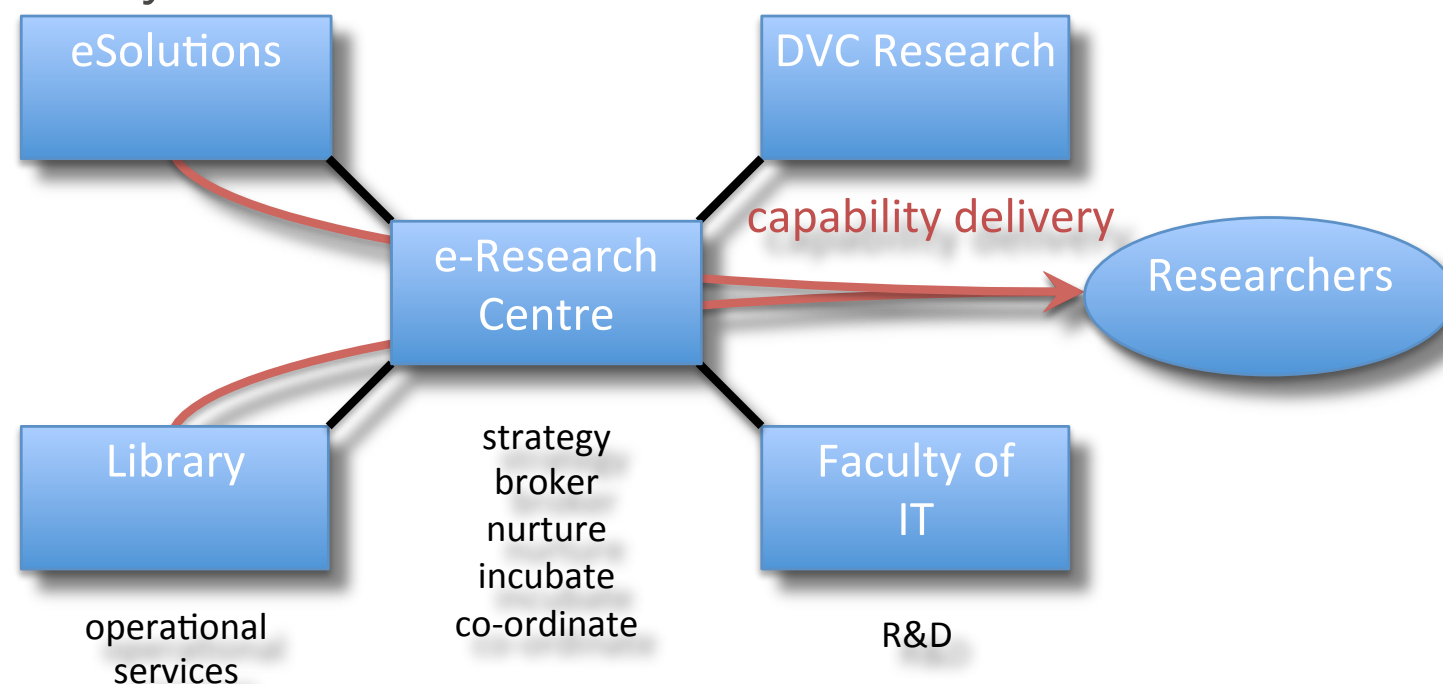
Improves effectiveness in delivery to the outputs

Lesson 3: Partnerships are critical

- *No one person/group can do this (eResearch) alone*
- Research <-> Technologist
- Research Office <-> IT Service Division
- University <-> University (and Research Institutes)

The Monash eResearch Model

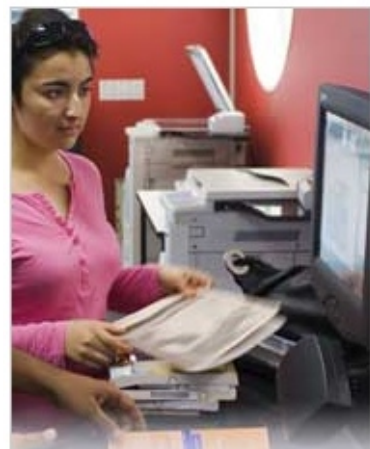
- The first (established by DVCR 2007) and still the most successful eResearch engagement model in Australia: **research-led** and based around strong **partnerships**
- The Monash e-Research Centre's primary role is to **partner** with research groups around the University, working with them to raise or **enhance** their research/e-Research capabilities to the next level, and then **connect** them to the most appropriate mechanisms or services to allow that research group to **sustain** that capability.



eResearch(ODVCR)-eSolutions Partnership

- Monash e-Research Centre is a **blended unit** of ~45 staff from ODVCR, Library and eSolutions (Monash's IT Service Division)
- Hub-spoke model with hub, and spokes embedded in some labs and faculties
- Functional reporting lines through to ProVC Research and Research Infrastructure
- Majority of staff organisationally in eSolutions
- Academic leadership represented by Director (ODVCR) - Prof Paul Bonnington (Faculty of Science, School of Mathematical Sciences)
- Shared management (ODVCR-eSolutions-Library) of Centre
 - *Steve Quenette (Deputy Director)*
 - *Wojtek Goscinski (Head of HPC)*
 - *Kheeran Dharmawardena (Manager RSS in eSolutions)*
 - *Anthony Beitz (General Manager)*
 - *David Groenewegen, Library*

Library: Research Data Management



Research data management

Monash University is committed to improving the way research data in all formats is created, stored, managed and disseminated. This website provides guidance on common research data management issues and highlights the portfolio of advisory and technical support services that are available to researchers.

