

Die Bilddatenbank der Photothek des Kunsthistorischen Instituts in Florenz MPI: eine Zope-basierte open-source-Lösung für die Präsentation von HiDA-Midas-Dokumenten im Internet

The image database of the photographic collection of the Kunsthistorisches Institut in Florenz MPI: a zope-based open-source solution for the presentation of HiDA-MIDAS documents in the Internet

Ruth Goebel
Talstr. 116
40217 Düsseldorf
Tel: 0211/343767, Fax: 0211/461825
E-Mail: rgoebel@fotostoria.de

Martina Hansmann
Kunsthistorisches Institut Florenz Max-Planck-Institut
Via G. Giusti, 44
I – 50121 Firenze
Tel.: 0039/055/249111, Fax: 0039/055/24911-55
E-Mail: m.hansmann@khi.fi.it, Internet: www.khi.fi.it

Zusammenfassung:

Im Rahmen eines durch die DFG finanzierten Projekts digitalisiert die Photothek des Kunsthistorischen Instituts Florenz 20.000 Originalaufnahmen aus ihren Beständen, um sie in Form einer webbasierten Bilddatenbank frei zur Verfügung zu stellen. Die Dokumentation der Metadaten erfolgt unter Verwendung des Datenbanksystems HiDA-MIDAS. In Zusammenarbeit mit der Photothek des Zentralinstituts für Kunstgeschichte in München hat die Photothek des KHI eine Webpräsentation entwickelt, die auf dem Open-Source-Tool ZOPE aufbaut, einem Applikations-server mit integriertem Content-Management-System. Für die Webdarstellung wird die hierarchische Struktur der HiDA-Daten in die ZOPE-eigene Datenbank implementiert, ohne die komplexe Struktur der Dokumente zu verändern. ZOPE bietet außerdem eine Reihe von Funktionen, die für webbasierte Datenbanken von besonderem Interesse sind. Zur Darstellung und Recherche der Bilder wurde das Open-Source-Tool Digilib an die Webdatenbank angeschlossen. Das System steht zur Nachnutzung und Weiterentwicklung für interessierte Partner zur Verfügung.

Abstract:

Within the framework of a project supported by the Deutsche Forschungsgemeinschaft, 20.000 original negatives of the Kunsthistorisches Institut Florenz are currently being digitized: They are now to become part of a freely accessible, web-based image database. For the documentation of the data we use the HiDA-MIDAS database system. In collaboration with the photographic collection of the Zentralinstitut für Kunstgeschichte in Munich, the Fototeca of the KHI has elaborated a web presentation of their HiDA-MIDAS documents that is based on the Open-Source-System ZOPE, an application server with integrated Content Management System. For this web presentation, the hierarchical structure of the HiDA data has been transferred into the ZOPE implemented database without changing the documents's original structure. In addition, ZOPE offers a series of functions that are particularly interesting for web-based database services. As an image viewer and study tool for images the Open-Source-Tool Digilib has been integrated into our web presentation. Interested partners are now invited to adapt this system and to contribute to it's further development.

Die seit der Gründung des Instituts im Jahr 1897 herangewachsene Photothek des KHI Florenz umfasst rund 580.000 Aufnahmen zur Kunst vor allem Ober- und Mittelitaliens. Zum weitaus größten Teil stammen diese Fotos von externen Photographen oder Bildarchiven; im Zentrum des Sammlungsaufbaus stand in der Vergangenheit stets die Zielvorstellung, einen Handapparat für die in Florenz arbeitenden ForscherInnen zur Verfügung zu stellen.

