

Stefan Wagenpfeil, Matthias Hemmje

Multimedia-Informationssysteme

Modul 63414, Lektion 1
Einführung, Begriffe und Überblick

Leseprobe

Fakultät für
**Mathematik und
Informatik**

Das Werk ist urheberrechtlich geschützt. Die dadurch begründeten Rechte, insbesondere das Recht der Vervielfältigung und Verbreitung sowie der Übersetzung und des Nachdrucks, bleiben, auch bei nur auszugsweiser Verwertung, vorbehalten. Kein Teil des Werkes darf in irgendeiner Form (Druck, Fotokopie, Mikrofilm oder ein anderes Verfahren) ohne schriftliche Genehmigung der FernUniversität reproduziert oder unter Verwendung elektronischer Systeme verarbeitet, vervielfältigt oder verbreitet werden. Wir weisen darauf hin, dass die vorgenannten Verwertungsalternativen je nach Ausgestaltung der Nutzungsbedingungen bereits durch Einstellen in Cloud-Systeme verwirklicht sein können. Die FernUniversität bedient sich im Falle der Kenntnis von Urheberrechtsverletzungen sowohl zivil- als auch strafrechtlicher Instrumente, um ihre Rechte geltend zu machen.

Der Inhalt dieses Studienbriefs wird gedruckt auf Recyclingpapier (80 g/m², weiß), hergestellt aus 100 % Altpapier.

Inhaltsverzeichnis

Inhaltsverzeichnis	III
Abbildungsverzeichnis	IV
1 L01 – Einführung, Begriffe und Überblick	5
1.1 Digitale Medien	6
1.1.1 Textformate	6
1.1.2 Bildformate	7
1.1.3 Videoformate	8
1.1.4 Audioformate	8
1.1.5 Kombinierte Formate	9
1.2 Multimedia	10
1.2.1 Multimedia-Kollektionen	10
1.2.2 Daten und Fakten	11
1.2.3 Multimedia-Objekte	12
1.2.4 Multimedia-Merkmale	14
1.3 Wissen und Information	16
1.3.1 Informationsbedürfnis und Informationsverhalten	18
1.3.2 Aufgaben des Information-Retrieval	22
1.4 Multimedia-Information-Retrieval	22
1.4.1 Multimedia-Informationssysteme	23
1.5 Übersicht der zentralen MMIR Modelle	26
1.5.1 Exact-Match-Modelle	26
1.5.2 Best-Match-Modelle	28
1.5.3 Ranking-Modelle	29
1.5.4 Vektorraum-Modelle	30
1.5.5 Weitere Modelle	32
1.5.6 Anfragekonstruktion	32
1.6 MMIR-Evaluation	33
1.7 Zusammenfassung	35
1.8 Lösung der Selbsttestaufgaben	36
Literaturverzeichnis	39

Abbildungsverzeichnis

Abbildung 1 - Ausschnitt des Aufbaus der Multimedialen Präsentation für die 1988 Ford New Car Announcement Show.	10
Abbildung 2 - Die BBC musste ihre Inhalte mehrfach neu archivieren und konvertieren.	11
Abbildung 3 - Begriffe im Kontext von Multimedia-Objekten	Fehler! Textmarke nicht definiert.
Abbildung 4 - Komponenten eines Information-Retrieval Systems	Fehler! Textmarke nicht definiert.
Abbildung 5 - Multimedia Merkmalsebenen eines Bildes	Fehler! Textmarke nicht definiert.
Abbildung 6 - Anwendungsspezifische Merkmale	Fehler! Textmarke nicht definiert.
Abbildung 8 - ZusaMultimediaienhang zwischen Daten, Wissen und Informationen.....	Fehler! Textmarke nicht definiert.
Abbildung 9 - Ein kognitives KoMultimediaunikationsmodell nach Belkin	Fehler! Textmarke nicht definiert.
Abbildung 10 - Bates Information Seeking and Searching Model (Bates 2002)	Fehler! Textmarke nicht definiert.
Abbildung 18 - Einteilung der Dokumentenmenge	Fehler! Textmarke nicht definiert.

1 L01 – Einführung, Begriffe und Überblick

Die Lehrveranstaltung „Multimedia-Informationssysteme“ besteht aus sieben aufeinander aufbauenden Lektionen. Zu Beginn jeder Lektion wird zunächst eine Übersicht über deren Aufbau gegeben, eine Einordnung in das übergeordnete Themenfeld dargestellt und die Lernziele festgelegt.

Insgesamt umfasst die Lehrveranstaltung die folgenden Lektionen:

- Lektion 01, Einführung und Überblick: Vorstellung des Themengebietes, Einführung wichtiger Begriffe und zentraler Konzepte.
- Lektion 02, Zeichenkodierung und Audioformate: Definitionen und Formate zur Kodierung und Speicherung von Text- und Audio-Informationen.
- Lektion 03, Bildformate und Videoformate: Definitionen und Formate zur Kodierung und Speicherung von Bild- und Video-Informationen.
- Lektion 04, Kollektionen, Archive, Zugriff und Distribution: Organisation und Verwaltung von Multimedia in Digitalen Bibliotheken, verteilten Datenbanken und zur Langzeitarchivierung.
- Lektion 05, Merkmalsrepräsentation und -extraktion: Repräsentation von Multimedia durch Merkmale und praxisorientierte Vorstellung der Extraktion von Merkmalen.
- Lektion 06, Metadatenbeschreibungsformate: Standards und Verfahren zur Beschreibung von Metadaten im Bereich Multimedia.
- Lektion 07, Interaktion und Retrieval: Vorstellung der Interaktion von Benutzenden und der typischen Anwendungsfälle im Bereich Multimedia Information Retrieval.

In der hier vorliegenden ersten Lektion erarbeiten sich die Studierenden einen Überblick über den Themenkomplex „Multimedia-Informationssysteme“ und das untrennbar damit verbundene Thema „Multimedia-Information-Retrieval“, welches als Basis für die nachfolgenden Kapitel dient. In sieben Abschnitten werden von den Teilnehmenden der Lehrveranstaltung nach einer allgemeinen Einleitung und Motivation (Abschnitt 1.1) die Begriffe „Multimedia“ (Abschnitt 1.2), „Information“ (Abschnitt 1.3), „Multimedia-Informationssysteme“ (Abschnitt 1.4), „Multimedia-Information-Retrieval“ (Abschnitt 1.5), „Retrieval-Modelle“ (Abschnitt 1.6) und „Evaluationsmethoden“ (Abschnitt 1.7) untersucht.

Die Studierenden erlangen als übergeordnetes Lernziel der Lehrveranstaltung „Multimedia-Informationssysteme“ ein umfassendes und detailliertes Verständnis über die Zusammenhänge und Prozesse im Bereich „Multimedia“. Da Multimedia aus modernen Anwendungen nicht mehr „wegzudenken“ ist und eine zentrale Rolle für nahezu jede Art von Software-Entwicklung spielt, stellt die Lehrveranstaltung somit auch eine wichtige Grundlage für praxis- und zukunftsorientiertes Arbeiten dar.

Mit Abschluss dieser ersten Lektion sollten die Teilnehmenden sicher im Umgang mit Begriffen, Konzepten und generellen Verfahrensweisen im Bereich „Multimedia-Informationssysteme“ und „Multimedia-Information-Retrieval“ sein; dies ist die Basis für die nachfolgenden Lektionen.

