

Suche in Multimediaarchiven von Kultureinrichtungen

Nicolas FLORES-HERR^{a,1}, Harald SACK^{b,1}, Klaus BOSSERT^c

^a *Acosta Consult, Zeielstrae 15 HH, 60318 Frankfurt, nf@acosta-consult.de*

^b *Senior Researcher, Hasso-Plattner-Institut fr Softwaresystemtechnik, Prof.-Dr.-Helmert-Strae 2-3, 14482 Potsdam, harald.sack@hpi.uni-potsdam.de*

^c *Acosta Consult, Zeielstrae 15 HH, 60318 Frankfurt, kb@acosta-consult.de*

Abstract. In diesem Kapitel werden Vorschläge für neue Suchparadigmen nach multimedialen Inhalten in Archiven von Kultureinrichtungen vorgestellt. Um die Notwendigkeit für eine Integration dieser neuen Technologien zu zeigen, werden zunächst Einschränkungen der klassischen katalogbasierten Bibliothekssuche im Zeitalter von immer weiter wachsenden Multimediastammungen beschrieben. Im Anschluss werden die Vor- und Nachteile zweier Suchparadigmen dargestellt, mit deren Hilfe in Zukunft für Wissenschaftler und Kulturschaffende die Suche nach multimedialen Inhalten erleichtert werden könnte. Zunächst werden die Perspektiven einer semantischen Suche auf Basis von Semantic-Web-Technologien in Bibliotheken beschrieben. Im Anschluss werden Suchmöglichkeiten für Multimediainhalte auf Basis von automatischer inhaltsbasierter Medienanalyse gezeigt. Das Kapitel endet mit einem Ausblick auf eine mögliche Vereinigung der beiden neuen Ansätze mit katalogbasierter Bibliothekssuche.

Keywords. Multimediastuche, Semantische Suche, Automatische Metadaterzeugung, Inhaltsbasierte Suche, Semantische Multimediastuche, Kultureinrichtungen, Bibliothekssuche, OPAC.

Einleitung

In Deutschland archivieren etwa 30.000 Kultureinrichtungen eine große Fülle von Multimediainhalten auf analogen und digitalen Trägern wie beispielsweise Büchern, Bildern, Audio-CDs, Tonbändern und Filmen. Durch effiziente Digitalisierungs- und Formatwandlungsverfahren ist es heutzutage möglich, in kurzer Zeit sehr große digitale Multimediaarchive aufzubauen. Hinzu kommen online publizierte multimediale Inhalte, die zahlreiche europäische Kultureinrichtungen aufgrund ihres erweiterten Sammelauftrags archivieren. Nun stehen Kultureinrichtungen in Anbetracht der kontinuierlich wachsenden Bestände vor der Herausforderung, ihren Nutzern eine einfache Suche und direkten Zugang zu ihren Multimediaarchiven zu ermöglichen.

Gleichzeitig sind in den letzten Jahren die Ansprüche der breiten Masse von Internetnutzern in Bezug auf die Suche nach und die Nutzung von Multimediaressourcen aufgrund der Popularität von Internetangeboten wie beispielsweise Google Images, Flickr, YouTube, Bing Video oder Wikimedia Commons deutlich gestiegen.

¹ Beide Autoren haben gleichermaßen zu der Publikation beigetragen. Dr. Nicolas Flores-Herr ist korrespondierender Autor: Nicolas Flores-Herr, Zeisselstr. 15 HH, 60318 Frankfurt, nf@acosta-consult.de.

Damit Kultureinrichtungen für Nutzer in Zukunft attraktiv bleiben, müssen sie vergleichbare internetbasierte Suchservices für ihre Multimediassammlungen anbieten, welche darüber hinaus auch den Ansprüchen von Wissenschaftlern und Kulturschaffenden genügen. Die Aufgabe der Kultureinrichtungen besteht nun darin, Technologien zu finden, mit denen diese Anforderungen erfüllt werden können.

Hierbei sollten nicht nur die Herausforderung massenhaft zu archivierender Inhalte im Fokus stehen, sondern auch die bisherigen Stärken von Kultureinrichtungen in neuem Licht betrachtet werden:

- Durch die Expertise von Bibliothekaren, Archivaren und Informationswissenschaftlern können in Kultureinrichtungen suchrelevante Metadaten von hoher Qualität, Vielfalt und Vollständigkeit erzeugt werden.
- Informationen zu Entitäten wie beispielsweise Personen, Körperschaften und Schlagworte können durch Kultureinrichtungen kuratiert werden.
- Staatliche Einrichtungen können aufgrund ihrer Beständigkeit eine nachhaltige Verfügbarkeit ihrer Ressourcen gewährleisten.
- Die Erschließung und Zugänglichmachung dieser Ressourcen in standardisierter und strukturierter Form schafft eine stabile Wissensbasis im Internet, die von Einrichtungen aus Wissenschaft und Kultur sowie Unternehmen gleichermaßen genutzt werden kann.

Neue Technologien zur Suche und Erschließung sollten Kultureinrichtungen darin unterstützen, diese Vorteile auszubauen und eine vertrauenswürdige, internetbasierte Wissensinfrastruktur zu erschaffen. In diesem Beitrag werden in den Abschnitten 2 und 3 mögliche Kandidaten für eine organische und lebendige Erweiterung der bisherigen Katalogsuche vorgestellt.

Der Beitrag gliedert sich in folgende Abschnitte:

1. *Aktueller Stand und Unzulänglichkeiten der Suche in Kultureinrichtungen auf Basis von Katalogen*
2. *Suche nach Multimediaressourcen auf Basis von Semantic-Web-Technologien*
3. *Multimediasuche auf Basis automatischer, inhaltlicher Erschließung*
4. *Ausblick Semantische Multimediasuche – eine Möglichkeit zur Nutzung automatischer Medienanalyse und von Semantic-Web-Technologien*

Zu Abschnitt 1: Bibliotheken haben mit der Ablösung der analogen Bibliothekskataloge durch einen Online Public Access Catalogue (OPAC) keine neuen Suchparadigmen für ihre Bestände eingeführt; vielmehr wurde die Zettelkastensuche mit ihren charakteristischen Merkmalen in digitaler Form reproduziert. Dies galt auch im Wesentlichen für den Erschließungsprozess zur Erzeugung von Metadaten für die Suche im Bibliothekskatalog. Der Erschließungsprozess in Bibliotheken beispielsweise erfolgt gemäß formalisierter Richtlinien², welche den Einsatz von automatischen Verfahren zur Inhaltsanalyse oder die Nutzung von semantischen Technologien bisher nicht explizit vorsehen.

In Abschnitt 1 wird beschrieben, welchen Herausforderungen Kultureinrichtungen bei der Erschließung und Suche nach digitalen Multimediaobjekten gegenüberstehen und wie die „klassische“ Archiv- oder Bibliothekssuche durch neue Suchverfahren erweitert werden kann.

² Regeln für den Schlagwortkatalog (RSWK) und Regeln für die alphabetische Katalogisierung (RAK)

Zu Abschnitt 2: Der Ausgangspunkt der Idee eines Semantic Web war die Vision von Tim Berners-Lee³, nach welcher Informationen im World Wide Web von Maschinen interpretiert und automatisch weiterverarbeitet werden. Dies beinhaltet beispielsweise Informationen über Relationen zwischen einzelnen Entitäten wie Personen, Orten und Organisationen, welche mithilfe von standardisierten Web-Ontologiesprachen wie RDF/OWL zu einer Wissensbasis modelliert werden können. Um eine semantische Suche⁴ auf standardisierte Weise zu implementieren und die Interoperabilität von Datenformaten zu gewährleisten, können die verschiedenen Komponenten des *Semantic Web Stack* [3] genutzt werden, der aus Technologien und Beschreibungssprachen besteht. Allerdings ist der Mehrwert jeder semantischen Suche wiederum von der Vertrauenswürdigkeit und Qualität der verwendeten Wissensbasis (z. B. Ontologie) und ihrer Bestandteile (Klassen, Instanzen und Relationen) abhängig. Weiterhin ist eine am Einsatzzweck (z. B. der Komplexität der Suche) ausgerichtete Angemessenheit der Modellierungstiefe von Daten von großer Bedeutung.

Die Modellierung von Wissensbasen und die Erzeugung der darin enthaltenen Daten gehen zum gegenwärtigen Stand der Technik noch mit erheblichem intellektuellem Aufwand einher. Aufgrund der ständig wachsenden Menge an sammlungsrelevanten Multimediainhalten ist es für Kultureinrichtungen derzeit kaum möglich, eine gleichbleibende und für eine semantische Suche hinreichend tiefe Modellierung einer Wissensbasis zu gewährleisten. Darüber hinaus berücksichtigen Erschließungsgeschäftsgänge von Kultureinrichtungen heutzutage zumeist nicht die technologischen Standards des Semantic Web. Dadurch wird eine Vernetzung von Archivinhalten mit Ressourcen des Internets und, damit verbunden, eine Anreicherung von Suchergebnissen erschwert.

In diesem Abschnitt wird beschrieben, auf welche Weise Semantic-Web-Technologien auf Multimediaarchive in Kultureinrichtungen angewendet werden können und worin die Einschränkungen bestehen.

Zu Abschnitt 3: Zwei Fachgebiete, die sich schon seit der Mitte des letzten Jahrhunderts einer computergestützten Charakterisierung von unstrukturierten Daten wie digitalen Bildern zugewendet haben waren „Information Retrieval“ und „Künstliche Intelligenz“ bzw. „Maschinelles Lernen“. Das Prinzip dieser Fachgebiete ist, dass durch eine computergestützte Berechnung eine explizit verfügbare maschinenlesbare Abstraktion (Merkmal bzw. Feature) eines digitalen Multimediaobjekts erzeugt wird. Aus informationstheoretischer Sicht besitzt diese Abstraktion eine geringere Dimensionalität als das ursprüngliche Medienobjekt. Entscheidend für eine Suche ist jedoch die Wahl der Berechnungsverfahren, mit deren Hilfe ein für das jeweilige Multimediaobjekt charakteristischer Satz von Merkmalen erzeugt werden kann. Jedes berechnete Merkmal kann als Basisvektor eines das Medienobjekt beschreibenden Merkmalsvektors interpretiert werden. Mithilfe dieser Merkmalsvektoren ist eine Multimediasuche auch ohne textuelle Metadaten oder Schlagworte eines Bibliothekskatalogs möglich. Allerdings kann Multimediasuche auf Basis von Merkmalsvektoren eine OPAC-basierte Suche in Kultureinrichtungen nicht vollständig ersetzen, weil es bis heute keine Verfahren gibt, die vollkommen automatisiert eine zuverlässige, regelwerkskonforme Erschließung durchführen können.

³ Welche 1999 vorgestellt und im Jahre 2001 im Scientific American publiziert wurde [1].

⁴ Es ist anzumerken, dass der Begriff „semantische Suche“ von „semantische Suche auf Basis von Semantic-Web-Technologien“ abzugrenzen ist. Letzterer, welcher Gegenstand des Kapitels ist, bezieht sich auf die Nutzung von Bestandteilen des Semantic Web Stacks des W3C. Siehe auch [2].

Die Möglichkeiten und Herausforderungen zur Suche von Multimediainhalten in Bibliotheken mithilfe von Technologien zur automatischen Inhaltsanalyse werden in Abschnitt 3 beschrieben.

Zu Abschnitt 4: Aufgrund der zahlreichen möglichen Anwendungen der Multimedia-suche auf Basis automatischer Inhaltsanalyse und der Semantischen Suche für Multimediaarchive in Kultureinrichtungen gibt es neue Forschungsansätze, diese Fachgebiete mit der klassischen Katalogsuche zu vereinigen. Der Abschnitt schließt mit einem kurzen Ausblick auf die Potenziale dieser Stoßrichtung ab.

Ein weiterer, viel beachteter und erfolgversprechender Ansatz der vergangenen Jahre war die Verschmelzung von Bibliothekssuche mit den Technologien des Web 2.0, also vor allem die Nutzerinteraktion betreffende Funktionalitäten. Dieser Trend in Richtung OPAC 2.0 wurde bereits in zahlreichen anderen Veröffentlichungen vorgestellt und diskutiert (siehe unter anderem [4] [5] [6] [7]). Aufgrund dessen und weil Konzepte des Web 2.0 in einigen Forschungsprojekten mit den in diesem Kapitel diskutierten Technologien kombiniert werden⁵, wird dieser Ansatz in Abschnitt 4 nur kurz beschrieben.

1. Aktueller Stand und Unzulänglichkeiten der Suche in Kultureinrichtungen auf Basis von Katalogen

1.1. Aktueller Stand der Katalogsuche in Bibliotheken

Sämtliche Sammlungen und Bibliotheken verzeichnen ihre Werke seit der Antike in Bibliothekskatalogen. Diese sind das Ergebnis von intellektueller und regelbasierter Erschließung, welche von qualifizierten Fachkräften (z. B. Bibliothekaren, Archivaren, Informationswissenschaftlern) durchgeführt wird. Noch in der zweiten Hälfte des 20. Jahrhunderts standen Karteikarten und Mikroformen im Mittelpunkt der Bibliothekssuche, welche wiederum Bestandteil von Zettelkatalogen bzw. Mikrofiche-Katalogen waren. Im Rahmen der Erschließung wurden für jeden Autoren, jeden Titel, jedes Schlagwort und eine Reihe weiterer Informationen eigene Karteikarten angelegt, mit deren Hilfe Nutzer im Bestand der Bibliothek recherchieren konnten.

