

Margret Plank, Uwe Rosemann

# Das Kompetenzzentrum für nicht-textuelle Materialien an der Technischen Informationsbibliothek

## 1 Einleitung

Audiovisuelle Medien, 3D-Modelle, 3D-Druckdaten, Grafiken, Forschungsdaten, Computersimulationen, Open Educational Resources (OER), Executable Papers<sup>1</sup>, Software-Stacks<sup>2</sup>, zusammengefasst unter der Bezeichnung nicht-textuelle Materialien, nehmen in Bedeutung und Nutzung für Wissenschaft, Forschung und Lehre kontinuierlich zu.

In starkem Kontrast zu der Bedeutung, dem Mengenwachstum und der Nutzung nicht-textueller Materialien steht deren mangelnde Recherchier- und Nutzbarkeit. Mit der technischen Erleichterung der Herstellung von nicht-textuellen Materialien und dem massenhaften Transfer von analogen Daten in digitale Formate steigt nicht nur die Menge dieser Materialien stark an, es wächst auch deren Heterogenität in Bezug auf Medientyp und -format. Zudem gibt es eine veränderte Nutzererwartung an die Such- und Präsentationsverfahren, die sich derzeit immer noch in erster Linie an textuellen Medien orientieren.

Was bedeutet dies für Bibliotheken? Es müssen geeignete Methoden, Verfahren, Werkzeuge und Infrastrukturen für einen professionellen Umgang mit nicht-textuellen Materialien entwickelt werden, damit der wachsende Bestand an derartigen Materialien so einfach publiziert, auffindbar und nutzbar gemacht werden kann, wie es heute bereits für textuelle Medien möglich ist.

Laut den Verfassern des Reports für die EU-Kommission „Riding the wave – How Europe can gain from the rising tide of scientific data“<sup>3</sup> wird es noch mindestens bis zum Jahr 2030 dauern, bis die für digitale nicht-textuelle Materialien erforderliche Infrastruktur umgesetzt wurde. Unter Beteiligung der führenden Infrastruktureinrichtungen gibt es in Deutschland bereits Ansätze zur Bewältigung der damit verbundenen Herausforderungen, wie z. B. der KII Report,<sup>4</sup> die bei der Gemeinsamen Wis-

---

1 Reproduzierbare wissenschaftliche Arbeiten, die mit den dazugehörigen Forschungsdaten- und Programmcode veröffentlicht werden.

2 Komplexe software-basierte Arbeits- und Forschungsumgebungen.

3 European Union: Riding the Wave. How Europe Can Gain from Rising Tide of Scientific Data. Final Report of the High Level Expert Group on Scientific Data. A Submission of the European Commission, October 2010. [cordis.europa.eu/fp7/ict/e-infrastructure/docs/hlg-sdi-report.pdf](http://cordis.europa.eu/fp7/ict/e-infrastructure/docs/hlg-sdi-report.pdf) (01.07.2014). Hier S. 4.

4 Kommission Zukunft der Informationsinfrastruktur (KII): Gesamtkonzept für die Informationsinfrastruktur in Deutschland. Empfehlungen der Kommission Zukunft der Informationsinfrastruktur im Auftrag der Gemeinsamen Wissenschaftskonferenz des Bundes und der Länder, April 2011.



schaftskonferenz der Länder (GWK) und im Wissenschaftsrat diskutiert werden. Hier werden klare Handlungsoptionen für Politik und Wissenschaftsorganisationen aufgezeigt, die aber noch zu keinen konkreten wissenschaftspolitischen Maßnahmen geführt haben. Aktuell wird von der GWK die Einsetzung eines Rats für Informationsinfrastruktur geplant, der zukünftige Strategien und Empfehlungen für Entwicklungsmaßnahmen formulieren soll.

Die Technische Informationsbibliothek (TIB) hat unterdessen mit der Gründung des Kompetenzzentrums für nicht-textuelle Materialien Fakten geschaffen. Im Kompetenzzentrum werden Werkzeuge und Portale entwickelt, die die Zugangs- und Nutzungsbedingungen für Audiovisuelle Medien, 3D-Modelle, Forschungsdaten etc. grundlegend verbessern und neue Formen der Nutzung vorhandener Bestände ermöglichen. Das Kompetenzzentrum unterstützt als kompetenter Ansprechpartner weitere Wissenseinrichtungen und Wissensanbieter in allen Fragen im Umgang mit nicht-textuellen Materialien und realisiert bedarfsorientiert Dienstleistungen für ihre Kunden.

Der Senat der Leibniz-Gemeinschaft hat in seiner Sitzung am 27.11.2008 einer Sonderevaluierung der geplanten Erweiterung der Technischen Informationsbibliothek (TIB) um ein Kompetenzzentrum für nicht-textuelle Materialien (KNM) zugestimmt, welche im Dezember 2009 durchgeführt wurde. Der Bewertungsbericht des Senatsausschusses Evaluierung bewertete das Vorhaben der Technischen Informationsbibliothek (TIB) im April 2010 positiv. Es sei, so der Senatsausschuss, „[...] von grundsätzlichem Interesse, zukunftsweisend und Erfolg versprechend. Die TIB Hannover bietet sowohl in technischer als auch in personeller, organisatorischer und infrastruktureller Hinsicht sehr gute Voraussetzungen für eine nachhaltige Entwicklung des geplanten Kompetenzzentrums.“<sup>5</sup> Der Senatsausschuss folgt in seiner Stellungnahme im Wesentlichen dem Bewertungsbericht (vgl. Stellungnahme des Senats der Leibniz-Gemeinschaft vom 7. Juli 2010).<sup>6</sup>

