

**Taipei  
2008**

# International Grid Symposium

## **e-Sciences and the Humanities: Opportunities and Challenges for Grid-based Research**

**David J. Bodenhamer  
The Polis Center**

**IGCS/Academia Sinica  
April 10, 2008**

- Digital humanities rapidly achieving distinct status but still amorphous in scope and direction:

“...study of what happens at the intersection of computing tools with cultural artefacts of all kinds....”

Centre for Computing in the Humanities, Kings College, University of London

- Significant e-Resources
  - Text archives
  - Data archives
  - Image archives
- Web-based production and access

- “e-Science is essentially a mode of working that is reliant on a distributed number of computing resources...which can be made to work together for the purposes of one research outcome.”

<http://www.methodsnetwork.ac.uk/escience/e-science.html>

- “... information, expertise, standards, policies, tools, and services that are *shared broadly across communities of inquiry but developed for specific scholarly purposes*: [it] is...more specific than the network itself, but...more general than a tool or a resource...for a particular project...or...a particular discipline.”

*Our Cultural Commonwealth* (ACLS Commission on Cyberinfrastructure, 2006)

- Digital libraries
- Initial goals focused on data
  - Creation
  - Access
  - Preservation
  - Exchange

- Rapid increase in digital materials (Project MUSE, JSTOR, databases, etc)
- On-line access
- Convenient search tools
- Digital archives and repositories (AHDS, D-space, etc.)
- Highly visible projects (Valley of the Shadow, Vision of Britain, etc.)

- e-Science came to mean grid
  - Data grid
  - Access grid
  - Computational grid
- Little demand for computational grid in the humanities



- Now, e-Science means grid-based collaboration
  - Technical collaboration (e.g., networks, exchange protocols, middleware, etc.)
  - Procedural collaboration (e.g., standards for access and use)
  - Scholarly collaboration
  - ICT-discipline or domain collaboration

- Application of intensive computing technologies and methods to research questions relating to the development of human society.
- Designed to manage and analyze the data deluge brought on by the new digital age in ways that are suitable for the humanities.
- Computer-based collaborations

- Most humanists still in e-Science 1.0 (digital libraries)
- Some humanists are moving into e-Science 1.5
- Few humanists are in e-Science 2.0
- Why?

- Approaches (themes)
- Questions
- Data
- Methods
- Tools
- Culture

- What themes engage modern humanists?
  - Interdisciplinary or multi-perspective
  - Interdependency
  - Intercultural and trans-cultural
  - Multi-scalar and inter-scalar
  - Visualizing complexity

- Modern humanists are interested in
  - Context
    - ✦ Non-linear
    - ✦ Fuzziness
  - Culture
    - ✦ Variation
    - ✦ Diffusion
    - ✦ Transmission

- Most humanities data still non-digital
- Wide range of data types
  - Archival records
    - ✦ Letters, diaries, newspapers, photographs, etc.
    - ✦ Government papers, censuses, maps, etc
  - Generated or collected data
    - ✦ Surveys, field samples
  - Multimedia
  - Artefacts

- What methods do they employ?
  - Discipline-based
    - ✦ Archival (document based)
    - ✦ Fieldwork (ethnography, oral history)
    - ✦ Hermeneutics (textual analysis and interpretation)
    - ✦ Statistical
  - Fusion



- Most humanists do not use computer-based analytical tools
- Most software used by humanists was not developed with humanities scholarship in mind
- A GIS example: the North American Religion Atlas ([www.religionatlas.org](http://www.religionatlas.org))

Most humanists still:

Work in isolation

Use traditional methods

Lack computing skills

Are not interdisciplinary

- Modern humanities more open to diverse approaches
- Strong shift toward interdisciplinary research
- Quantitative data and methods more common
- More collaborative

- Digitization
- Use of e-Tools
  - Information management
  - Data integration
  - Textual analysis
  - Statistical analysis
- Visualization

- Framing suitable e-problems
- Evidence
  - Incomplete and missing data
  - Ambiguous data
  - Conflicting typologies/semantic standards
  - Language barriers
  - Legal access
  - Metadata

- Disciplinary knowledge base
  - Theory/models/methods
  - Best practices
- Technology
  - Expensive
  - Steep learning curve
  - Too little interaction with ICT experts

- Spatio-temporal integration
- Epistemology
  - Nonlinear
  - Non-expert knowledge
  - Choice and customization

- Access Grid
- Resource Grids
  - Data Grid
  - Text Grid (e.g., TextGrid, Germany)
  - Image Grid (e.g., USC Digital Archive)
  - Spatial Grid (e.g., HGIS-Europe)



- DARIAH (Digital Research Infrastructure for the Arts and Humanities, UK)
- Virtual Knowledge Studio (Netherlands)
- CATCH (Continuous Access to Cultural Heritage, Netherlands)

- New collaboratories that allow
  - Retrieval
  - Contextualization
  - Hypothesis building
  - Flexible narration
  - Integration into knowledge networks
- Virtual Research Environments (VRE)
  - GRID+Web 2.0

# Opportunities for humanities grid

- Life paths
- Networks
- Virtual reality
- Gaming
- Simulations
- Deep mapping

- The *human* humanities grid
  - Opportunities for interchange and experiment
  - Creation of knowledge base
  - Development of problem statements
  - Working out of research design
  - Building viable partnerships and a sustainable community