Heutzutage sind Bibliothekskataloge nahezu flächendeckend in elektronischer Form als OPAC verfügbar, mit deren Hilfe Nutzer auch über das WWW im Bestand der jeweiligen Kultureinrichtung recherchieren können. Die Medienschließung durch Bibliothekare erfolgt auch im digitalen Zeitalter gemäß tradierter Regelwerke, welche institutionenübergreifend festlegen, in welcher Weise die Ansetzung von Namen von Personen, Orten, Organisationen sowie der Schlagworte geschehen soll. Aus diesen Gründen ist der Informationsgehalt von digitalen Katalogeinträgen nahezu identisch mit den vormals gebräuchlichen Karteikarten oder Mikroformen.

1.2. Probleme bei einer Katalogsuche in Multimediabeständen

Traditionell verzeichnen Bibliotheken hauptsächlich Bücher und Periodika, doch haben viele Einrichtungen auch multimediale Inhalte in analoger Form wie beispielsweise Tonträger, Bildtonträger und Fotos in ihrem Bestand. Im Zuge der Digitalisierung von Büchern und Multimediaformaten sowie aufgrund eines erweiterten Sammelauftrags⁶

⁵ Unter anderem im Projekt CONTENTUS im Rahmen des THESEUS-Forschungsprogramms.

⁶ Zum erweiterten Sammelauftrag der Deutschen Nationalbibliothek siehe [8].

archivieren einige Kultureinrichtungen mittlerweile im großen Maßstab multimediale Inhalte auch in digitaler Form. Die Herausforderung für die bibliothekarische Erschließung im Zeitalter der Digitalisierung besteht nun einerseits in der

- a) immer größer werdenden Menge an gesammelten Objekten und andererseits in
- b) der Beschreibung der nur schlecht oder gänzlich unstrukturierten Datenbereiche, insbesondere von Multimediaobjekten.

Trotz dieser Herausforderungen haben sich die Regelwerke zur Erschließung und, damit verbunden, auch die Bibliothekssuche als Ganzes in den letzten Jahren nicht grundlegend geändert. Für den OPAC-Nutzer spiegelt sich dieser Umstand in der Suchoberfläche von Bibliotheken wieder, der in der Regel über folgende Sucheinstiege erfolgt:

1. Autor
2. Titel
3. Körperschaft
4. Schlagwort
5. ISBN
6. Verlag
7. Erscheinungsjahr.

Bei der Erschließung einer elektronischen Ressource mit großen unstrukturierten Bereichen (z. B. Hörbuch oder Dokumentarfilm) können aus Zeitgründen nicht alle darin vorkommenden Entitäten wie Personen, Orte, Organisationen und Schlüsselwörter erfasst werden. Darüber hinaus liefern OPACs meist keine Angaben zur inneren Struktur eines Multimediaobjekts, (z. B. Kapitel, Seitenzahl oder Zeile im Falle eines Buchs oder Zeitangabe im Falle von Audioinhalten und audiovisuellen Medien), welche aus den Suchtreffern als direkte Einstiegspunkte in das jeweilige Medium dienen könnten.

Um Nutzern den Zugang zu Bibliotheksinhalten zu erleichtern, reichern daher seit einigen Jahren ausgewählte Einrichtungen ihre Kataloge mit digitalen Inhaltsverzeichnissen, Volltexten, Miniaturbildern (Thumbnails) und externen Links an⁷. Dies eröffnet Bibliotheksnutzern die Möglichkeit, in einem gewissen Rahmen inhaltsbasiert in digitalen Werken zu suchen. So können Nutzer schon heute in einigen digitalen Bibliotheken⁸ eine Volltextsuche durchführen, welche die angereicherten Elemente einbezieht, und so mit dem Suchergebnis beispielsweise den direkten Einstiegspunkt zum betreffenden Textabschnitt mitliefert. Eine flächendeckende, vollständige inhaltliche Erschließung von Multimediaobjekten, welche über die klassische Katalogisierung und die gegenwärtig eingesetzte Kataloganreicherung hinaus geht, erfolgte bisher allerdings nicht.

Dies ist zum Teil darin begründet, dass große Sammlungen von digitalisiertem Kulturgut heutzutage in der Hand von Unternehmen sind, die Technologien für eine automatische Erschließung in großem Maßstab einsetzen können. So ist gegenwärtig Google Bücher [11] mit über 15 Millionen Werken die mit Abstand größte Sammlung digitalisierter Bücher mit durchsuchbaren Volltexten. Im Vergleich dazu hat in Frankreich das Projekt Gallica (Bibliothèque numérique de la Bibliothèque nationale de France) – einem der größten Digitalisierungsprojekte europäischer Kultureinrichtungen – bisher nur etwa 150.000 Bücher digitalisiert [12] und zum Teil erschlossen. Auf der

⁷ Zur Kataloganreicherung bei der Deutschen Nationalbibliothek siehe [9]. Weitere Literaturhinweise sind in [10] enthalten.

⁸ Einige Beispiele für angereicherte Kataloge finden sich u. a. auf den Webseiten der Deutschen Nationalbibliothek, Herzogin Anna Amalia Bibliothek, Bayerische Staatsbibliothek und Herzog August Bibliothek.

gesamten europäischen Ebene wird die Anzahl der digitalisierten Bücher auf 1,2 Millionen geschätzt [13].

Ein weiteres Problem sind die in den Bibliotheken gebräuchlichen Datenformate und Übertragungsprotokolle, welche nicht auf gängigen Internetstandards basieren und somit für die Nutzung der Kataloginformationen durch Partner außerhalb des Kulturbetriebs hohe technische Hürden aufbauen.

Auch in den Archiven von Sendeanstalten, welche eine zentrale Bedeutung für die Bewahrung audiovisueller Kulturgüter haben, findet sich eine ähnliche Situation wie in Bibliotheken. Durch die Vielzahl an Redaktions- und Produktionssystemen sowie Speichermedien ist in den Sendeanstalten oft eine heterogene Mischung technischer Systeme mit jeweils eigenen Metadatenformaten wie MXF, BMF, FESAD und FESADneu entstanden. Inzwischen gibt es zwar bereits Rechercheoberflächen, die eine parallele Suche verschiedener Archivsysteme ermöglichen (zum Beispiel das System Medienbroker [13], welches zurzeit vom Bayerischen Rundfunk, dem Mitteldeutschen Rundfunk und seit Neuestem auch vom Rundfunk Berlin-Brandenburg eingesetzt wird), jedoch sind diese nach wie vor auf dem Konzept der kataloggestützten Suche aufgebaut und wegen ihrer Komplexität meist nur von ausgebildeten Archivaren und Dokumentaren effizient zu nutzen. Somit werden auch bei Sendeanstalten die Potenziale inhaltsbasierter Suchtechnologien durch die aktuell verwendeten Systeme bei Weitem nicht ausgeschöpft.

1.3. Probleme bei Erschließung von digitalen Multimediabeständen am Beispiel Musik

Im Rahmen der klassischen Bibliothekerschließung ist die Bereitstellung eines inhaltsbasierten Zugriffs auf Bibliotheksressourcen mit beträchtlichem Arbeitsaufwand verbunden. So muss für jedes neue Schlagwort eine eigene Karteikarte angelegt werden. Für den Fall, dass vorhandene Schlagworte das Werk hinreichend beschreiben, müssen diese gemäß RSWK in einer bestimmten vorgeschriebenen Reihenfolge aufgenommen und sogenannte Schlagwortketten gebildet werden. Diese Form der Erschließung ist nicht nur zeitaufwändig, sondern setzt auch bei den Nutzern der Bibliothek ein hohes Maß an Spezialwissen voraus.

Um speziell Multimediabestände zu organisieren und zugreifbar zu machen, wurde eine ganze Reihe von Fachthesauri und verschiedener Regelwerke (zum Beispiel „RAK-Musik“ für Musikalien) entwickelt. Hierbei wurde berücksichtigt, dass klassische Sucheinstiege wie Titel und Autor der Veröffentlichung teilweise nicht vorhanden sind bzw. für die Nutzer keinen wirklichen Mehrwert bei der Suche darstellen. Beispielsweise können bei einer Suche nach Musik Katalogerweiterungen wie Einheits-sachtitel [15] von großem Nutzen sein, um unterschiedliche Schreibweisen einer Komposition auf ein Werk zurückzuführen. Aber auch hier ist, ebenso wie bei Schlagwortketten, für die Suche eine gute Kenntnis der Erschließung notwendig, die in der Regel bei Bibliotheksnutzern nicht vorausgesetzt werden kann. Somit ist nicht nur die Suche nach einem Musikstück, sondern auch dessen Katalogisierung mit Schwierigkeiten verbunden. Zum Beispiel kann ein Musikbibliothekar für eine regelkonforme Titelaufnahme einer einzelnen Musik-CD fallweise bis zu mehreren Stunden benötigen.

1.4. Chancen durch eine Erweiterung der Katalogsuche durch Semantic Web und automatische Inhaltsanalyse

Zusammengefasst sind wesentliche Probleme der Katalogisierung von digitalen Multimediabeständen die aufwändige Struktur- und Inhaltsanalyse einer immer größer wer-

denden Menge von Multimediainhalten. Mit elaborierten Strategien zur Kataloganreicherung wird zwar versucht, diesem Problem entgegenzuwirken und eine inhaltsbezogene Suche zumindest teilweise zu ermöglichen. Aber hier besteht das Problem, dass, im Gegensatz zur bestehenden Erschließung, die Kataloganreicherung weder inhaltlich noch technisch als Teil von Regelwerken definiert ist. Somit ist zum Beispiel nicht gewährleistet, dass die zusätzlich entstehenden Suchinhalte institutionenübergreifend in standardisierten Formaten genutzt werden können.

Umgekehrt könnten durch die konsequente Nutzung von W3C-Standards [16] – insbesondere des Semantic Web Stacks [3] [17] – bestehende Kataloginformationen einfacher mit externen Quellen vernetzt und Nutzern außerhalb des Kulturbetriebs zur Verfügung gestellt werden.

2. Suche nach Multimediaressourcen auf Basis von Semantic-Web-Technologien

2.1. Kurze Definition der Semantischen Suche auf Basis von Semantic-Web-Technologien

Im Gegensatz zum traditionellen Information Retrieval, bei dem auf der Basis textueller Daten eines (Text-)Dokumentenkorpus ein Suchindex aufgebaut wird, über den ein effizienter und zielgerichteter Zugriff ermöglicht wird (schlüsselwortbasierte Suche, Keyword-based Search), erfolgt die semantische Suche auf Basis von Semantic-Web-Technologien durch die im Korpus enthaltenen Instanzen, Relationen und Klassen. Auch aus multimedialen Daten lassen sich deskriptive, inhaltsbeschreibende Metadaten auf textueller Basis gewinnen, die sich für die schlüsselwortbasierte Suche nutzen lassen. Ein generelles Problem natürlichsprachlicher Textdokumente oder deskriptiver, textueller Metadaten liegt in der Mehrdeutigkeit der natürlichen Sprache auf unterschiedlichen semantischen Abstraktionsebenen. Ein Wort (Term) selbst kann unterschiedliche Bedeutungen besitzen (Homonymie), so bezeichnet beispielsweise der Term „Bank“ sowohl eine Sitzgelegenheit als auch ein Finanzinstitut. Eine klassische schlüsselwortbasierte Suche nach dem Begriff „Bank“ resultiert in Ergebnisdokumenten, in denen der Term „Bank“ in unterschiedlichen Bedeutungen verwendet wird, von denen nicht alle der vom Benutzer intendierten Bedeutung entsprechen. Aber selbst wenn auf eine einheitliche Bedeutung geschlossen werden könnte, so kann ein Wort auch in unterschiedlichem Kontext und mit unterschiedlicher Absicht (Pragmatik) vom Autor des Dokuments verwendet worden sein, die mit den Absichten des suchenden Benutzers nicht übereinstimmen muss.

Ebenso können Dokumente mit synonymen Wörtern, die denselben Begriff mit unterschiedlichen Termen beschreiben, über eine einfache schlüsselwortbasierte Suche nicht gefunden werden, da inhaltsrelevante Dokumente das eigentliche Suchwort nicht enthalten müssen. Neben solchen Synonymen können Begriffe auch mit Metaphern und anderen sprachlichen Ausdrucksmitteln umschrieben werden, ohne dass der inhaltlich damit bezeichnete Suchbegriff in einem relevanten Dokument auftauchen muss – somit stehen die inhaltliche Bedeutung (Semantik) eines Dokuments und nicht nur die darin verwendeten Zeichenketten im Zentrum einer inhaltsbasierten Suche.

Zur semantischen Suche müssen Informationsressourcen (Dokumente) mithilfe semantischer Metadaten annotiert werden, die die Bedeutung der Inhalte selbst formalisieren und kodieren, damit diese maschinell gelesen und korrekt interpretiert werden können. Die Repräsentation der Semantik erfolgt mithilfe sogenannter „Ontologien“ und wird im folgenden Abschnitt an einem kurzen Beispiel erläutert.

