HRA

A Virtual Research Infrastructure for Transcultural Research

International Symposium on Grid Computing,
Taipe 22. April 2009
Peter Gietz
Agenda

- The Cluster of Excellence „Asia and Europe“
- Design and building bricks of the Heidelberg Research Architecture
- Options for a central repository and its gridification
The Cluster of Excellence

- "Asia and Europe in a Global Context"
- Part of the German Federal Excellence Initiative
- "to establish internationally visible, competitive research and training facilities"
- A cluster of over 50 interdisciplinary projects
- Analysing cultural interactions between Asia and Europe in past and present
The Cluster of Excellence

- Has a dedicated agenda:
  - Finding and analysing flows of concepts, practices and institutions
  - That take place within shifting asymmetries
  - Globalisation is not a new phenomenon, but has happened since the beginning of mankind
  - The Global or transculturality is a phenomenon of its own
  - Everything is part of this global process
  - People are a medium of such flows

- More information available at:
  http://www.vjc.uni-hd.de
Heidelberg Research Architecture

- HRA is a virtual research infrastructure
- It consists of two main sections:
  - a database architecture that can be utilised for research projects such as the Translingual Concepts Database (TCD) and the Transcultural Images Database (TID);
  - the IT infrastructure required for a modern work environment, providing tools for close interaction within the Cluster's Research Areas and other scholarly publics.
HRA Requirements

Large amount of different databases already exist

- Language resources
- Image resources
- Bibliographical resources
- Music, Films, etc. etc.

Most Cluster project will have the need for such or new data bases
HRA the overall Architecture
Titel: Bapuji on eternal sleep
Künstler: Chowdhury, Sudhir
Gattung: 
Ort: 
Zusammenhang: 
Gebäude: 
Serientitel: 
Serientitel (Umschrift): 
Serie: 
Serie (Schriftzeichen): 
Detail: 
Element: 
Künstler: 
Autor: 
Pool-Inhalte/Kontakt: 
Suchen: 
Profisuche: 
Neues Bild: 
Hilfe: 
HeidiCON: 
Download: 
Editor: 
Detail: 
Aktivitäten: 

Flows in images

Start-up project „Archiving Mothers and Fathers of the Nation in Europe and Asia“
HyperImage Client (editor)
Window for manual defining parts of an image ("layers")
Web Services Interface

HyperImage

WS Client

SOAP / WSDL

WS Server

HeidIcon

EasyDB

SOAP / WSDL

WS Server

Fedora

DB
<table>
<thead>
<tr>
<th>Title</th>
<th>Year</th>
<th>Publication</th>
<th>Volume</th>
<th>Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Entry “zhengzi”</td>
<td>1999</td>
<td>Cihai 韩海</td>
<td></td>
<td>4164-4165</td>
</tr>
<tr>
<td>Concluding Remarks</td>
<td>1992</td>
<td>Habermas and the Public Sphere</td>
<td></td>
<td>462-479</td>
</tr>
<tr>
<td>Entry “zhengzi”</td>
<td>1990</td>
<td>Hanyu da cidian</td>
<td></td>
<td>424</td>
</tr>
<tr>
<td>Entry “zhengzi”</td>
<td>1979</td>
<td>Cihai</td>
<td></td>
<td>3355</td>
</tr>
<tr>
<td>Entry “zhengzi”</td>
<td>1965</td>
<td>Cihai 韩海</td>
<td></td>
<td>2307</td>
</tr>
</tbody>
</table>
QuotationFinder

- QuotationFinder is a software for automated text comparison.
- It provides an efficient way to find intertextual references, such as citations, plagiarisms and allusions.
- The source text can either be compared to another text or a set of texts.
- In order to carry out an efficient search, an index of the latter has to be available.
- For that reason QuotationFinder also contains a tool to create comprehensive full-text indexes.
Multimedia content

Moving objects

- Description of multimedia content using MPEG-7 standard
- Mark elements within a frame and follow it over time
- Scene-annotation, text transcription, searchable multilingual subtitles
- Evaluation of ELAN (Multimedia Annotator, Max Planck Institute for Psycholinguistics Nijmegen)
ELAN (MPI Psycholinguistics) Complex annotations on video and audio resources
Fedora consists of a Core Repository Service
- exposes its functionality via Web services
- object-oriented data model,
- no predefined hierarchy
- Separation between Front and Back End
- Open source software
- Is conceived to be the central repository for HRA
Fedora and Grids

- There are activities to integrate Fedora into Storage Grids
  - iRODS storage module for Fedora
  - iRODS is a grid storage software that interprets rules to decide how the system is to respond to various requests and conditions
- Within the D-Grid Project TextGrid, the current Grid Middleware will be integrated with Fedora
- The cluster is eager to evaluate the results of these activities
eSciDoc

- eSciDoc is an eResearch environment
- developed specifically for use by scientific and scholarly communities
- to collaborate globally and interdisciplinary.
- It comprises a set of services and solutions
- Based on Fedora
- Could be a valuable enhancement for HRA
What do we want to do with Ontologies?

- find integrated knowledge
- produce new knowledge
- provide evidence for new hypotheses
- verify or challenge old hypotheses
What do we want to do with Ontologies?

➤ There could be an overall ontology which includes classes and relation types that all cluster projects will need, e.g.,
  - classes like person, time, location, event
  - relations like hasType, wasInfluencedBy, isIdentifiedBy, isComposedOf

➤ Every cluster project using the ontology technologies could build a sub ontology fulfilling their very special needs,
What to start with?

CIDOC CRM

- is an international standard (ISO 21127:2006)
- Is a extensible ontology for concepts and information in the field of cultural heritages
- Meant for museums (any kind of), archives and libraries – also for information exchange
- Creates a formal semantic for information integration
- Supports standard technologies like RDF and OWL
Thank you!

Questions?

gietz@asia-europe.uni-heidelberg.de