Auf Grund der positiven Stellungnahme des Senats der Leibniz-Gemeinschaft zum Kompetenzzentrum für nicht-textuelle Materialien (KNM) im Rahmen der turnusmäßigen Evaluierung der Technischen Informationsbibliothek (TIB) am 24. November 2011 kam der Senat der Leibniz-Gemeinschaft am 24. November 2011 zu der folgenden Beurteilung: „Die Empfehlungen der Sonderevaluierung eines ersten Konzeptes durch den Senat der Leibniz-Gemeinschaft konnten überzeugend umgesetzt und dadurch das nun vorgelegte Konzept zum KMO deutlich weiterentwickelt werden. Die Erweiterung der TIB um ein Kompetenzzentrum für multimediale Objekte wird als

---

[www.leibniz-gemeinschaft.de/fileadmin/user\\_upload/downloads/Infrastruktur/KII\\_Gesamtkonzept.pdf](http://www.leibniz-gemeinschaft.de/fileadmin/user_upload/downloads/Infrastruktur/KII_Gesamtkonzept.pdf) (01.07.2014).

<sup>5</sup> Leibniz Gemeinschaft. Der Senat: Stellungnahme zum Konzept für die Erweiterung der Technischen Informationsbibliothek (TIB) Hannover um ein Kompetenzzentrum nicht-textuelle Materialien, 07. Juli 2010. [www.leibniz-gemeinschaft.de/fileadmin/user\\_upload/downloads/Evaluierung/Senatsstellungen/Senatsstellungnahme-TIB-KNM-2010.pdf](http://www.leibniz-gemeinschaft.de/fileadmin/user_upload/downloads/Evaluierung/Senatsstellungen/Senatsstellungnahme-TIB-KNM-2010.pdf) (01.07.2014). Hier B-2.

<sup>6</sup> Vgl. Leibniz, Stellungnahme (wie Anm. 5).

sinnvoll und erfolgversprechend bewertet.<sup>7</sup> Der Senat beschloss, dass zusätzliche Personal- und Sachmittel zum Auf- und Ausbau des Kompetenzzentrums zur Verfügung gestellt werden sollen.

Der Gründung des Kompetenzzentrums an der TIB ging die Schließung des Instituts für Wissen und Medien (IWF) in Göttingen voraus. In der Stellungnahme des Senats der Leibniz-Gemeinschaft vom 7. Juli 2010 wird empfohlen, den Filmbestand der IWF Wissen und Medien gGmbH i. L. von rund 11.500 veröffentlichten Werken in die Technische Informationsbibliothek (TIB) zu integrieren. Laut Stellungnahme ist die Technische Informationsbibliothek (TIB) die adäquate Einrichtung, den Filmbestand der IWF Wissen und Medien gGmbH i. L. zu erhalten, zu integrieren und besser zugänglich zu machen (Stellungnahme des Senats der Leibniz-Gemeinschaft vom 7. Juli 2010).<sup>8</sup>

## 2 Vorarbeiten

Die TIB hat die Bedeutung nicht-textueller Materialien für Wissenschaft und Lehre frühzeitig erkannt und sich bereits vor der Gründung des Kompetenzzentrums mit Aktivitäten und in Projekten engagiert.

Ziel des durch die Deutsche Forschungsgemeinschaft geförderten Projektes *CODATA*<sup>9</sup> (Laufzeit 01.08.2006–30.04.2009) war es, die Publikation und Registrierung wissenschaftlicher Daten als eigenständigen, fächerübergreifenden Wissenschaftsservice bei der TIB zu etablieren. Als Pilotanwendung diente dabei die Geowissenschaften/Erdsystemforschung.

Das durch die Deutsche Forschungsgemeinschaft geförderte Projekt *Virtuelle Fachbibliothek Chemie II*<sup>10</sup> (Laufzeit 01.01.2008–31.03.2010) befasste sich mit der Entwicklung von Konzepten und Werkzeugen zur aktiven Unterstützung des Nutzers bei der Erfassung, Erschließung und Verknüpfung von Datenbeständen mittels Information Extraction Methoden (u. a. chemische OCR) nach individuellen Kriterien, die in der Schaffung von personalisierten Wissensräumen für die Chemie münden sollten.

Das 2011 beendete Projekt *PROBADO*<sup>11</sup> wurde 5 Jahre lang von der Deutschen Forschungsgemeinschaft als Leistungszentrum für Forschungsinformation gefördert.

<sup>7</sup> Leibniz Gemeinschaft. Der Senat: Stellungnahme zur Technischen Informationsbibliothek Hannover (TIB), 24. November 2011. [www.leibniz-gemeinschaft.de/fileadmin/user\\_upload/downloads/Evaluierung/Senatsstimmungen/Senatsstimmnahme-TIB-2011.pdf](http://www.leibniz-gemeinschaft.de/fileadmin/user_upload/downloads/Evaluierung/Senatsstimmungen/Senatsstimmnahme-TIB-2011.pdf) (01.07.2014). Hier S. 3.

<sup>8</sup> Vgl. Leibniz, Stellungnahme (wie Anm. 5).

<sup>9</sup> Technische Informationsbibliothek (TIB): *CODATA*. [www.tib-hannover.de/de/forschung-und-entwicklung/abgeschlossene-projekte/codata/](http://www.tib-hannover.de/de/forschung-und-entwicklung/abgeschlossene-projekte/codata/) (01.07.2014).

<sup>10</sup> TIB: *Virtuelle Fachbibliothek Chemie II*. [www.tib-hannover.de/de/forschung-und-entwicklung/abgeschlossene-projekte/virtuelle-fachbibliothek-chemie-ii/](http://www.tib-hannover.de/de/forschung-und-entwicklung/abgeschlossene-projekte/virtuelle-fachbibliothek-chemie-ii/) (01.07.2014).