Die Neuerwerbungen werden seit 1993 mit dem Datenbanksystem HiDA-Midas erfasst. Damit war ein erster Schritt zur Digitalisierung der Metadaten getan. Doch wurde die Datenbank lediglich als eigenständiger, interner Katalog ohne direkte Verknüpfung zu den Bildern selbst geführt. 2002 hat die DFG den Antrag des KHI auf Digitalisierung unserer Originalaufnahmen bewilligt. Ziel war der Ausbau der Datenbank zu einer webfähigen Bilddatenbank, um hierdurch sowohl den Zugang zu unseren Materialien als auch die Qualität der Recherche verbessern zu können. Für die inhaltliche Erschließung der Digitalisate bot die schon bestehende Dokumentation die Grundlage. Der mit dem Projekt verbundene Übergang von einer vor allem intern genutzten Textdatenbank zu einem Instrument, das der Veröffentlichung und Bereitstellung unserer Fotodokumentation im Internet dient, hat allerdings einen Perspektivwechsel mit sich gebracht: In den Mittelpunkt des Interesses rückte nun die Erfassung der institutseigenen Spezialbestände. Dies geschah sowohl aus urheberrechtlichen Gründen als auch aufgrund einer Abwägung von Kosten und Nutzen der Digitalisierung. Die über Jahrzehnte angekauften Fotos anderer Archive stehen dem Benutzer natürlich weiterhin zu Studienzwecken zur Verfügung. Doch hat die Frage der Nutzungsrechte und -möglichkeiten der Bilder im Internet unsere Arbeitsstrategien einschneidend verändert. Angesichts der Verfügbarkeit eines Mediums, mit dem Bilder jederzeit ortsungebunden zugänglich sind, müssen in der Arbeit der Photothek bislang gültige Prioritäten kritisch überprüft werden. So verliert die Bereitstellung von Aufnahmen fremder Archive im gleichen Maß an Bedeutung, in dem auch diese ihre Ressourcen im Web verfügbar machen. Webbasierte Bilddatenbanken bieten dabei nicht nur neue Möglichkeiten zur Divulgation und Vernetzung von Bildmaterialien, sondern sie werden aus unserer Sicht außerdem zunehmend zu einem Arbeitsinstrument, das neue und gerade für Wissenschaftler hochinteressante Möglichkeiten der Bildrecherche bietet. Neben die traditionelle analoge Photothek des KHI wird in der Zukunft deshalb eine weltweit verfügbare digitale Photothek mit eigenem Profil treten.

Im September 2002 hat das auf zwei Jahre angelegte, im Oktober 2004 abgeschlossene Projekt die Arbeit aufgenommen. 20.000 Aufnahmen sind digitalisiert, in der Datenbank erschlossen und zur langfristigen Sicherung archiviert worden. Die digitalisierten Fotos umfassen u.a. die Kampagnen des langjährigen Institutsphotographen Artini, bedeutende Negativfondi wie die Aufnahmen toskanischer Architektur, Malerei und Plastik in den Archiven Bazzecchi und Laurati, die vor dem zweiten Weltkrieg durchgeführte Dokumentation der florentinischen Urbanistik von Hilde Lotz sowie de Giovannis detaillierte Aufnahmen der Ober- und Unterkirche von S. Francesco in Assisi. Eine Analyse des Erhaltungszustandes dieser Fondi erwies, dass große Teile des Negativarchivs im Erhalt gefährdet waren. Die Digitalisierung sollte daher nicht nur der verbesserten Verbreitung des Materials, sondern auch der langfristigen Bewahrung der Bildinformation dienen. Vor diesem Hintergrund fiel die Entscheidung, zur Sicherung der maximalen Information nach dem Negativ zu scannen und die Auflösung jeweils so anzupassen, dass die Bildinformation der Negative möglichst verlustfrei festgehalten werden konnte.

Die MPG ermöglichte die Anschaffung eines Imacon-High-End-Scanners, der diese Qualitätsstandards erfüllt und zudem das gesamte Spektrum unserer Negativformate und -Materialien verarbeiten konnte. Die Masterscans haben als TIFF im Schnitt ein Volumen von 100 MB, die überarbeiteten Arbeitskopien (ebenfalls im TIFF-Format) eine Größe von 20 MB. Zur Langzeitarchivierung werden die Bilddateien bei der Gesellschaft für wissenschaftliche Datenverarbeitung (GWDG) in Göttingen abgelegt, außerdem auf dem Server des Max-Planck-Instituts für Plasmaphysik in Garching.

Für die Bereitstellung einer Hida-Midas-Bilddatenbank im Web gab es bereits eine Softwarelösung, doch hätte - neben der Zahlung der Lizenzen - auf jeden Fall eine kostenintensive Anpassung an unsere Bedürfnisse erfolgen müssen. Nach unserer Überzeugung ist eine solche Investition in ein proprietäres System aus politischer und ökonomischer Sicht langfristig wenig sinnvoll, so dass wir

eine Open-Source-Lösung bevorzugt haben. In dieser Situation bot es sich an, mit dem Zentralinstitut für Kunstgeschichte in München zusammenzuarbeiten: Dort war für die Web-Präsentation des ebenfalls in HiDA erschlossenen historischen Farbdiagramm-Archivs ein Open-Source-Tool auf der Basis von ZOPE im Aufbau. Wir haben dieses System in Absprache mit dem ZI kostenfrei übernommen und die nächsten Ausbauschnitte finanziert. Mit dieser Entscheidung für eine Open-Source-Neuentwicklung bot sich gleichzeitig die Möglichkeit, die Recherche und vor allem die Darstellung von Suchergebnissen für unsere Bedürfnisse zu optimieren. Dabei kam uns die Tatsache zugute, dass ZOPE nicht nur eine Reihe von Funktionen (wie eine objektorientierte Datenbank mit Indizes, Kataloge usw.) mitbringt, sondern auch eine komplette Content-Management-Webumgebung für die Darstellung von Informationen im Netz. So war unser Zeit- und Kostenaufwand insgesamt sehr überschaubar, ohne dass wir auf Sonderwünsche verzichten mussten. ZOPE ist ein System, das permanent von verschiedenen Entwicklern ausgebaut wird. Dabei entstehen Open-Source-Erweiterungen, die für unsere Zwecke interessant sind (etwa Anwendungen, die auf geographischen Informationssystemen beruhen): Sie können mit vertretbarem Aufwand an unsere Bedürfnisse angepasst werden, weil sowohl das Grundsystem als auch die Erweiterungen im Quellcode zur Verfügung stehen. Als Vorteil erweist sich dabei auch die Tatsache, dass es in objektorientierten Systemen wie ZOPE technisch recht einfach ist, Erweiterungen anzuschließen.

