

PROBADO 3D - Integration von 3D-Objekten in Digitale Bibliotheken. Ein Dienstleistungsangebot für Architektur und Ingenieurwesen zur Erschließung und Bereitstellung von Multimediadokumenten

Ina Blümel, Dr. Irina Sens
TIB Hannover

„Je stärker wissenschaftliches Arbeiten auf IuK-Technologien ausgerichtet ist, desto eher bringen wissenschaftliche Erkenntnisse neue Ausdrucksformen jenseits von Text und Bild hervor. Experimentelle Untersuchungsreihen liefern komplexe und große Datensätze, erzeugen Modelle, Animationen und Simulationen. Diese müssen nachvollziehbar und reproduzierbar sein und sie müssen deshalb adäquat elektronisch publiziert werden können. [...] Hinzu kommt, dass sich bisherige Such- und Präsentationsverfahren in erster Linie an textuellen Dokumenten orientieren. Mittlerweile kommen weitere multimediale Formen ins Spiel (Abbildungen, Fotos, Videos, Speech sowie verschiedene Hybridformen), für die eine Erweiterung der Funktionalität notwendig ist. Selbstverständlich sind für die o.g. neuen Verfahren der Wissenschaftskommunikation die Suchverfahren zu adaptieren.“ [1]

Der Auszug aus dem Bericht „Neuausrichtung der öffentlich geförderten Informations-einrichtungen“ der Bund-Länder-Kommission beschreibt einen Strukturwandel in der Informationsvermittlung wissenschaftlicher Bibliotheken und erklärt das verstärkte Interesse an Verfahren und Werkzeugen für die Handhabung von Multimediainhalten in Bibliotheken, deren Dienstleistungen bisher weitestgehend auf Informationsbestände beschränkt sind, die in textueller Form vorliegen. Um den aktuellen Anforderungen der Fachinformationsvermittlung standzuhalten ist es unabdingbar, Dienstleistungskonzepte anzupassen und die neuen Dokumenttypen, z.B. 3D-Modelle der Architektur, nahtlos in die Bearbeitungsprozesse digitaler Bibliotheken zu integrieren. Das von der Deutschen Forschungsgemeinschaft seit Februar 2006 als Leistungszentrum für Forschungsinformation geförderte Projekt PROBADO (Prototypischer Betrieb allgemeiner Dokumente; <http://www.probado.de>) entwickelt Verfahren für die automatische Erschließung, Verfügbarkeit, Recherche und Bereitstellung dieser Medientypen, und zwar exemplarisch für 3D-Graphik und Musik mit der Möglichkeit, PROBADO auch auf andere Medien zu erweitern. Zunächst soll PROBADO in die Angebote der TIB Hannover und BSB München integriert und später auch dezentral verfügbar gemacht werden können. Neben TIB und BSB sind drei Informatik-Lehrstühle der Universitäten Bonn, Graz und Darmstadt Projektpartner. Die Fertigstellung des PROBADO-Dienstes ist für 2011 geplant.

3D-OBJEKTE

Was aber erschwert die Handhabung von multimedialen Objekten und warum erschließen sich die Inhalte nicht mit konventionellen Methoden? Wird beispielsweise eine 3D-Modelldatei im herkömmlichen Texteditor geöffnet, sind lediglich Zahlen zu

