



Neil Clarke  
Manager, e-Research Centre  
13 November 2006

# LaRDS – Monash's Large Research Data Store

briefing to Data Management Seminar 13 Nov 2007

# Introducing LaRDS

- Monash's petabyte Research Data Store

1 petabyte =  $10^{15}$  byte

1 petabyte = 1 billion megabyte

1 petabyte  $\approx$  250,000 DVDs

- Production service from Jan 2008

- Pilot services since 2006

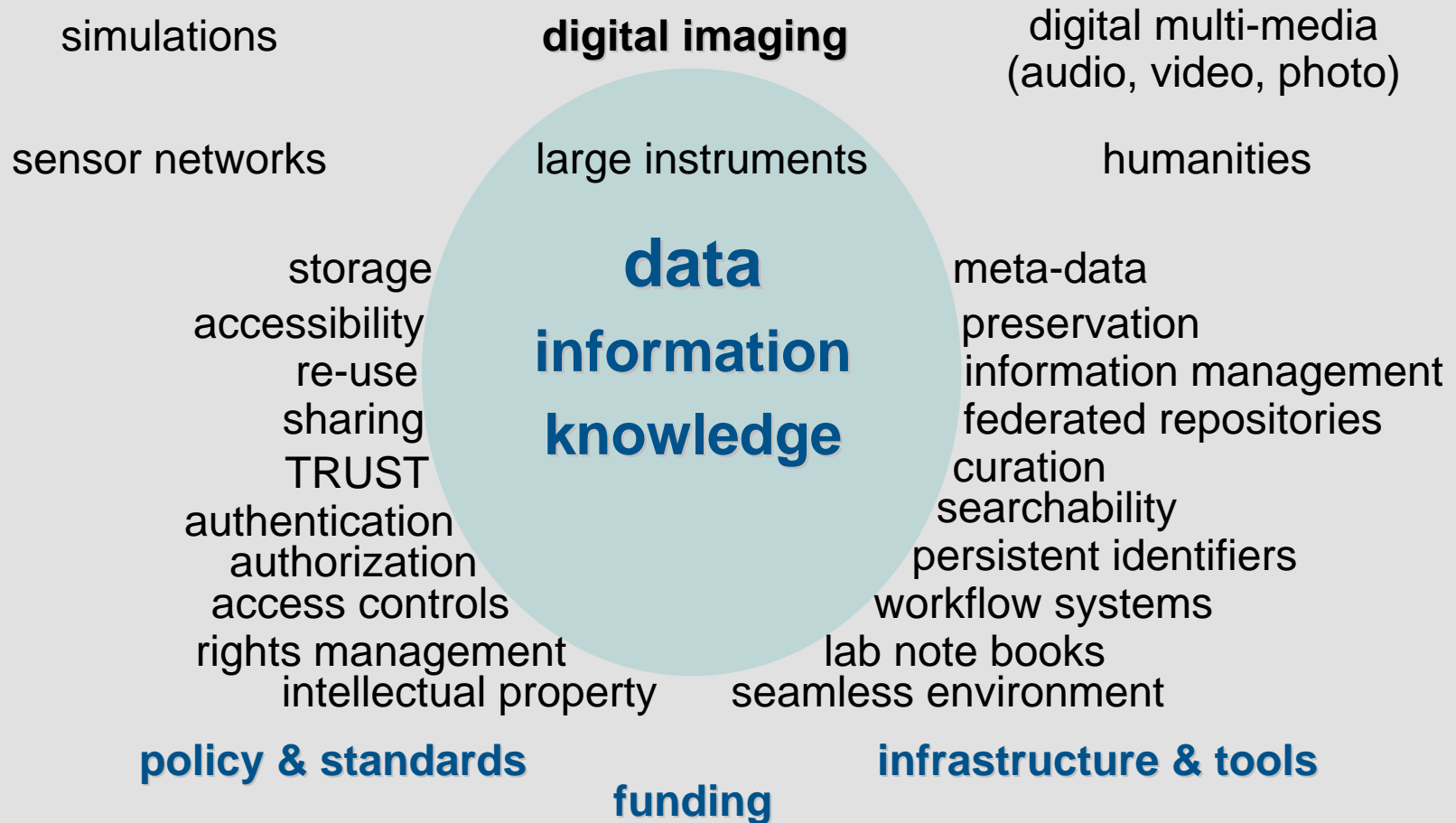
# LaRDS pilot users include

- Shoah Visual History Archive
- Medicine (various) including
  - Australian Mouse Brain Map
- Climate research
- Engineering (various) including
  - Fluid mechanics
- School of Music – cultural archives
- Biological sciences
- Synchrotron science
- Monash HPC users (various)

# Motivation: the 'Data Deluge'

- High resolution digital imaging
- Digitized data and information of all types
- HPC simulations
- Research collaboration
- Most e-Research themes track back to data

# “e-Research” is all about **data**



# LaRDS storage modes

- **On-line disk with backup**
- **Near-line tape archive**
  - Functions 'like' disk
  - All tape held within robot tape library
  - Of the order of ~minute to retrieve
- **Copies of all data held in physically diverse data centres**

# Capacity

- Disk capacity 2008: ~200 TB  
1 TB =  $10^{12}$  byte = 1000 GB
- Tape library capacity: 1.6 PB  
1 PB =  $10^{15}$  byte = 1000 TB
- Expanding over time as needed  
aim: keep capacity ahead of demand
- High bandwidth to network and HPC facilities

# Continuum of service interfaces

- C:\ drive backup
- Novell Netware shared drive: R:\ drive
- NFS (unix network file system)
- GridFTP (ARCS data mover fabric)
- SRB (Storage Resource Broker)
- SQL (database server)
- others as needed

# LaRDS two-tier allocation model

- Generous basic allowance per faculty
  - by formula based on *research activity* (measured by Research Income, Academic Staff and HDR Students)
  - funded centrally via the normal central support services charging mechanism
  - allocated within faculty by ADR or delegate
- Additional capacity for ‘very high achievers’ on a fee-for-service (once-off charge) basis

# Basic allocations 2008, growing over time

| File capacity at least (TB):        | Disk | Tape | Total |
|-------------------------------------|------|------|-------|
| Art & Design                        | 2    | 5    | 7     |
| Arts                                | 12   | 32   | 44    |
| Business & Economics                | 11   | 29   | 40    |
| Education                           | 5    | 13   | 18    |
| Engineering                         | 18   | 48   | 66    |
| Information Technology              | 6    | 16   | 22    |
| Law                                 | 2    | 5    | 7     |
| Medicine, Nursing & Health Sciences | 65   | 173  | 238   |
| Pharmacy                            | 6    | 16   | 22    |
| Science                             | 13   | 35   | 48    |
| Gippsland                           | 6    | 16   | 22    |
| DVC-R (MUARC, etc.)                 | 4    | 11   | 15    |

# Additional capacity charges 2008

- Once-off charge
- On-line disk capacity \$5.60 / GB
- Near-line tape archive \$2.40 / GB
- Ideally, no-one will need to pay this
  - basic allowance should be adequate
- Charge rates reduce over time

# Further information

- LaRDS web site:  
[http://www.monash.edu/eresearch/  
activities/lards.html](http://www.monash.edu/eresearch/activities/lards.html)
- LaRDS Service Manager:  
[Neil Clarke](#)

# Questions & Answers ...

---