Der wesentliche Teil der für das Projekt erfolgten spezifischen Programmierung ist faktisch am wenigsten sichtbar. Die Daten werden aus dem Hida-Export-Format (ULD) mittels eines Importskriptes in die ZOPE-eigene Datenbank (ZODB) eingelesen. Dem Nutzer stehen zum Browsen eine auf dynamischen Abfragen basierende hierarchische Ortssuche sowie eine Künstlersuche zur Verfügung, außerdem die gezielte Suche in einzelnen Indizes. Die Hida-eigene Struktur komplexer hierarchischer Dokumente zu Kunstwerken und Werkzusammenhängen, an die fotospezifische Informationen in sog. Fotogruppen angehängt sind, wurde dabei respektiert, um unsere Zusammenarbeit mit dem Marburger DISKUS-Verbund nicht in Frage zu stellen. Dieses Importskript wurde zunächst auf der Grundlage der bestehenden Daten der Fotothek des Zentralinstituts für Kunstgeschichte entwickelt und dann unserer sehr viel breiter angelegten Dokumentation angepasst. Damit hat dieser Teil der Programmierung einen Stand erreicht, von dem auch andere Anwender von Hida-Midas profitieren dürften, die nach Darstellungsmöglichkeiten ihrer Daten im Internet suchen. Als Open-Source Programmierung stehen das Importskript und alle weiteren Module zur weiteren Nutzung frei.

Wie sieht es nun mit der Darstellung der Daten im Web aus?

Hida-Midas verwaltet im strengen Sinn keine Bilder, sondern dokumentiert Kunstwerke. Daraus resultiert die bereits erwähnte komplexe und hierarchische Struktur der Dokumente, die in der Darstellung der Daten berücksichtigt werden muss. In der gezielten Suche kann einerseits mit einer Freitexteingabe, andererseits mit Listenindizes operiert werden, wodurch sich die Recherche deutlich flexibilisiert hat. Dabei können nicht nur verschiedene Indizes miteinander verknüpft werden, sondern mittels bool'scher Operatoren auch Begriffe innerhalb eines Index. Dem Benutzer steht eine Suche nach Künstler, Standort und Bauwerk zur Verfügung, außerdem bietet das System ikonographische Suchmöglichkeiten wie die gezielte Recherche nach dargestellten Personen oder eine strukturierte Suche nach Bildinhalten, mit der Möglichkeit, diese auf bestimmte ikonographische Klassen einzuschränken. In der Datumssuche wird mit mathematisch definierten Zeiträumen operiert, so dass auch Teilmengen eines Zeitabschnitts Treffer liefern.

Eine der wesentlichen Aufgaben bestand darin, die komplexe, bisweilen auch unübersichtliche hierarchische Struktur der Hida-Midas-Dokumente adäquat darzustellen. Die Dokumente werden nicht mehr in flachen Trefferlisten, sondern in einer Kombination aus Trefferliste und aufklappbarer Strukturansicht in einem eigenen Fenster angezeigt. Hierdurch kann der Benutzer seine Suchergebnisse zügiger und effizienter sichten und erfährt schon in der Kurzanzeige, dass sich weitere Informationen hinter den zunächst erscheinenden Angaben befinden.

Die Hida-Dokumente werden in der Vollanzeige übersichtlicher und strukturierter ausgegeben. Die Ausgabe erfolgt nicht einfach feldweise, also als vollständige Liste der Einträge im Datenblatt, sondern sie wurde deutlich gestrafft und strukturiert. Dies ist technisch mit einer Reihe im

Hintergrund laufender Abfragen verbunden, die an das System einige Ansprüche an Rechenleistung und Performance stellen, aus Sicht einer größeren Nutzerfreundlichkeit aber notwendig erschienen.

