

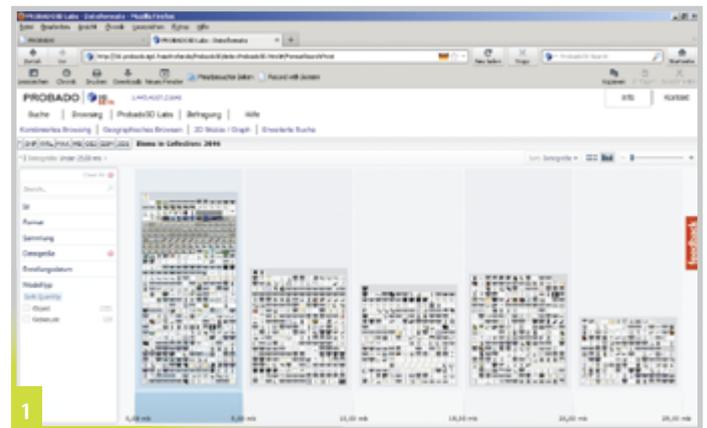
3D@Technische Informationsbibliothek (TIB)

DAS PROJEKT PROBADO ZEIGT EFFIZIENTE MÖGLICHKEITEN ZUR SUCHE IN 3D-DATEN

Immer mehr digitale Wissensobjekte finden Eingang in Bibliotheken. Innerhalb der Future Internet Initiative ist die TIB Mitglied der Arbeitsgruppe Internet of Contents, die Fragestellungen rund um die Erschließung und Nutzung von nichttextuellen, insbesondere audiovisuellen Daten in den Blick nimmt. Am Beispiel des Projekts PROBADO zeigen die Autorinnen, wie 3D-Modelle in die bibliothekarische Prozesskette integriert werden können.

Die Technische Informationsbibliothek (TIB) in Hannover hat als Deutsche Zentrale Fachbibliothek für Technik sowie Architektur, Chemie, Informatik, Mathematik und Physik die Aufgabe, für diesen Fächerkanon wissenschaftliche Informationen umfassend bereitzustellen. GetInfo, das Onlineportal der TIB für technisch-naturwissenschaftliche Fach- und Forschungsinformationen, bündelt den Zugang zu führenden Fachdatenbanken, Verlagsangeboten und Bibliothekskatalogen mit integrierter Volltextlieferung. Aktuell sind knapp 180 Millionen Metadatensätze durchsuchbar. GetInfo bietet Kunden aus Wissenschaft und Wirtschaft damit einen weltweit einzigartigen Zugang zu technisch-naturwissenschaftlichen Fachinformationen.

Aktuell beeinflussen folgende Rahmenbedingungen die Entwicklung digitaler Bibliotheken: Erste Herausforderung ist die Diversifizierung digitaler Wissensobjekte, so dass zu den klassischen Publikationen immer mehr digitale, audiovisuelle Medien und weitere nichttextuelle Materialien und Forschungsdaten hinzukommen. Zudem haben die Kunden wachsende Anforderungen an Informationsdienste, analog zu ihren allgemeinen Interneterfahrungen hinsichtlich schnellem, ortsunabhängigem und barrierefreiem Infor-



mationszugriff sowie ihrem Umgang mit Web 2.0-Diensten. Außerdem sind das Semantische Web und, als dritter maßgeblicher Faktor, der allgemeine Anstieg des zu verarbeitenden, verfügbaren und vernetzbar zu machenden Informationsvolumens¹ zu nennen. All dies führt zum Information Overload beim Benutzer, so dass er aus den zur Verfügung gestellten Informationen nicht mehr den optimalen Nutzen ziehen kann. Auch von politischer Seite werden Infrastrukturen zum Management von wissenschaftlich generierten und konsumierten Daten und Informationen gefordert.²

Aus diesen Anforderungen resultiert die TIB-Strategie, Verknüpfungen und Suchmöglichkeiten über alle Typen digitaler Wissensobjekte, also der im Kontext wissenschaftlichen Arbeitens produzierten

Medientypen und Formate, im Portal GetInfo bereitzustellen. Da es weder vorgesehen noch leistbar ist, alle Objekte bei der TIB physisch vorzuhalten, vernetzt GetInfo verteilte Angebote und Archive. Dauerhafte Verfügbarkeit wird durch persistente Identifier gewährleistet, dabei handelt es sich um dauerhafte Bezeichner, die zur Zitierung und Verlinkung von elektronischen Ressourcen verwendet werden. Die TIB vergibt DOI-Namen für Forschungsdatensätze und nichttextuelle Materialien unter dem Dach des Vereins DataCite.