sehen, siehe Abb. 1 oben. Diese repräsentieren Koordinaten der Punkte im Raum, zwischen denen sich ein Gitter aufspannt, welches das Modell darstellt. Inhalte sind nicht zu erkennen, also, ob das Modell ein Bauteil ist, z.B. aus dem Maschinenbau, aus dem Fachbereich Archäologie stammt oder ob es sich um ein Architekturmodell handelt, bspw. ein Gebäude. Im Zeichenprogramm ist sofort zu erkennen: Architektur, Gebäude, drei Stockwerke. Muss also jedes zu erschließende 3D-Modell vom sachkundigen Bibliothekar im entsprechenden CAD-Programm geöffnet, besehen und manuell erschlossen werden? Täglich werden tausende 3D-Modelle im Architekturkontext produziert und bei einer Veröffentlichung von den Autoren in den seltensten Fällen umfassend mit Metadaten annotiert.

```
mt111b index_1modell101.mt1
# object object #67
g object_67
v 4.18077 26.93720 -2.40000
v 4.18077 19.73720 -2.40000
v -4.51923 19.73720 -2.40000
v -4.51923 5.78720 -2.40000
v -14.51923 5.78720 -2.40000
v -14.51923 26.96720 -2.40000
v -11.23013 26.96720 -2.40000
v -11.23013 26.66720 -2.40000
v -1.91923 26.66720 -2.40000
v -1.91923 26.93720 -2.40000
v 4.18077 26.93720 -2.39000
v 4.18077 19.73720 -2.39000
v -4.51923 19.73720 -2.39000
v -4.51923 5.78720 -2.39000
v -14.51923 5.78720 -2.39000
v -14.51923 26.96720 -2.39000
v -11.23013 26.96720 -2.39000
v -11.23013 26.66720 -2.39000
v -1.91923 26.66720 -2.39000
v -1.91923 26.93720 -2.39000
# 20 vertices
vn 0.00000 0.00000 -1.00000
vn 0.00000 0.00000 1.00000
vn 0.00000 1.00000 0.00000
vn 1.00000 0.00000 0.00000
```

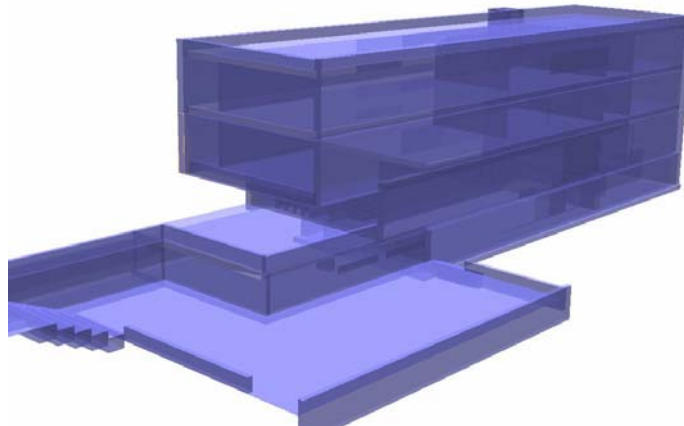


Abb. 1: 3D-Objekt im Texteditor und im 3D Programm

ERSCHLIESSUNG

Ein Hauptziel des Projektes ist die Minimierung manueller Katalogisierungsarbeit. Wie kommt man also möglichst automatisch an Information in 3D-Modellen? Bei der Einpflege von 3D-Modellen ins System werden zunächst technische Aspekte wie Modellgröße und Höhenrichtung ausgelesen, Modelle in das von den Suchmaschinen erkennbare Format konvertiert, Anzeigeformate generiert und störende Elemente wie z.B. Bodenplatten entfernt. Das sieht der Benutzer nicht unbedingt in den Metadaten, ist im Erschließungsprozess aber wichtig für die spätere Durchsuchbarkeit.

Weitere Indexdaten werden durch inhaltsbasierte Erschließung gewonnen. Zum einen können formbeschreibende Merkmale hinsichtlich ihrer Bedeutung für eine bestimmte Klasse von manuell vorklassifizierten Objekten bewertet werden, wobei Wissen aus der Annotation von Modellen extrahiert und für das Retrieval genutzt wird. Innerhalb des Projekts existiert eine annotierte Datenbank für architekturenspezifische Bauelemente, mit deren Hilfe neu zu erschließende Modelle klassifiziert werden können. Bei der Suche nach Komponenten innerhalb eines Gebäudemodells bietet es sich zusätzlich an, die oftmals aufgrund des Modellierungsprozesses vorhandene Strukturierung des Modells auszunutzen. Innerhalb einer solchen Strukturierung werden geometrische Primitive wie Dreiecke und Polygone zu semantischen Einheiten wie Bauelemente, Gebäudeabschnitte, Stockwerke usw. gruppiert. Durch Extraktion der einzelnen Einheiten wird die Suche nach einzelnen Bauelementen innerhalb eines Gebäudes möglich, siehe Abb. 2 links. Für den Interpretationsprozeß werden weitere semantische Schichten definiert. Zum einen sind dies physikalisch-logische, wie die Gebäudestruktur, oder z.B. der Gebäudetyp unter Berücksichtigung von Topologie, siehe Abb. 2 unten, Bruttogeschoßfläche und anderer relevanter Eigenschaften. Für technische Details wird auf [2][3] verwiesen.