Guidelines

- > [Data management planning](#)
- > [Ownership, copyright and IP](#)
- > [Ethical requirements](#)
- > [Durable formats](#)
- > [Storage and backup](#)
- > [Sharing data and controlling access](#)
- > [Documentation and metadata](#)

Quick links

- > [Contacts](#)
- > [FAQs](#)
- > [News](#)
- > [Policies](#)
- > [Resources](#)
- > [Services](#)

Managing your research data



MONASH University
Library

Sharing Data for Research

Lesson 4: Understand what “success” is

- Service delivery success vs Research success

Measuring Success

- ~1/3 of ARC and NHMRC grants (external research income) engage with Monash's eResearch Centre for advanced e-Research capabilities to enable the new research possibilities (2007: ~0)
- Rise coincides with Monash's rise through the Go8 to 2nd in ARC/ NHMRC
- Linked to research **impact**
- Monash e-Research Centre KPI's:
 - *new research methods leading to new discoveries and impact*
 - *external research and research infrastructure funding*
 - *raising Monash's international reputation*
 - *attracting world-class researchers to Monash*

Lesson 5: Understand Researchers

- Strong academic leadership of the eResearch programme: CRITICAL
 - *Professorial level*
- Language and Communication:
 - *i.e. “Service Division” vs “Capability (Research) Partner”*

Engaging with Researchers

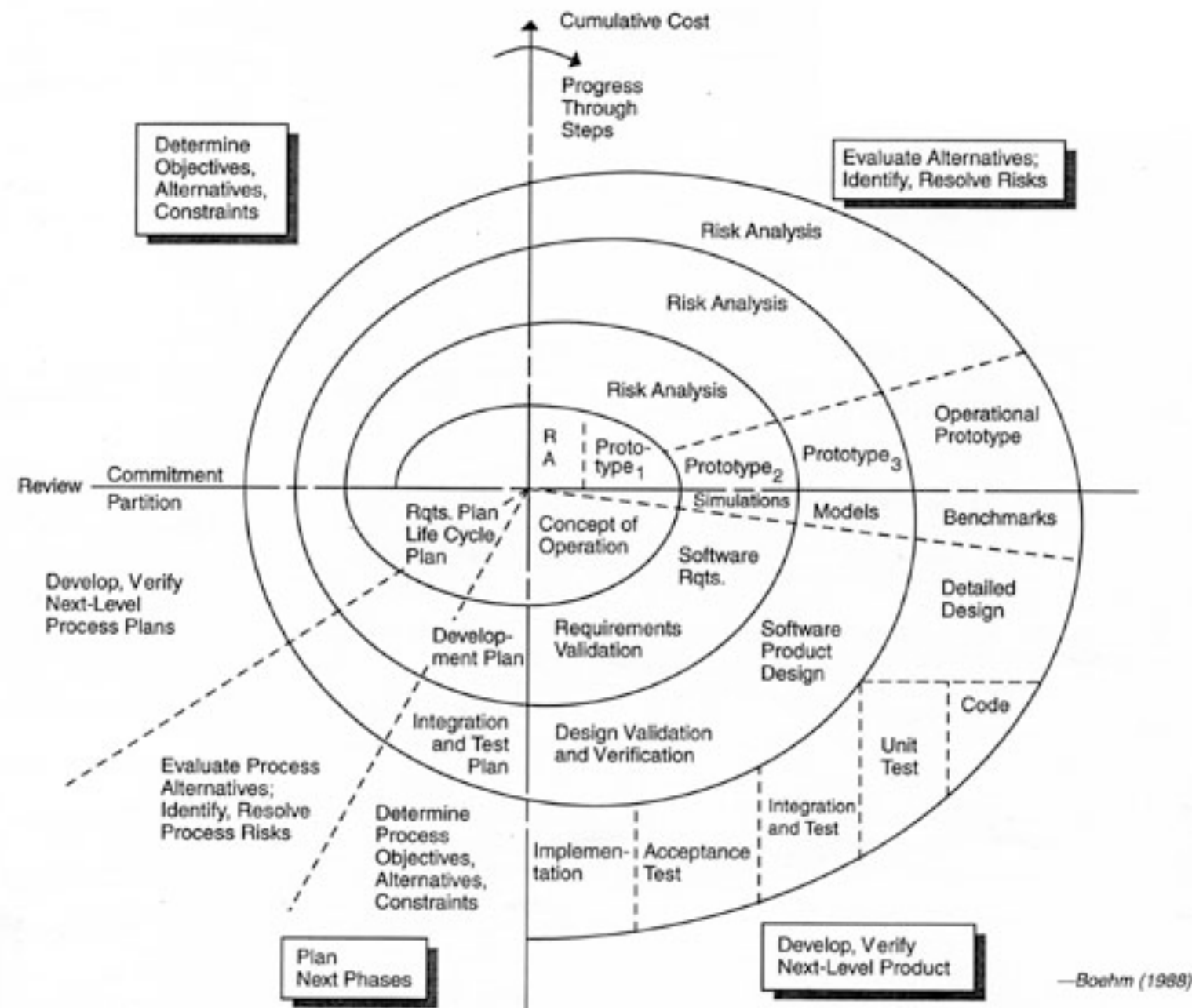
- They thrive on ambiguity – service providers try to remove ambiguity =
breakdown in communication and understanding
- **Analytical** versus **Interpretive** communication
- The three T's of an Innovation Environment: Talent, Technology and
Tolerance (for ambiguity and risk)
- Motivations: cost of entry needs to be **low** and benefits need to be **high**