<sup>11</sup> *PROBADO*. [www.probado.de/](http://www.probado.de/) (01.07.2014).

Neben dem Bibliothekspartner Technische Informationsbibliothek (TIB) waren für den Anwendungsbereich 3D-Modelle drei Informatik-Lehrstühle der Universitäten Bonn, Graz und Darmstadt sowie das Fraunhofer-Institut für Graphische Datenverarbeitung IGD Projektpartner. Ziel des PROBADO-Projekts war die Integration von nicht-textuellen Materialien in die bibliothekarische Prozesskette, bestehend aus Akquise, Erschließung, Nachweis, Bereitstellung und Auslieferung. PROBADO ist ein substantieller Bestandteil der Strategie der Technischen Informationsbibliothek (TIB), Verknüpfungen und Suchmöglichkeiten über alle Typen digitaler Dokumente im Portal GetInfo bereitzustellen.

Im Mai 2005 hat die TIB die weltweit erste *DOI-Registrierungsagentur* für wissenschaftliche Daten gegründet. Der DOI-Service der TIB ist für ihre Sammelgebiete in Deutschland erster Ansprechpartner für die Registrierung von Forschungsdaten und weiteren nicht-textuellen Materialien. Der DOI Service steht wissenschaftlichen Einrichtungen (wie Forschungsinstituten, Bibliotheken, Universitätsfakultäten etc.) zur Verfügung, die eine eigene Infrastruktur zur Speicherung ihrer Objekte betreiben.<sup>12</sup> Unter maßgeblicher Beteiligung der TIB wurde 2009 das internationale Konsortium *DataCite* gegründet, mit dem Ziel, einen weltweit einheitlichen Zugang zu Forschungsdaten zu ermöglichen.<sup>13</sup>

### 3 Ziele

Im Kompetenzzentrum für nicht-textuelle Materialien werden die erforderlichen Konzepte, Werkzeuge und Strukturen zur Sammlung, Erschließung und Bereitstellung sowie die Standardisierung und Archivierung nicht-textueller Materialien entwickelt. Diese sind im Einzelnen:

- systematische Sammlung nicht-textueller Materialien aus dem Fächerspektrum der TIB
- Entwicklung einer technischen Infrastruktur zur Verwaltung von nicht-textuellen Materialien
- Entwicklung medienpezifischer Portale und Werkzeuge unter Einsatz von Technologien aus der automatischen Medienanalyse sowie der visuellen und semantischen Suche
- Verknüpfung von nicht-textuellen Materialien mit weiteren Forschungsinformationen wie z. B. Volltexten und Forschungsdaten über das Fachportal GetInfo/Ein|Portal
- Entwicklung von bedarfsorientierten Dienstleistungen
- Vergabe von Digital Object Identifier (DOI) als persistente Identifikatoren

<sup>12</sup> TIB: DOI-Service. [www.tib-hannover.de/de/dienstleistungen/doi-service/](http://www.tib-hannover.de/de/dienstleistungen/doi-service/) (01.07.2014).

<sup>13</sup> DataCite. [www.datacite.org/](http://www.datacite.org/) (01.07.2014).

- Langzeitarchivierung der gesammelten und nachgewiesenen nicht-textuellen Materialien
- Aufbau von Netzwerken, Gremienarbeit, Community-Building auf nationaler und internationaler Ebene
- Unterstützung von Medienproduzenten bei der Publikation und Nutzbarmachung nicht-textueller Materialien
- Forschungs- und Entwicklungsaufgaben, die sich aus der Aufgabenstellung des visuellen Erschließens und des visuellen Retrievals ableiten.

## 4 State of Play

Die TIB sammelt auf der Basis von Marktanalysen nicht-textuelle Materialien, wobei der Fokus derzeit auf AV-Medien, 3D-Modellen und Forschungsdaten liegt. So hat die TIB im Herbst 2010 in Zusammenarbeit mit der Firma Infratest eine detaillierte Nutzer- und Marktanalyse durchgeführt, die qualitative und quantitative Informationen darüber enthält, wer in Deutschland audiovisuelle Medien in welchem Umfang und in welcher Art produziert, bereitstellt bzw. benötigt. Befragt wurden 100 Medienzentren und 326 Produzenten. Als Ergebnis der Befragung liegen der TIB je ein Handbuch Produzenten und Medienzentren wissenschaftlicher Filme aus Naturwissenschaft und Technik vor, die Aufschluss über Bestände, Erschließung und Verfügbarkeit der Medien geben.

Auf dieser Basis werden digitale AV-Medien wie zum Beispiel Computervisualisierungen, Lernmaterialien, Simulationen, Experimente, Interviews, Vorlesungs- oder Konferenzaufzeichnungen gesammelt. Komplettiert wird der Bestand der AV-Medien durch den Ende 2012 auf die TIB übertragenen IWF-Filmbestand samt Rechten. Dieser umfasst rund 11.500 veröffentlichte Werke. Es handelt sich um einen historisch gewachsenen Filmbestand mit Produktionsjahren, die teilweise in den 1910er Jahren liegen. Bei dem Bestand handelt es sich in der Regel um Unikate und Originale, die auf unterschiedlichsten Trägerformaten analog und digital vorliegen. Das Angebot lässt sich in drei Filmarten unterteilen: Forschungsfilm, Hochschulunterrichtsfilm und Dokumentationsfilm. Rund 60 % der Sammlung sind digitalisiert worden. Nach Fächern differenziert, setzt sich der IWF Filmbestand wie folgt zusammen: Biologie (27 %), Ethnologie (19%), Medizin (11 %), Technik (10 %), Geschichte (5 %), Geowissenschaften (3 %), Physik (7 %), Psychologie (4 %), Chemie (3 %), sonstige Fächer bzw. fachlich nicht zuordenbar (11 %).