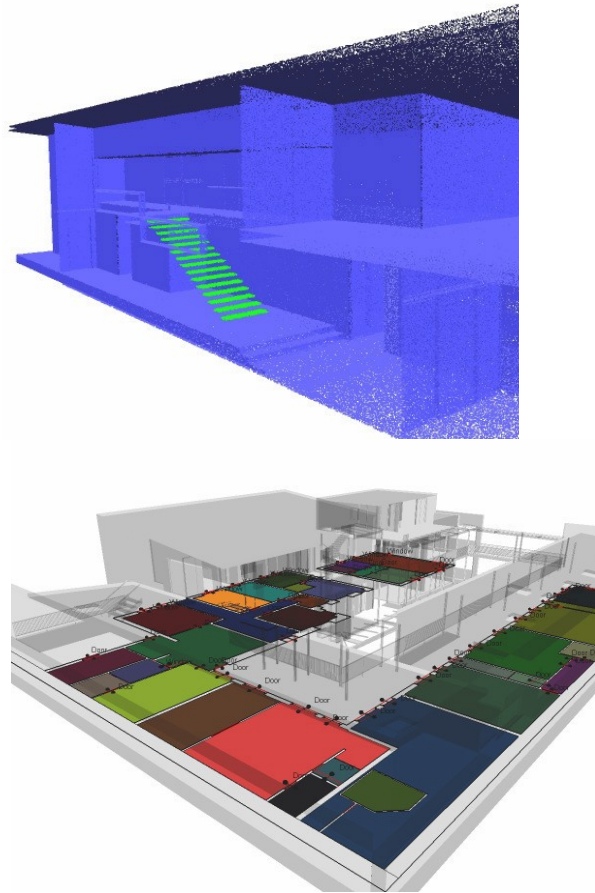


Abb. 2: Inhaltsbasierte Erschließung. Oben: Finden von Bauelementen. Unten: Erkennen von Geschossen, Räumen, Raumverbindungen und Repräsentation im Graphen.

METADATEN

Im Architekturkontext kommt den Metadaten für 3D-Modelle eine besondere Bedeutung zu, da sie neben der Annotation und Klassifikation die Möglichkeit bieten, Modelle zu anderen Themen, beispielsweise gebaute Architektur oder ausgelobte Wettbewerbe in Relation zu setzen. Das kann für den Nutzer in der Praxis einen erheblichen Mehrwert bei der Ergebnisanzeige darstellen.

Die PROBADO-3D-Metadaten sind in drei Gruppen organisiert: Die zentrale Entität aus Gruppe 1 ist „Modelinfo“, die als abstrakte Darstellung eines 3D-Modells beschrieben werden kann. Die Entität „Modelfile“ stellt technische Metadaten des konkreten Modells zur Verfügung. Meistens liegt ein Modell in unterschiedlichen Formaten vor, die wie Instanzen eines 3D-Modells behandelt werden können. Somit kann

„Modelinfo“ mehrfache Relationen zum „Modelfile“ haben. Ein detailliertes Gebäudemodell enthält normalerweise weitere Modelle, die extrahiert und separat als Modell nachgewiesen werden können. Analog ist die Schachtelung von „Modelinfo“ möglich.

Gruppe 2 umfaßt Informationen zu Personen oder Körperschaften, die mit 3D-Modellen (Erstellung, Eigentum, Urheber), mit „Object“ (Architekt, Designer) oder mit „Event“ (z.B. wettbewerbsauslobende Institution, Institut des Seminars) in Beziehung stehen. Gruppe 3 enthält optionale Entitäten, die erweiterte Informationen über ein Modell liefern, und zwar „Object“ (gebaute Architektur, Konstruktion oder Designobjekt), „Place“ (Ortsangabe), „Event“ (Ereignis, z.B. Seminar oder Wettbewerb) und „Subject“ (Begriff, Klassifikation). Die Gruppen 2 und 3 sind an das bibliothekswissenschaftliche Datenmodell für bibliographische Metadaten FRBR [4] angelehnt. Innerhalb des Metadatenschemas finden sich fast alle 15 Dublin-Core-Elemente [5] als Attribute oder Entitäten wieder. Das Schema ist so angelegt, dass es Metadaten von 3D-Modellen anderer Fachbereiche als Architektur aufnehmen kann. Zusätzliche Attribute können aufgenommen werden, ohne dass das Datenschema erweitert werden muß.