Die ZOPE-basierte Präsentation der Bilddaten wurde ab Sommer 2004 um ein weiteres Open-Source-Tool namens DIGILIB ergänzt, das u.a. am Max-Planck-Institut für Wissenschaftsgeschichte in Berlin und der Universität Bern entwickelt wird. Digilib ist ursprünglich für die Digitalisierung von Büchern entwickelt worden, findet inzwischen aber zusehends auch für Bildmaterialien Verwendung. Die digitalen Bilder werden mit Hilfe der Datenbank auf dem Digilib-Server angesprochen. An diesem Punkt eröffnen sich dem Benutzer verschiedene Möglichkeiten, mit dem Bild zu arbeiten. Es stehen komfortable Zoom-Werkzeuge zur Verfügung, mit denen der hohe technische Aufwand unserer Digitalisierung voll genutzt werden kann. Außerdem besteht die Möglichkeit, im Bild Marker zu setzen und Bilder und Bildausschnitte durch feste URLs zu referenzieren.

Die technische Grundlage sieht kurz gefasst wie folgt aus: Bei einer Suchanfrage liefert der Server dem Nutzer immer nur den jeweils angesteuerten, aus dem 20 MB großen TIFF-Bild ‚on the fly‘ errechneten Bildausschnitt (Gesamtansicht oder Zoom), und zwar als PNG-Datei von etwa 550 x 400 Pixel (200-250 KB). Da immer nur die tatsächlich notwendige Datenmenge geliefert wird, ist das System im Web äußerst performant. Zudem stellt dieses Verfahren einen gewissen Schutz für unsere Bilder dar, da eben immer nur PNGs mit der genannten Größe auf den eigenen Rechner gespeichert werden können, nie aber das ganze Bild im TIFF-Format. Digilib zielt also nicht primär darauf ab, Bilder zum Download zur Verfügung zu stellen und deren weitere Nutzung und Handhabung dem Benutzer zu überlassen. Doch hat der Nutzer den Vorteil, dass er jederzeit kostenfrei auf eine webbasierte Arbeitsumgebung mit zahlreichen Funktionen zugreifen kann, ohne sein lokales System belasten zu müssen.

Digilib ist ein System, in dem alle Anfragen an den Bildserver ausschließlich über URLs gesteuert werden. Das hat den Vorteil, dass der Nutzer durch Marker oder sogar durch eine Bearbeitung aktiv in die Bilder eingreifen kann. Diese Eingriffe werden nicht serverseitig gespeichert und stellen damit für unsere Hardware keine Belastung dar. Sie werden vielmehr beim erneuten Aufrufen der URL aus dem zugrunde liegenden Bild (TIFF) errechnet. Dieses System, Bilder, Bildausschnitte und Veränderungen über URLs anzu steuern zu können, bietet damit ganz neue Möglichkeiten, Bilder zu zitieren und zu referenzieren: Sorgt man für entsprechend stabile Domainadressen, sind die URLs in Digilib und ZOPE prinzipiell zitierfähig.

Zukunftsperspektiven:

Das KHI plant, den Inhalt seiner Bilddatenbank kontinuierlich auszubauen. Die Bilder aus den neuen Fotokampagnen der Phototek, die einerseits der Erstellung von Referenzaufnahmen bekannter Werke, und andererseits der Dokumentation neuer Forschungsmaterialien gewidmet sind, sollen auch den nicht in Florenz arbeitenden KollegInnen zeitnah über das Netz zugänglich gemacht werden. Die Nutzung von Digilib zielt dabei auf eine neue Qualität in der Bildrecherche und Bildanalyse, die über die Möglichkeiten der Konsultation eines konventionellen Abzugs weit hinausgehen, wie die bis ins Detail präzise Zitierfähigkeit und die komfortablen Zoommöglichkeiten.

Die Webpräsentation des Zentralinstituts für Kunstgeschichte in München und des KHI Florenz steht programmatisch für eine Nachnutzung zur Verfügung. Unser Engagement für eine Open-Source-Lösung war wesentlich dadurch motiviert, dass unsere Webdatenbank durch einen größeren Kreis weiterverwendet werden kann. Nach unserer Überzeugung sollte öffentliche Förderung vorrangig in den Aufbau und Unterhalt von Open-Source-Lösungen investiert werden. Denn der Rückgriff auf einen offenen Standard stellt sicher, dass die Pflege und damit letztlich die Existenz von Systemen nicht von einem einzigen Dienstleister oder Anbieter abhängig ist. Vielmehr können Open-Source-Tools langfristig gemeinsam mit anderen Interessenten weiterentwickelt werden. Dabei können und sollten sich in Anpassung an unterschiedliche Bedürfnisse unterschiedliche Funktionalitäten ausbilden. Der Einsatz graphischer Elemente und visueller Hilfestellungen wie Landkarten und Pläne beispielsweise steckt noch in den Anfängen, und es steht zu hoffen, dass diese Anwendungen mit weiteren Partnern aufgebaut werden können.