Lesson 6: Understand how to build the unknown

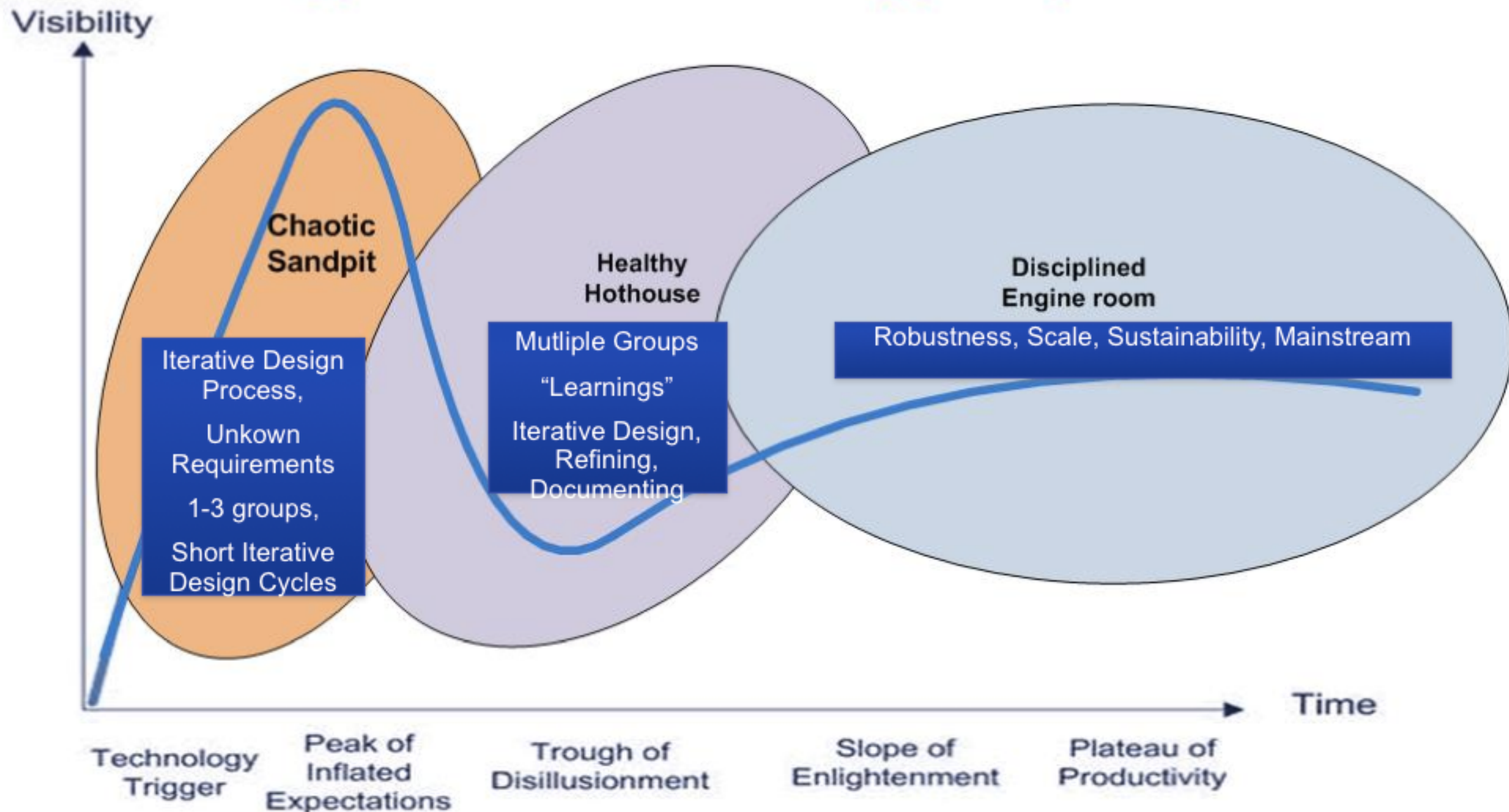
- Research is about “doing things we don’t know how to do”
- Researchers-Technologists partner to go on the “journey of the unknown together”
- IT Requirements are not understood at the beginning of the “journey”
- The Journey is Iterative and Perpetual
 - *difficult to contract*

Spiral Development – Journey is iterative

- “Agile” projects



e-Research Innovation Pathways in Monash



myTARDIS - Connecting scientific instrument data to people

The screenshot shows the myTARDIS website interface. At the top, there's a navigation bar with 'MyTardis', 'Home', 'About', 'Data', 'Stats', 'Register', and 'Log In'. The main content area is titled 'Experiment' and 'X-ray crystal structure of the streptococcal specific phage lysin PlyC' by Sheena McGowan, Ashley Buckle, and James Whisstock. It shows a sample diffraction image (PlyCB: plycb0001.img) and a list of 2 datasets: PlyCB (100 files, 3.8 GB) and PlyC (411 files, 6.5 GB). Below the datasets, there's a 'Publication (PNAS)' section with a text abstract about bacteriophages and lysins. At the bottom, it mentions 'This entry contains two diffraction datasets:' and lists 'PlyCB (PDB ID: 4F87)' with a corresponding 3D molecular structure visualization.

- Started from scratch 2007, with no particular plan
- Now solving the problem of archival, sharing and citation of raw data
- Easy sharing, access and publication of terabyte datasets
- More than 2 million files, 20 terabytes collected automatically for access - and growing
- 8 organisations have deployments that collect synchrotron data, mass spectrometry, electron microscopy, quantum physics, HPC processing or medical imaging data



Experiment

X-ray crystal structure of the streptococcal specific phage lysin PlyC

Sheena McGowan , Ashley Buckle , James Whisstock

2

531

10.3 GB

26th October 2012

Public

Describe

[Show Description](#)

DOI for this data: 10.5072/03/50171CC7EEC1DSample diffraction image (PlyCB: plycb0001.img);Publication (...)

[Description](#)[Metadata](#)[Sharing](#)[Transfer Datasets](#)

Institution Monash University

Licensing Unspecified

Download All

[ZIP](#)[TAR](#)Share &
Publish

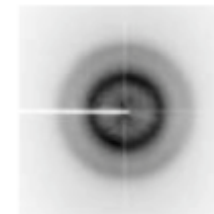
Download

Browse

2 Datasets

[Download Selected](#)

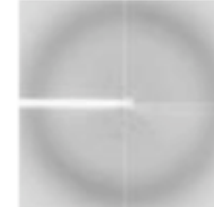
Just start typing to filter datasets based on descriptions



PlyCB

120

3.8 GB



PlyC

411

6.5 GB

Integrate..

Provide..