Neben dem Auffinden von Daten mit zielgruppen- und objektorientierten innovativen und intuitiven Suchverfahren soll die gefundene Information nach relevanten Kernelementen analysiert und visualisiert werden können. Um insbeson-

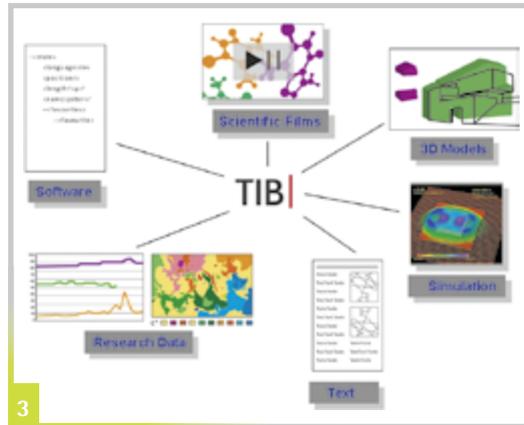
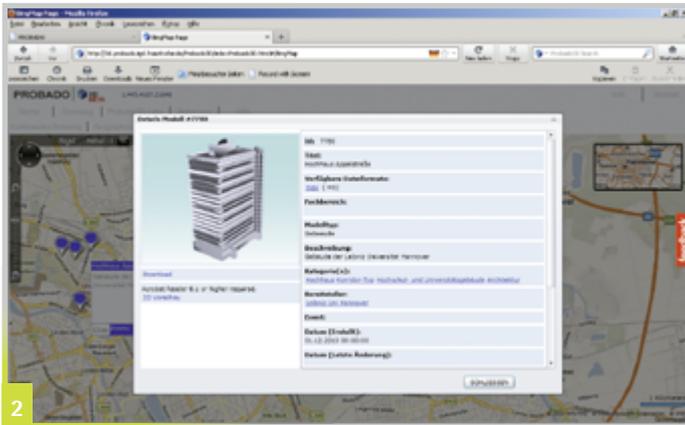
dere bei den nichttextuellen Materialien Nachweis und Zugang grundlegend zu verbessern, hat die TIB ihre Expertise in diesem Bereich gebündelt und erweitert ihr Service- und Forschungsspektrum, indem sie ein Kompetenzzentrum für nichttextuelle Materialien (KNM) aufbaut.³ In diesem Kontext ist eine verstärkte Zusammenarbeit unter anderem mit dem L3S zum Betreiben angewandter Forschung im Bereich Multimedia

PROBADO Projekts für den Bereich 3D ist die Integration von 3D Modellen in die bibliothekarische Prozesskette, bestehend aus Akquise, Erschließung, Nachweis, Bereitstellung und Auslieferung. Neben dem Anwendungspartner TIB waren drei Informatik-Lehrstühle der Universitäten Bonn, Graz und Darmstadt sowie das Fraunhofer IGD Projektpartner, wobei die Kompetenzen im Bereich 3D auch nach Abschluss des Projekts als ver-

von bestimmten Merkmalen Ausgangsvoraussetzung. Ein erster Vorverarbeitungsschritt dient dazu, Invarianz- und Robustheitseigenschaften zu erzeugen, zum Beispiel Skalierung auf einen bestimmten Maßstab oder einheitliche Orientierung. Um die Vielzahl von CAD-Formaten zu unterstützen, müssen diese Daten vollautomatisch in ein einheitliches Format konvertiert werden. Dieses dient als Eingabe für die verschiedenen Indexer

Abbildung 1
Browsing-Ansicht

Abbildung 2
Ergebnisansicht eines Suchtreffers in PROBADO 3D



Retrieval geplant. Innerhalb der Future Internet Initiative ist die TIB Mitglied der Arbeitsgruppe Internet of Contents, die Fragestellungen rund um die Erschließung und Nutzung von nichttextuellen, insbesondere audiovisuellen Daten in den Blick nimmt.

Verschiedene Kooperationsprojekte mit Beteiligung der TIB bewegen sich in diesem Kontext. Neben den aktuell laufenden Projekten »Entwicklung eines AV-Portals« mit dem Hasso-Plattner-Institut in Potsdam und »Visueller Zugang zu Forschungsdaten« mit dem Fraunhofer IGD und der TU Darmstadt ist insbesondere das kürzlich beendete Projekt PROBADO zu nennen. PROBADO wurde fünf Jahre lang von der Deutschen Forschungsgemeinschaft als Leistungszentrum für Forschungsinformation gefördert. Ziel des

traglich geregelte Kooperation weiter gebündelt werden.