1.1 *Digitale Medien*

Die automatisierte Verarbeitung und Verwaltung von digitalen Medien und Daten stellt eine der Hauptaufgaben unserer Zeit dar. Heutige Benutzende nehmen sowohl die Rolle des Konsumenten als auch des Produzenten von digitalen Medien ein, bspw. indem auf Social-Media Bilder, Videos oder Kommentare geteilt werden. Hierzu werden Anwendungen bereitgestellt, die die digitale Repräsentation von Inhalten komplett für den Benutzenden kapseln und intuitive, einfach zu bedienende Benutzungsoberflächen bereitstellen. Benutzende müssen sich typischerweise nicht mehr darum kümmern, Inhalte in bestimmte digitale Repräsentationen zu konvertieren oder zu transformieren, all das geschieht im Hintergrund. Dennoch ist es für den Datenaustausch zwischen Anwendungen oder Benutzenden essenziell, Standards einzuhalten, Formate festzulegen und somit auch eine Basis für professionelle Anwendungen zu bilden.

Die Kongressbibliothek der USA hat vor Jahren ein Projekt ins Leben gerufen, dass die Standardisierung von digitalen Inhalten unterstützen soll und sammelt seitdem die gängigsten Standards und ihre Formatbeschreibungen [1]. Details über die einzelnen Standards und ihre technischen Unterschiede finden sich in Lektion 2. Im Folgenden werden dennoch bereits jetzt einige typische Beispiele für die unterschiedlichen Formate digitaler Medien angeführt:

1.1.1 *Textformate*

Die meisten Textformate sind heutzutage nicht mehr menschenlesbar und benötigen spezielle Anwendungen, um sie in verständlicher Form darzustellen. Glücklicherweise sind diese Anwendungen standardisiert und jeder kann selbstverständlich Word- oder PDF-formatierte Dokumente öffnen. Allerdings muss klar sein, dass das maschinenlesbare Format erheblich von dem angezeigten Format abweicht, denn die reine Textinformation wird hier durch Metainformationen über Formatierungen, Verlinkungen oder eingebettete andere Multimedia-Objekte angereichert. Die Semantik der Formatierung stellt ebenfalls wieder ein Multimedia-Merkmal dar, denn ein fett formatierter Text hat ggfs. eine wichtigere Bedeutung als ein normal formatierter Text.

Gängige Formate sind:

PDF: Das Portable Document Format [2] wurde von Adobe spezifiziert und konserviert Schriftarten und Layout-Informationen. Es wurde zum Quasi-Standard für Dokumentaustausch.

DOC/DOCX: Dieses Format findet seinen Ursprung in Microsoft Word und basiert mittlerweile auf einem standardisierten XML-Schema, dem Open Office XML [3]. Dokumente werden typischerweise mit Microsoft Word oder Open Office editiert, allerdings kann aufgrund der XML-Struktur auch eine Bearbeitung in vielen anderen Anwendungen erfolgen.

TXT: Das TXT-Format beinhaltet typischerweise Plain-Text-Dateien, in denen das Format keine Rolle spielt, keine Layout-Möglichkeiten vorhanden sind, stattdessen die reine Information transportiert wird. Dieses Format wird häufig auch für technische Zusammenhänge genutzt, um bspw. Listen in Komma-getrennter Art zu beschreiben (CSV-Datei, „comma separated value“) und somit einen einfachen, standardisierten Austausch zu ermöglichen.

HTML: Die Hypertext Markup Language [4] wurde entwickelt, um Inhalte für das Internet bereitzustellen und auf einfache Art Formatierungen und Verlinkungen vorzunehmen. Hierbei wird ein Plain-Text-Dokument mit zusätzlichen Markups (z. B. `<h1></h1>`) angereichert, welche dann vom Browser entsprechend interpretiert und formatiert werden.

XML: Die eXtensible Markup Language [5] ist ein Format, mit dem der automatisierte Austausch von Informationen und die standardisierte Beschreibung domänenspezifischer Dokumente ermöglicht wird. XML-Dokumente sind wohlgeformt (d. h. jedes öffnende Tag wird auch wieder geschlossen) und überprüfbar (d. h. ihre Struktur kann gegen ein Schema validiert werden).

Json: Das Javascript Object Notation Format [6] wurde als Ergänzung zu XML eingeführt und soll den technischen Austausch von Informationen optimieren. Hierbei wird auf die Überprüfbarkeit von Gültigkeit verzichtet und die XML-Tags werden durch eine syntaktische Struktur aus diversen Klammern ersetzt. Dies führt zu deutlich schlankeren Dokumenten für den Datenaustausch.

1.1.2 Bildformate

Bilder sind eines der am häufigsten genutzten Multimedia-Inhaltsformate. In den letzten Jahren hat sich eine Reihe von Standard-Formaten für die Speicherung von Bildern entwickelt:

TIF/TIFF: Das Tag Image Format [7] wurde bereits 1986 definiert und wird mittlerweile von Adobe verwaltet. TIFF-Dateien verfügen über verlustfreie Kompression, können Metadaten beinhalten und werden häufig im Bereich Optical Character Recognition (OCR), Desktop Publishing oder beim Scannen von Dokumenten eingesetzt.

PNG: Dateien im Portable Network Graphic Format [8] unterstützen ebenfalls verlustfreie Kompression und basieren auf Farbpaletten (24 oder 32 bit). Außerdem können Informationen zur Transparenz gespeichert werden. PNG-Dateien werden häufig zum Veröffentlichenden von Bildern genutzt.

GIF: Das Graphic Interchange Format [9] wurde 1987 von CompuServe eingeführt und reduziert Bilder auf eine 256-farbige Palette, die typischerweise zu sehr kleinen Dateigrößen führt. GIF-Dateien werden immer häufiger durch PNG-Dateien ersetzt.

JPG/JPEG: Die Joint Photographic Expert Group [10] hat das JPEG Format 1992 als komprimiertes Format eingeführt, bei dem die Kompressionsrate variabel eingestellt werden kann. Somit entstehen verschieden große Zieldateien, was den Austausch von Bildern erheblich vereinfacht.

RAW [11]: Dieses Format wird häufig von Digitalkameras genutzt, um verlustfreie Bilder im Originalformat zu produzieren.

Die folgende Tabelle zeigt die wichtigsten Formate im Überblick anhand eines Beispiel-Bildes, das mit einer Fuji-XT3-Kamera aufgenommen wurde (RAW-Format, 6000x4000 Pixel, 300 dpi, 50.6 MB):

Format	Dateigröße	Kompression	Transparenz	Farbtiefe
TIFF	72 MB	Verlustfrei	Ja	32 bit
PNG	45,8 MB	Verlustfrei	Ja	48 bit
GIF	15,7 MB	Verlustfrei	Ja	8 bit
JPG (Q12)	23.1 MB	Verlustbehaftet	Nein	24 bit
JPG (Q10)	10.6 MB	Verlustbehaftet	Nein	24 bit
JPG (Q2)	2.5 MB	Verlustbehaftet	Nein	24 bit

Zusätzlich zu diesen Formaten können auch Meta-Informationen, wie z. B. durch das EXIF-Format spezifiziert als Merkmale von Bildern betrachtet werden.

1.1.3 Videoformate

Im Bereich der Videoformate gibt es ebenfalls eine Vielzahl gängiger Formate und Speicherstrukturen. Grundsätzlich lassen sich jedoch folgende gängigen Formate unterscheiden:

AVI: Das Audio Video Interlaced Format [12] wurde von Microsoft 1992 eingeführt und enthält verschiedene Streams unterschiedlicher Daten. Typischerweise sind dies mindestens die Video- und die Audio-Spur, die unterschieden werden können.

MOV: Das Quick Time File Format [13] wurde von Apple propagiert und ist ein Container-Format, in dem mehrere Spuren von Audio- und Video-Inhalten enthalten sein können.

MP4: Das MPEG-4 Part 14 Format [14] ist ein modernes und plattformübergreifendes Container-Format, das zusätzlich standardisiert Annotationen, Rechtemanagement und Kompression für Audio und Video unterstützt. Ein moderner Codec zur Kompression ist der H.264-Algorithmus.

