Thoughts on a Distributed Web-Portal for World-Wide Collaboration Among Architectural Archives and Historians

Bernd Kulawik, Bern, Switzerland

Why? – From the researcher's perspective

- information needed for research is distributed among many archives, often internationally
- usually, research starts with known sources
- researchers have to contact every institution (they know of)
- since 1991 there is the Web, invented to connect research
- but most of the information stored in archive databases is still not available online or only through separated search, not through search engines like google.
- limited funding for searching, travelling, collecting information
- much of the research work is done multiple times by researchers

Why? – From a archives' perspective

- since the 1980s many archives use databases
- most of these databases are based on different software and concepts and therefore not compatible
- usually, these older databases are not „fit“ for the WWW
- if they are online, their content is not or only partially searchable through the web (ttw) via search-engines
- the concepts for data storage and search are very different
- access for researchers is limited or non-existent
- users usually cannot contribute their information, i.e. it has to be collected by archivists themselves again

Requirements: Structure

- institutions should be able to maintain their own web-space
- this distributed storage should appear to users as a single one
- access to different areas has to be regulated in several levels
- information should be stored in a structured, common way: …

Requirements: Structure: ordering principles

- structure should be „flat“ and understandable to users
- single archival objects should be treated as „objects“
- access to these data objects should be like in the „real world“
- objects should have their own, readable URL
- non-objects like menta-information should be linked to all relevant objects
- information based on interpretation should not be used as an ordering principle, because it is subject to change

Requirements: Technology

- data storage should be mirrored in several locations
- separate data storage for separate institutions preferable
- information should be accessable from a single entrance point
- all information should be accessable over the Web
- all information should be connected via links and Unique IDs
- structure should not be based on interpretation
- information should be managed through a structured workflow
- information should be virtually collectable in „pools“
- discussion and annotation of information should be possible

Requirements: Software

- free and open source software to avoid dependencies
- software should not be platform / operating system dependent
- only open web standards should be used, no proprietary types
- software should allow distritubed computing and storage
- an active, open and world-wide community of developers and service providers
- the software should be easy to use and require very short training for users (i.e. one day or less)

Requirements: Community

- institutions and their members as well as natural persons should be allowed to become members and contributors of information
- a workflow process has to be organised, inside institutions as well as for single members, to handle information
- all members should be allowed to contribute entries
- membership should be based on real names, not pseudonyms
- a committee should be formed to prepare and make decisions; its work should be as open as possible and web-based

Web Application Server ZOPE + CMS *Plone*

- ZOPE = Z Object Publishing Environment
  - free and open source, object-oriented
  - written in Python and C = platform independent
- allows human readable URLs and paths to data objects
- contains its own webserver and object-oriented database
- information from relational databases can be integrated
- ZEO (= ZOPE Enterprise Objects) allows distributed storage
- Plone: large community, award-winning, standard conform
  - allows frames to „run“ applications like data viewers ttw

Existing Portals: AAE and MACE

- Architecture Archives Europe: architecturalarchives.net:
  - restricted to European archives
  - web-site completely contained within a small „frame“
  - content not accessable and findable from the web
- MACE = Metadata for Architectural Contents in Europe
  - concentrates on metadata for (unstructured) information
  - metadata-set seems not to be translatable
  - all information again „locked“ inside a „frame“
  - many information from outside the portal, i.e. can disappear