Die Verwendung semantischer Metadaten zur Realisierung einer inhaltsbasierten, semantischen Suche kann das traditionelle, textbasierte Information Retrieval auf folgende Weise unterstützen [18]:

- Das Erzielen von Suchergebnissen mit größerer Genauigkeit und Vollständigkeit (High Precision und High Recall)
- Sinnvolle und zielgerichtete Präzisierung und Erweiterung von Suchergebnissen (Query String Refinement)
- Herleitung von implizit vorhandenen, verdeckten Informationen (Inferencing)
- Herstellung von inhaltsbezogenen Querverweisen und Assoziationen (Cross Referencing)
- Nutzung von semantischen Beziehungen zur Visualisierung und Navigation durch den Such- oder Ergebnisraum der Suche (Facettierte Suche, Browsing, Explorative Suche)

2.2. Nutzung formaler Wissensrepräsentationen (Ontologien) für die Suche in Kultureinrichtungen

Wie eingangs erwähnt, kann die Repräsentation der Semantik mithilfe sogenannter „Ontologien“ erfolgen. Der Begriff *Ontologie* stammt ursprünglich aus der Philosophie und zählt zur Disziplin der *Metaphysik*, die sich primär mit dem *Sein*, dem *Seienden* als solchem und mit den fundamentalen Typen von *Entitäten* beschäftigt. Im Gegensatz zur *Erkenntnistheorie* (*Epistemologie*) beschreibt die *Ontologie* die *Welt*, wie sie tatsächlich ist, und nicht, wie sie uns gefiltert durch unsere *Sinnesorgane* und durch unsere *persönliche Erfahrung* erscheint. In der *Informatik* reduziert sich der *Ontologiebegriff* auf eine rein technische Sichtweise und bezeichnet eine „explizite, formale Spezifikation einer gemeinschaftlichen Konzeptualisierung“, d. h. ein abstraktes Modell (*Konzeptualisierung*), das alle relevanten Begriffe innerhalb eines Wissensgebiets (*Domäne*) und deren Beziehungen untereinander abbildet, wobei die Bedeutung aller Begriffe vollständig definiert werden muss (*explizit*) in einer *maschinenlesbaren Form* (*formal*) und *Konsens* unter den kommunizierenden Parteien über die Bedeutung der *Ontologie* herrschen muss [19]. Hierbei muss angemerkt werden, dass im Gegensatz zur *Philosophie* in der *Informatik* *subjektives Wissen* durchaus Teil einer *Ontologie* sein kann.

Einfache Beispiele für *Ontologien* sind z. B. *Thesauri*, also *Wörterbücher*, in denen inhaltliche Zusammenhänge zwischen einzelnen Begriffen aufgezeigt werden, wie z. B. *Ober- und Unterbegriffe*, *Spezialisierungen* und *Verallgemeinerungen*, *Synonyme* und *assoziativ verknüpfte Begriffe*. Speziell zur Beschreibung des *Archivbestands* von *Kultureinrichtungen* bestehen bereits seit *Langem* spezielle (*kontrollierte*) *Vokabularien* und *Thesauri*, die eine *eindeutige inhaltliche Beschreibung* der *kulturellen Assets* erlauben, um gezielt auf diese (bzw. auf *Informationen* über diese) zugreifen zu können. Ein prominentes Beispiel aus diesem Bereich ist das *Conceptual Reference Model CIDOC-CRM* [20], ein *konzeptuelles Modell* zur Beschreibung von *Ressourcen* und *Prozessen* in *Archiven* und *Museen*. Es wurde vom *International Council of Museums (ICOM)* entwickelt und liegt seit 2006 als *ISO-Standard* vor (*ISO 21127*). Ein weiterer *Beschreibungsstandard* ist der von der *Library of Congress* in *Kooperation* mit der *Society of American Archivists (SAA)* entwickelte *EAD (Encoded Archival Description) Standard* [21] [22]. *EAD* dient zur *Erstellung* sogenannter „*Findhilfen*“ in *Archiven*. Als „*Findhilfen*“ werden in diesem Zusammenhang *indexartige Strukturen* verstanden,

Da Kultureinrichtungen institutionenübergreifend strukturierte Datenbestände zu Entitäten sowohl erzeugen als auch kuratieren, kann durch Nutzung von Semantic-Web-Formaten dieses Wissen in das WWW integriert werden. So werden im bis zum Jahr 2012 laufenden Projekt „Gemeinsame Normdatei (GND)“ [25] die Normdaten der deutschen Bibliotheksverbände zusammengeführt. Bereits jetzt werden im Rahmen eines Testbetriebs diese Daten teilweise öffentlich zur Nutzung bereitgestellt [27]. Die GND umfasst die Personennormdatei (PND), die Gemeinsame Körperschaftsnormdatei (GKD) und die Schlagwortnormdatei (SWD). In der PND werden natürlichen Personen sowie literarische Gestalten und Kunstfiguren erfasst mit dem Ziel, Personen eindeutig für die Verwendung in Bibliothekskatalogen zu identifizieren. Darüber hinaus werden unterschiedliche Namensschreibweisen zusammen mit biografischen und bibliografischen Daten erfasst. Die GKD ist die im deutschen Sprachraum relevante Normdatei für in- und ausländische Körperschaftsnamen. Hierzu zählen öffentliche und privatwirtschaftliche Institutionen sowie Veranstaltungen. Neben der Angabe von synonymen Namen und Bezeichnungen werden geografische Angaben (z. B. Firmensitz) und bibliografische Informationen bereitgestellt. Die SWD stellt ein terminologisch kontrolliertes und strukturiertes Vokabular bereit, das eine Vielzahl einzelner Fachgebiete abdeckt. Ein Teil der Schlagworte verfügt sogar über Verweise auf Synonyme, Ober- und Unterbegriffe sowie verwandte (assoziierte) Begriffe. Die einzelnen Dateien der GND sind auch als RDF-kodierte Wissensrepräsentation im Semantic Web verfügbar [28] und damit direkt zur Unterstützung einer semantischen Suche geeignet.

Um z. B. Personen eindeutig als semantische Entitäten referenzieren zu können, kann die PND genutzt werden. Sollen Personen in einen semantischen Suchindex aufgenommen werden, erfolgt dies durch die Aufnahme einer Referenz auf einen Datensatz innerhalb der PND, die die betreffende Person repräsentiert. Soll etwa ein literarisches Werk in einen semantischen Suchindex aufgenommen werden, können die Namen der Autoren in der PND ermittelt und durch die entsprechenden (eindeutigen) Identifikatoren ergänzt werden. Die durch die PND zur Verfügung gestellten synonymen und alternativen Namen können genutzt werden, um eine Sucheingabe eindeutig der gesuchten Entität zuzuordnen.

2.4. Beispiel für Semantische Suche in Multimediaarchiven

Eines der bekanntesten Projekte zur Suche von digitalem Kulturgut ist die Europeana [29], die europäische digitale Bibliothek. Das Ziel von Europeana ist, den Benutzer in die Lage zu versetzen, die digitalen Ressourcen europäischer Museen, Bibliotheken und Archive über ein zentrales Suchportal zu erforschen. Zahlreiche weitere Projekte tragen dazu bei, den über die Europeana verfügbaren Datenbestand zu mehren, so z. B. das European Film Gateway [30] (Filme, Filmplakate, Fotografien, Drehbücher), das Archives Portal Europe [31] (APENet, Nationalarchive), EU Screen [32] (Fernsehübertragungen) oder aber auch das Musical Instruments Museum Online (MIMO) [33]. Mitte 2010 umfasste der Datenbestand der Europeana über 14 Millionen digitalisierte Bücher, Landkarten, Fotos, Gemälde, Filme und Musikstücke.

Die Europeana stellt einen Prototyp zur semantischen Suche über einen Teil der in der Europeana verfügbaren Datenbestände bereit, der Semantic-Web-Technologien und Linked Data [34] nutzt, um Assoziationen zwischen Entitäten herzustellen und um Suchbegriffe in einen Kontext zu stellen. So zeigt der Prototyp der Europeana mittels einer Autovervollständigungsoption dem Nutzer während der Eingabe des Suchbegriffs inhaltliche Alternativen in semantischer Relation zu den aktuell verfügbaren Datenressourcen an (siehe Abbildung 2).

This is a research prototype of Europeana's semantic search engine.
Enter a search term, for example: Egypt, Rembrandt, window.

q lisa search

artefact view all 14 results ▶

Portret van Lisa Langerhorst-Ver...
schilderij Sluifjers, Jan

PORTRAIT DE MONA LISA (1479-1528...
DITE LA JOCONDE DI SER PIERO DA VINCI ;
VINCI Léonard de (dit) LEONARDO

Portret van Lisa Langerhorst-Ver...
schilderij Sluifjers, Jan

collection

Shiller Fine Art, Lisa

concept

1617-18 leerjongen bij Antoni LI...

Hier schildert hij de Mona Lisa.

location view all 4 results ▶

Lisa
river

Lisabon
city


Lisart
inhabited place

person view all 133 results ▶

Lisa

Lisa

Lisalba
Italian artist 1675 1775



Louvre
11,327 artworks

Abbildung 2: Die semantische Suche in der Europäischen Digitalen Bibliothek Europeana

Ein weiteres Beispiel für die Umsetzung einer semantischen Suche ist das finnische MuseuSuomi, dessen Funktionsweise und Architektur in einer Veröffentlichung von Hyvönen *et al.* [35] aus dem Jahr 2004 beschrieben wird.

Zur Gewährleistung einer inhaltsbasierten semantischen Suche muss die Bedeutung der Dokumenteninhalte mithilfe formaler Wissensrepräsentationen (z. B. Ontologien) explizit gemacht werden. Kultureinrichtungen stehen zu diesem Zweck bereits existierende Vokabularien und Wörterbücher zur Verfügung, wie z. B. der CIDOC-CRM- oder der EAD-Standard. Von besonderer Bedeutung sind dabei Katalogentitäten, wie sie in der GND zusammengestellt sind. Über diese können personen- und schlagwortbasierte semantische Suchverfahren auf einfache Weise implementiert werden.

3. Multimediasuche auf Basis automatischer inhaltlicher Erschließung

Im folgenden Abschnitt soll der Einsatz von Verfahren zur automatisierten inhaltlichen Erschließung von Multimediaobjekten in Kultureinrichtungen diskutiert werden. Um dem Leser einen Überblick über mögliche Verbesserungen für die Suche in Multimediaarchiven zu geben, wird zunächst die Medienanalyse in Grundzügen beschrieben. Diese einführenden Abschnitte richten sich an Leser, die sich bisher nur wenig mit automatischen Verfahren zur Inhaltsanalyse von Medien auseinandergesetzt haben. Die in diesem Rahmen verwendeten Literaturangaben sind als Einstiegspunkte in die Thematik zu verstehen und haben nicht den Anspruch auf Vollständigkeit.

3.1. Hintergrund

Aufgrund der verstärkten Nutzung von multimedialen Inhalten in den vergangenen Jahren ist eine einhergehende Ausweitung der Suchverfahren von textuellen Inhalten auf Multimediendateiformate zu beobachten. Weil Multimediadateien (Text, Bilder, Audio, Video usw.) überwiegend aus unstrukturierten Informationen (Binary Large Objects) bestehen, deren automatische Analyse in großen Sammlungen bislang eine schwierige Aufgabe darstellt, verwenden viele Multimediaportale größtenteils manuell erzeugte bzw. nutzergenerierte textuelle Metadaten wie Bildunterschriften oder Umgebungstext (Google Images) oder Tags (YouTube, Flickr) für ihre Suche. Die auf diese Weise manuell erzeugten textuellen Metadaten sind in vielen Fällen nicht hinreichend ausführlich und präzise, um einzelne Multimediaobjekte in großen Sammlungen zu finden. Sofern Nutzer von Kultureinrichtungen für ihre Recherche präzisere Such- und Filtermöglichkeiten benötigen, ist eine höhere Erschließungstiefe notwendig. Dies ist prinzipiell durch eine eingehende intellektuelle Erschließung möglich⁹. Allerdings ist dies mit einem enormen Zeitaufwand verbunden: Um beispielsweise eine Stunde audiovisuelles Material durch einen Experten annotieren zu lassen, sind etwa vier Stunden manueller Arbeit notwendig. Mit Blick auf die Fülle von Material in Filmarchiven und Rundfunkanstalten ist diese Form der Erschließung nicht praktikabel.

Eine weitere Einschränkung der manuellen Erschließung ist die Tatsache, dass eine Neuerschließung des gesamten Bestands, zum Beispiel aufgrund geänderter Nutzeranforderungen an die Suche, nicht skaliert. Es ist beispielsweise vorstellbar, dass Nutzer eines Tages nach der Eigenschaft „Stimmung bzw. Atmosphäre eines Fotos“ Fotosammlungen durchsuchen möchten. Sofern eine Analyse dieser Eigenschaft bisher nicht Teil des Erschließungsprozesses war, muss der gesamte Fotobestand erneut manuell verschlagwortet werden, was mit hohen Kosten und Aufwänden verbunden ist.

