Arts and Humanities e-Science

Tobias Blanke
tobias.blanke@kcl.ac.uk
## E-Science Definitions

<table>
<thead>
<tr>
<th>Definition</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>‘e-Science is about the global collaboration in key areas of science and</td>
<td>John Taylor, Director General of UK Research Councils</td>
</tr>
<tr>
<td>the next generation of computing infrastructure that will enable it’</td>
<td></td>
</tr>
<tr>
<td>‘the development and deployment of a networked infrastructure and culture</td>
<td>Sheila Anderson, Director AHeSSC</td>
</tr>
<tr>
<td>through which resources – (...) – can be shared in a secure environment,</td>
<td></td>
</tr>
<tr>
<td>and in which new forms of collaboration can emerge, and new and advanced</td>
<td></td>
</tr>
<tr>
<td>methodologies explored’</td>
<td></td>
</tr>
</tbody>
</table>
E-Science: Building bridges

Using networks to connect resources

- Grids to allow virtual computing across “admin domains”
  - Virtual digital libraries, virtual museums, virtual observatories

- Technology that was first adopted in sciences...
Grand challenges everywhere

- Ongoing growth of corpora due to major digitisation projects (OCR, OMR, etc.)
- Digital recordings of current human developments
- Computational methods to deal with inconsistent data
- Reluctance to work together in research ??
Complexities of Humanities Data

- Qualitative human based data needs novel methods of selection (unstructured)
- Few standard formats or interfaces
- Semantics barrier: complexity and context dependency of research material
Established in 1996
Evolution in 2003
- Managing Executive
- AHDS Literature, Languages and Linguistics
  - AHDS History
  - AHDS Archaeology
  - AHDS Performing Arts
  - AHDS Visual Arts
These resources are free for educational and private use
Generally available online
Museum of London Archaeological Archive

New Survey of London Life and Labor, 1929-1931

London College of Fashion: The Woolmark Company

Imperial War Museum

Designing Shakespeare
Collections

- Highly diversified in terms of types and sizes
- Require specialised knowledge
- Multimedia: Images, text, death metal music
- Quantum leap: 2 TB in 2005, incoming video collection 30 TB, …
Working with data –
E-Science in the Arts and Humanities
Virtual Workbenches
Virtual Vellum - Sheffield

- Generic image-collaboration environment
  - large volume image datasets
  - high-resolution image files (> 8K x 6K pixels)
- Grid-enabled
  - Data Grid using Storage Resource Broker
- Access Grid integration
- Froissart Manuscript Project provides initial image dataset, digitally photographed by Colin Dunn, Scriptura Ltd

Prototype manuscript viewer by Colin Dunn, Scriptura Ltd
Images © Stonyhurst College, Lancashire and Scriptura Ltd
E-Curator - UCL

• New methodology for recording the surface detail and colour quality of a range of object types and materials

• Examine how the resulting datasets could be transmitted, shared and compared

• To begin to build expertise in the use and transmission of 3D scan data as a curatorial tool

An example view from a colour 3D data set, image courtesy of Arius 3D and the Royal Ontario Museum, Canada.
Collaboration in Performance Arts
Data Services for Associated Motion capture User Categories - CultureLab
Musicology

• **musicSpace (Southampton):**
  - Access to Heterogeneous Music Resources
  - Semantic Web

• **Purcell Plus (Goldsmith):**
  - Music and text retrieval to support musicologists
Research Infrastructures

DARIAH- Digital Research Infrastructure for the Arts and Humanities
to facilitate long-term access to, and use of, all European humanities and cultural heritage digital information
Fedora demonstrator

- Exemplary digital archive
- Collections at the Nordisk Forskningsinstitut (NFI) as well as reference works taken from the ECHO collections at MPIWG
- Fedora and Grid integration (flat files)
I. repository rests on grid
   - data perspective:
     repository mounts a data grid
   - the underlying "grid" could be any utility, including Amazon 3S, EC2 or GData

IIa. repository ties in grid
   - digital objects are distributed in grid
   - processing objects is distributed

IIb. repository embeds utilities
   - insular embedding of dedicated infrastructure utilities

III. gridified repository
   - repository services are distributed
   - ad hoc repository federations, etc.
   - repository *is* a repository utility, consisting of a number of dedicated utilities
Service-oriented Architecture

- Quality Assurance
- Authoring
- Depositing
- Search
- Browse

Solutions
- GUI-Tools
- GoogleMap
- OnRevToolkit
- weitere Solution

Application Services
- Depositing

Intermediary Services
- Item Validator
- Duplicate Detection

Basic Services
- Context
- Item
- Container
- Search

Core Infrastructure
- Fedora
- Migrara
- PostgreSQL
- Lucene

Scholarly Workbench
Interconnected set of Web Services
Thank you

Interested in collaborating?