Darüber hinaus werden digitale 3D-Modelle mit Fokus auf den AEC-Sektor (*Architecture, Engineering, and Construction*) sowie der akademischen Gemeinschaft gesammelt. Dies betrifft in erster Linie Bauteile und Einrichtungsgegenstände sowie Gebäudemodelle, beispielsweise aus den Phasen Entwurfsplanung, Visualisierung oder Konstruktion.

Analog zur Akquisestrategie der Technischen Informationsbibliothek (TIB) müssen diese Materialien die allgemein von der TIB an wissenschaftliche Informationen gestellten Qualitätsanforderungen erfüllen. Die nicht-textuellen Materialien werden

- nach internationalen Standards erschlossen
- möglichst unter Creative Commons Lizenzen bereitgestellt
- ggf. langzeitarchiviert
- mit einem eindeutigen Zitierlink versehen
- über medienspezifische Online-Portale verfügbar und mit innovativen Suchmöglichkeiten (z. B. visuelle und semantische Suche) auffindbar gemacht
- mit weiteren Forschungsinformationen (z. B. Volltexten und Forschungsdaten) verknüpft.

Darüber hinaus werden die Metadaten unter einer CC0-Lizenz bereitgestellt

Für die Verwaltung und Archivierung von AV-Medien ist in 2011 ein Media-Asset-Managementsystem (MAM) in die Systemumgebung der Technischen Informationsbibliothek (TIB) integriert worden. Das Media-Asset-Managementsystem (MAM) erlaubt eine professionelle Erfassung von nicht-textuellen Materialien, die Distribution in unterschiedliche Kanäle und Portale, verfügt über eigene Transcoder, die alle gängigen Codecs beherrschen, und erstellt Statistiken. Darüber hinaus sind Workflows und Schnittstellen für den Im- und Export von AV-Medien vom Media-Asset-Managementsystem (MAM) zu den TIB Portalen entwickelt worden. Das Metadatenchema zur standardisierten Erfassung nicht-textueller Materialien basiert auf dem aktuellen DataCite-Schema und wurde um einige Elemente, die zu einer detaillierten Beschreibung beispielsweise eines AV-Mediums benötigt wurden, ergänzt. Das Metadatenchema steht den Mediengebern online zur Verfügung.<sup>14</sup>

Damit die Technische Informationsbibliothek (TIB) ihren Nutzern nicht-textuelle Materialien über ihre Portale anbieten kann, schließen die Mediengeber eine Lizenzvereinbarung mit der TIB ab und legen die Nutzungsbedingungen fest. Zur Auswahl stehen einfache Lizenzvereinbarungen sowie unterschiedliche Creative-Commons-Lizenzen. Die Technische Informationsbibliothek (TIB) empfiehlt ausdrücklich die Open-Access-Lizenz „CC-Namensnennung – Deutschland 3.0“: Diese sieht am wenigsten Restriktionen für die Nutzung in Forschung und Lehre vor und gewährleistet gleichzeitig, dass der Urheber genannt werden muss. Standard-Lizenzvereinbarungen sind von der Technischen Informationsbibliothek (TIB) entwickelt worden und stehen online für die Mediengeber bereit.<sup>15</sup>

Bei hoher Relevanz für Wissenschaft und Lehre und geeigneter technischer Qualität werden die nicht-textuellen Materialien langzeitarchiviert. Zur Gewährleistung

<sup>14</sup> TIB: Inhalte bereitstellen. [www.tib-hannover.de/de/dienstleistungen/kompetenzzentrum-fuer-nicht-textuelle-materialien-knm/inhalte-bereitstellen/](http://www.tib-hannover.de/de/dienstleistungen/kompetenzzentrum-fuer-nicht-textuelle-materialien-knm/inhalte-bereitstellen/) (01.07.2014).

<sup>15</sup> TIB, Inhalte (wie Anm. 14).

der nachhaltigen Verfügbarkeit und Benutzbarkeit von nicht-textuellen Materialien ist die Implementierung eines Langzeitarchivierungssystems eine essenzielle Voraussetzung. Die Technische Informationsbibliothek (TIB) betreibt ein professionelles Archivierungssystem, welches innerhalb des Leibniz-Bibliotheksverbundes Forschungsinformation (Goportis) auch von ihren Partner-Bibliotheken Deutsche Zentralbibliothek für Medizin (ZB MED) und Leibniz-Informationszentrum Wirtschaft (ZBW) genutzt wird. Die Basis des Produktivsystems zur Langzeitarchivierung der Technischen Informationsbibliothek (TIB) basiert auf dem Produkt „Rosetta“ der Firma Ex Libris.<sup>16</sup>

Soweit sie den Nutzern zur Ansicht oder zum Download zur Verfügung gestellt werden dürfen, erhalten die nicht-textuellen Materialien einen eindeutigen Zitierlink (DOI-Namen).<sup>17</sup> Die DOI-Registrierung erfolgt über die API-Schnittstelle von DataCite. Zusätzlich zur DOI-Registrierung der Filme bietet das AV-Portal einen zeitbasierten Zitierlink an. Mit Hilfe des offenen Standards Media Fragment Identifier (MFID) wird für jedes Filmsegment ein zitierbarer DOI angezeigt.