Die stetig zunehmende Bedeutung von mit Computer Aided Design (CAD) erstellten 3D-Modellen, beispielsweise aus den Fachbereichen Maschinenbau, Bauingenieurwesen und Architektur, steigerte im Laufe der letzten Jahre die Nachfrage nach effizienten Möglichkeiten zur Suche in 3D-Daten. Mit heterogenen Planungsaufgaben steigen Anzahl und Komplexität der Modelle, gleichzeitig benötigen Benutzer geeignete Werkzeuge zur Suche in 3D-Modellbeständen, um sie für eigene CAD-Entwürfe heranzuziehen.

Um fachlich individuelle Suchstrategien zu berücksichtigen und eine inhaltsbasierte Suche und Visualisierung anzubieten, ist die Analyse der Daten auf das Vorkommen

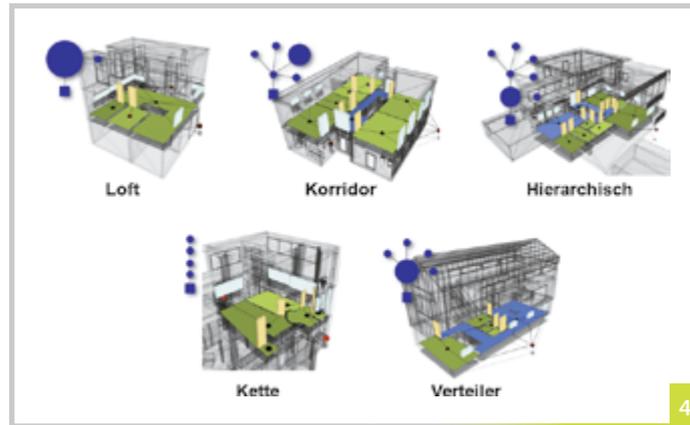
sowie für die Generierung der Vorschauformate. Der folgende Schritt erzeugt aus den Rohdaten so genannte Deskriptoren oder Indexdaten, zum Beispiel globale oder lokale geometrische Beschreibungen, die als Vektoren repräsentiert sind. Erst die Indexdaten werden durch Algorithmen miteinander auf Ähnlichkeit verglichen. Metadaten können durch high-level Indexdaten wie Information über die Geschossigkeit oder Quadratmeter von Gebäudemodellen bzw. generell mit maschinellem Lernen erzeugt werden, wozu auf Publikationen unter www.probado.de verwiesen wird.

Ausgangspunkt für die Implementierung verschiedener Suchmöglichkeiten war eine Anforderungsanalyse, die zu Beginn des Projekts in Form einer Befragung von Architek-

Abbildung 3
TIB Strategie – Verknüpfung und Suche über alle Typen digitaler Dokumente im TIB-Portal GetInfo

- 1 »The Diverse and Exploding Digital Universe«, IDC-Whitepaper, März 2008, siehe <http://www.emc.com/collateral/analyst-reports/diverse-exploding-digital-universe.pdf>.
- 2 Riding the wave. How Europe can gain access from the rising tide of scientific data. Final report of the High Level Expert Group on Scientific Data. A submission to the European Commission. October 2010, siehe <http://cordis.europa.eu/fp7/ict/e-infrastructure/docs/hlg-sdi-report.pdf>.
- 3 Ein Konzept der TIB zum Kompetenzzentrum für multimediale Objekte wurde bereits im Dezember 2009 auf Antrag der Gemeinsamen Wissenschaftskonferenz (GWK) und in Verantwortung der Leibniz-Gemeinschaft sonderevaluiert und positiv bewertet, siehe <http://www.wgl.de/?nid=ssn&nidap=&print=0>.

Abbildung 4
Basistypen globaler Gebäudestrukturen



dungen zwischen Räumen in Form von Türen, Fenstern und Treppen. Dabei wurde zunächst untersucht, wie bestimmte Raumkonfigurationen innerhalb einer Datenbank von durch RVGs repräsentierten Gebäuden gefunden werden können. Darüber hinaus wurde ein Verfahren entwickelt, das eine schnelle und effektive Ähnlichkeitsbestimmung von zwei RVGs ermöglicht und somit die Grundlage für Browsing und Klassifika-