PRAXIS

Warum werden 3D-Modelle in der Architektur nachgefragt? Mit Computer Aided Design (CAD) erstellte 3D-Modelle sind aus dem Planungsalltag von Architekten und Ingenieuren nicht mehr wegzudenken. Wo früher analog mit Tusche auf Papier gezeichnet und Modelle in der Werkstatt gebaut wurden, entwirft und plant der Architekt heute mit dem Computer. Im Baubereich gibt es 3D-Modelle ganzer Städte, Gebäudemodelle, Modelle technischer Bauteile und 3D-Objekte zur Visualisierung von Innen- und Außenräumen. Im Planungsprozeß wird beispielsweise immer weniger einem Fachbuch oder Zeitschriftenartikel mit einer Baubeschreibung gesucht, die dann in CAD umzusetzen ist, sondern nach den Modellen selbst, also konkreten Lösungen eines bestimmten Problems. Das Finden von 3D-Modellen anhand individueller Parameter erfordert über die rein textbasierte Suche hinausgehende Möglichkeiten, z.B. das Suchen mittels einer Skizze.

RECHERCHE

Für die 3D-spezifische Suche gewährt das PROBADO-System sowohl klassisch textbasierte als auch mehrere visuelle Suchmöglichkeiten. Die Suche mittels Raumverbindungsgraphen benutzt die Raumkonfiguration bzw. Topologie von Gebäudemodellen. Diese Informationen können für eine 2D-Anfrage mittels am Bildschirm eingegebenen Graphen verwendet werden. Auch dreidimensionale Anfragen nach Gebäuden oder Bauelementen bzw. Objekten können interaktiv skizziert werden. Dabei werden die Ergebnisse noch während des Zeichnens zurückgeliefert, so daß das Volumen so lange modifiziert werden kann, bis das gewünschte Suchergebnis vorliegt, siehe Abb. 3 oben. Alternativ kann der Architekt CAD-Software, die für PROBADO mittels PlugIn als Eingabeinterface verwendet werden kann, zur Erstellung des Anfragemodells benutzen. Die Komponentensuche wird gebraucht, um Bauelemente wie Dächer,

Treppen etc. innerhalb von Gebäudemodellen bzw. als einzelne Modelle in der Datenbank zu finden. Die Suchmaschine benutzt Query-by-example, d.h. der Benutzer kann Modelle in das System hochladen, um ähnliche Komponenten zu finden. Bei der Ergebnisanzeige wird dem Benutzer eine Weboberfläche zur Interaktion bereitgestellt. Relevance Feedback des Benutzers kann vom System wiederum für die verfeinerte Suche und Erschließung der Modelle verwendet werden. Zweidimensionale Ergebnisanzeigen visualisieren nachvollziehbar ähnliche Modelle als Cluster auf einer Fläche, siehe Abb. 3 unten. Die Suchen im 3D-Repository werden in [6] weiter veranschaulicht.

