# Infrastruktur zur Erhaltung und Nutzung des kulturellen Erbes im Informationszeitalter

## An Infrastructure for Cultural Heritage in the Information Age

Robert Casties, Dirk Wintergrün

Max-Planck-Institut für Wissenschaftsgeschichte
Wilhelmstr. 44, D-10115 Berlin
Tel.: 030/22667-342, Fax: -340
E-mail: itgroup@mpiwg-berlin.mpg.de, Internet: http://itgroup.mpiwg-berlin.mpg.de

**Zusammenfassung:**

Wir präsentieren Elemente einer elektronischen Arbeitsumgebung für Forscher in den Geisteswissenschaften. Diese Arbeitsumgebung unterstützt die interaktive Arbeit mit verschiedenen Arten von Quellen in Form von digitalen Images, Volltexten oder Datenbanken. Teile dieser Umgebung werden im Rahmen des EU-Projekts ECHO (European Cultural Heritage Online) entwickelt und verwendet.

**Abstract:**

We are presenting elements of an electronic working environment for scientists working in the humanities. This environment provides means for interactive work on different kinds of sources in form of scanned images full texts and databases. Parts of this environment are developed and used in the framework of the ECHO (European Cultural Heritage Online) project by the EU.

We imagine a working environment, exploiting the new information technologies, for the scientist of the future, working in the field of history. This workbench will provide access to numerous kinds of sources and other types of information. It will furthermore provide the instruments necessary to work with this information in new and innovative ways.

The scientist will be able to manage the sources necessary for his or her work:
- Accessing databases with bibliographic information
- Acquiring electronic and classical sources directly from his desktop
- Managing and building digital collections
- Sharing sources with his partners via the internet
- Disseminating his or her electronic publication on the Net, linked to the sources used for the research done.

Tools on the desktop will provide the following functions:
- Digitized images of source material can be viewed and shared in high quality, adapted to scientific needs.
- Language technologies will help to work with sources written in a variety of different languages.
- Specialized electronic dictionaries will be available online
- Online databases will support access to archival material
- An electronic document archive helps to organize publications

The European Initiative "European Cultural Heritage Online" (ECHO) aims to develop and to distribute open source tools necessary to reach this goal. We will present the vision of this project and the overall architecture and tools, which are currently under development in the framework of this project. (Christoph Liess will present the ECHO project in a related talk.)

Already now, a series of tools are available and used by researchers at our institute and its cooperation partners:

**digilib -- a versatile image viewing environment for the internet.**
- digilib supports a wide range of image formats and viewing options on the server side while only requiring an internet browser with JavaScript and a low bandwidth internet connection on the client side.
- digilib enables very detailed work on an image as required by scientists with elaborate zoom features like an option to show images on the screen in their original size.
- digilib facilitates cooperation of scientists over the internet and novel uses of source material by image annotations and stable references that can be embedded in URLs.
- digilib is also part of the ALCATRAZ framework of image, text and annotation tools. It is Open Source Software jointly developed by the MPIWG and the University of Bern and used by the LINEAMENTA project of the Bibliotheca Hertziana and many others.

**The Berlin Viewing Environment (BVE)**
- A flexible environment for displaying and working with books, manuscripts and other text-related content on the Internet.
- Morphological analyses for different languages enable automatic links to many specialized dictionaries and cross-links within the corpus.
- Text functions only need minimal XML markup that can be easily produced from most formats. Scanned page images can be linked to the text or used alone if no fulltext is available.
- All functions are accessible with a web browser from anywhere.

**Content Storage System**
- The MPIWG develops a new system for the storage and management of digital content for the ECHO project and other projects of the institute.
- The storage server is accessible like a network volume or through a web interface.
- Integral part of the storage system is a schema and mechanism for storing and accessing metadata in XML format for all objects in the storage system. All metadata information can be searched and used internally and externally.
- The system is integrated with digilib and BVE. Image and text functionality is available for any applicable document.
- A web frontend allows access to all functions and all data from anywhere with a simple web browser.