3.2. Metadaten aus Feature-Extraktionsverfahren

Die automatisierte Erschließung von Multimediainhalten und deren Suche ist seit einigen Jahrzehnten Forschungsgegenstand des Fachgebiets Multimedia Information Retrieval¹⁰. Multimedia Information Retrieval ist eine Methodologie, die entwickelt wurde, um Multimediadatenbanken *inhaltsbasiert* nach Objekten durchsuchen zu können, also

⁹ Zum Beispiel bietet die Plattform Pandora Musikempfehlungen an, die auf den individuellen Musikgeschmack der Nutzer abgestimmt sind. Die Empfehlungen werden auf Basis von vorab intellektuell erzeugten Metadaten generiert, welche aus „The Music Genome Project“ stammen.

¹⁰ Ein Überblick über aktuelle Forschungsprojekte ist unter [37] zu finden.

beispielsweise in einer Fotodatenbank eine Suche nach Aufnahmen des Eiffelturms zu ermöglichen.

Voraussetzung sind hierfür Metadaten auf Basis von Merkmalen, welche die semantische Lücke [38] [39] zwischen einer Beschreibung des Mediums in Binärdarstellung und dessen Beschreibung durch einen Menschen („dies ist eine Aufnahme des Eiffelturms“) verringern¹¹. Daher wird bei der Berechnung von Features in der Regel versucht, Sinnesorgane und menschliche Wahrnehmung nachzuahmen. Das automatische Erkennen von Informationen in visuellen Medien ist jedoch bis heute weder gelöst noch abschließend erforscht, es gilt als eine der größten Herausforderungen der Informationsverarbeitung [40]. In den aktuellen Ansätzen kommen Verfahren aus dem Bereich Mustererkennung zum Einsatz, welche in Medienobjekten Regelmäßigkeiten, Wiederholungen, Ähnlichkeiten oder Gesetzmäßigkeiten erkennen können.

Um relevante Metadaten für eine inhaltsbasierte Suche automatisch zu erzeugen, müssen zunächst aus Multimediaobjekten bestimmte Merkmale maschinell extrahiert werden. Bei einer Audiodatei können das zum Beispiel Größen mit geringem semantischem Gehalt wie Maximalpegel oder Signal-Rauschverhältnis sein. Es gibt allerdings auch Verfahren, die Metadaten mit höherem semantischem Abstraktionsgrad erzeugen und beispielsweise die Stimmung eines Musikstücks oder einer Fotografie untersuchen. Für jeden Medientyp gibt es eine Vielzahl von Extraktionsverfahren, welche Metadaten mit unterschiedlichem Abstraktionsgrad bzw. semantischem Gehalt erzeugen [41]. In Abhängigkeit des semantischen Gehalts werden die extrahierten Merkmale in Low-, Mid- und High-Level-Features (siehe Abschnitt 3.3) unterteilt. Dies ist allerdings keine streng wissenschaftliche Einordnung, sondern ein in der Fachgemeinde gewachsener Sprachgebrauch. Damit einhergehend vergrößert sich die semantische Lücke zur Repräsentation des Mediums in Binärdarstellung.

Verfahren zur Extraktion von Merkmalen mit niedrigem semantischem Gehalt sind häufig statistische Messgrößen (z. B. vorherrschender Farbwert bzw. Färbung eines Bildes). Dies ist abzugrenzen von Verfahren zum Auslesen technischer Metadaten, welche textuelle Bereiche einer Multimediadatei untersuchen.

Computergestützte Verfahren zur Extraktion von Merkmalen mit hohem semantischem Gehalt zielen darauf ab, multimediale Inhalte „wie ein Mensch“ zu verstehen und zu beschreiben [41] [42]. Um Maschinen in diese Lage zu versetzen, müssten Verfahren entwickelt werden, die bestimmte Aspekte menschlicher Wahrnehmung nachahmen können – was bislang aber nur innerhalb abgeschlossener Domänen befriedigende Ergebnisse erzielt. Zum Beispiel würden Menschen ein Bild mit Früchten auf einem Tisch und in Schalen in der Regel als Stillleben erkennen und beschreiben. In gleicher Weise können Maschinen durch Lernverfahren in die Lage versetzt werden, diese spezifischen Merkmale im Bild als „Stillleben“ zu identifizieren. Nutzer wären nun in der Lage in einer Bilddatenbank nach dem Merkmal „Stillleben“ zu suchen, ohne dass die Ressource zeitaufwändig manuell erschlossen werden muss. Allerdings setzt dieses Verfahren eine entsprechende Menge annotierter Trainingsdaten voraus – andere Merkmale des Konzepts „Stillleben“ können ohne vorhergehendes Lernen nicht maschinell erkannt werden. Dies ist mit aktuellen Verfahren nur auf einem sehr kleinen Bilddatenbestand möglich. Darüber hinaus ist die Genauigkeit dieser Verfahren zu gering und für einen Einsatz in Multimediaarchiven noch nicht geeignet.

¹¹ Die verbleibende Kluft zwischen der Semantik von extrahierten Merkmalen (Features) und der menschlichen Wahrnehmung von Inhalt oder Ähnlichkeit wird häufig als Semantische Lücke bezeichnet.

Allerdings besteht selbst unter Annahme von optimalen Merkmalsextraktionsverfahren ein weiteres Problem. Um bei obigem Beispiel zu bleiben, können Maschinen (ebenso wie Menschen) ein digitales Bild sowohl als „Ölgemälde“ als auch als „Kunstwerk“ beschreiben, da es bei der Extraktion von Merkmalen mit hohem semantischem Gehalt viele gleichberechtigte Deutungsmöglichkeiten gibt. Darüber hinaus ist die Extraktion mancher Merkmale mit hohem semantischem Gehalt aus Multimediaobjekten subjektiv, zum Beispiel ob das Foto einer Landschaft „künstlerisch wertvoll“ ist oder nicht. Diese prinzipiell vorhandene Ambiguität von Ressourcen und Subjektivität bei der Extraktion von Semantik hat zur Folge, dass Merkmalsextraktionsverfahren mit hohem semantischem Gehalt nur in sehr spezifischen Kontexten genutzt werden können.

Maschinengestützte Verfahren zur Analyse unstrukturierter Inhalte waren zu Anfang stark von Methodiken des Fachgebiets Maschinelles Sehen (Computer Vision [43]) beeinflusst. Maschinelles Sehen wiederum entstand zeitgleich mit mathematischen Modellen von Sinnesorganen (z. B. visuelle Systeme von Säugetieren) in den 1950er und 1960er Jahren [44] [45]. Aufgrund dessen ist es nicht überraschend, dass einige Methoden, welche in den frühen 1980ern für Maschinelles Sehen entwickelt wurden, auch für die Erzeugung von Low-Level-Features verwendet werden (z. B. Kantendetektion [46] [47] [48]).

Mittlerweile ist es möglich, eine Vielzahl an Features aus Medien zu extrahieren und sinnvoll für eine Suche einzusetzen. Zum Beispiel ermöglicht Google Images [49] eine Suche nach Bildfärbung, einem aus Internetbildern extrahierten Low-Level-Feature. Weiterhin analysiert die Suchmaschine blinkx [50] Videos aus dem Internet nach einer Vielzahl von Merkmalen wie beispielsweise Szenenwechsel oder gesprochene Schlüsselwörter [51]. An dieser Stelle ist anzumerken, dass eine kulturaffine Suche möglicherweise – aufgrund abweichender Nutzeranforderungen – die Extraktion anderer Merkmale aus Multimediaressourcen erfordert.

Doch bevor ein möglicher Einsatz von Automatismen zur inhaltlichen Erschließung in Multimediaarchiven von Kultureinrichtungen diskutiert wird, soll zunächst eine zweckdienliche Charakterisierung der Multimediasuche auf Basis automatischer inhaltlicher Erschließung vorgenommen werden.

3.3. Charakterisierung und Abgrenzung der Multimediasuche auf Basis automatischer inhaltlicher Erschließung

Eine *Multimediasuche auf Basis automatischer inhaltlicher Erschließung* verfügt über folgende Charakteristika:

- **Medienformate mit großen unstrukturierten Datenabschnitten.** Der Nutzer sucht nach Inhalten einer Kultureinrichtung, die nicht ausschließlich aus Textdokumenten bestehen. Basis der Suche, im Gegensatz zu einer text- oder schlüsselwortbasierten Suche nach Textdokumenten, sind nicht die textuellen Anteile der Multimediaobjekte, sondern deren unstrukturierte Binärteile, welche durch geeignete Verfahren analysiert werden. Die Voraussetzung hierfür ist, dass ein Multimediaobjekt durch ein automatisches Verfahren in sinnvoller Weise analysiert und beschrieben werden kann¹². Darüber hinaus kann eine medienübergreifende Multimediasuche durchgeführt werden, deren Ergebnisse mehrere Medientypen, zum Beispiel Video-, Audio- und Bildobjekte, beinhalten können.

¹² So können nach heutigem Stand der Forschung beispielsweise verschlüsselte Multimediaobjekte durch automatische Verfahren ohne vorige Dekodierung nicht in sinnvoller Weise analysiert werden.

- **Ein automatisierter inhaltlicher Erschließungsschritt.**
- **Textbasierte Suche nach Inhalten.** Mittels automatischer Erschließung ist es möglich, auch in Medien, die zunächst keinen maschinenlesbaren Text enthalten, nach Strings oder Schlüsselbegriffen zu suchen (z. B. Videos und Sprachaufnahmen). Das oben genannte Beispiel der OCR liefert beispielsweise aus texthaltigen, gescannten, analogen Bildvorlagen maschinenlesbaren Text, der mittels einer Volltextsuchmaschine gefunden werden kann. Dieser Text kann darauf folgend noch auf weitere Merkmale untersucht werden. Zum Beispiel können, ebenfalls automatisiert, Textpassagen bestimmten Schlüsselbegriffen zugeordnet und verschlagwortet werden. Analog kann gesprochener Text in Videos und Sprachaufnahmen durch Speech-to-Text-Verfahren verarbeitet werden. Bekannte Beispiele für Suchmaschinen auf Basis von automatisch extrahiertem Volltext sind die Buchsuchmaschinen von Amazon (mit der Funktion „Search Inside“) [52], Google Book Search [53] sowie das Project Gutenberg [54] [55].
- **Inhaltsbasierte Suche.** Unter einer inhaltsbasierten Suche [57] [58] [56] versteht man die Abfrage von Multimediainhalten aus einer Datenbank auf Basis extrahierter Merkmale im Unterschied zu einer schlüsselwortbasierten Suche. Bei einer inhaltsbasierten Suche wird die Multimediadasammlung auf Basis von Merkmalen (Foto mit blauer Bildfärbung) oder Beispielobjekten (z. B. Suche ein Foto, das ähnlich zum Beispielfoto ist) durchgeführt. Hierbei werden die gesuchten Merkmale mit Merkmalen von Objekten der Multimediadasammlung verglichen. Je nach Suchverfahren bzw. Filtereinstellung können auf diese Weise Objekte in der Multimediadasammlung identifiziert werden, die zum Beispiel eine Ähnlichkeit oder Gegensätzlichkeit zum Beispielobjekt besitzen. Ein bekanntes Beispiel für Content Based Retrieval ist das sogenannte Query-by-Humming [59] [60]: Ein Nutzer singt eine Melodie eines Liedes in ein Mikrofon, die Melodie wird als Merkmal extrahiert und eine Musiksammlung wird nach Musikstücken mit gleicher oder ähnlicher Melodie durchsucht. Ein weiteres Beispiel ist die Suche nach Bildern mit ähnlichen Merkmalen (Content Based Image Retrieval). Ein Nutzer besitzt ein Foto mit einem Sonnenuntergang, dessen Merkmale mithilfe automatischer Inhaltsanalyseverfahren extrahiert werden: dies könnten zum Beispiel eine rote Bildfärbung und graduelle Farbverläufe sein. Auf diese Weise wird die Bilddatenbank nach Fotos mit roter Bildfärbung und graduellen Farbverläufen durchsucht und möglicherweise andere Fotos mit Sonnenuntergängen gefunden. Es ist hierbei möglich, dass der Nutzer gewichten kann, welches Merkmal er für seine inhaltsbasierte Ähnlichkeitssuche nutzen möchte (rote Bildfärbung oder gradueller Farbverlauf).
- **Suche auf Basis von High-Level-Features.** Auf Basis geeigneter automatisch extrahierter Merkmale mit hohem semantischem Gehalt (High-Level-Features) kann eine Suche auf Basis von Entitäten (z. B. Person, Ort, Organisation oder Werk) oder semantischer Konzepte („romantischer Sonnenuntergang“, „humorvoller Text“) durchgeführt werden. In vielen Fällen müssen zur Extraktion von High-Level-Features Algorithmen zunächst mithilfe von annotierten Datensätzen trainiert werden [61] (überwachtes Lernen). Das bedeutet, ein Algorithmus lernt aus einer Grundgesamtheit von Objekten, die eine Gemeinsamkeit besitzen (z. B. ein Testset aus Bildern einer Fotosammlung, welche von Nutzern als „romantisch“ beschrieben wurden) oder eine Entität (z. B.