MXF: Das Material Exchange Format [15] definiert Strukturen für den Austausch von audiovisuellem Material. Es wird häufig von Produktions- oder Broadcastfirmen eingesetzt.

Weitere Videoformate und deren technische Hintergründe finden sich in Lektion 2.

1.1.4 Audioformate

Die Aufzeichnung von Audio-Informationen erfolgt meistens über eine Analog-Digital-Konvertierung. Hierbei wird eine analoge Quelle zunächst in eine Abfolge digitaler Messpunkte überführt,

die dann für die Speicherung in ein konkretes Format umgewandelt werden. Hierbei gibt es unterschiedliche Verfahren, die gängigsten sind:

AIFF: Das Audio-Interchange Format [16] wurde von Apple als verlustfreies, unkomprimiertes Format eingeführt. Die Sample-Rate gibt die Qualität der Aufzeichnung an und kann variiert werden.

WAV: Das WAVE Format [17] stellt das korrespondierende Format von Microsoft dar.

MP3: Das MPEG 2 Layer III Audio Encoding [18] ist ein Kompressionsformat, das es dem Audio-codec erlaubt, Signale, die Menschen typischerweise nicht akustisch erfassen oder differenzieren können, zu entfernen. MP3-Dateien sind sehr klein und werden daher häufig für den Austausch und das Streamen von Musik eingesetzt.

MIDI [19]: Hierbei handelt es sich um ein spezielles Format zum Abspeichern von Meta-Informationen von Musikstücken in Form der einzelnen gespielten Noten. Midi-Dateien sind sehr klein und können beim Abspielen Sound-Erzeugungssysteme nutzen, um einen möglichst realistischen Klang zu erzeugen.

1.1.5 Kombinierte Formate

In den letzten Jahren werden digitale Formate häufig kombiniert. Bilder oder Videos werden in Dokumente eingebettet, Videos enthalten textuelle Beschreibungen, Social-Media Posts werden mit Audio-Dateien kombiniert, usw. Eine klassische und scharfe Trennung zwischen den einzelnen digitalen Formaten verwischt mehr und mehr und führt zu neuen Herausforderungen im Bereich der digitalen Verarbeitung – und letztendlich auch zum Begriff „Multi-Media“.

Selbsttestaufgabe 1.1:

Welche Multimedia-Formate kennen Sie?

Selbsttestaufgabe 1.2:

Was ist der Hauptunterschied zwischen HTML und XML?

Selbsttestaufgabe 1.3:

Welche der Eigenschaften trifft auf JSON-Dokumente zu?

- Austausch von Informationen
- inhaltliche Prüfung aufgrund von Schemata ist möglich
- Überprüfung der strukturellen Gültigkeit ist möglich
- Dokumente sind kürzer als XML-Dokumente
- E-Mail-Nachricht
- JSON basiert auf XML-Tags

Selbsttestaufgabe 1.4:

Welche Bildformate unterstützen Transparenz?

1.2 Multimedia

Eine der ersten Erwähnungen des Begriffs „Multimedia“ geht zurück auf die späten 1980er-Jahre, als die Firma Ford auf der „1988 Ford New Car Announcement Show“ in Detroit ihre Automobile in einer – im wahrsten Sinn des Wortes – multimedialen Präsentation vorstellte (siehe Abbildung 1).



Abbildung 1 - Ausschnitt des Aufbaus der multimedialen Präsentation für die 1988 Ford New Car Announcement Show.

Seit dieser Zeit hat sich vieles verändert und eine gute Begriffsdefinition für „Multimedia“ wird von Tay Vaughan [20] wie folgt gegeben: *„any combination of text, graphic, art, sound, animation, and video that is delivered by computer or other electronic or digitally manipulated means“*.

Dieser enorm weit gefasste Rahmen zeigt somit unmittelbar die Wichtigkeit der Technologien, mit denen sich diese Lehrveranstaltung beschäftigt. Um es kurz zu sagen: Multimedia ist überall!

1.2.1 Multimedia-Kollektionen

Die dauerhafte Nutzbarkeit von digitalen Medien führt dazu, dass diese in Kollektionen von Multimedia-Objekten organisiert werden. Der Begriff „Kollektion“ wird hierbei nicht nur verwendet, um zu beschreiben, dass eine gewisse Anzahl an digitalen Medienobjekten verwaltet werden soll,

eine Kollektion zeichnet sich zusätzlich dadurch aus, dass bestimmte beschreibende Metadaten zusätzlich zu den eigentlichen digitalen Medienobjekten abgelegt werden. Diese zusätzlichen Metadaten stellen dann auch die Basis für Multimedia-Informationssysteme (MMIS) dar, in denen nach digitalen Medienobjekten gesucht werden kann. Ohne diese Metadaten müsste der Benutzende jedes einzelne Medienobjekt öffnen, durchsehen und somit herausfinden, ob es seiner Suchanfrage entspricht. Die Metadaten einer Kollektion werden typischerweise verwendet, um einen sog. Index zu erstellen, der den inhaltlich gezielten Zugriff über eine effiziente Suche (engl. *Searching*) oder Exploration (engl. *Browsing*) ermöglicht. Dies ist auch die Grundlage für das Multimedia-Information-Retrieval, also den inhaltlich gezielten Zugriff über die Suche oder Exploration anhand von Merkmalen, die nicht mehr von Menschen, sondern von Systemen analysiert und verwaltet werden.

1.2.2 Daten und Fakten

Multimedia ist schon seit langer Zeit ein wichtiges Medium. Dies führt in verschiedenen Bereichen zu permanenten Aktivitäten, um Multimedia-Objekte dauerhaft und in verwendbarem Zustand zu konservieren. Beispielsweise geht das Archiv der British Broadcasting Company (BBC) auf die frühen 1890er Jahre zurück [21]. Seit dieser Zeit wurden zahllose Projekte durchgeführt, um den immer neu entstehenden digitalen Formaten Rechnung zu tragen und das Archiv auf einem möglichst modernen Stand zu halten (siehe Abbildung 2). Aktuell umfasst das Archiv der BBC 165 Petabyte Speicher, jede Woche entstehen ca. 120 Terabytes an neuem Material.



Abbildung 2 - Die BBC musste ihre Inhalte mehrfach neu archivieren und konvertieren.

Diese Zahlen zeigen auch, dass „Multimedia“ hochgradig relevant in der heutigen Zeit ist. Und rein zahlentechnisch kommt es noch „schlimmer“, denn durch den Einzug der sozialen Medien verwischte die Grenze zwischen klassischen Produzenten und Konsumenten, so dass nunmehr jeder Benutzende selbst Inhalte erzeugt, verteilt, kommentiert, verändert oder konsumiert. Jährlich werden somit 1.2 Trillionen digitaler Fotos und Videos aufgezeichnet, mittlerweile über 85 % auf Smartphones [22]. Bereits 2020 überschritt Facebook die Marke von einer Milliarde aktiver Benutzer pro Monat, Instagram erreichte 500 Millionen tägliche Benutzer seines „Story“ Features und bis Ende 2023 werden 3.4 Milliarden monatlich an aktiven Social-Media-Benutzenden erwartet [23]. Der weltweite Media- und Entertainment-Markt hat ein jährliches Volumen von über 2.1 Trillionen US-Dollar [24].

Hinzu kommen die Fortschritte im Detailgrad der digitalen Repräsentation von Multimedia-Objekten; bspw. können aktuelle Kameras, wie die Sony a7R IV/35 [25], bis zu 61 Megapixel Auflösung produzieren, und selbst Smartphones, wie das Xiaomi Redmi Note 10 Pro [26] verschieben diese Grenze auf 108 Megapixel.