Die TIB bietet ihren Nutzern seit Ende 2011 einen medienübergreifenden Zugang über das zentrale TIB-Fachportal GetInfo an. Dort sind die Metadaten von AV-Medien, 3D-Modellen und Forschungsdaten aller kooperierenden Wissensanbieter durchsuchbar, die Medien selbst können entweder in den objektspezifischen Repositorien der Technischen Informationsbibliothek (TIB) gehostet werden oder in den verteilten Archiven verbleiben. Die nicht-textuellen Materialien werden über GetInfo recherchierbar gemacht, mit weiteren Forschungsinformationen verknüpft und den Nutzern (soweit zu diesem Zweck lizenziert) zum Download zur Verfügung gestellt. Einzelne Sequenzen, Trailer oder Thumbnails werden als Vorschau eingestellt. Die Medien können zudem über mobile Endgeräte abgerufen werden.

Durch die Beteiligung von Kooperationspartnern sollen Metadaten unter CC0-Lizenz einer sinnvollen Bereitstellung in den jeweils relevanten Communities zugeführt werden. Die TIB stellt hierfür von ihr lizenzierte Metadaten und Vorschaudateien Partnern, wie beispielsweise der Europeana, der Deutschen Digitalen Bibliothek und dem Deutschen Filminstitut, sowie vielen weiteren Einrichtungen zur Verfügung. Ein weiterer Ausbau der Kooperationen, beispielsweise mit Fachinformationsdiensten, wird vorangetrieben.

Im Kompetenzzentrum werden zudem medienspezifische Portale und Werkzeuge entwickelt, die in die TIB Systemarchitektur eingepasst werden. Diese medienspezifischen Portale zeichnen sich durch Such- und Präsentationsverfahren aus, die auf einer automatischen Medienanalyse und Metadatengenerierung basieren. Bereits existierende Ansätze aus der Forschung werden dabei in die Praxis übertragen. So

<sup>16</sup> GOPORTIS: Digitale Langzeitarchivierung. [www.goportis.de/kompetenzen/versorgung-mit-wissenschaftlichem-content/digitale-langzeitarchivierung.html](http://www.goportis.de/kompetenzen/versorgung-mit-wissenschaftlichem-content/digitale-langzeitarchivierung.html) (01.07.2014).

<sup>17</sup> TIB: DOI-Service (wie Anm. 12).

beispielsweise eine automatisierte Videoanalyse mit Szenen-, Sprach-, Text- und Bilderkennung sowie eine semantische Anreicherung und Verknüpfung der Daten.

## Beispiel für Produktivsysteme: TIB|AV-Portal

Angesichts der steigenden Zahl von wissenschaftlichen AV-Medien und der Notwendigkeit, diese möglichst sogar auf Segmentebene zu erschließen, gibt es einen großen Bedarf an Lösungen für eine automatische Erschließung. Die Technische Informationsbibliothek (TIB) entwickelt gemeinsam mit dem Hasso-Plattner-Institut für Softwaresystemtechnik an der Universität Potsdam (HPI) ein Portal für AV-Medien.<sup>18</sup> Das AV-Portal optimiert den Zugang zu und die Nutzung von wissenschaftlichen Filmen aus Technik und Naturwissenschaften, wie z. B. Computeranimationen oder Vorlesungs- und Konferenzaufzeichnungen. Das wesentliche Merkmal des Portals ist die Kombination von State of the Art Multimedia-Analyseverfahren und Linked Open Data Verfahren für die semantische Analyse und das Retrieval.<sup>19</sup> Die Entwicklung orientiert sich an den 2010 durchgeführten Anforderungs- und Nutzungsanalysen.<sup>20</sup> 2011 ist ein teilfunktionaler Prototyp des AV-Portals entwickelt worden, 2012–2013 erfolgte die Weiterentwicklung und der Betabetrieb des Systems. Seit April 2014 ist das Portal online.<sup>21</sup>

Eines der wesentlichen Merkmale des AV-Portals ist die semantische Analyse, Suche und Exploration der im Portal hinterlegten AV-Medien.<sup>22</sup> Im Projekt wurden bekannte Analyseverfahren adaptiert bzw. weiterentwickelt wie z. B. Form-, Sprach- oder Strukturerkennung, um zusätzliche inhaltserschließende und strukturelle Metadaten zu erzeugen, die den Nutzer bei seiner Suche unterstützen. Die Nutzer können durch die Anreicherung der Daten mit semantischen Informationen Verbindungen zwischen den AV-Medien erkennen und darin explorativ navigieren.

Nach dem Einstellen eines neuen AV-Mediums sowie der dazugehörigen autoritativen Metadaten in das Portal wird das Medium zunächst automatisch in Schlüsselszenen segmentiert und Keyframes für einen visuellen Index erzeugt. Nach dieser strukturellen Analyse wird das Medium einer intelligenten Zeichenerkennung (ICR)

**18** TIB: AV-Portal. [www.tib-hannover.de/de/dienstleistungen/kompetenzzentrum-fuer-nicht-textuelle-materialien-knm/av-portal/](http://www.tib-hannover.de/de/dienstleistungen/kompetenzzentrum-fuer-nicht-textuelle-materialien-knm/av-portal/) (01.07.2014).

**19** Plank, Margret u. Janna Neumann: TIB's Portal for Audiovisual Media. New Ways of Indexing and Retrieval. In: IFLA Journal (2014) H. 1. S. 17–23.

**20** Plank, Margret: Kompetenzzentrum für multimediale Objekte (KMO). Abschlußbericht einer Anforderungsspezifikation für ein AV-Portal für wissenschaftliche Filme. [www.tib-hannover.de/fileadmin/kmo/Abschlussbericht\\_Studie\\_AVPortal.pdf](http://www.tib-hannover.de/fileadmin/kmo/Abschlussbericht_Studie_AVPortal.pdf) (19.07.2014).