Dataset

Dataset 3

From the experiment: [Experiment 0000](#)[Edit Dataset Description](#)

1

7

126.0 MB

A few diffractions from your dataset

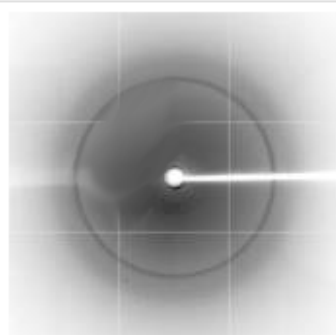


Image 1

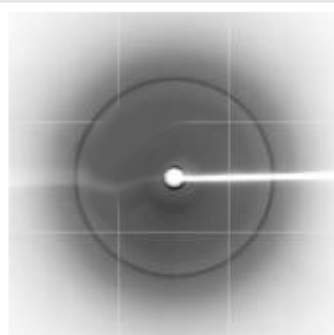


Image 2

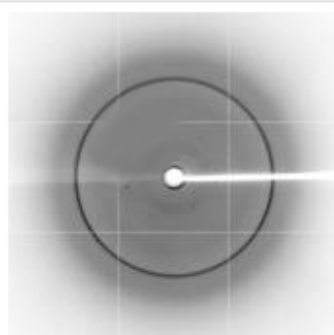


Image 3

Datafile Information

Datafile Metadata

For the file: [Image 1](#) (18.0 MB)

TRD datafile		Edit
Detector Distance	399.999 mm	
Detector S/N	928	
Direct Beam X-Pos	158.004 mm	
Direct Beam Y-Pos	158.083 mm	

360 Files

[Download Selected Files](#)

Search: Enter part or all of a filename, then press enter

[Upload Files](#)Select: [All](#) / [None](#)

☐	Image 1.img (18.0 MB)	   
☐	Image 2.img (18.0 MB)	   
☐	Image 3.img (18.0 MB)	   
☐	Image 4.img (18.0 MB)	   
☐	Image 5.img (18.0 MB)	   
☐	Image 6.img (18.0 MB)	   
☐	Image 7.img (18.0 MB)	   

Sharing

Users

Users who have a share in this experiment:

Username	Name	Permissions
mytardis	Steve Androulakis	Read Edit Owner
oded	Oded Kleifeld	Read
bosco	Bosco Ho	Read
synchrotron	Synchrotron Test	Read Edit Owner

[Change User Sharing](#)

Groups

Groups who have a share in this experiment:

There are currently no groups with access to this experiment.

[Change Group Sharing](#)

Links

This experiment is private. A temporary link can be created by its owner(s) and privately shared for direct access.

Temporary access links provide full access to recipients regardless of an experiment's public status.

Temporary Link	Expiry	Granted By
Right click.. copy link	15th November 2012	mytardis
Right click.. copy link	7th December 2012	mytardis

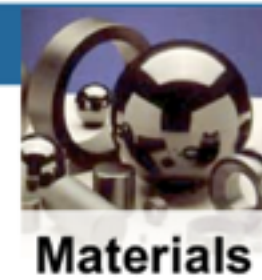
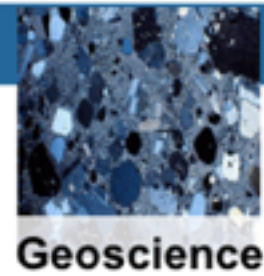
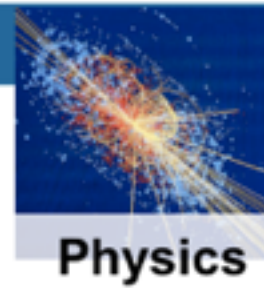
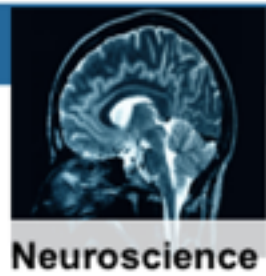
[Create New Temporary Link](#)

Creative solutions

- Solutions are not always “designed/engineered” – they are often “grown” via an **iterative** process
- If researchers engage all the way along the iterative process -- they will engage at the end
- The research challenges are unique and therefore the solutions are often “unique” - no one size fits all
- Leveraging existing underlying infrastructure wherever possible (storage, network, VMs, compute etc)

Lesson 7: Leverage, leverage and leverage

- Leverage existing capabilities (e.g. Professional institutional IT service delivery.
 - Leverage global infrastructure
 - Leverage organisational structures
 - Leverage expertise across the institution
-
- If it is not in the right form....TRANSFORM it.
-
- State/sell the leveraging from the researchers point of view (not the service point of view)



National &
Characterisation
Council



State
(Victorian)