Abbildung 5
Spezifikation eines Raumverbindungsgraphen zur Suche

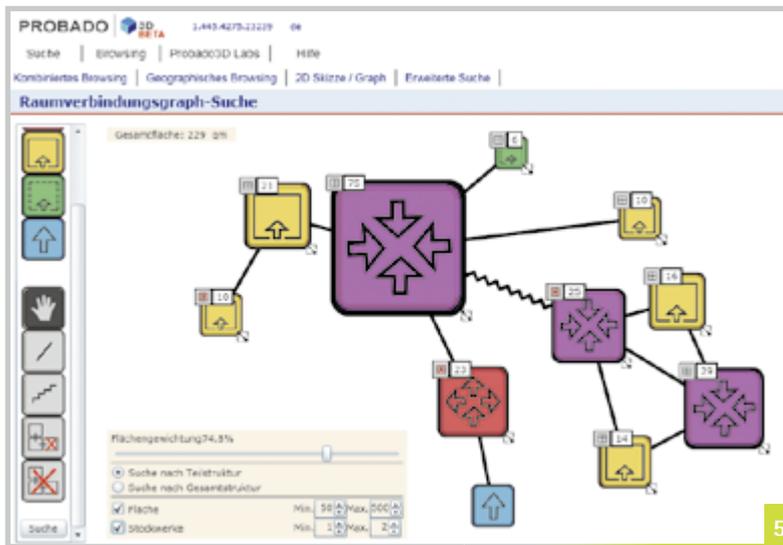


Abbildung 6
Gebäudemodell mit extrahiertem Raumverbindungsgraph des ersten Stockwerks

ten hinsichtlich ihrer bevorzugten Suchmöglichkeiten durchgeführt wurde. Hierbei wurden drei wichtige Modalitäten identifiziert, nämlich die Suche anhand von Metadaten (feldbasiert oder im Volltext), die formbasierte Suche nach Inneneinrichtungs- oder Umgebungsgegenständen, sowie die topologische Suche nach Gebäudemodellen anhand der enthaltenen Raum- und Stockwerksstrukturen.

Ausgehend von Vorarbeiten im Bereich der formbasierten Suche nach 3D-Modellen wurden Methoden erarbeitet, die bestehende Suchverfahren an die besonderen Erfordernisse des Bereichs der Architektur anpassen. Hierzu wurde zunächst ein Verfahren entwickelt, das basierend auf einer

Objektcharakterisierung anhand lokaler Formdeskriptoren unter Zuhilfenahme von Algorithmen aus dem Bereich des maschinellen Lernens verbesserte Suchergebnisse gegenüber klassischen Methoden aufweist. Darüber hinaus wurde untersucht, wie sich der neu entwickelte Ansatz hinsichtlich der in der Architektur eine zentrale Rolle spielenden form- oder funktionsbasierten Klassifikationsschemata von Inneneinrichtungsgegenständen verhält.

Zur Repräsentation der Raum- und Stockwerksstrukturen von Gebäuden wurde das Konzept der Raumverbindungsgraphen (RVGs) eingeführt. Deren Knoten charakterisieren attributierte Räume, die Kanten stehen für Verbindungen

tion darstellt, um insbesondere bei 3D-Modellen mit wenigen Metadaten gezielt nach Inhalten zu recherchieren.

Für PROBADO 3D wurde die Benutzeroberfläche für Anfragen und Ergebnisvisualisierung mittels der Rich Internet Application Technologie Silverlight realisiert, da hiermit vielfältige Interaktionsmöglichkeiten insbesondere für die inhaltsbasierten Anfragen realisiert werden konnten. PROBADO 3D unterstützt neben der klassischen Stichwortsuche unterschiedlichste Anfragemöglichkeiten, zum Beispiel ein webbasiertes 3D Modellierungswerkzeug, das es dem Benutzer ermöglicht, unmittelbar die Auswirkung seiner Modellieraktionen auf die Ergebnismenge zu sehen.

Der Benutzer kann außerdem eine RVG Suche durch das Skizzieren eines Anfragegraphen starten. Gebäudemodelle, die ein reales Gebäude repräsentieren, können über eine geographische Suche basierend auf Bing Maps entsprechend visualisiert werden. PROBADO 3D unterstützt das Browsing durch den Modellbestand auf Basis der Metadaten für Sammlung, Bereitsteller, Dateiformat und Kategorien. Das Ergebnis selbst kann auch als Eingabe für eine inhaltsbasierte Anfrage (query-by-example) verwendet werden. Als Plugin-Unterstützung für Produkte von Drittanbietern wurde eine Unterstützung für Google Sketchup entwickelt, mit der das erstellte Modell direkt aus Sketchup heraus als Anfrageobjekt für eine Objektsuche benutzt werden kann.