Multimedia findet Anwendung in nahezu jedem Bereich der Informationstechnologie. Im medizinischen Sektor sind Röntgenbilder, MRT- oder CT-Verfahren, aber auch die Echtzeitaufzeichnung von EKG- oder EEG-Signalen Vertreter für Multimedia-Objekte. Der Broadcast-Bereich (TV, Radio, Streaming), der Gaming-Sektor (Spiele, Konsolen), der Kommunikations-Bereich (Messenger, Video-Konferenzen, Kollaborationsplattformen) – all dies sind nur einige Beispiele für den Einsatz von Multimedia.

Sie sehen, „Multimedia“ ist eine der wichtigsten und relevantesten Technologien unserer Zeit, mit der jeder Mensch direkt oder indirekt in Berührung ist. Die fundierte und strukturierte Bereitstellung von Lösungen in diesem Umfeld ist eine zentrale Aufgabe der Software-Entwicklung und der sichere Umgang mit den zugehörigen Konzepten und Begriffen ein zentraler Baustein für erfolgreiche Projekte.

In den folgenden Abschnitten werden nun die Grundlagen und zentralen Begriffe im Themenbereich „Multimedia“ eingeführt.

1.2.3 *Multimedia-Objekte*

Aufgrund der Vielzahl von technologischen Teilbereichen, Formaten, Medienobjekt-Typen (wie z. B. Bild, Audio, Video, Text), ist es wichtig, von vornherein eine Reihe von Begriffen einzuführen, die im Verlauf der Lehrveranstaltung verwendet werden.

Multimedia an sich ist immer die Repräsentation eines realen Vorgangs. Das Bild oder Video, das Sie aufnehmen, der Text den Sie schreiben, all das hat einen Bezug zu Objekten oder einem Kontext von Objekten und Vorgängen der realen Welt. Hierbei ist es unerheblich, ob der Text ein reales Ereignis dokumentiert oder fiktive Ereignisse beschreibt, denn auch die Vorstellungskraft eines Autors/einer Autorin oder die künstlerische Leistung eines Malers/einer Malerin oder Komponisten/Komponistin stellt einen real existierenden Vorgang dar. Daher lassen sich die folgenden

Begriffe einführen und abgrenzen (siehe auch **Fehler! Verweisquelle konnte nicht gefunden werden.**):

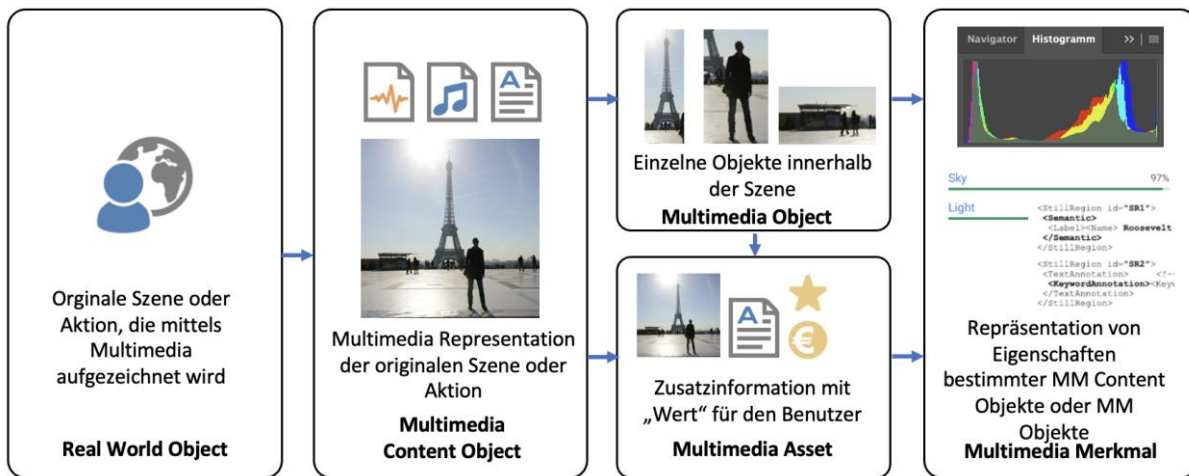


Abbildung 3 - Begriffe im Kontext von Multimedia-Objekten

Real-World-Objekt: Dies ist die originale physische Szene aus Objekten und Vorgängen in der realen Welt, Aktion, die mittels Multimedia digital repräsentiert und aufgezeichnet wird.

Multimedia-Content-Objekt: Eine multimediale Repräsentation der originalen Szene, des Vorgangs oder der Aktion in einem meist standardisierten Format. Hier wären beispielsweise eine Audio-Datei, ein Video, Foto oder eine textuelle Beschreibung zu verorten.

Multimedia-Objekt: Innerhalb eines Multimedia-Content-Objektes können verschiedene einzelne digitale Objekte als Repräsentation von Objekten der realen Welt identifiziert werden. Dies können bei einem Text die Abschnitte, Sätze oder Buchstaben sein, bei einem Bild die gezeigten Personen oder Vorgänge, bei Videos die einzelnen Szenen und ebenfalls die gezeigten Personen oder Vorgänge, oder bei Audio bestimmte zeitbasierte Informationen, und vieles mehr.

Multimedia-Asset: Jede Art von Multimedia-Content-Objekt oder Multimedia-Objekt kann für die Benutzenden einen „Wert“ erhalten. Dieser kann tatsächlich monetär zu betrachten sein (d. h., wenn Lizenzgebühren für die Nutzung des Multimedia-Content-Objektes anfallen, wie bspw. auf Streaming-Plattformen); er kann aber auch virtuell sein, weil die Benutzenden bspw. ein Foto zu einem Album „Lieblingsbilder“ hinzufügen oder einen Social-Media-Post mit „Like“ markiert haben.

Multimedia-Merkmal: Dies sind Repräsentationen von Eigenschaften bestimmter Multimedia-Content-Objekte oder Multimedia-Objekte, die Informationen auf verschiedenen Ebenen und in verschiedenen Detailgraden beschreiben (siehe auch Abschnitt **Fehler! Verweisquelle konnte nicht gefunden werden.**). Häufig werden Merkmale auch „Feature“ genannt.

Neben diesen einführenden Begriffen aus dem Bereich „Multimedia“ ist auch ein grundlegendes Verständnis des Fachgebietes „Information Retrieval“ nötig, um die nachfolgenden Abschnitte einordnen zu können. Auch, wenn für diesen Bereich mehrere Abschnitte in dieser Lehrveranstal-

tung vorgesehen sind, so soll die folgende, grundlegende Übersicht bereits den Rahmen abstecken und einige zentrale Begriffe einführen (siehe **Fehler! Verweisquelle konnte nicht gefunden werden.**).