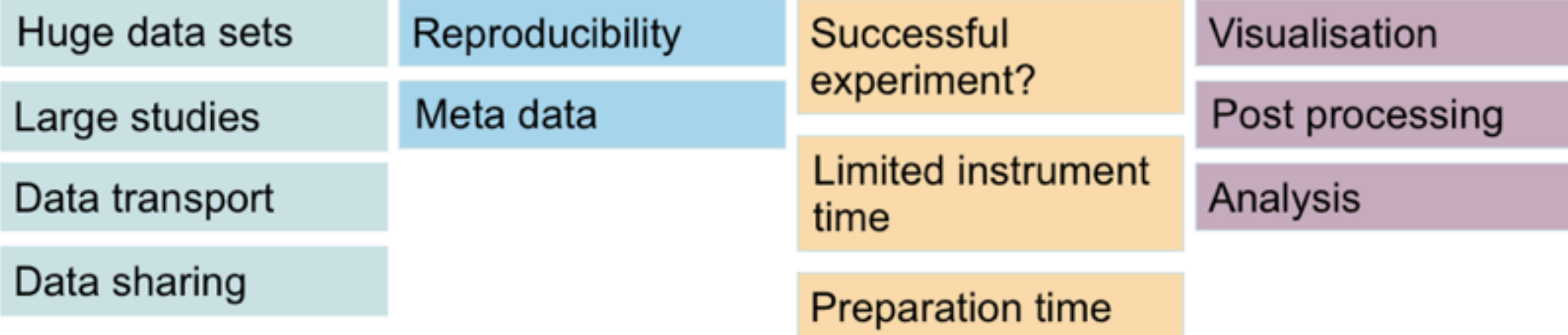
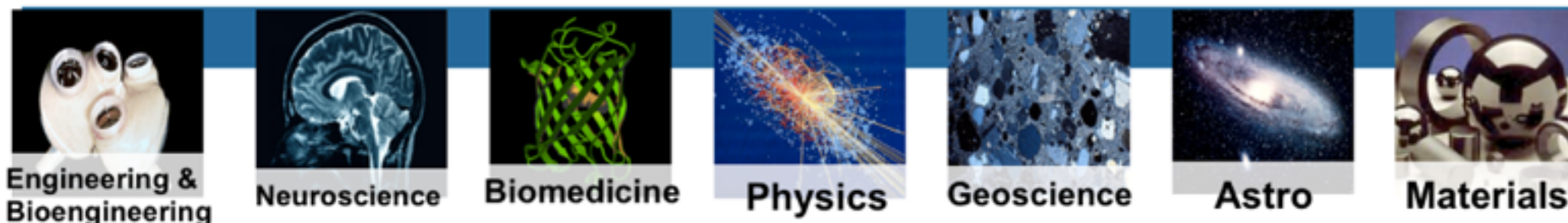
MCE

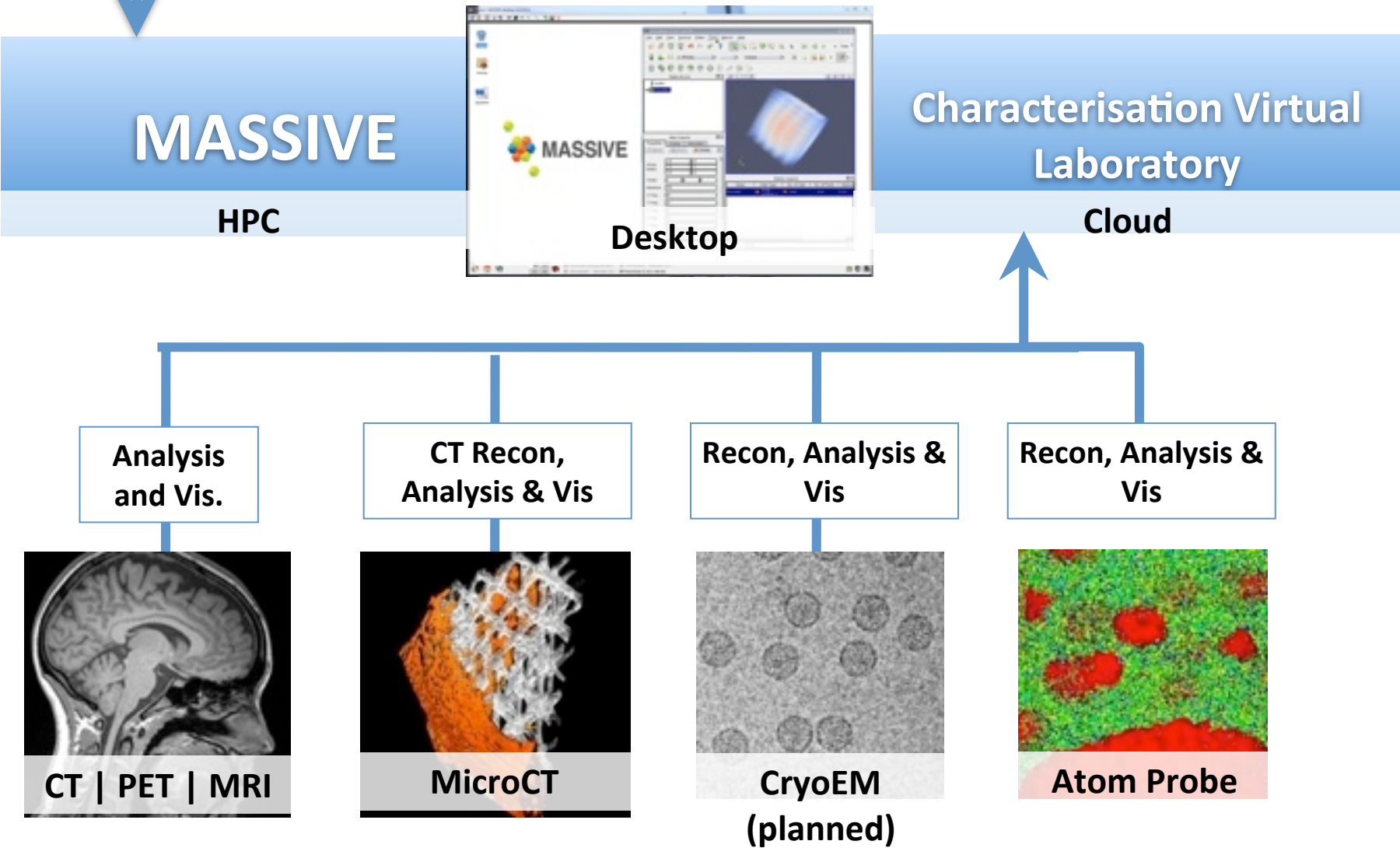
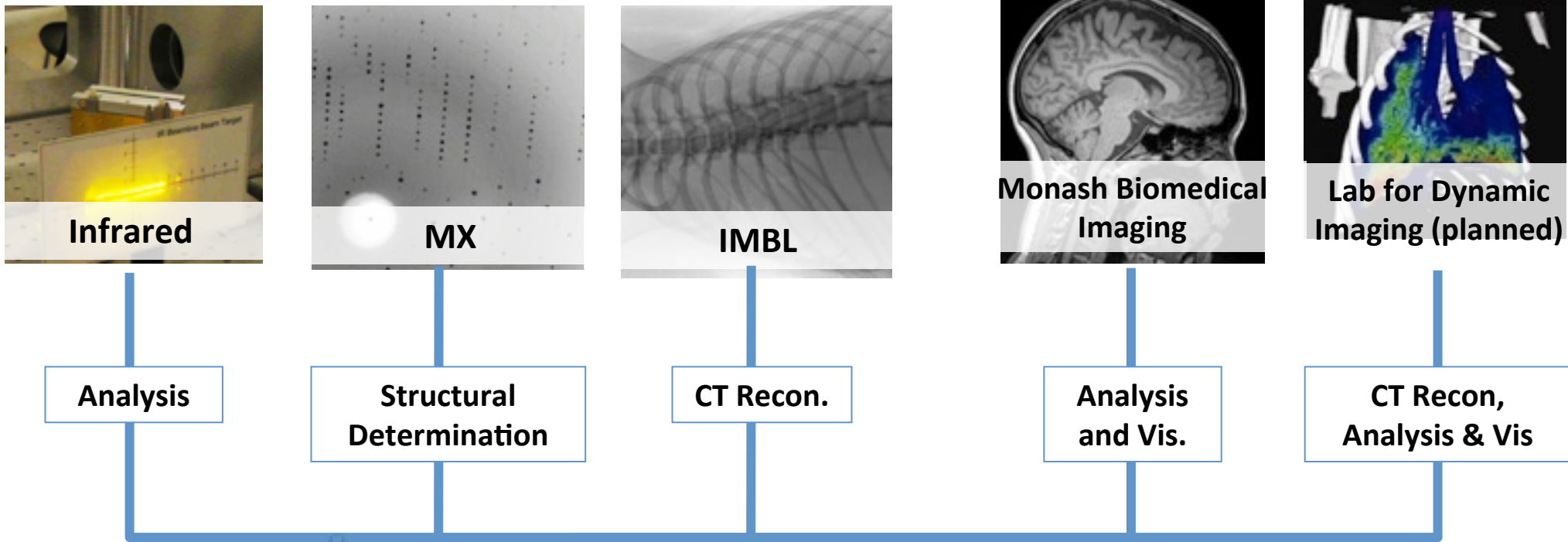
MM

