In the next 5 years, we’ll produce more scientific data than in human history.

How can we deal with this data deluge?
Evolving Research Data Lifecycle

1. Conduct Experiment
2. Collect Data
3. Analyse
4. Publish

 Conduct Experiment
  
 Collect Data

Analyse

Publish
e-Research

A set of technologies to support collaborative networked science.

Some of the Challenges

- Acquiring data from instruments
- Storing and managing large quantities of data
- Processing large quantities of data
- Searching and discovering
- Sharing research resources and work spaces between institutions
- Collaborating
- Publishing research
ARCHER
A secure and seamless collaborative Research Env:
- Acquisition
- Access
- Analysis
- Annotation
- Metadata Mgmt
- Workflows
- Publish

Collaboration Repositories  Computational Grids
Changes in Data Analysis

Analysis Challenges
• Legacy
• Disparate Systems
• Proprietary solutions
• Collaboration

Emerging Technologies
• High Performance Computing and Distributed Processing (Nimrod)
• Workflow Engines (Kepler)
• Seamless and scalable security (Shibboleth)
• Collaborative Workspaces and Tools (MAMS IAM & VO)
• Collaborative Filtering (Social Networking Technology)
Kepler Workflow

- **PN Director**
  - Configure SRB server settings here

- **SRB Server**
  - SRB
  - Configure SOAP server port here

- **Cima Configuration**
  - Configure Experiment timeout here

- **Cima Parcel Switch**

- **Process New Run**
  - Process Double Data
  - Cima Binary File Writer
  - Cima Image Converter
  - Cima Cam Feeds Getter

- **List camera feeds to store here**
Challenges in Publishing Research

• Legal
• Ownership
• Searching and Discovery
• Preservation
• Object and metadata format
• Limiting Access
• Publication Time
• Validating/Modifying/Extending Research
Changes in Publication

1. Collaborative Repository (e.g. ARCHER)
2. Validate Research
3. Extend Research
4. Publication Repository (e.g. ARROW)
5. Publish (document, dataset, workflow, ...)

Note: This diagram illustrates the collaborative and publication processes in research environments, emphasizing the importance of archiving and publishing research outputs using repositories like ARCHER and ARROW.
Conclusion

ARCHER will allow Researchers to:

• Better deal with their data deluge
• Collaborate within and between research institutions; and
• Be more effective and efficient in their research