**21** TIB|AV-Portal<sup>Beta</sup>: Startseite. [av.getinfo.de](http://av.getinfo.de) (01.07.2014).

**22** Sack, Harald u. Margret Plank: AV-Portal. The German National Library of Science and Technology's Semantic Video Portal. In: ERCIM News (2014) H. 96. S. 33–34. [ercim-news.ercim.eu/images/stories/EN96/EN96-web.pdf](http://ercim-news.ercim.eu/images/stories/EN96/EN96-web.pdf) (19.07.2014).



zugeführt, welche im Video sichtbare Buchstaben und Zeichen erkennt. Zudem wird, basierend auf der Audiospur des AV-Mediums, mittels automatischer Spracherkennung eine Transkription generiert. Mit Hilfe der Visual Concept Detection werden visuelle Bildinhalte entsprechend vordefinierter fachspezifischer und fächerübergreifender Kategorien wie z. B. Landschaft, Maschine, Zeichnung, Animation und Vorlesung klassifiziert.<sup>23</sup> Anschließend werden alle auf diese Weise zusätzlich generierten Metadaten mit einer linguistischen und semantischen Analyse verarbeitet. Hierbei werden Entitäten identifiziert, disambiguiert und z. B. im Falle der deutschen Version auf die Gemeinsame Normdatei (GND) gemappt.<sup>24</sup> Die Ergebnisse der semantischen Analyse werden unter anderem für eine inhaltsbasierte facetthierarchische Suche verwendet. Das Leistungsspektrum des zu implementierenden AV-Portals spiegelt somit einen aktuellen Stand der Forschung und Technik wider.

## Beispiel für Produktivsysteme: Portal PROBADO 3D

Im Projekt PROBADO<sup>25</sup>, das durch die Deutsche Forschungsgemeinschaft gefördert wurde, wurden Werkzeuge entwickelt, die es wissenschaftlichen Bibliotheken wie der Technischen Informationsbibliothek (TIB) erlauben, wissenschaftlich relevantes multimediales Material genauso selbstverständlich wie textuelle Fachinformation in ihre Prozessketten zu integrieren. An der TIB wurde mit PROBADO 3D, Teil des „Leistungszentrums für Forschungsinformation“ PROBADO, Expertise zur Erschließung, Recherche und Bereitstellung von 3D-Modellen gebündelt.

Mit Computer-Aided Design (CAD) erstellte 3D-Modelle sind im Baubereich sowohl aus der Praxis als auch aus der Forschung und Lehre nicht mehr wegzudenken. Anzahl und Komplexität von Modellen steigen durch umfassenden Einsatz aktueller CAD-Anwendungen und heterogener Planungsaufgaben stark an. Gleichzeitig steigen die Erwartungen der Nutzer an geeigneten Zugangs- und Präsentationsmodalitäten zur Suche in und Nutzung von solchen Beständen. Um das gezielte Auffinden von 3D-Modellen zu ermöglichen, bietet PROBADO 3D ein Onlineportal auf der Basis von inhaltsbasierten Erschließungstechnologien an. Es können unterschiedliche Modellinhalte identifiziert werden und automatisch beschreibende Daten erstellt

---

**23** Hentschel, Christian, Ina Blümel u. Harald Sack: Automatic Annotation of Scientific Video Material Based on Visual Concept Detection. In: Proceedings of the 13th International Conference on Knowledge Management and Knowledge Technologies. Hrsg. von Stefanie Lindstaedt u. Michael Granitzer. New York: ACM 2013.

**24** Steinmetz, Nadine u. Harald Sack: Cross-Lingual Semantic Mapping of Authority Files. Präsentation auf der Konferenz „Semantic Web in Libraries“. SWIB 2013, Hamburg.

**25** PROBADO (wie Anm. 11).

werden. Hinzu kommen visuelle Suchmöglichkeiten wie z. B. eine intuitive Suche mit 2D- und 3D-Skizziertool und eine Ergebnislieferung schon während des Zeichnens.<sup>26</sup>

Der Zugriff erfolgt per Web-Interface oder direkt aus der CAD-Umgebung. Im Portal lassen sich Einzelmodelle und Sammlungen integrieren, sowie externe Datenbanken anbinden. Die Zitierfähigkeit von 3D-Modellen wird durch die Vergabe von Digital Object Identifier Names (DOI) gewährleistet. Darüber hinaus kann eine parallele Recherche nach 3D-Modellen und anderen Medientypen im TIB-Portal GetInfo erfolgen.

Die im PROBADO-Projekt entwickelten Werkzeuge für die Erschließung und Bereitstellung von 3D-Modellen gehen im Vergleich zu konkurrierenden Angeboten weit über den State of the Art hinaus, vgl. Marktstudie PROBADO 3D.<sup>27</sup> Besonders hervorzuheben sind hierbei unter anderem die Funktionalitäten Modell-Verarbeitungskette, struktur- und formbasierte Erschließung und Suche, Retrievalverbesserung mittels Benchmarks inklusive Portalentwicklung PROBADO 3D.<sup>28</sup> Die Forschungs- und Entwicklungsarbeiten wurden basierend auf Benutzerbefragungen durchgeführt.