MB



Institution
(Monash
University)





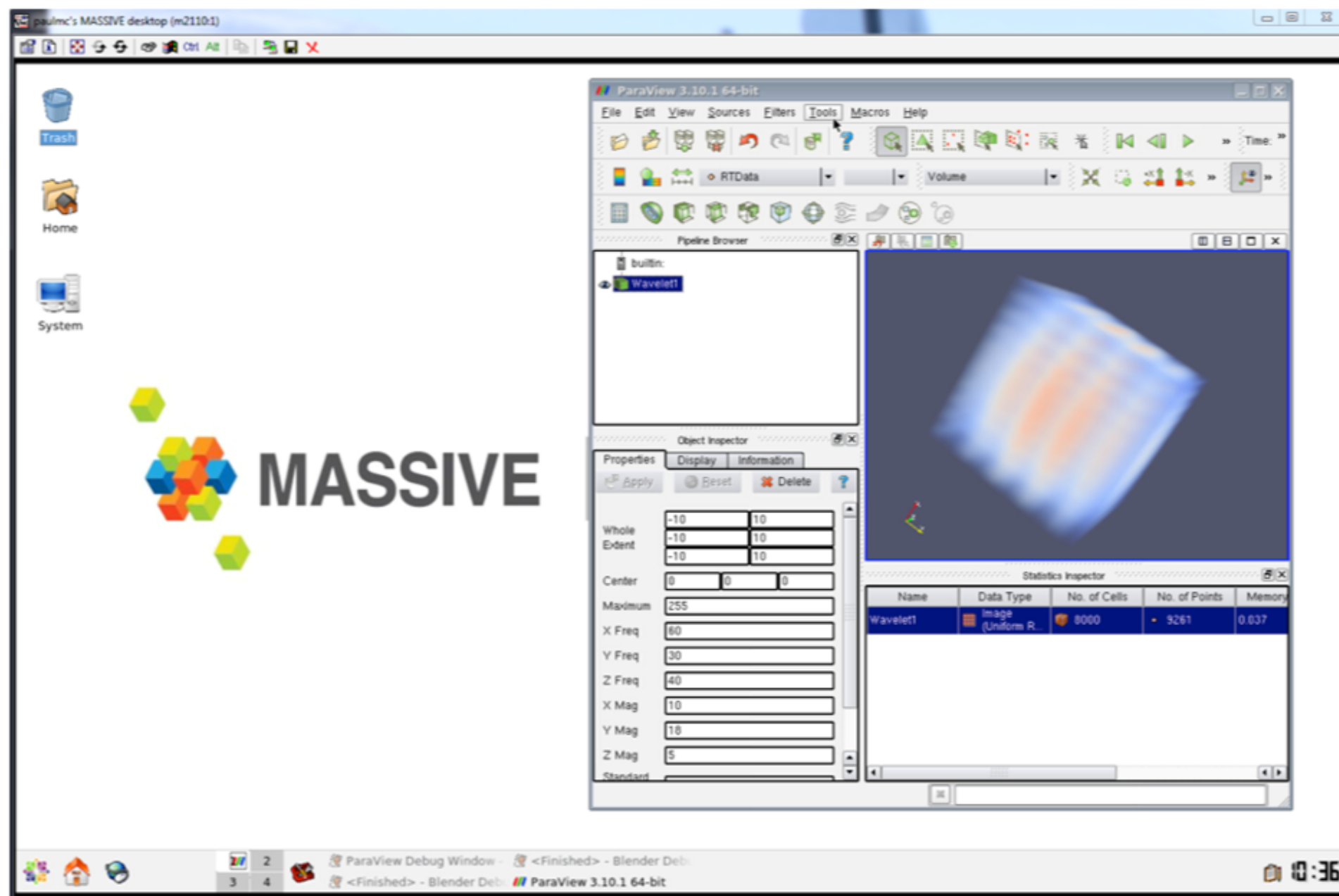


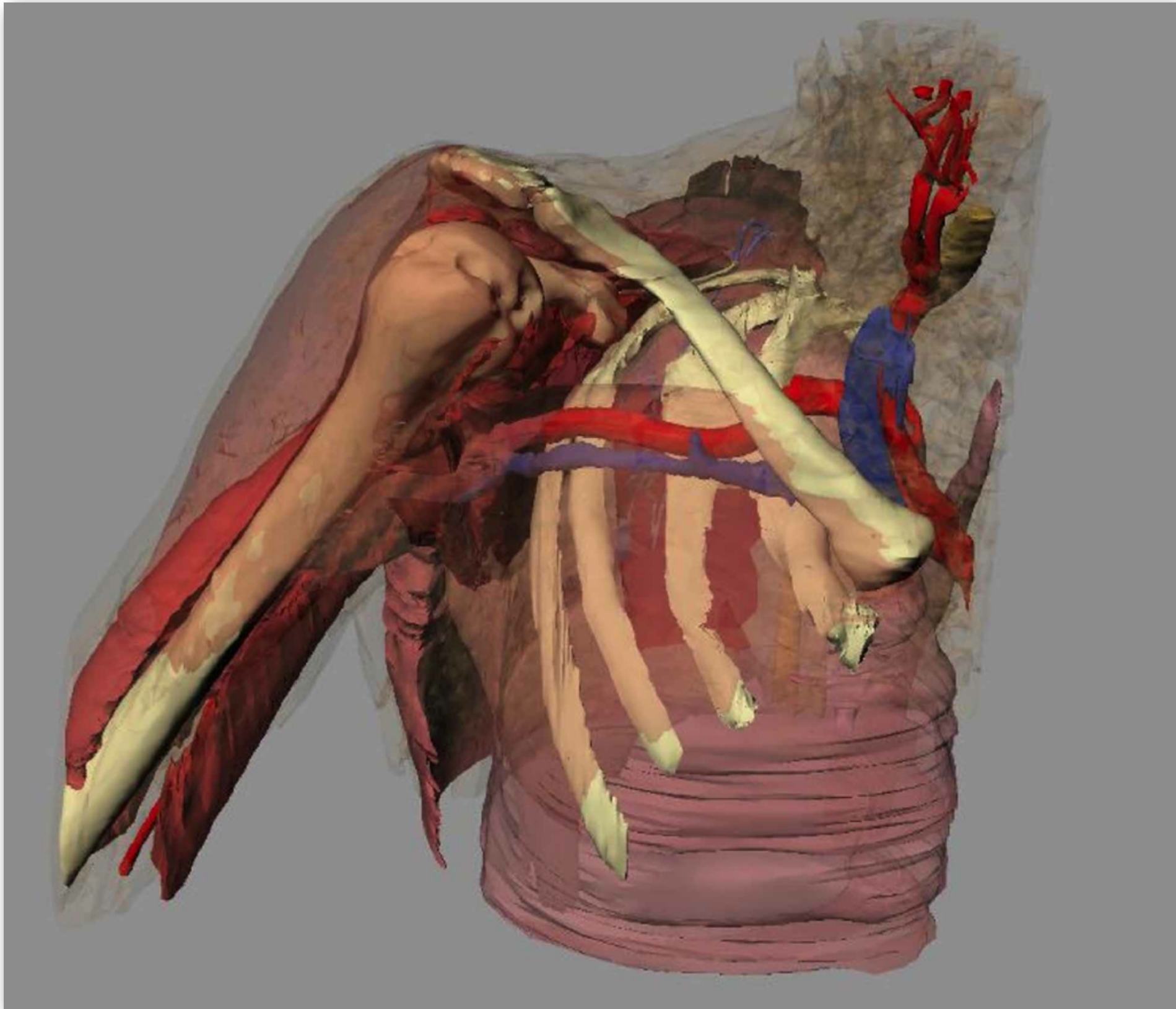
Lesson 8: Make the “right” things easy

MASSIVE Interactive Desktop

- Characterisation and Visualisation tools:

- Paraview
- Amira
- Drishti
- ImageJ
- Seg3D
- VolView
- MayaVi
- Etc.





Lesson 9: Think globally, because researchers do!

- Collaborate globally - partner beyond your state and country
 - *e.g. Monash's partnerships with Oxford eResearch Centre, UCSD*

International Engagement



SC11
Seattle, WA

Connecting communities
through HPC

SC Is the International Conference for
High Performance Computing, Networking,
Storage and Analysis

SEARCH

Conference Dates: November 12-18, 2011 Exhibition Dates: November 14-17, 2011