Abhängig von der Anfrageart bietet PROBADO 3D unterschiedliche Möglichkeiten der Ergebnispräsentation: Für die inhaltsbasierten Suchen können die Elemente der Ergebnismenge im 2D Raum angeordnet werden, sodass ähnliche Ergebnisse als Cluster dargestellt werden. Neben der »klassischen« Ergebnispräsentation wurde auch eine 3D Vorschau auf Basis von PDF mit eingebettetem 3D Inhalt integriert. Neben der weiten Verbreitung des Adobe Readers spricht besonders die Eignung für Langzeitarchivierung (ISO-Standard PDF/A) für PDF als Vorschauformat. Im Falle einer Anfrage mittels eines RVG wird bei den Treffern zusätzlich der Anfragegraph innerhalb des 2D Grundrisses des Gebäudes visualisiert.

Nach Abschluss einer internen Testphase wurde die erste Version von PROBADO 3D Mitte des Jahres 2010 der Öffentlichkeit im Internet zugänglich gemacht. Basierend auf Nutzerrückmeldungen während des Betriebsjahres wurde das

Web-Frontend an zahlreichen Stellen hinsichtlich seiner Nutzerfreundlichkeit überarbeitet. Um eine weitere Verbreitung der Anwendung zu erreichen, wurde der einfache Wechsel zwischen einer deutsch- und englischsprachigen Version der Webseite implementiert.

Der in PROBADO 3D online verfügbare Datenbestand umfasst aktuell rund 8000 Modelle. Etwa ein Viertel der Modelle wird auf PROBADO-internen Servern vorgehalten. In einer Einverständniserklärung werden die Bereitstellung und Art der Nutzung (zum Beispiel mit Creative-Commons-Lizenz) festgelegt. Die verbleibenden drei Viertel der Modelle sind auf externen Servern verfügbar. Der Benutzer erhält in PROBADO 3D Metadaten, Vorschaudateien und den jeweiligen Link zur Modellvorschau auf dem externen Server. Der Zugriff zum Original wird vom jeweiligen Bereitsteller individuell (Authentifizierung, Bezahlung, etc.) geregelt. Die TIB hat PROBADO 3D in ihr Dienstleistungsangebot für den Baubereich aufgenommen und in GetInfo als externe Quelle eingebunden. Zur parallelen Recherche nach 3D-Modellen und anderen Medientypen werden aktuell 3D-Modell-Meta- und Vorschau-dateien in den GetInfo-Index integriert. Außerdem werden ausgewählte Datensätze in die Europeana integriert, eine digitale Bibliothek, die das kulturelle Erbe Europas in digitaler Form via Internet zugänglich machen soll.

Bei der Weiterentwicklung von PROBADO 3D sind zum einen kontextbasierte Fragestellungen interessant, also die Entwicklung von Methoden, die Architekten beim Planungs- und Konstruktionsprozess unterstützen, etwa in Form von Vorschlägen, welcher Gebäudetyp hinsichtlich Form und Aufteilung geeignet ist, in



Ina Blümel

Jahrgang 1975, wissenschaftliche Mitarbeiterin an der Technischen Informationsbibliothek, ist seit 2011 stellvertretende Leiterin des Kompetenzzentrums Nichttextuelle Materialien. Kontakt: ina.bluemel@tib.uni-hannover.de



Dr. Irina Sens

Jahrgang 1965, ist seit 1999 stellvertretende Direktorin der Technischen Informationsbibliothek. Kontakt: irina.sens@tib.uni-hannover.de

eine bestehende Umgebungsbebauung ästhetisch und konzeptionell integriert zu werden. Hierzu könnte insbesondere für parametrische Modelle der im Building Information Modeling (BIM) zukünftig unterstützte Standard der Industry Foundation Classes (IFC) mit seinen oftmals vorhandenen Blockhierarchien als einheitliche Modellrepräsentation interessant sein. Außerdem ist eine über die Migration in neue Browserversionen hinausreichende Weiterentwicklung des Frontends hin zu einer PlugIn-freien Version auf Basis von WebGL und Html5 langfristig sinnvoll.

PROBADO wurde auf nationalen beziehungsweise internationalen Konferenzen und Veranstaltungen präsentiert sowie in Veröffentlichungen behandelt (siehe Auflistung unter www.probado.de). Bei der CeBIT 2011 in Hannover wurde PROBADO 3D auf dem Gemeinschaftsstand des Landes Niedersachsen einer breiten Öffentlichkeit vorgestellt.