## Weitere Projekte

Die TIB führt ständig weitere innovative Projekte mit Schwerpunkt auf digitale Angebots- und Zugriffsformen für nicht-textuelle Materialien durch, die der Weiterentwicklung des Bibliotheks- und Informationswesens dienen. Einige Beispiele:

Das Projekt *VisInfo* (Visueller Zugang zu Forschungsdaten)<sup>29</sup> wurde über drei Jahre (2010–2012) im Rahmen des SAW-Verfahrens durch die Leibniz-Gemeinschaft gefördert.<sup>30</sup> Die Technische Informationsbibliothek (TIB) nahm die Projektleitung wahr, Partner waren GRIS (Graphisch-Interaktive Systeme) Darmstadt und das Fraunhofer-Institut für Graphische Datenverarbeitung *IGD*. Das Projekt *VisInfo*, Visueller Zugang zu Forschungsdaten, beschäftigte sich mit der Entwicklung von neuen, visuellen Verfahren für die Suche und Darstellung von Forschungsdaten in Digitalen Bibliotheken. Ziel des Projektes war die Entwicklung und prototypische Umsetzung von

**26** Berndt, René, Ina Blümel u. Raoul Wessel: PROBADO3D – Towards an Automatic Multimedia Indexing Workflow for Architectural 3D Models. In: Publishing in the Networked World. Transforming the Nature of Communication. Proceedings of 14th International Conference on Electronic Publishing. Helsinki: ELPUB 2010. S. 79–88. [elpub.architexturez.net/system/files/pdf/107\\_elpub2010.content.pdf](http://elpub.architexturez.net/system/files/pdf/107_elpub2010.content.pdf) (19.07.2014).

**27** Marktstudie PROBADO 3D: Suche und Erschließung von 3D Modellen für den Baubereich. Engage 2011. Hier S. 3.

**28** TIB: Publikationen. [www.tib-hannover.de/de/dienstleistungen/kompetenzzentrum-fuer-nicht-textuelle-materialien-knm/publikationen/](http://www.tib-hannover.de/de/dienstleistungen/kompetenzzentrum-fuer-nicht-textuelle-materialien-knm/publikationen/) (01.07.2014).

**29** TIB: *VisInfo*. [www.vis-info.info/](http://www.vis-info.info/) (01.07.2014).

**30** TIB: *VisInfo*- Visueller Zugang zu Forschungsdaten. [www.tib-hannover.de/de/forschung-und-entwicklung/abgeschlossene-projekte/visinfo-visueller-zugang-zu-forschungsdaten/](http://www.tib-hannover.de/de/forschung-und-entwicklung/abgeschlossene-projekte/visinfo-visueller-zugang-zu-forschungsdaten/) (19.07.2014)

innovativen Ansätzen für den interaktiven, graphischen Zugang zu Forschungsdaten, um diese optimal im Information Retrieval Prozess darstellbar und durchsuchbar zu machen. Der Fokus des Projektes lag auf der inhaltsbasierten, visuellen Suche sowie der Präsentation von Forschungsdaten aus dem Bereich der Erd- und Umweltwissenschaften. Ziel war, neben der metadatenbasierten Suche, auch die unmittelbare Suche in den Forschungsdaten. Die Anforderungen aus Sicht des Nutzers hinsichtlich des Suchprozesses sowie der Präsentation von Suchergebnissen wurden intensiv untersucht und in den Entwicklungsprozess rückgekoppelt.

Im Rahmen des durch EU Mittel geförderten Projekts *DURAARK*<sup>31</sup> (Projektlaufzeit: 01.02.2013–31.01.2016) wird ein Langzeitarchivierungssystem für Architekturinhalte entwickelt, das die Datensuche und -zugang auf verschiedenen semantischen Ebenen ermöglicht und dabei die Möglichkeiten derzeit genutzter Metadatenschemata bei weitem übertrifft. Gleichzeitig bietet es sichere und zukunftsfähige Datenspeicherung, indem es durch digitalen Verfall entstehende Probleme angeht.

Das durch die Deutsche Forschungsgemeinschaft geförderte Projekt *RADAR*<sup>32</sup> (Projektlaufzeit: 01.09.2013–31.08.2015) hat zum Ziel, eine Forschungsdateninfrastruktur aufzubauen und zu etablieren, um damit das in vielen wissenschaftlichen Disziplinen noch fehlende Forschungsdatenmanagement zu fördern. Damit soll RADAR einen wichtigen Beitrag zu einer besseren Verfügbarkeit, nachhaltigen Bewahrung und verbesserten Publikationsfähigkeit von (eigenständigen) Forschungsdaten leisten. Projektpartner sind das FIZ Karlsruhe – Leibniz-Institut für Informationsinfrastruktur (FIZ), das Karlsruher Institut für Technologie (KIT), Steinbuch Centre for Computing (SCC), Ludwig-Maximilians-Universität München (LMU), Fakultät für Chemie und Pharmazie sowie das Leibniz-Institut für Pflanzenbiochemie (IPB). Als generische, disziplinübergreifende Dienstleistung verfolgt RADAR einen zweistufigen Ansatz, der ein (nicht disziplinspezifisches) Einstiegsangebot zur Archivierung von Forschungsdaten (auch im Kontext der guten wissenschaftlichen Praxis) und ein höherwertiges Angebot zur Datenarchivierung mit integrierter Datenpublikation beinhaltet. Das Einstiegsangebot mit einer formatunabhängigen Archivierung und einem minimalen Metadatensatz richtet sich dabei vor allem an Wissenschaftler aus den „small sciences“, die Forschungsanträge stellen und darin die Möglichkeit zur Archivierung ihrer Daten gemäß von vorgegebenen Haltefristen (z. B. 10 Jahre nach DFG-Empfehlung) haben.

Das durch die Deutsche Forschungsgemeinschaft geförderte *Kompetenzzentrum für Forschungsdaten aus Erde und Umwelt (Komfor)* (01.07.2011–30.06.2014) ist als Bindeglied zwischen wissenschaftlichen Einrichtungen, Verlagen, Bibliotheken und einem bestehenden Archivnetzwerk für Daten aus Erd- und Umweltforschung geplant. Projektpartner sind: Zentrum für Marine Umweltwissenschaften, Universität

---

31 DURAARK: Scope of Project. [duraark.eu/](http://duraark.eu/) (01.07.2014).

32 TIB: RADAR – Research Data Repository. [www.tib-hannover.de/de/forschung-und-entwicklung/projekte/radar-research-data-repositorium/](http://www.tib-hannover.de/de/forschung-und-entwicklung/projekte/radar-research-data-repositorium/) (01.07.2014).