Eiffelturm) beinhalten, gemeinsame Merkmale zu extrahieren. Auf Basis dieser gemeinsamen Merkmale ist ein Algorithmus in der Lage, in einer nicht annotierten Sammlung Objekte zu identifizieren, die ein spezifisches semantisches Konzept oder eine bestimmte Entität enthalten.


3.4. Beispiele für eine Suche nach Bildern mit und ohne Text auf Basis automatischer Inhaltsanalyse

Da es eine Vielzahl von Methodologien und Verfahren gibt, kann im Folgenden nur ein kleiner Ausschnitt der Möglichkeiten zur automatischen Inhaltsanalyse von Bildern wiedergegeben werden.

Im Falle von digitalisierten Büchern, Zeitungen und Zeitschriften ist es technisch möglich, Artikel, Inhaltsverzeichnisse, Kapitel oder Abbildungen weitestgehend automatisiert zu extrahieren. Mithilfe dieser gewonnenen Strukturmetadaten können sehr effizient Einstiegspunkte erzeugt werden, die ein Nutzer direkt an die gesuchte Stelle des Werks leitet. Wie bereits in Abschnitt 1 dargestellt wurde, können bestehende Bibliothekskataloge mit diesen Strukturmetadaten angereichert werden, um eine inhaltsbasierte und zielgerichtete Suche zu ermöglichen.

Es ist darüber hinaus möglich zu erkennen, ob es sich bei dem Bild um ein Foto, eine Zeichnung oder um eine Grafik handelt. Mithilfe dieser automatisch extrahierten Merkmale können bereits heute Nutzer von Google Images ihre Suchergebnisse filtern.

Es gibt darüber hinaus Verfahren zur Logoerkennung, die Hinweise auf Entitäten, zum Beispiel bestimmte Produkte oder Organisationen, im Bild liefern. Diese Verfahren sind zwar marktreif, funktionieren aber nur mit einer eingeschränkten Menge von Logos, mit denen das Verfahren zuvor trainiert wurde. Die Verfahren sind nicht in der Lage, das Merkmal „Logo“ an sich in einem Bild zu erkennen. Weiterhin kommt Logoerkennung vor allem im Industriebereich zum Einsatz, zum Beispiel bei automatischen Sortiermaschinen von Postsendungen oder bei der Erkennung von Pfandflaschen.

Aus sämtlichen digitalen Bildinhalten mit Text in Form von Rastergrafik können mittels eines OCR-Verfahrens maschinenlesbare Texte extrahiert werden. Dieser kann wiederum für eine Entitätenextraktion sowie Klassifikation des Bildes bei einer Suche genutzt werden. In Abhängigkeit der Qualität des Digitalisats und des OCR-Verfahrens wäre eine automatische Anreicherung von Bildern um Entitäten theoretisch möglich. So könnten beispielsweise digitale Fotos mit Textgehalt (siehe ) automatisch verschlagwortet oder entsprechenden Konzepten („School“, „Chemistry“) zugeordnet werden.

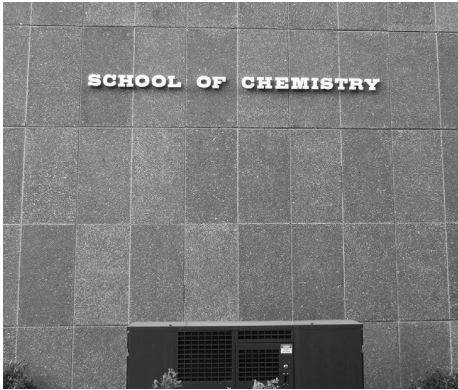


Abbildung 3: Foto mit Textgehalt

Auch wenn Dienste wie Google Goggles [62] Texterkennung für Mobiltelefone anbieten, ist es nach heutigem Stand der Technik immer noch schwierig, mit hoher Verlässlichkeit in einer großen abwechslungsreichen Fotosammlung Text in unbekannter Sprache und Schriftart vor beliebigem Hintergrund zu erkennen. Es kann allerdings davon ausgegangen werden, dass mit dem Aufkommen neuer, robuster Verfahren diese Möglichkeit zur automatischen Erschließung großer Fotosammlungen genutzt werden kann.

Mittels Gesichtserkennungs- und allgemeiner Entitätenerkennungsverfahren können Personen und Entitäten (z. B. Eiffelturm) in einem Bild erkannt werden. Da Kultureinrichtungen bereits über Datensätze zu Entitäten [25] [63], insbesondere Personen, verfügen, könnten diese Verfahren ein nicht erschlossenes Bild automatisch mit einem entsprechenden Datensatz vernetzen¹³. Zum Beispiel können Fotos, auf welchen Sehenswürdigkeiten wie der Eiffelturm abgebildet sind, mit einem entsprechenden Normdatensatz verbunden werden. Sofern Personen auf dem Foto abgebildet sind, könnten diese mit Katalogdatensätzen wie der PND [66] verlinkt werden, wenn Verfahren zur Personenidentifikation entsprechend mit der Größe des Bestands skalieren. Allerdings werden durch die automatisierte Identifikation von Personen und die Bereitstellung dieses Wissens im Internet etliche Fragen bezüglich des Schutzes der Privatsphäre aufgeworfen, welche gesondert geklärt werden müssen.

Auch wenn beispielsweise Google Goggles unter anderem eine Erkennung von Sehenswürdigkeiten und Kunstwerken anbietet (welches die Datenbasis einschränkt, indem es zusätzlich GPS-Daten des Mobiltelefons einbezieht), kann davon ausgegangen werden, dass die genannten Verfahren bisher nicht die nötige Reife für die Verwendung in großen Multimediensammlungen besitzen [67] [68].

Selbige Einschränkung gilt in noch größerem Maße für Verfahren zur Extraktion von High-Level-Features, welche Bildern semantische Konzepte (z. B. Sport, Freizeit) und Stimmungen zuordnen können. Die Güte dieser Bildanalyseverfahren ist stark von dem jeweiligen extrahierten High-Level-Feature und der Bildsammlung abhängig. Es ist ebenso möglich, ein extrahiertes semantisches Konzept mit einem entsprechenden Schlagwort (z. B. aus der Schlagwortnormdatei [69]) des Katalogs zu vernetzen.

¹³ Ein ähnlicher, vielbeachteter Ansatz ist die Vernetzung von Bildern mit Ontologien und kontrollierten Vokabularen auf Basis von Features. [64] [65]

3.5. Beispiele für eine Audiosuche auf Basis automatischer Inhaltsanalyse

Ebenso wie bei der automatischen Analyse von Bildinhalten gibt es bei Audiomedien zahlreiche Verfahren zur Inhaltsanalyse, welche Merkmale mit unterschiedlichem semantischem Gehalt identifizieren können. Mit wachsendem semantischem Gehalt sinkt tendenziell die Güte der Verfahren bei großen Sammlungen mit verschiedenartigen Audioinhalten (E-Musik, U-Musik, Sprachaufnahmen, Natur- und Tiergeräusche, Klangkunst usw.). Verfahren zur automatischen Inhaltsanalyse von digitalen Audioinhalten können in der Regel auch auf die Tonspur von audiovisuellen Medien angewendet werden.

Sofern der digitale Audiomitschnitt Musikpassagen, Sprache und auch nicht-sprachliche Bereiche wie Geräusche enthält, ist es möglich, mittels automatischer Verfahren die entsprechenden Übergänge zu erkennen und mit Einstiegspunkten (Marker) zu versehen. Auf diese Weise gelangt der Nutzer direkt zu seinem Suchergebnis in der Audioaufnahme.

Aus digitalen Audioaufnahmen mit Sprachinhalt können mittels Speech-to-Text- oder Speech-to-Syllable-Verfahren Text- respektive Silbentranskripte erstellt werden. Es ist weiterhin möglich, verschiedene Sprecher zu unterscheiden und das Geschlecht des jeweiligen Sprechers zu identifizieren. Die Qualität der Transkripte und die Erkennung von Sprechern sowie deren Geschlecht ist allerdings stark von den Aufnahmebedingungen und der Klangqualität abhängig. In der Regel ist die Qualität der Sprachanalyseverfahren für Studioaufnahmen gut (mit abwechselnd redenden Sprechern ohne Hall oder Echo), bei sogenannten Field Recordings (mit Störgeräuschen wie beispielsweise Wind oder Hintergrundmusik) können nach heutigem Stand der Technik gesprochene Sätze nur bruchstückhaft erkannt werden. Bei historischen Aufnahmen mit Störgeräuschen wie zum Beispiel Rauschen oder Knacksen ist die Erkennungsrate ebenfalls niedrig. Speziell aus den Transkripten können mithilfe semantischer Textanalyseverfahren, Entitäten und Konzepte extrahiert werden. Analog zur automatischen Analyse von Bildinhalten können diese Entitäten mit entsprechenden Entitäten in Bibliothekskatalogdatensätzen, sofern vorhanden, verlinkt werden.

Aus *Musikinhalten* können Merkmale wie beispielsweise musikalische Parameter wie Rhythmus, Tonart, Melodie extrahiert werden, psychoakustische Größen wie Lautheit, Impulshaltigkeit usw. oder subjektive Stimmungsparameter wie „rockig“, „entspannt“. Auf diese Weise kann der Nutzer eines Musik- oder Lautarchivs nach Merkmalen suchen, die bisher auf Basis einer regelkonformen Katalogisierung nur wenig oder kaum berücksichtigt wurden.

3.6. Chancen und Herausforderungen bei der Nutzung von Verfahren zur automatischen Inhaltsanalyse in Kultureinrichtungen

Welche Verfahren für eine spezielle Einrichtung möglicherweise infrage kommen, hängt von einer Vielzahl von Faktoren ab. Eine wesentliche Rolle bei der Auswahl der Verfahren spielt die Art der Sammlung, also ob die Multimediaobjekte einer speziellen Wissensdomäne angehören oder thematisch nicht zusammenhängen. Weiterhin sollten die Anforderungen der Nutzer berücksichtigt werden, um zu klären, welche Such- und Filtermöglichkeiten nützlich sind und welche nicht. Viele Produkte [70] im Bereich automatischer Verfahren zur automatischen Inhaltsanalyse haben Anforderungen an die technische und personelle Infrastruktur, welche nicht von allen Kultureinrichtungen erfüllt werden können. Beispielsweise nutzen zahlreiche Kultureinrichtungen proprie-

täre Bibliothekssoftware, welche die Integration von kommerziellen Lösungen für Multimediaanalyse schwierig gestaltet.

Eine weitere Markteintrittshürde für Verfahren zur automatischen Inhaltserschließung ist die Integration der neuen Technologien in bestehende Geschäfts- und Erschließungsprozesse von Kultureinrichtungen. Bisherige Regelwerke zur Erschließung berücksichtigen den Einsatz von Verfahren zur automatischen Inhaltserschließung nicht und somit sind diese Technologien, sofern sie im Rahmen von Kataloganreicherungsgeschäftsgängen genutzt werden, ein Zusatz, aber kein integraler Bestandteil bibliothekarischer Erschließung.

Insbesondere die Nutzung von Verfahren zur Extraktion von Merkmalen mit hohem semantischem Gehalt skalieren nicht mit der Größe der Multimediasammlungen, sondern können nur auf einer eingeschränkten Wissensdomäne eingesetzt werden. Beispielsweise eröffnet theoretisch die automatische Extraktion von semantischen Konzepten viele neue Suchmöglichkeiten. In der Praxis können die meisten Verfahren, welche Inhalte mit hohem semantischem Gehalt extrahieren nicht auf große Sammlungen angewendet werden.

Zusammengefasst kann die katalogbasierte Multimediasuche in Kultureinrichtungen mittels Verfahren zur automatischen Inhaltsanalyse generierter Metadaten um viele neue Such- und Filtermöglichkeiten erweitert werden. Hierbei ist anzumerken, dass Automatismen eine Erschließung durch Experten in Kultureinrichtungen keinesfalls ersetzen können; Unter anderem, weil die Qualität der erzeugten Metadaten mit wachsendem semantischem Abstraktionsgrad sinkt und viele Verfahren zunächst aufwändig trainiert werden müssen (überwachtes Lernen), um zuverlässig Ergebnisse in der gewünschten Qualität zu liefern. Allerdings ist es vorstellbar, dass durch die intensive und konsequente Nutzung von automatischen Verfahren Personalressourcen frei werden, um bibliothekarische Erschließungsgeschäftsgänge stärker auf die Verbesserung der Informationsqualität von Katalogmetadaten, insbesondere Normdaten, auszurichten. Weiterhin ist es möglich, bestehende Erschließungsprozesse effizienter zu gestalten. Dies ist für Kultureinrichtungen von Vorteil, die eine große Zahl von unstrukturierten Multimediaressourcen effizient verarbeiten und durchsuchbar machen müssen.

4. Ausblick Semantische Multimediasuche – Perspektiven für die Nutzung automatischer Medienanalyse und von Semantic-Web-Technologien

Grundlage einer semantischen Multimediasuche ist die vorangegangene und in Abschnitt 3 bereits beschriebene inhaltliche Erschließung über geeignete Multimediaanalysetechniken. Ergebnis dieser inhaltlichen Erschließung sind in der Regel textuelle, deskriptive Metadaten, deren Zuordnung temporal (für zeitabhängige Medientypen Video oder Audio), spatial (für Medien mit einem räumlichen Bezug, d. h. Text oder Bild) oder spatiotemporal (d. h. zeit- und raumbezogen, wie z. B. in Video) erfolgt. Um dies zu gewährleisten, geht der inhaltlichen Erschließung zunächst eine strukturelle Analyse der Multimediaressourcen voran, die eine zusammenhängende Ressource in kohärente Unterstrukturen zerlegt, die separat inhaltlich analysiert und annotiert werden können.