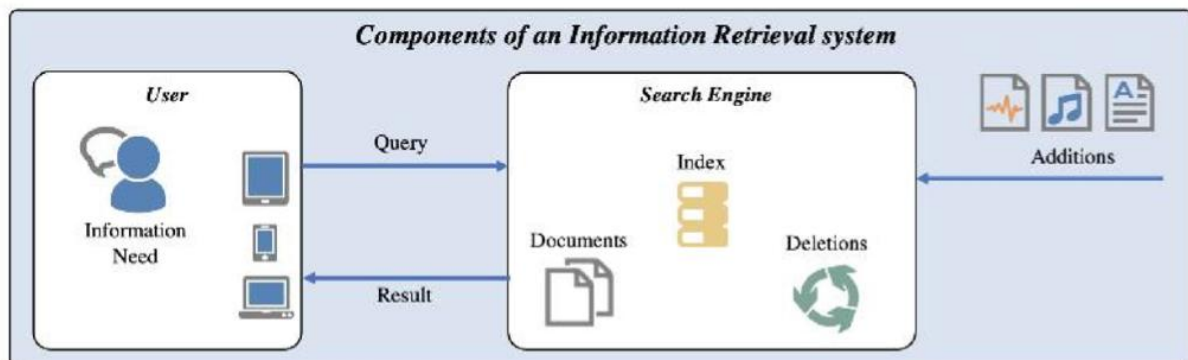


Abbildung 4 - Komponenten eines Information-Retrieval Systems

Das „Information Retrieval“ und die zugehörigen „Informationssysteme“ werden typischerweise eingesetzt, um das Informationsbedürfnis eines Benutzenden zu befriedigen. Hierzu werden z. B. Anfragen an eine Suchmaschine (engl. *Search Engine*) gestellt, die dem Benutzenden dann in Form von Ergebnissen (engl. *Result*) präsentiert werden. Dieses Verhalten kennt jeder von einer klassischen Suchmaschinen-Suche. Im Hintergrund findet jedoch eine Reihe von weiteren Prozess-Schritten statt: die Search-Engine erzeugt automatisiert permanent einen sog. Index, der einen optimierten Zugriff auf die Repräsentationen von Merkmalen ermöglicht, um somit die eigentlichen Ergebnisse (Dokumente/Medienobjekte) schnell finden zu können. Dies wird dadurch erschwert, dass permanent neue Dokumente/Medienobjekte in der Kollektion des Informationssystems hinzukommen, verwaltet und in den Index eingefügt werden müssen, und dass auch bestehende Dokumente/Medienobjekte entfernt werden können. Das klassische „Information Retrieval“ fokussiert hier auf die Erstellung von Index-Strukturen und die Modellierung von Anfrage- und Ergebnisrepräsentationen, wohingegen der Begriff „Informationssystem“ zusätzlich die Prozesse der Verwaltung, des Hinzufügens und Löschens von Multimedia-Content-Objekten, Multimedia-Objekten und Multimedia-Merkmalen beinhaltet. Im weiteren Verlauf der Lehrveranstaltung werden die folgenden Abkürzungen verwendet:

MMIR: Multimedia-Information-Retrieval

MMIS: Multimedia-Informationssystem

MM: Multimedia

Die Basis für jede Art von MMIR und MMIS sind MM-Merkmale und deren Repräsentation. Diese werden im folgenden Abschnitt in den Gesamtüberblick eingeordnet.

Selbsttestaufgabe 1.5:

Erklären Sie den Unterschied zwischen einem Multimedia-Objekt und einem Multimedia-Asset.

1.2.4 Multimedia-Merkmale

Die Extraktion von MM-Merkmalen aus MM-Content-Objekten und MM-Objekten stellt die Basis für MMIR und MMIS dar. Hierbei sind verschiedene Abstraktionsebenen von Merkmalen zu unterscheiden. Anhand eines typischen Fotos lassen sich diese Merkmalsebenen recht gut erkennen (siehe **Fehler! Verweisquelle konnte nicht gefunden werden.**). Es ist jedoch zu berücksichtigen, dass andere MM-Content-Objekte, wie z. B. Video, Audio oder Text natürlich auch andere Merkmale ausprägen.

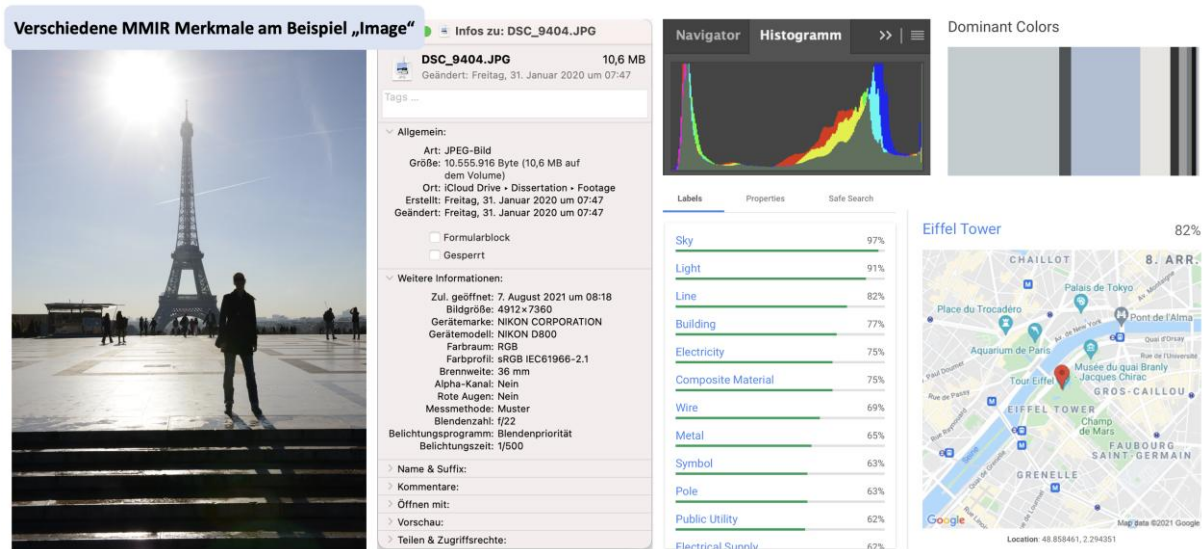


Abbildung 5 – Multimedia-Merkmalsebenen eines Bildes

Merkmale können in den folgenden Abstraktionsebenen entstehen:

Technische Merkmale: Unter diese Kategorie fallen Merkmale, die typischerweise nicht sofort menschenverständlich sind, wie z. B. die binäre Repräsentation von Dateiformaten, Steuersignale, oder auch mathematische Berechnungen, wie Histogramme von Fotos.

Allgemeine Merkmale: Diese Merkmale beschreiben Eigenschaften des MM-Content-Objektes, wie Erstellungsdatum, Dateigröße, Name der Datei, aber auch Basis-Metadaten eines bestimmten Dateiformates (z. B. EXIF-Daten bei Bildern), aus denen dann Informationen über die eingesetzte Kamera oder auch bereits Geo-Informationen ausgelesen werden können.

Format-spezifische Merkmale: Häufig können generisch Informationen aus einer bestimmten Art von MM-Content-Objekt ausgelesen werden. Für alle Bilder lassen sich bspw. immer die dominanten Farben ermitteln, unabhängig davon, was auf dem Bild zu sehen ist und mit welcher Kamera es aufgenommen wurde.

Inhalts-basierte Merkmale: Dieser Bereich wird häufig durch Techniken aus dem Gebiet der Mustererkennung unterstützt. Hier geht es darum, zu ermitteln, was bspw. auf einem Foto zu sehen ist. Es werden Personen, Gegenstände, Rahmen, Objektgrenzen usw. erkannt oder berechnet.

Semantische Merkmale: Hier wird einem inhaltsbasierten Merkmal eine Bedeutung zugewiesen, bspw. wird eine erkannte Person als „Sandra“ und ein erkanntes Objekt als „Eiffelturm“ beschrieben. Es kann außerdem eine räumliche Beziehung, wie „steht vor“, eingeführt werden. Im Text-Umfeld wären semantische Merkmale bspw. die automatische Erkennung des Themas eines Textes.

Die Extraktion und Repräsentation von Merkmalen werden in der zweiten Lektion im Detail behandelt. Wichtig ist jedoch bereits jetzt, dass Merkmale häufig auch einen konkreten Bezug zu Problemdomänen oder Anwendungsbereichen aufweisen können. **Fehler! Verweisquelle konnte nicht gefunden werden.** zeigt dies am Beispiel des DICOM-Standards [27], der u. a. für die Beschreibung von Mammographie-Bildern verwendet werden kann.