Bremen (Koordination), Deutsches Klimarechenzentrum GmbH, Hamburg, Alfred-Wegener-Institut für Polar- und Meeresforschung, Bremerhaven, Helmholtz-Zentrum Potsdam Deutsches GeoForschungsZentrum, Potsdam, Deutsches Zentrum für Luft- und Raumfahrt, Oberpfaffenhofen. Allgemeines Ziel ist die nachhaltige Verbesserung von Datenverfügbarkeit und -qualität. Konkret sollen nachhaltige und verlässliche Wege zur Publikation wissenschaftlicher Daten geschaffen werden, die den Qualitätsstandards wissenschaftlichen Publizierens entsprechen. Dazu ist beabsichtigt, eine aus der Sicht der Nutzer einheitliche Instanz zu schaffen, welche wissenschaftliche Projekte, Institute, Forschergruppen oder auch einzelne Wissenschaftler in allen Fragen des Datenmanagements begleitet – von der Projektplanungsphase über Datenerhebung, Qualitätssicherung, Registrierung und Langzeitarchivierung bis zur Publikation von Daten. Als technisches Kernstück soll eine entsprechende webbasierte personalisierbare Serviceplattform aufgebaut werden.<sup>33</sup>

Mit dem voranschreitenden Wandel hin zu einer kollaborativen digitalen Wissenschaft beschäftigt sich das *Open Science Lab* der TIB.<sup>34</sup> Hier sollen neuartige Webanwendungen wie beispielsweise Social Software, kollaborative Plattformen, Wikis und Stack Exchange für Forscherinnen und Forscher in enger Abstimmung mit überregionalen Benutzer-Communities erprobt und weiterentwickelt werden. Die kontinuierliche Entwicklung neuer Werkzeuge und Praktiken in allen oben genannten Bereichen wirft komplexe neue Probleme auf, für die das Team OpenScience Lösungen entwickelt. Die TIB ist Gründungsmitglied des strategischen Forschungsverbundes Science 2.0 des Leibniz Forschungsverbundes.<sup>35</sup>

## 5 Zusammenfassung und Ausblick

Ziel der TIB ist es, ihren Nutzern einen intelligenten und nahtlosen Zugriff auf Daten, Informationen und Wissen zu ermöglichen und große Informationsräume strukturiert und geeignet zugänglich zu machen. Es besteht ein großer Bedarf an Konzepten, Werkzeugen und Strukturen, welche dazu beitragen, nicht-textuelle Materialien für die Nutzer verbessert erschließen und bereitstellen zu können. Die TIB fokussiert sich im Kompetenzzentrum für nicht-textuelle Materialien derzeit auf die Entwicklung entsprechender Lösungen für die Medientypen AV-Medien und 3D-Modelle. Die TIB hat es sich darüber hinaus zum Ziel gemacht, durch die Vergabe von DOI-Namen den Zugang zu Forschungsdaten über das Internet zu erleichtern, die Akzeptanz von For-

<sup>33</sup> TIB: KomFor – Kompetenzzentrum für Forschungsdaten aus Erde und Umwelt. [www.tib-hannover.de/de/forschung-und-entwicklung/projekte/komfor-kompetenzzentrum-fuer-forschungsdaten-aus-erde-und-umwelt/](http://www.tib-hannover.de/de/forschung-und-entwicklung/projekte/komfor-kompetenzzentrum-fuer-forschungsdaten-aus-erde-und-umwelt/) (01.07.2014).

<sup>34</sup> TIB: Das Open Science Lab der TIB. <http://blogs.tib-hannover.de/opensciencelab/> (01.07.2014).

<sup>35</sup> ZBW – Deutsche Zentralbibliothek für Wirtschaftswissenschaften Leibniz-Informationszentrum Wirtschaft: Leibniz-Forschungsverbund Science 2.0. [www.leibniz-science20.de/](http://www.leibniz-science20.de/) (01.07.2014).

scheidungsdaten als eigenständige, zitierfähige wissenschaftliche Objekte zu steigern und somit die Einhaltung der Regeln guter wissenschaftlicher Praxis zu gewährleisten.

Für den schnellen und erfolgreichen Transfer von Forschungsansätzen, beispielsweise aus dem Multimedia Retrieval, der semantischen Suche und der automatischen Erschließung, in die Praxis der digitalen Bibliothek, sind kollaborative Forschungsprojekte zwischen Infrastruktureinrichtungen und Forschungseinrichtungen die Voraussetzung. Wie bereits im KII Report gefordert, gehört hierzu eine gezielte Ausweisung entsprechender Förderprogramme.<sup>36</sup>

Die kontinuierliche Zunahme der Bedeutung und Nutzung nicht-textueller Materialien für Wissenschaft, Forschung und Lehre stellt zudem neue Anforderungen an die Informationskompetenz der unterschiedlichen Nutzergruppen. Um dem gerecht werden zu können, müssen Spezialisten für die Erschließung, die Recherche, die Bereitstellung und die (Langzeit-)Archivierung nicht-textueller Materialien ausgebildet werden. Hierfür müssen sowohl die entsprechenden Hochschulen und Ausbildungsgänge gefördert werden und Ausbildungsinhalte erweitert bzw. angepasst werden als auch die Mitarbeiter und Mitarbeiterinnen der wissenschaftlichen Bibliotheken systematisch fortgebildet werden.

---

**36** Kommission, Gesamtkonzept (wie Anm. 4), S. B39.