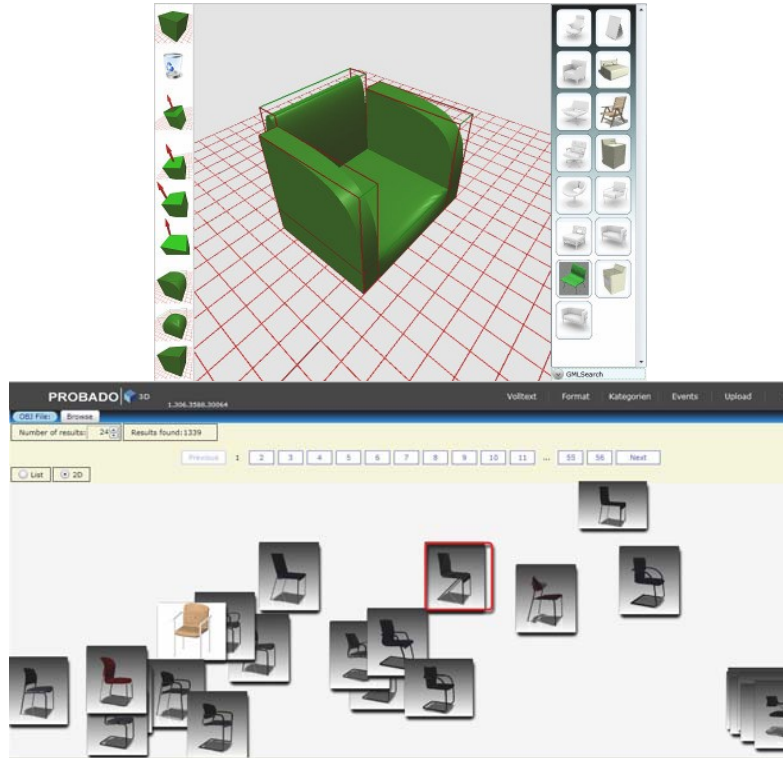


Abb. 3: Suche per 3D-Skizze mit Ergebnissen in sequentieller Darstellung sowie Ergebnisanzeige in 2D-Darstellung

SYSTEM

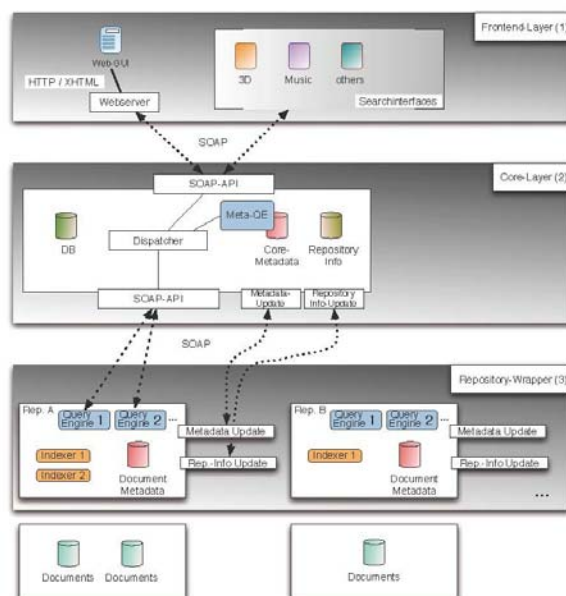


Abb. 4: Aufbau des PROBADO Frameworks

Abbildung 4 gibt einen Überblick über die PROBADO Systemarchitektur. Das Modell besteht aus drei Schichten: (1) Benutzerschnittstellen, (2) Systemschicht und (3) Repositories. Service Oriented Architecture (SOA) wird verwendet, um die verschiedenen Schichten des Systems zu verbinden. PROBADO-Services können ausgetauscht werden und jeweils in unterschiedlicher Programmiersprache implementiert sein. Die Kommunikation wird durch die Web Service Description Language (WSDL) beschrieben. Schicht 1 ermöglicht dem Benutzer, sowohl inhaltsbasierte als auch klassisch textbasierte Suchanfragen zu stellen. Inhaltsbasierte Suchen können nur über die jeweiligen Multimedia-Dokumenttypen gestellt, textbasierte Anfragen über alle Dokumente abgegeben werden. Anfragen können sowohl über PROBADO-Suchseiten als auch von anderen Bibliothekssuchportalen gestellt werden. Schicht 2 nimmt Anfragen entgegen und ordnet sie den jeweiligen Suchmaschinen in den Repositories zu. Schicht 3 umfasst die einzelnen Repositories spezifischer Multimediadokumenttypen mit den Rohdaten, den Indexdaten, den Metadaten und den dokumentenspezifischen Suchmaschinen. Aktuell gibt es zwei PROBADO-Repositories: eines für 3D-Modelle und eines für Musikdokumente. Die Dokumente müssen nicht notwendigerweise in den Repositories vorgehalten werden. Analog zur Situation bei Forschungsdaten im Bereich Erde und Umweltwissenschaften ist ein kooperatives Modell zwischen PROBADO bzw. Bibliotheken für den Nachweis und die Bereitstellung und externen Datenzentren für das physische Speichern der Daten vorstell- und umsetzbar [7] [8]. Weitere Informationen zu Systemarchitektur und Repositories sind in [9] [10] dargestellt.