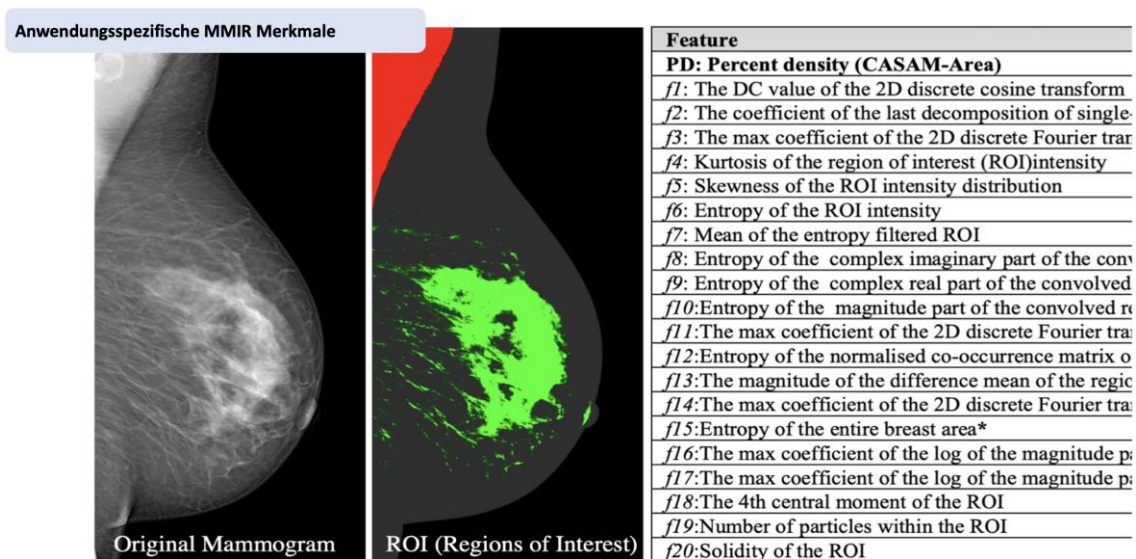


Abbildung 6 - anwendungsspezifische Merkmale

An diesem Beispiel wird außerdem bereits deutlich, dass Merkmale (z. B. das Merkmal f10) häufig eine domänenspezifische Semantik besitzen. Wird ein MMIR unterstützt oder ein MMIS verwendet, so muss dafür Sorge getragen werden, dass die Benutzenden ein grundlegendes Verständnis der gespeicherten Daten besitzen. Ohne Kenntnis, welche domänenspezifischen Merkmale es gibt, kann der MMIR-Prozess deutlich komplizierter werden. An dieser Stelle werden dann Techniken zur Merkmalsrepräsentation, Anfragekonstruktion und semantischen Suche eingesetzt, um das MMIR zu optimieren.

Selbsttestaufgabe 1.6:

Welche Ebenen von Merkmalen im Bereich Multimedia kennen Sie?

1.3 Wissen und Information

Der Begriff der Information ist zentral für die Einordnung und das Verständnis des Information-Retrieval (IR). Wie in vielen wissenschaftlichen Disziplinen gibt es auch hier eine Menge an verschiedenen Definitionen. Eine allg. Definition kann bspw. in gängigen Lexika gefunden werden,

wie z. B. die Definition von Merriam Webster [28]: „*the communication or reception of knowledge or intelligence*“. Information wird somit als das Empfangen von Wissen umschrieben.

Auf dieser Basis ist es insofern wichtig, direkt eine Abgrenzung zu dem eng verwandten Begriff des Wissens und auch von digitalisierten Daten als Grundlage des IR an sich vorzunehmen. Der Duden [29] definiert Daten bspw. als: „(durch Beobachtungen, Messungen, statistische Erhebungen u. a. gewonnene) [Zahlen]werte, (auf Beobachtungen, Messungen, statistischen Erhebungen u. a. beruhende) Angaben, formulierbare Befunde“. Wissen hingegen ist ein wesentlich komplexeres Konzept, das eine enge Verbindung zu weiteren Konzepten, wie u. a. zu den Fähigkeiten und Kompetenzen eines Menschen aufweist. Als eine Vertreterin der Wirtschaftswissenschaften vertritt Rowley in der vielzitierten Publikation [30] die Anschauung, dass Wissen auf Basis der Definition von Daten und Information abgeleitet wird. Wissen wird dabei als eine Kombination aus Daten, Informationen und persönlichem Kontext verstanden und dazu genutzt, Entscheidungen zu treffen. In diesem Umfeld ist es Konsens, dass die folgende Verkettung gilt:

Daten -> Wissen -> Information -> Entscheidung / Lösung

In dieser Lehrveranstaltung wird die Definition von Kuhlen Verwendung finden, da sie zum einen aus dem Blickwinkel der Informationswissenschaften [31] verfasst wurde und sich zum anderen einer breiten Akzeptanz in wissenschaftlichen Gemeinschaften erfreut. Kuhlen stellt eine Definition aus semiotischer (Wissenschaft der Zeichensysteme, u. a. Sprache) Sicht auf und definiert Information als „*die Teilmenge von Wissen, die von einer bestimmten Person oder einer Gruppe in einer konkreten Situation zur Lösung von Problemen benötigt wird und häufig nicht vorhanden ist.*“ Kuhlen unterscheidet Wissen von einem Zustand einer Person, der unabhängig von Kommunikationspartner/innen ist. Information fasst er im Gegensatz als einen Zustand der Kommunikation zwischen Personen auf. Er stellt damit fest, dass die Information eine Teilmenge von Wissen ist. Aus seiner Sicht gilt: Daten sind die Basis, um Wissen zu erzeugen, und aus Wissen können wieder Daten entstehen. Informationen hingegen leiten sich wiederum aus Wissen ab, können aber auch wieder zu Wissen werden (vgl. **Fehler! Verweisquelle konnte nicht gefunden werden.**). Kuhlen prägte auch den bekannten Spruch „Information ist Wissen in Aktion“.

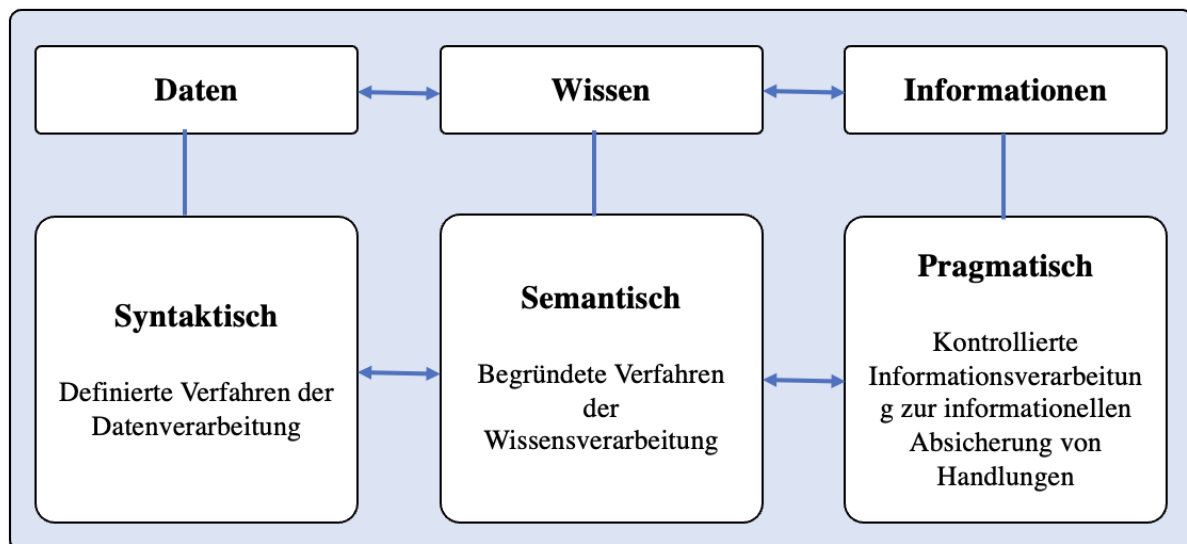


Abbildung 7 - Zusammenhang zwischen Daten, Wissen und Informationen

Kuhlen definiert, dass Daten rein auf syntaktischer Ebene angesiedelt sind. Daten sind somit mit Regelwerken beschreibbar. Wissen jedoch umfasst über die Definition von Daten hinaus semantische Zusammenhänge. Und Information ist zu guter Letzt Wissen, das z. B. handlungs-, entscheidungs- oder problemlösungsrelevant ist und somit Bestandteil der Lösung bzw. Beantwortung eines Informationsbedürfnisses eines Informationssuchenden. Wissen ist also dann und deshalb als Information zu betrachten, weil es für Informationssuchende neu/unbekannt und in der Regel relevant für ein zu Grunde liegendes Informationsbedürfnis ist.