Liegen deskriptive, textuelle Metadaten zeit- und ortsbezogen zu den jeweiligen Multimediaressourcen vor, müssen diese mithilfe geeigneter formaler Wissensrepräsentationen kodiert werden, damit sie sowohl untereinander als auch mit bereits vorhandenen semantischen Entitäten in Bezug gesetzt werden können. Erst wenn hinter den textuellen Metadaten semantische Entitäten stehen, die den betreffenden Inhalt eindeutig (formal) repräsentieren, ist eine zielgenaue semantische Suche möglich. Im Fol-

genden soll die Integration multimedialer Inhalte in eine formale Wissensrepräsentation auf Basis von Katalogdaten und automatischer Multimediaanalyse anhand von grundlegenden Beispielen aufgezeigt werden.

4.1. Integration multimedialer Inhalte in eine formale Wissensrepräsentation auf Basis von Katalogmetadaten und automatischer Multimediaanalyse

Im folgenden Abschnitt wird beispielhaft die semantische Annotation multimedialer Inhalte mithilfe formaler Wissensrepräsentationen auf der Basis zugrunde liegender Katalogmetadaten und Medienanalysedaten an verschiedenen Medientypen gezeigt. Dieser Abschnitt zielt darauf ab, die Perspektiven für Multimediasuche in Kultureinrichtungen durch eine Vereinigung der in den Abschnitten 1 bis 3 vorgestellten Ansätze aufzuzeigen. Wo es sinnvoll erschien, wurden Möglichkeiten kollaborativer Ansätze erwähnt.

4.1.1. Audio (Musik)

Eine Schallplatte aus einem Musikarchiv soll semantisch annotiert werden. Der Prozess startet mit der Digitalisierung des Quellmaterials, d. h. der Abtastung der Schallplatte nebst Anfertigung eines Scans des Plattencovers. Die Digitalisate selbst werden Gegenstand verschiedener automatischer Analyseverfahren, wie z. B. OCR, angewandt auf das Plattenlabel und das Plattencover (siehe auch Abschnitt 3.4). Dadurch können automatisch Informationen zu Titel, Komponist, darbietendem Künstler, etc. gewonnen werden.

Anschließend werden deskriptive, textuelle Metadaten mit geeigneten semantischen Entitäten verbunden (Named Entity Recognition). In unserem Fall können dazu Katalogmetadaten, wie z. B. die PND, GKD und die SWD herangezogen werden. Der Komponisten- oder Künstlernamen kann mit der PND abgeglichen werden und ein bestehender Plattenvertrag mit einer Produktionsfirma oder einem Verlag erlaubt die Verknüpfung mit der GKD. Weitere erkannte Begriffe aus dem Titel oder dem Begleittext zum Stück können mit der SWD verlinkt werden. Dabei werden die genannten Normdaten als formale Wissensrepräsentation, z. B. als Linked-Data-Ressource bereitgehalten. Oft ist dabei eine einfache, auf Textvergleich basierende Zuordnung möglich. Allerdings entstehen bei Namensgleichheit Mehrdeutigkeiten, die in einem anschließenden Disambiguierungsschritt aufgelöst werden müssen. Zur Disambiguierung werden weitere Metadaten aus dem direkten Kontext der zu verknüpfenden Textdaten herangezogen (d. h. der umgebende Text und die darin erkannten Entitäten), um so auf statistischer Basis eine Entscheidung zu treffen.

Neben der automatischen Analyse können auch Experten oder einfache Benutzer Metadaten zu den digitalisierten Medien beisteuern. Während Experten aufgrund ihrer Fachkenntnis autoritative Metadaten vergeben, die für eine gezielte Informationssuche von besonderer Bedeutung sind, können Benutzer auch persönliche Metadaten in Form von einfachen Schlagworten (Tags) oder Kommentaren beisteuern (nicht-autoritative Metadaten), die für eine anschließende personalisierte Suche verwendet werden können. Eine semantische Verknüpfung mit geeigneten Entitäten kann dabei entweder wie im Falle der automatisiert gewonnenen Metadaten erfolgen, oder aber der Benutzer wählt bei der Eingabe seiner Metadaten bereits simultan eine oder mehrere semantische Entitäten mit aus. Unabhängig von der semantischen Analyse kann ein digitalisiertes Musikstück auch mit weiteren Stücken verlinkt werden, die ähnlich klingen oder weitere ähnliche Eigenschaften aufweisen, um eine ähnlichkeitsbasierte Suche zu ermöglichen.

Abschließend werden die vorhandenen Inhalte durch semantische Multimediale Suchinterfaces den Benutzern zugänglich gemacht.

4.1.2. Audio (Sprache)

Ebenso wie Musikaufnahmen müssen archivierte Sprachaufnahmen zuerst digitalisiert werden, bevor sie einer automatischen Analyse zugänglich gemacht werden können. Eventuell vorhandenes Begleitmaterial wird ebenso wie das Äußere des Aufzeichnungsmediums digital (fotografisch oder per Scan) erfasst. Um eine Sprachaufnahme einer automatischen Analyse unterziehen zu können, müssen zuerst qualitätsverbessernde Maßnahmen (Filterung) getroffen werden, mit denen Hintergrund- und Störgeräusche weitgehend entfernt werden, um eine hinreichend gute Erkennungsqualität bei möglichst hoher Verlässlichkeit gewährleisten zu können. Die Audiodatei wird dann automatisch in inhaltlich kohärente kleinere Teilabschnitte segmentiert, die durch Intensitätsabfälle, wie z. B. Sprechpausen, voneinander getrennt sind. Diese Segmente werden dann automatisch mit maschinengenerierten Metadaten versehen, die beschreiben, ob das Segment Sprachinhalte enthält, welches Geschlecht der Sprecher hat, oder wer der Sprecher ist. Segmente, die Sprache enthalten, werden, wie in Abschnitt 3.5 beschrieben, mit Speech-to-Text- bzw. Speech-to-Syllables-Verfahren analysiert. Mithilfe der vorhandenen Metadaten (Sprecher, Geschlecht, etc.) wird diese Analyse genau auf die Quellen abgestimmt. Liegt eine Referenzdatenbank vor, besteht die Möglichkeit, bedeutende Sprecher zu erkennen und diese automatisch mit der PND zu verlinken. Soweit möglich werden auch wichtige Entitäten (Orte, Körperschaften, etc.) erkannt und entsprechenden Normdaten wie z. B. SWD (Personen, Körperschaften, Orte) oder GKD/KND (Körperschaften) zugeordnet. Da eine automatische Texterkennung auf der Basis heutiger Technologien nicht fehlerfrei arbeitet, könnten die erkannten Metadaten im Rahmen einer kollaborativen Erschließung von Expertengruppen korrigiert und erweitert werden. Insbesondere kann auf diese Weise ein vom System als „unbekannt“ bezeichneter Sprecher von Experten identifiziert und korrekt benannt werden. Die erzeugten Wort- und Silbenrepräsentationen der Quelle können nun mithilfe eines Suchalgorithmus auf Grundlage von Fuzzy-Logic-Verfahren von Nutzern durchsucht werden. Diese Suchen liefern die originalen (oder umformatierten) Audiodateien und nutzen dabei die vorhandenen Metadaten bzw. den transkribierten Text, um die Ergebnisse in einer verständlichen und leicht zu navigierenden Form anzuzeigen.

4.1.3. Bild mit Text

Printmedien, wie z. B. Bücher, Dokumente und Zeitungen aus Archiven und Bibliotheken müssen zunächst mithilfe eines optischen Scanners digitalisiert werden, bevor eine automatische Analyse starten kann. Das vom Scanner gelieferte Bildmaterial wird zunächst einer Bildqualitätsoptimierung unterzogen. Dieser Schritt hat zwei Ziele: zum einen die möglichst originalgetreue Wiedergabe der Quelle für menschliche Leser und zum anderen die Optimierung der Bilddatei für optische Texterkennungsverfahren (OCR). Die Metadatengenerierung unterteilt sich in zwei separate Schritte: Zunächst wird jede Seite entsprechend ihres Layouts segmentiert. Bei Zeitungen müssen dabei Artikel, die sich über mehrere Spalten erstrecken, anhand ihrer Überschrift und des umgebenden Leerraums erkannt werden. Der gedruckte Text selbst wird durch ein OCR-Verfahren erkannt. Auch wichtige Strukturen, wie etwa Abbildungen und Inhaltsverzeichnisse werden so unterschieden. Danach werden aus dem erkannten Text wichtige Entitäten, also z. B. Personen, Orte oder Körperschaften sowie Schlagworte extrahiert

und mit korrespondierenden semantischen Entitäten aus PND, SWD und GKD/KNL mithilfe von Data-Mining-Verfahren verlinkt (autoritative Metadaten). Darauf aufbauend können die gescannten Medien mit anderen Wissensbasen (z. B. Wikipedia) oder semantischen Lexika (z. B. Wordnet, Germanet) verlinkt werden. Dem so entstandenen Wissensnetz können anschließend Benutzer eigene Inhalte und Metadaten, wie z. B. beschreibende Tags hinzufügen (nicht-autoritative Metadaten). Autoritative und nicht-autoritative Metadaten müssen dabei stets separat behandelt werden. So können etwa Zeitzeugen mit ihrem Detailwissen über Personen und Ereignisse wertvolle Beiträge zum Wissensbestand des Systems leisten, ohne eventuell die hohe Qualität und Integrität der bestehenden Faktenbasis durch weniger gesichertes Wissen zu beeinträchtigen.

Eine semantische Suche über die gescannten Textdatenbestände kann dann sowohl über eine volltextbasierte Suche als auch über eine medienspezifische, spezialisierte Textsuche erfolgen. So kann allgemein in Printmedien – bei entsprechender Strukturerschließung – nach Standardtext, Bildüberschriften, Bildunterschriften oder in Legenden gesucht werden. Im speziellen Fall von Büchern beinhaltet dies die Suche in Inhaltsverzeichnis, speziellen Kapiteln und Vor- und Nachworten. Für Zeitungen kann ein Stöbern in den enthaltenen Bildern und Abbildungen angeboten werden sowie eine Suche in Artikeln und in Extraktionen einzelner Artikel.

4.1.4. Bild (ohne Text)

Zunächst muss das Bildmedium (z. B. Druck, Fotografie, Gemälde oder Lichtbild) mithilfe eines geeigneten optischen Abtastgeräts erfasst und digitalisiert werden. Die eigentliche Medienanalyse und Metadatengenerierung unterscheidet dabei zunächst die Extraktion sogenannter Low-Level-Merkmale, wie z. B. Helligkeits- und Farbverteilung oder Kanten- und Kantenrichtungsverteilung. Diese können für das Gesamtbild bzw. aus einzelnen für das Bild signifikanten Regionen ermittelt werden. Low-Level-Merkmale dienen direkt der Kategorisierung und dem Clustering der Bildinformation über maschinelle Lernverfahren und ermöglichen so die Implementierung einer Suche auf Basis der visuellen Ähnlichkeit der Bilder untereinander (siehe auch Abschnitt 3.4). Durch die Extraktion von High-Level-Merkmalen werden Entitäten erkannt (z. B. Personen, Orte, Körperschaften, Schlagwörter) und Metadaten für die Suche aus deren semantischen Beschreibungen generiert. Erkannte Entitäten werden automatisch mit entsprechenden Normdaten (PND, GKD, SWD) verbunden. Durch diese Verbindungen können die Inhalte zu anderen Wissensquellen (z. B. Wikipedia oder Brockhaus) oder zu semantischen Lexika (Wordnet, Germanet) ermittelt werden. Zusätzlich können Benutzer eigene Inhalte und Metadaten zu den jeweiligen erfassten Medien hinzufügen, um auf diese Weise die automatisch ermittelten Metadaten zu ergänzen bzw. zu korrigieren. Eine semantische Suche in den derart erschlossenen und annotierten Bilddaten bietet neben der direkten Suche nach Schlüsselwörtern, die auf den annotierten Entitäten abgebildet werden, zusätzlich die Möglichkeit, zu einem gegebenen Bild möglichst (visuell) ähnliche Bilder zu ermitteln (Query by Example).