INTEGRATION

Der modulare Aufbau des Systems erleichtert die Integration in Bibliotheksportale. An der TIB als zentraler Fachbibliothek für Technik sowie Architektur, Chemie, Informatik, Mathematik und Physik soll PROBADO 3D als zusätzliches Dienstleistungsangebot für Architekten und Ingenieure über „GetInfo, das Fachportal für Technik und Naturwissenschaften“ erreichbar sein. Dabei sollen die Metadaten aus PROBADO über die textuelle Suche aus GetInfo durchsucht und Ergebnisse zusammen mit textuellen Dokumenten angezeigt werden. Zum Download und für PROBADO-spezifische Suchmöglichkeiten wird der Benutzer auf die PROBADO-Seiten weitergeleitet. Die Integration in Bibliotheken wird auch in [11] veranschaulicht.

In PROBADO ist außerdem die Vergabe von Digital Object Identifiern (DOIs)-Namen für multimediale Objekte vorgesehen. Mit DOIs werden digitale Objekte zitierfähig. Der Architekt kann also eine dauerhafte Referenzadresse für sein 3D-Modell erhalten. Die TIB ist seit 2005 DOI-Registrierungsagentur für wissenschaftliche Primärdaten (Forschungsdaten), hat ihr Angebot mittlerweile aber auf alle Arten nicht-textuelle wissenschaftliche Inhalte ausgeweitet, siehe [7] [8].

REFERENZEN

- [1] „Neuausrichtung der öffentlich geförderten Informationseinrichtungen“ Abschlussbericht der Bund-Länder-Kommission-Arbeitsgruppe „Zukunft der Fachinformation“, Heft 138, S. 7f, 20, 2006, ISBN 3-934850-85-5.
- [2] Wessel R., I. Blümel, and R. Klein: The Room Connectivity Graph: Shape Retrieval in the Architectural Domain. In Proceedings of The 16-th International Conference in Central Europe on Computer Graphics, Visualization and Computer Vision, 2008, ISBN 978-80-86943-15-2.
- [3] R. Wessel, I. Blümel, R. Klein: A 3D Shape Benchmark for Retrieval and Automatic Classification of Architectural Data. EUROGRAPHICS 2009 Workshop on 3D Object Retrieval, March 2009.
- [4] <http://www.ifla.org/VII/s13/frbr/frbr.pdf>.
- [5] <http://dublincore.org>.
- [6] I. Blümel, R. Wessel, H. Krottmaier: The PROBADO Framework: A Repository for Architectural 3D-Models. In Browsing Architecture, Metadata and Beyond, Fraunhofer IRB Verlag, 2008, ISBN 978-3-8167-7770-0.
- [7] J. Brase: Using digital library techniques - Registration of scientific primary data in Research and advanced technology for digital libraries, Springer LNCS 3232, 2004, ISBN 3-540-23013-0.
- [8] J. Brase, J. Klump: Zitierfähige Datensätze: Primärdaten-Management durch DOI. In Wissenschaftskommunikation der Zukunft in Schriften des FZ Jülich - Reihe Bibliothek, Band 18, 2007, ISBN 978-3-89336-459-6.

- [9] Krottmaier, H., F. Kurth, T. Steenweg, H. J. Appelrath, and D. Fellner: PROBADO – A Generic Repository Integration Framework. In Proceedings of Research and Advances Technology for Digital Technology: 11th European Conference – ECDL 2007, Budapest, Springer 2007, ISBN 978-3-540-74850-2.
- [10] I. Blümel, J. Diet, H. Krottmaier. Integrating Multimedia Repositories into the PROBADO Framework. Third International Conference on Digital Information Management (ICDIM), 2008.
- [11] I. Blümel, I. Sens Das PROBADO-Projekt: Integration von nichttextuellen Dokumenten am Beispiel von 3DObjekten in das Dienstleistungsangebot von Bibliotheken. S. 79 ZfBB, Heft 2, 2009, Klostermann, Frankfurt am Main.