Nachdem nun der Begriff „Information“ eingeführt wurde, gilt es auch, den Begriff des „Retrievals“ zu betrachten. Lexika belegen diesen Begriff mit unterschiedlichen Beschreibungen, wie „die Informationsgewinnung“, „Wiederauffinden von Informationen“ oder auch „Informationswiedergewinnung“. Naheliegend ist zumindest, dass der Begriff der Information und dessen Verarbeitung eine maßgebliche Rolle spielen. Nach einer bekannten Definition von Baeza-Yates [32] ist IR „*part of computer science which studies the retrieval of information (not data) from a collection of written documents. The retrieved documents aim at satisfying a user information need usually expressed in natural language*“. Das IR beschreibt damit generell gesprochen einen Prozess der automatisierten Unterstützung beim gezielten Zugriff auf Information. Formen oder besser gesagt „Informationsstrategien“ des gezielten Zugriffs auf Information sind dabei die Suche nach Informationen über Anfragen (engl. *Searching*) oder Explorationen (engl. *Browsing*). Ein Prozess ist hierbei eine zunächst ganz allgemein geordnete Menge von Verarbeitungsschritten, die sukzessive abgearbeitet werden. Im IR wird so ein Prozess typischerweise durch Anwender (engl. *User*) initiiert, weil diese einen Bedarf nach relevanter Information haben. Etwas konkreter, aber sehr ähnlich der Definition von Baeza-Yates, definieren Manning et al. in die allgemeine Anforderung an das IR: „*Information-Retrieval (IR) is finding material (usually documents) of an unstructured nature (usually text) that satisfies an information need from within large collections (usually stored on computers)*“. Das folgende Kapitel wird zunächst noch einmal darauf eingehen, welche Bedeutung dem Informationsbedürfnis im MMIR zukommt.

1.3.1 Informationsbedürfnis und Informationsverhalten

Aus den obigen Definitionen wird deutlich, dass das Informationsbedürfnis eines Benutzenden Auslöser für die Anwendung eines IR-Prozesses ist: „*finding material [...] that satisfies an information need*“.

Manning et al. [33] beschreiben „*an information need is the topic about which the user desires to know more, and is differentiated from a query, which is what the user conveys to the computer in an attempt to communicate the information need*“. Das Informationsbedürfnis, das über einen IR-Prozess aufgelöst werden soll, entspricht demnach dem Gegenstand der Betrachtung, über welchen die Benutzenden eines IR-Prozesses in Unsicherheit sind bzw. ihr Wissen erweitern möchten. Aus der angeführten Beschreibung geht weiterhin hervor, dass das Informationsbedürfnis in Form einer Anfrage (engl. *Query*) formuliert und an den IR-Prozess übergeben werden kann. Informationsbedürfnis und Query müssen jedoch als zwei separate Eigenschaften betrachtet werden. Die Query ist hierbei nur das Transportmittel, um dem Informationsbedürfnis Ausdruck zu verleihen, d. h. das Informationsbedürfnis muss in eine Query transformiert werden.

Die Übersetzung bzw. die Formulierung der Query unterliegt dabei immer einer gewissen Unge-
wissheit oder Vagheit. Insbesondere, da es potenziell viele Möglichkeiten gibt, sich auszudrücken,
oder es gar nicht einfach möglich ist, den Bedarf in Worte zu fassen. Vor allem, wenn er vom
Menschen an eine Maschine kommuniziert (übersetzt) werden muss.

Ein bekanntes, grundlegendes Modell, welches das Informationsbedürfnis und die Interaktion eines Menschen mit einem Informationssystem beschreibt, ist das sog „Unnormale Wissenszustand“ (engl. *Anomalous States of Knowledge (ASK)* von N. Belkin [34] (siehe auch **Fehler! Verweisquelle konnte nicht gefunden werden.**). ASK ist ein problemorientiertes Modell und basiert auf der Annahme, dass ein menschlicher Erzeuger/Erzeugerinnen von Wissen und menschliche Nutzende dieses Wissens effizient miteinander austauschen möchten. Diese Annahme schränkt damit bewusst das Problemfeld auf die menschliche Kommunikation folgendermaßen ein:

- 1) Der Nutzende eines IR-Prozesses erkennt ein Informationsbedürfnis und übergibt eine entsprechende Suchanfrage an das System, um das Bedürfnis zu befriedigen.
- 2) Die Aufgabe des IR-Prozesses ist es, dem Benutzenden ein Ergebnis zu liefern, das höchstwahrscheinlich dem Bedürfnis entspricht. Ergebnis können hier beliebige Multimedia-Content-Objekt sein.
- 3) Der Benutzende untersucht das gelieferte Ergebnis, womit das Informationsbedürfnis zu einem bestimmten Grad (= Relevanz) gedeckt wird oder auch gar nicht adressiert werden konnte.

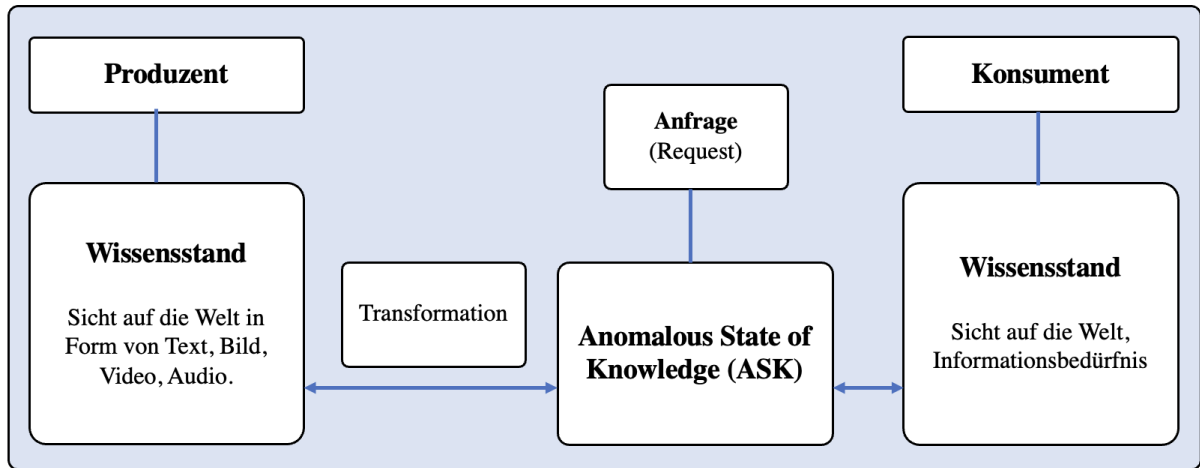


Abbildung 8 - Ein kognitives Kommunikationsmodell nach Belkin

Das Kommunikationsmodell, das ASK beschreibt, ist schematisch in **Fehler! Verweisquelle konnte nicht gefunden werden.** dargestellt. Auf der linken Seite befindet sich hierbei der Erzeuger/die Erzeugerin oder Produzent/Produzentin, welche ihren Wissensstand und ihre Sicht auf die Welt in Form von Text manifestiert haben. Auf der rechten Seite hingegen befindet sich der Benutzende oder Konsument/Konsumentin mit einem abweichenden Wissenszustand, welcher sich in einer Frage (Request) manifestiert.