About | Technical Program | Communities | Exhibits | Social Media | News & Press | Registration, Travel & Hotels

- Overview
- SCinet
 - SCinet Research Sandbox
- Important Dates
- Committees
- Sponsoring Societies
- Steering Committee
- Who to Contact
- SC History

SCinet

During the advance conference, SCinet's infrastructure, trademarks, and a variety of other resources are available to attendees.

Designated by the Center for Energy's SCinet, SCinet's network provides performance and reliability for the SC community.

In previous years, SCinet's network has provided performance and reliability for the SC community.

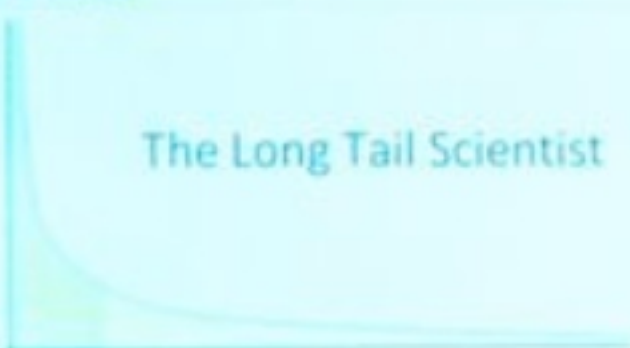
SCinet's network provides performance and reliability for the SC community.





