

DFG-Projekt DIGIPLAN

Vierteljahresbericht
April 2007

Inhalt: 1. Projektstand / 2. Workflow (Änderungen) / 3. Sonstige Aktivitäten / 4. Ausblick

Die [Deutsche Forschungsgemeinschaft](#) finanziert seit April 2006 die Digitalisierungskampagne DIGIPLAN, die in drei Jahren nahezu den gesamten Bestand des „Architekturmuseums der Technischen Universität Berlin in der Universitätsbibliothek“ (Namenswechsel zum 1.1.2007) digital faksimilieren und in hoher Auflösung (3.000 Pixel Bildbreite) über das Internet zur Verfügung stellen soll.

Die Geschichte der Dokumentation unseres Bestandes und die Intentionen des Digitalisierungsprojekts sind bereits in den beiden früheren Vierteljahresberichten ausführlich geschildert worden. Im folgenden werden der Ablauf der Digitalisierung und die dabei zu Tage getretenen Probleme ausführlicher als bislang geschildert – nicht nur, um der Berichtspflicht zu genügen, sondern um die Erfahrungen zum Nutzen anderer Digitalisierungsprojekte zu veröffentlichen.

1. Projektstand

Zwischen Februar 2006 und April 2007 wurden auf dem eigenen [CRUSE](#)-Scanner ca. 6.700 Nutzeroaufträge und Überformatblätter gescannt.

Der ab August 2006 tätige Dienstleister scannte zwischen September 2006 und April 2007 ca. 13.700 Blatt.

Insgesamt lagen so Anfang April **20.400 Scans** vor.

2. Workflow

Die Digitalisierung des Architekturmuseums erfolgt aufgrund der Empfindlichkeit des Materials nicht in der Reihenfolge der Inventarnummern, sondern schrank- bzw. schubweise. Die in der Datenbank vorhandenen Daten zu einem Schub werden mit Hilfe eines mittels WLAN mit dem Server verbundenen Laptops im Depot kontrolliert. Die Datenkontrolle erfolgt also im Zuge des Aushebens für den Transport zum Scanner. Die kontrollierten Daten werden anschließend dem Dienstleister, der den überwiegenden Teil der Scans erzeugt, zusammen mit den zu scannenden Plänen übergeben. Der Dienstleister übergibt die Scans auf externen Festplatten an das Architekturmuseum.

Beide Scanner erzeugen zusammen bis zu 320 GB Daten (180 Dienstleister, 140 Architekturmuseum) pro Monat. Die Scans werden wöchentlich in die MuseumPlus-Datenbank und in die Online-Datenbank importiert, so dass der Arbeitsfortschritt der Allgemeinheit zeitnah zur Verfügung steht.

2.1. Datensicherung

Jeder Scan wird als Masterimage im TIFF-Format (verlustfrei LZW-komprimiert; Farbprofil ECI-RGB) gespeichert. Aufgrund der stark variierenden Blattgrößen sind die Dateien zwischen 1 und 600 MB groß (erwartetes Gesamtdatenvolumen des Projekts ca. 8 TB). Die Sicherung der Masterimages erfolgt dreifach:

- ▶ Speicherung der Tagesproduktion auf DVD. Zusätzlich werden nach der optischen und technischen Kontrolle der Scans die dabei verbesserten Dateien auf DVD gebrannt. Alle DVDs, die älter als vier Wochen sind, werden im Keller eines 1 km entfernten Bibliotheksgebäudes gelagert.
- ▶ Speicherung auf drei Wechselfestplattenreihen im Wochenrhythmus. Alle vollen Festplatten der Sekundären Plattenreihe werden ebenfalls im Keller des 1 km entfernten Bibliotheksgebäudes gelagert.
- ▶ Planung: Wöchentliche kummulierende Sicherung im Speichernetzwerk (SAN) des Rechenzentrums der Technischen Universität Berlin

Aus den Masterimages werden zahlreiche JPG-Größen für verschiedenen Zwecke (interne Arbeitsdaten, Herausgabe an Nutzer, Online-Datenbank, Synchronisation mit anderen Bilddatenbanken) berechnet.

2.2 Qualitätsprobleme des Dienstleisters

Die im ersten Dreivierteljahr der Projektzeit aufgetretenen Qualitätsprobleme der Dienstleistungs-Firma betrafen folgende Bereiche: Produktionsausfall und mangelnde Lieferqualität.

2.2.1 Produktionsausfall:

- ▶ Die Produktion begann vier Wochen nach dem vereinbarten Starttermin 1. Juli 2006.
- ▶ Die versäumte Validierung der Scanoptik zu Projektbeginn führte zu einem Produktionsausfall vom 22. August bis zum 8. September 2006. Die vor dem 21. August erstellten 600 Scans waren unscharf und mussten wiederholt werden.
- ▶ Ein weiterer Produktionsstopp erfolgte aufgrund dringender Workflow-Überarbeitungen vom 19. Februar bis 12. März 2007

Der bisherige Produktionsausfall beträgt sechs Wochen (mindestens 3.000 Blatt).

2.2.2 Lieferqualität:

- ▶ September 2006 bis Januar 2007: von den 7.500 erzeugten Scans war jeder 20. Scan (5 Prozent) fehlerhaft (versehentliche Löschung, zerschnitten, unscharf).
- ▶ Oktober 2006: Der Dienstleister übergibt die Scans entgegen der vertraglichen Vereinbarung nicht getaggt, da er sich außerstande erklärt hat, einen validen Tagging-Algorithmus zu entwickeln. Stattdessen taggt die TU die Scans seit dem mit einem von Herrn Dr. Nägelke geschriebenen VBA-Tool.
- ▶ März 2007: 966 Scans wurden statt in 300 dpi Auflösung in 200 dpi an uns ausgeliefert. Zudem ca. 2.500 Scans mitunter auf 800 dpi interpoliert und dann unter Aufgabe der Proportionen (!) auf 200 dpi gesetzt.

3. Sonstige Aktivitäten: *Workshop im Januar 2007*

Ende Januar führten wir einen mit über 100 Teilnehmern erfreulich rege besuchten Workshop durch, auf dem ein intensiver Informationsaustausch über Digitalisierungs- und Archivierungsfragen stattgefunden hat. Um diesen Austausch weiter zu befördern, wurde eine Liste mit allen Namen und Emailadressen an die Teilnehmer verschickt.

4. *Ausblick*

Im Sommer und Herbst 2007 sind weitere Workshops und Tagungspräsentationen geplant, die der Verbreitung und Validierung der im DFG-Projekt DIGIPLAN gemachten Erfahrungen dienen.