4.1.5. Audiovisuelle Daten

Audiovisuelle Daten aus Multimediaarchiven können analog als optisch belichtetes Filmmaterial oder als elektronisch auf Magnetband aufgezeichnetes Videomaterial vorliegen. Beide Varianten müssen mit unterschiedlichen Abtastverfahren digital erfasst werden. Der Video-Stream wird zunächst automatisch in inhaltlich kohärente Segmente, d. h. in Kapitel, Szenen, Shots und Subshots hierarchisch unterteilt. Ein

Kapitel enthält mehrere inhaltlich zusammenhängende Szenen. Eine Szene besteht aus mehreren inhaltlich zusammenhängenden Shots. Ein Shot bezieht sich auf eine zusammenhängende Folge von Einzelbildern, deren Anfang und Ende durch einen Schnitt oder eine andere Art der Überblendung gekennzeichnet sind. Subshots sind Teile von Shots, in denen bestimmte (inhaltliche) Objekte im Bild zu sehen sind und durch Eintritt bzw. Verlassen des Bildbereichs durch das Objekt gekennzeichnet sind. Alternativ kann auch eine Segmentierung anhand der Audiospur der Videodaten vorgenommen werden. Diese kann dann wieder intensitätsgesteuert bzw. auf Grundlage der Identifikation von Sprechern erfolgen. Technische Metadaten, wie z. B. Zeitindizes, Zeitdauern, (Band- oder Film-)Länge des korrespondierenden Originalmaterials werden automatisch ermittelt. Besonders interessante Bildregionen (z. B. mit Personen) werden aus den Einzelbildern des Videostreams in gleicher Weise wie bei den zuvor beschriebenen Bilddaten extrahiert und für die Suche und Archivierung relevante Low-Level-Merkmale (Helligkeits- und Farbverteilung sowie weitere bildbeschreibende Merkmale) werden berechnet. Durch die Extraktion von High-Level-Merkmalen – hier sowohl aus der vorhandenen Bildinformation als auch aus der Audioinformation) werden Entitäten (z. B. Personen, Orte, Körperschaften) automatisch erkannt. Zusätzlich können Einzelbilder des Videostreams auch Textinformationen beinhalten. Um diese effizient aus dem Video-Stream zu ermitteln, werden zunächst repräsentative Einzelbilder (Key-Frames) bestimmt, die Kandidaten mit potenziellen Textinhalten herausfiltern. Diese Textkandidaten werden mithilfe geeigneter Vorverarbeitung und Filterung qualitativ optimiert und von einer OCR-Software erkannt.

Erkannte Entitäten werden automatisch mit semantischen Entitäten aus den Normdaten verknüpft (PND, GKD, SWD). Die erzeugten semantischen Metadaten werden in ein audiovisuelles Suchsystem integriert und mit einem semantischen Netzwerk verknüpft. Die Ergebnisse der automatischen Analyseverfahren können von Experten mithilfe einer automatisch erzeugten Zusammenfassung überwacht und ggf. korrigiert werden. Erst danach werden die automatisch erzeugten Metadaten für die Verwendung im Suchindex vorbereitet und mit anderen Wissensquellen verlinkt. Zusätzlich können Benutzer eigene Inhalte und Metadaten zu den jeweiligen erfassten Medien hinzufügen, um auf diese Weise die automatisch ermittelten Metadaten zu ergänzen bzw. zu korrigieren. Eine semantische Suche in den derart erschlossenen und annotierten audiovisuellen Daten bietet neben der direkten Suche nach Schlüsselwörtern, die auf die annotierten Entitäten abgebildet werden, zusätzlich die Möglichkeit, zu einem gegebenen AV-Objekt möglichst (visuell) ähnliche AV-Objekte zu ermitteln (Query by Example). Da alle semantischen Metadaten einen direkten Zeitbezug besitzen, kann dadurch auch eine Suche innerhalb der jeweiligen AV-Medien durchgeführt werden und exakt die Szenen ermittelt werden, die die gesuchten Inhalte zeigen. Dabei ist der jeweilige Kontextbezug durch die zuvor durchgeführte strukturelle Analyse in Kapitel, Szenen, Shots und Subshots zu berücksichtigen.

4.2. Explorative Suche – Perspektiven für Kultureinrichtungen

Aktuelle Suchmaschinen orientieren sich am klassischen Suchparadigma, d. h. um eine gesuchte Informationsressource zu finden, gibt der Benutzer in seiner Suchabfrage einen für die gesuchte Informationsressource charakteristischen Suchbegriff ein. Im Falle des textbasierten Information Retrievals werden Textdokumente zurückgeliefert, die den Suchbegriff tatsächlich enthalten. Der Benutzer geht also von der Annahme aus, dass die zu findenden Informationsressourcen den verwendeten Suchbegriff bein-

halten bzw. damit annotiert wurden. Prinzipiell muss der Benutzer daher in etwa wissen, was er überhaupt sucht bzw. finden möchte.

Bei der Suche in einem Bibliothekskatalog sucht der Benutzer nach Autoren, Titeln, Verlagen oder Schlagworten. Entweder kennt er den Namen des betreffenden Autors bzw. den Buchtitel, oder aber er versucht sich an einer thematischen Zuordnung des von ihm gesuchten Werkes und schlägt diese im Schlagwortkatalog nach, in dem den Informationsressourcen von autoritativer Stelle, d. h. vom Autor, dem Verleger oder dem Bibliothekar Schlagwörter zugeordnet wurden. Da aber Benutzer und Schlagwortautor unterschiedlicher Auffassung über die zutreffende Zuordnung von Schlagwörtern sein können, ist diese Variante der Suche nicht immer zielführend.

Anders ist die Situation, wenn der Benutzer nicht genau weiß, was er überhaupt genau sucht. Wenn er sich z. B. erst einmal einen Überblick über die zu einem Themengebiet vorhandenen Informationsressourcen bzw. über den gesamten Suchraum verschaffen möchte. Oft ist dies aufgrund der vorhandenen Dokumentenmenge in einem digitalen Archiv oder gar dem WWW unmöglich. In der traditionellen Bibliothek dagegen hat der Benutzer die Möglichkeit, die Bücherregale selbst zu betrachten, in denen die vorhandenen Informationsressourcen entsprechend einer vorgegebenen Systematik eingeordnet wurden. So kann er innerhalb eines Themengebiets „herumstöbern“ und dabei zufällig auf Bücher stoßen, die ihn interessieren, wobei er sich dessen zuvor eventuell gar nicht bewusst war. Diese Möglichkeit der „zufälligen und glücklichen Entdeckung“ wird im Englischen als „Serendipity“ bezeichnet. Es geht dabei darum, neue Suchergebnisse zu entdecken, nach denen der Benutzer zunächst gar nicht gesucht hatte. Diese Art der intrinsisch motivierten Erkundung des Suchraums ist uns aus unserem täglichen Leben vertraut und wird im Gegensatz zur zuvor beschriebenen zielgerichteten Suche als „explorative Suche“ bezeichnet [71]. Um eine explorative Suche zu realisieren, muss das Wissen um die Inhalte und Zusammenhänge der zu durchsuchenden Dokumente mithilfe geeigneter Domain-Ontologien und Linked-Data-Ressourcen repräsentiert werden.

Über thematisch passende Domain-Ontologien werden Klassenzugehörigkeiten und Beziehungen von Klassen untereinander repräsentiert, die sowohl zwischen den Instanzen als auch zwischen den Klassen explizite und implizite inhaltliche Verbindungen herstellen. Diese Verknüpfungen werden dazu genutzt, die traditionelle Informationssuche zu erweitern und Wege entlang dieser Verbindungen aufzuzeigen, mit deren Hilfe der vorhandene Suchraum zielgerichtet erkundet werden kann [18] [72]. Insbesondere wenn eine zielgerichtete Suchabfrage nicht direkt zum Erfolg führt, d. h. wenn zu einem Suchbegriff keine direkten Suchtreffer gefunden werden konnten, da die betreffende Datenbasis keine Dokumente zu diesem Suchbegriff enthält, kann die explorative Suche thematisch naheliegende Suchtreffer liefern, die im Zusammenhang mit dem ursprünglichen Suchbegriff stehen. So erhält der Nutzer selbst dann eine mit großer Wahrscheinlichkeit nützliche Information zurück, wenn eine herkömmliche Suche keine Suchergebnisse präsentieren kann. Mithilfe der explorativen Suche wird daher die Informationssuche selbst zu einem Sucherlebnis, angetrieben durch die Neugierde des Benutzers, die sie auf der anderen Seite aber auch zu beflügeln vermag.

5. Fazit

Die Einführung neuer Paradigmen für eine internetbasierte Suche in großen Multimedia-sammlungen sollen für Endnutzer den Zugang zu Multimediaobjekten deutlich verbessern und die Position von Kultureinrichtungen als Wissensbasis im Internet stärken.

Ein großes Problem für die Qualität einer flächendeckenden Erschließung in Kultureinrichtungen sind große, stetig wachsende Multimediasammlungen, welche überwiegend aus teil- oder unstrukturierten Ressourcen bestehen. Um Nutzern die Suche nach Wissen in großen multimedialen Datenbeständen zu ermöglichen, ist eine ausführliche Inhaltsanalyse jeder Ressource erforderlich. Aufgrund des hohen Aufwands kann dies nicht durch eine Intensivierung und Ausweitung intellektueller Erschließung gelöst werden. In diesem Beitrag wird vorgeschlagen, die Inhaltsanalyse von Multimediaobjekten durch automatische Verfahren zu beschleunigen. Neben neuen Suchmöglichkeiten, die mittels einer einfachen Katalogsuche bisher nicht darstellbar waren (wie beispielsweise Ähnlichkeitssuche oder Suche nach Low-Level-Features), ist ein wesentlicher Vorteil der automatischen Inhaltsanalyse die Möglichkeit zur effizienten Identifikation von Entitäten, welche mit Normdaten zu Personen, Schlagworten, Orten oder Organisationen vernetzt werden können.

In einem nächsten Schritt können automatisch erzeugte Metadaten und intellektuell erzeugte Katalogmetadaten in eine gemeinsame Ontologie integriert werden. Sofern dies unter Berücksichtigung der gängigen W3C-Standards geschieht, ist die institutionenübergreifende Vernetzung von Wissensbasen – insbesondere mit Einrichtungen außerhalb des Kulturbetriebs – einfacher als auf Basis der bisher genutzten Technologien und Standards. Allerdings berücksichtigen Erschließungsgeschäftsgänge in Kultureinrichtungen heutzutage zumeist nicht die technologischen Standards des W3C. Aus diesem Grund können suchrelevante Metadaten aus Bibliothekskatalogen – und damit einhergehend auch die verbundenen Multimediainhalte – bisher kaum durch die Tools des Semantic Web genutzt werden. Dies beschränkt gegenwärtig eine internetbasierte Suche im Wesentlichen auf eine (Kultur-)Portalsuche.

Durch Nutzung von Automatismen könnte der Schwerpunkt von Kultureinrichtungen zukünftig in der Erzeugung von Metadaten mit hoher Informationsqualität liegen, welche durch regelbasierte intellektuelle Erschließung, Kuratierung, Vertrauenswürdigkeit, Verfügbarkeit und Persistenz gegenüber vielen Internet-Ressourcen ein Alleinstellungsmerkmal besitzen würden. Darüber hinaus kann, wie in Abschnitt 4 beschrieben wird, die semantische Vernetzung und Anreicherung multimedialer Ressourcen nicht nur durch institutionsgebundene Bibliothekare, Archivare und Informationswissenschaftler erfolgen, sondern auch durch die Einbeziehung externer Experten realisiert werden.

Sofern künftige Erschließungsgeschäftsgänge Technologien des Semantic Web Stack nutzen und Inhalte in der Linked Open Data Cloud [73] verfügbar gemacht werden, kann das Netzwerk aus Kultureinrichtungen zu einer vertrauenswürdigen Quelle internetbasierten Wissens werden [75] [76] [76]. Auf diese Weise könnten die Ressourcen von Kultureinrichtungen durch eine Vielzahl von internetbasierten Suchmaschinen und in unterschiedlichen Kontexten genutzt werden.

Literatur

- [1] T. Berners-Lee, J. Hendler & O. Lassila, *The Semantic Web*, Sci Am, 2001.
- [2] K. Reichenberger, *Kompendium Semantische Netze*; Springer, Berlin, S. 17-18, 2010.