MANCHESTER
The University of Manchester

The Long Tail Scientist



Professor Carole Goble FREng
carole.goble@manchester.ac.uk
University of Manchester, UK

10/11/2011 14:49:21 - previously 10/11/2011 - Science 44/45 2011/12 Sept 2011

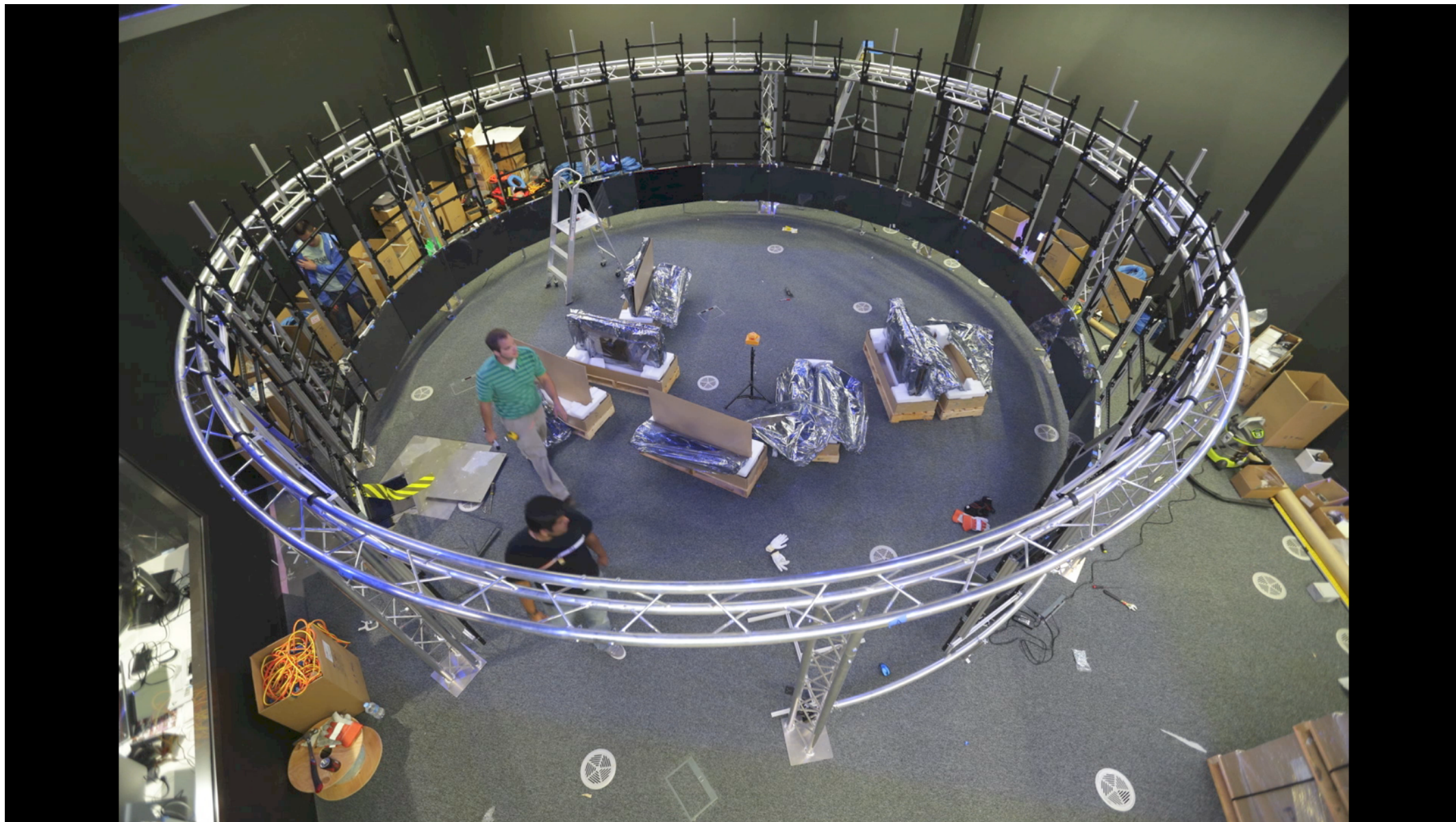


Lesson 10: (The most important)

People are more important than technology

- Monash eResearch Centre
 - *2008: 6 staff*
 - *2013: 45 staff*
 - growth largely through transformation and leveraging
- The “3rd” Career Path: eResearch Specialist
 - *Institutional Structures*
- Computers only do what they are told to do
 - *Research is the journey into the unknown*
 - *e.g. Visualisation*

Completing the “loop”....





Knowledge Discovery

There are known knowns; there are things we know we know.

We also know there are known unknowns ... we know there are some things we do not know.

But there are also unknown unknowns – the ones we don't know we don't know.

Donald Rumsfeld (2002)



Figure 6. CAVE2 running in full 3D mode. The visualization shows reconstructed brain vessels and cortical tissue gathered from MRI data, supplemented with artificially generated microvasculature. Effective display of the finer details of this dataset is possible thanks to CAVE2 resolution.



Figure 8. CAVE2 can run in Hybrid mode, to juxtapose high-resolution 2D and 3D datasets side-by-side. The right side shows a stereoscopic, head-tracked rendering of planetary-scale Martian terrain data. The left side shows high-resolution pictures taken by the Curiosity rover along with 2D topographical maps of Mars.

New Horizons Monash CAVE2™

84 million pixels Arrived Oct 2013

The CAVE2™ System was developed by the Electronic Visualization Laboratory (EVL) at the University of Illinois at Chicago. CAVE2 is a trademark of the University of Illinois Board of Trustees. (Photo: Lance Long, EVL)