- [3] I. Horrocks et al.; Semantic Web Architecture: Stack or Two Towers?, *Principles and Practice of Semantic Web Reasoning*, Band 3703, pp. 37–41, 2005.
- [4] P. Danowski & L. Heller, Library 2.0 – The future of the library?, *BIBLIOTHEK Forschung und Praxis* 31 (2), 2007, 130–136.
- [5] O. Flimm, Die open-source-software OpenBib an der USB Köln – Überblick und Entwicklungen in Richtung OPAC 2.0, *BIBLIOTHEK Forschung und Praxis* 31 (2), 2007, 185–192.
- [6] T. Sierra, J. Ryan & M. Wust, Beyond OPAC 2.0: Library Catalog as Versatile Discovery Platform, *Code4Lib Journal*, 1, 2007, <http://journal.code4lib.org/issues/issue1>, abgerufen am 13.02.2011.
- [7] D. Pattern, OPAC 2.0 and beyond, *32nd UKSG Annual Conference and Exhibition*, 2009, <http://www.slideshare.net/daveyp/opac-20-and-beyond>, abgerufen am 13.02.2011.
- [8] Deutsche Nationalbibliothek: *Netzpublikationen – Einführung und Sammlung*, http://www.d-nb.de/netzpub/sammlung/ueber_np.htm, abgerufen am 13.02.2011.
- [9] Deutsche Nationalbibliothek: Service – *Kataloganreicherung*, <http://www.d-nb.de/service/zd/kataloganreicherung.htm>, abgerufen am 18.02.2011.
- [10] Seite „Kataloganreicherung“. *Wikipedia*, Die freie Enzyklopädie. Bearbeitungsstand: 11.11.2010, 19:12, <http://de.wikipedia.org/w/index.php?title=Kataloganreicherung&oldid=81394826#Literatur>, abgerufen am 13.02.2011.
- [11] Aktueller Begriff: Bibliotheken im Internet – Die Google-Buchsuche, *Wissenschaftlicher Dienst des Deutschen Bundestages*, Nr. 43/07, 2007, http://www.bundestag.de/dokumente/analysen/2007/Bibliothek_im_Internet.pdf, abgerufen am 12.02.2011.
- [12] M. Hauer & R. Diedrichs; Kataloganreicherung in Europa, *BuB – Forum Bibliothek und Information*, 5, 394-397, 2010. [http://www.agi-imec.de/internet.nsf/26efb65f701b0871c125751a00413614/3f191bb231fd57ec1257749004a9e7d/\\$FILE/Kataloganreicherung_in_Europa_2010_c.pdf](http://www.agi-imec.de/internet.nsf/26efb65f701b0871c125751a00413614/3f191bb231fd57ec1257749004a9e7d/$FILE/Kataloganreicherung_in_Europa_2010_c.pdf), abgerufen am 12.02.2011.
- [13] Aktuelle Zahlen aus „Der Staat spart, Google digitalisiert“, Spiegel Online am 26.03.2011 <http://www.spiegel.de/netzwelt/netzpolitik/0,1518,753229,00.html>, abgerufen am 27.03.2011.
- [14] M. Sieburg, Präsentation: *Der Medienbroker – Von der Diplomarbeit am IRT zur strategischen Enterprise Application*, 2010, <http://www.irt.de/webarchiv/showdoc.php?z=NDA5MiMxMDA1MjEzI3BkZg==>, abgerufen am 02.02.2011.
- [15] M. Payer, RAK-WB: Skript, Kapitel 7.1: „*Verschiedene Titel einer Schrift – Einheitssachtitel*“, Fassung vom 06.10.2002, <http://www.payer.de/rakwb/rakwb07.htm>, abgerufen am 13.02.2011.
- [16] World Wide Web Consortium: *About W3C Standards*, <http://www.w3.org/standards/about.html>, abgerufen am 13.02.2011, 2011.
- [17] World Wide Web Consortium: *Standards – Semantic Web*, <http://www.w3.org/standards/semanticweb/>, abgerufen am 13.02.2011, 2011.
- [18] H. Sack, NPbibSearch: An Ontology Augmented Bibliographic Search; *Proc. of SWAP 2005, the 2nd Italian Semantic Web Workshop*, Trento, Italy, December 14-16, 2005, CEUR Workshop Proceedings, ISSN 1613-0073, 2005.
- [19] T. Gruber, Toward principles for the design of ontologies used for knowledge sharing, *International Journal of Human-Computer Studies* 43, (5-6), 1995, 907-928.
- [20] International Council of Museums, *Scope Definition of the CIDOC Conceptual Reference Model*, Fassung vom 07.01.2003, <http://www.cidoc-crm.org/scope.html>, abgerufen am 13.02.2010.
- [21] Society of American Archivists – *EAD Roundtable, Encoded Archival Description – What is EAD?*, <http://www.archivists.org/saagroups/ead/aboutEAD.html>, abgerufen am 13.02.2010.
- [22] The Library of Congress, *Standards – EAD*, <http://www.loc.gov/ead/index.html>, abgerufen am 13.02.2011.
- [23] DBpedia community project, *About DBpedia*, <http://wiki.dbpedia.org/About>, abgerufen am 13.02.2011.
- [24] C. Bizer et al., *Linked data on the web*, Proceedings of the 17th International Conference on World Wide Web (WWW), ACM, 1265–1266, 2008.
- [25] *Linked Data – Connect Distributed Data across the Web*, <http://linkeddata.org/home>, abgerufen am 15.02.2011.
- [26] Deutsche Nationalbibliothek, *Standardisierung – Gemeinsame Normdatei (GND)*, <http://www.d-nb.de/standardisierung/normdateien/gnd.htm>, abgerufen am 13.02.2011.
- [27] Deutsche Nationalbibliothek: Service – *Datendienst „Bibliographische Dienstleistungen“*, <http://www.d-nb.de/service/zd/datendienst.htm>, abgerufen am 13.02.2011.
- [28] Deutsche Nationalbibliothek: Service – *Linked-Data-Service*, http://www.d-nb.de/hilfe/service/linked_data_service.htm, abgerufen am 13.02.2011.
- [29] Europeana, *think culture*, <http://www.europeana.eu/portal/aboutus.html>, abgerufen am 15.02.2011.
- [30] About European Film Gateway, <http://www.europeanfilmgateway.eu/about.php>, abgerufen am 15.02.2011.
- [31] Über APEnet, http://www.apenet.eu/index.php?option=com_content&view=article&id=47&Itemid=61&lang=de, abgerufen am 15.02.2011.
- [32] About EUscreen, http://www.euscreen.eu/?page_id=2, abgerufen am 15.02.2011.
- [33] About MIMO, <http://www.mimo-project.eu/content/about-mimo>, abgerufen am 15.02.2011.

- [34] Europeana: ThoughtLab: *New ways of searching and/or browsing*, http://www.europeana.eu/portal/thoughtlab_semanticsearching.html, abgerufen am 18.02.2011.
- [35] E. Hyvönen et al., Finnish Museums on the Semantic Web: *The user's Perspective on MuseumFinland*; 2004, <http://www.archimuse.com/mw2004/papers/hyvonon/hyvonon.html>, abgerufen am 15.02.2011.
- [36] M. S. Lew et al., Content-Based Multimedia Information Retrieval: State of the Art and Challenges; *ACM Transactions on Multimedia Computing, Communications and Applications* **2** (1), 1–19, 2006.
- [37] CHORUS+, *Projects*, <http://www.ist-chorus.org/Projects>, abgerufen am 15.02.2011.
- [38] J. S. Hare et al., Mind the Gap: Another look at the problem of the semantic gap in image retrieval, *Multimedia Content Analysis, Management and Retrieval*, San Jose, California, USA, pp. 607309-1, 2006.
- [39] Y. Kompatsiaris & P. Hobson, *Semantic Multimedia and Ontologies*, Springer, Luxemburg, 2008.
- [40] S. F. Chang, The holy grail of content-based media analysis, *IEEE Multimedia* **9** (2), 2002, 6–10.
- [41] M. R. Naphade & T. S. Huang, Extracting semantics from audio-visual content: the final frontier in multimedia retrieval, *IEEE Trans. on Neural Networks* **13** (4), 2002, 793-810.
- [42] A. Hauptmann et al., Can High-Level Concepts Fill the Semantic Gap in Video Retrieval? A Case Study With Broadcast News, *Multimedia, IEEE Trans. on Multimedia* **9** (5), 2007, 958–966.
- [43] D. H. Ballard & C. M. Brown, *Computer Vision*, Prentice-Hall, 1982.
- [44] D. H. Hubel & T. N. Wiesel, Quantitative analysis of cat retinal ganglion cell response to visual stimuli, *Vision Research* **5** (11), 1965, 583–601.
- [45] R. W. Rodieck, Receptive fields of single neurons in the cat's striate cortex, *J Physiol* **148**, 1959, 574–591.
- [46] H. G. Barrow & J. M. Tenenbaum, Interpreting line drawings as three-dimensional surfaces, *Artificial Intelligence* **17** (1-3), 1981, 75–116.
- [47] W. R. Levick, Receptive fields and trigger features of ganglion cells in the visual streak of the rabbit's retina, *J Physiol (Lond)* **188**, 1967, 285–307.
- [48] M. van Wyk, W. R. Taylor & D. I. Vaney, Local edge detectors: a substrate for fine spatial vision at low temporal frequencies in rabbit retina, *J Neurosci* **26** (51), 2006, 13250–63.
- [49] Google, *Erweiterte Bildersuche*, http://www.google.de/advanced_image_search, abgerufen am 15.02.2011.
- [50] Blink, *Video Search Technology*, <http://www.blinkx.com/video-technology>, abgerufen am 15.02.2011.
- [51] Blink Technology: *Under the Hood*, http://issuu.com/smussencom/docs/under_the_hood_2009, abgerufen am 15.02.2011.
- [52] Amazon.de, *Blick ins Buch*, <http://www.amazon.de/exec/obidos/tg/browse/-/14224751>, abgerufen am 15.02.2011.
- [53] D. Lewandowski, Google Buchsuche. Bücher kostenlos zum Download, *Password* **10/2006**, 36, <http://www.durchdenken.de/lewandowski/doc/pw102006.pdf>, abgerufen am 15.02.2011.
- [54] About Project Gutenberg, <http://www.gutenberg.org/wiki/Gutenberg>About>, abgerufen am 15.02.2011.
- [55] Project Gutenberg, *Liste der deutschsprachigen Bücher*, <http://www.gutenberg.org/browse/ languages/de>, abgerufen am 15.02.2011.
- [56] M. Lew et al., Content-based Multimedia Information Retrieval: State of the Art and Challenges, *ACM Transactions on Multimedia Computing, Communications, and Applications*, 1–19, 2006.
- [57] R. Datt et al., Image Retrieval: Ideas, Influences, and Trends of the New Age, *ACM Computing Surveys* **40** (2), 2008, Artikel 5.
- [58] H. Müller et al., A review of content-based image retrieval systems in medical applications—clinical benefits and Future Directions, *International Journal of Medical Informatics* **73** (1), 2004, 1–23.
- [59] R. J. McNab et al., The New Zealand Digital Library MELody index, *D-Lib Magazine*, 1997.
- [60] R. B. Dannenberg et al., A comparative evaluation of search techniques for query-by-humming using the MUSART testbed. *Journal of the American Society for Information Science and Technology* **58** (5), 2007, 687–701.
- [61] A. Vailaya et al., Image Classification for Content-Based Indexing, *IEEE Transactions on Image Processing* **10** (1), 2001, 117–130.
- [62] Google Goggles, <http://www.google.com/mobile/goggles/>, abgerufen am 15.02.2011.
- [63] Normdatensatz des Pariser Eiffelturms, <http://d-nb.info/gnd/4173343-5>, abgerufen am 15.02.2011.
- [64] M. Vacura et al., Describing low-level image features using the comm ontology; http://www.eurecom.fr/~troncy/Publications/Troncy_Vacura-ictp08.pdf, abgerufen am 15.02.2011.
- [65] D. Stan and, I. K. Sethi, Mapping low-level image features to semantic concepts, <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.89.3426&rep=rep1&type=pdf>, abgerufen am 15.02.2011.
- [66] Deutsche Nationalbibliothek: Standardisierung – *Personennamendatei (PND)*, <http://www.d-nb.de/standardisierung/normdateien/pnd.htm>, abgerufen am 15.02.2011.
- [67] N. Pinto, D. D. Cox & J. J. DiCarlo, Why is Real-World Visual Object Recognition Hard?, *PLoS Comput Biol* **4** (1) e27, 2008. <http://www.ploscompbiol.org/article/info:doi/10.1371/journal.pcbi.0040027>, abgerufen am 15.02.2011.
- [68] R. Jenkins & A. M. Burton; 100% Accuracy in Automatic Face Recognition. *Science* **25** **319** (5862), 2008, 435.

- [69] Deutsche Nationalbibliothek: Standardisierung – Schlagwortnormdatei (SWD), <http://www.d-nb.de/standardisierung/normdateien/swd.htm>, abgerufen am 15.02.2011.
- [70] Zum Beispiel Autonomy Virage, <http://promote.autonomy.com/components/pagenext.jsp?topic=PRODUCT::VIRAGE>, abgerufen am 15.02.2011.
- [71] G. Marchionini, Exploratory search: from finding to understanding, *Commun. ACM* **49** (4), 2006, 41-46.
- [72] J. Waitelonis & H. Sack, Towards Exploratory Video Search by Using Linked Data, *Proc. of 2nd IEEE International Workshop on Data Semantics for Multimedia Systems and Applications (DSMSA2009)*, December 14-16, 2009, San Diego, California, 2009.
- [73] S. Gradmann, Knowledge = Information in Context: on the Importance of Semantic Contextualisation in Europeana. *Europeana White Papers*, 2010. http://www.version1.europeana.eu/c/document_library/get_file?uuid=cb417911-1ee0-473b-8840-bd7c6e9c93ae&groupId=10602, abgerufen am 17.02.2011.
- [74] K. Bossert, N. Flores-Herr & J. Hannemann, CONTENTUS: Technologien für digitale Bibliotheken der nächsten Generation; *Dialog mit Bibliotheken* **21** (H. 1), 2009, 14-20. http://www.d-nb.de/netz-pub/info/pdf/dialog_2009_1_bossert.pdf abgerufen am 18.02.2011.
- [75] J. Nandzik et al., CONTENTUS – Towards Semantic Multimedia Libraries, World Library and Information Congress: 76th IFLA General Conference and Assembly, 2010.
- [76] B. Litz et al., Towards Personalized Semantic Multimedia Search for Digital Libraries; PATCH: *3rd International Workshop on Personalized Access to Cultural Heritage*, 2011.