Das Informationsbedürfnis und die Formulierung der Suchanfrage werden somit als zwei getrennte Konzepte betrachtet, denn die Suchanfrage versucht „lediglich“, den Bedarf der Benutzenden abzubilden. Modelle, welche die iterative Interaktion der Benutzenden mit dem IR-Prozess beschreiben, werden thematisch dem sog. Information Seeking oder auch Informationsverhalten zugeordnet. Hierbei können sechs Merkmale identifiziert werden [35]:

Merkmale	Systemdesign
Start (engl. Starting): Aktivitäten, die der Initiierung der Suche zugeordnet werden können.	Wird ein neues Projekt gestartet, sollen relevante Referenzen zum Thema verfügbar sein, die z. B. Schlüsselstudien referenzieren, einen Überblick verschaffen oder als Basis einer Verkettung dienen. Falls derlei Informationen nicht initial vorhanden sind, sollten diese über das IR verfügbar gemacht werden.
Verketten (engl. Chaining): Folgen von Referenzen (z. B. Fußnoten oder Literaturverzeichnis)	Das einfache Verfolgen von Referenzen soll unterstützt werden.

Durchstöbern (engl. Browsing): vage Suche im Bereich des potenziellen Interesses	Unterstützung von semistrukturierter Suche, z. B. über Namen von Autor/innen, Konferenzen oder Schlagworte.
Abgrenzung (engl. Differentiation): Ausnutzen von Unterschieden zwischen dem bereits untersuchten Material, um einen Überblick über die Eigenschaft und Qualität des Materials zu bekommen	Werden unterschiedliche Quellen in die Suche aufgenommen, kann eine Einteilung in relevante und weniger relevante Quellen hilfreich sein, um die Menge an verfügbarem Material zu verdichten.
Überwachung (engl. Monitoring): durch Überwachen einer bestimmten Quelle, über alle Neuerungen informiert bleiben	Das System sollte eine Liste an Quellen und Änderungen bereitstellen.
Extrahieren (engl. Extracting): systematische Analyse einer Quelle nach relevantem Material	

Ein Modell von Bates stellt den Menschen und seinen Werdegang in das Zentrum des Information-Seeking-Prozesses und betrachtet diesen Vorgang unter Berücksichtigung von Erkenntnissen aus der Kunde des Menschen (Anthropologie): „*Looking at us as a species that exists physically, biologically, socially, emotionally, and spiritually, it is not unreasonable to guess that we absorb perhaps 80 percent of all our knowledge through simply being aware, being conscious and sentient in our social context and physical environment*“ [36]. Bates erkennt verschiedene Modi, die Einfluss auf das Verhalten bei der Suche nehmen. Ein Kernelement ihres Modells ist dabei, dass Informationen nicht nur aktiv, sondern auch passiv über die Zeit aufgenommen werden (siehe auch **Fehler! Verweisquelle konnte nicht gefunden werden.**):

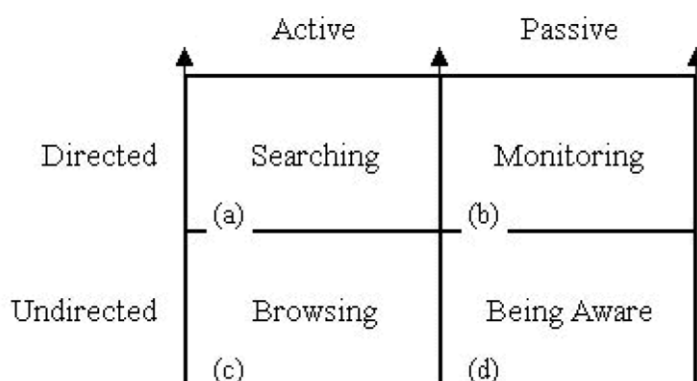


Abbildung 9 - Bates Information Seeking and Searching Model [36]

Bewusstsein (engl. Awareness): das, was wir aufnehmen und lernen über passives Verhalten

Überwachen (engl. *Monitoring*): Überwachen und Stöbern sind sich ergänzende Modi. Überwachen ist direkt und passiv, während Stöbern ungerichtet und aktiv betrieben wird.

Stöbern (engl. *Browsing*): Es gibt kein spezifisches Informationsbedürfnis, sondern es findet eine aktive Suche nach neuen Informationen statt.

Suchen (engl. *Searching*): ein aktives Bestreben, Antworten zu offenen Fragen zu finden oder ein Verständnis für einen neuen Sachverhalt zu bekommen.

Eine Studie von Hemminger et al. [37] zeigt, welchen Einfluss die allgegenwärtige Verfügbarkeit internetbasierter Angebote auf das Information-Seeking-Verhalten von akademischen Wissenschaftlern hat. Das Ergebnis zeigt, dass mit zunehmendem Angebot dieser Dienste diese auch zunehmend in Anspruch genommen werden. In Konsequenz führte dies zu weniger Besuchen einer Bibliothek und fast vollständiger Kommunikation über das Internet. Das Information Seeking hat somit einen großen Einfluss auf das IR und setzt das Verhalten der Benutzenden eines IR-Prozesses in den Vordergrund.

Zusammenfassend lässt sich hier sagen, dass ein Informationsbedürfnis typischerweise durch ein Informationsverhalten befriedigt werden kann. Unabhängig vom zugrunde liegenden Modell führt letzteres entweder zur aktiven Formulierung einer Anfrage oder zum Browsen/Filtern von möglichen Ergebnismengen, um das Informationsbedürfnis zu befriedigen. Dieses Grundprinzip stellt nunmehr die Basis für Information-Retrieval dar.

Selbsttestaufgabe 1.7:

Bringen Sie die folgenden Begriffe in eine sinnvolle Reihenfolge: Wissen | Entscheidung | Daten | Information | Lösung.

Selbsttestaufgabe 1.8:

Was versteht man unter dem Begriff „Anomalous State of Knowledge“?

1.3.2 Aufgaben des Information-Retrieval

In der IR-Literatur hat sich für die Umsetzung der Suche auch die Bezeichnung des „Ad Hoc Retrievals“ eingebürgert. Das Ad Hoc Retrieval wird als eine der übergeordneten *Aufgaben* (engl. *Retrieval Task*) im IR behandelt. Manning et al. [33] beschreiben den Charakter einer Retrieval Task als: „*the task executed by the information system in response to a user request. It is basically of two types: ad hoc and filtering*“. In einer Retrieval Task werden also gemeinhin die beiden Aufgaben des *Ad Hoc Retrievals* und des *Filterings* unterschieden:

- Das *Ad Hoc Retrieval* ist die Standardaufgabe des IR: Die Nutzenden eines IR-Prozesses spezifizieren hierbei ein Informationsbedürfnis mittels einer Suchanfrage, welche ein automatisiertes Verfahren startet. Die Ausgabe des Verfahrens besteht dabei in einer klassischen IR-Anwendung in einer Menge relevanter Dokumente mit textuellem Inhalt. Die Bezeichnung *Ad Hoc* leitet sich daraus ab, dass die Anzahl möglicher Fragen fast unbegrenzt ist.

- Das *Filtering* beschreibt die Aufgabe, bei der ein relativ statisches Informationsbedürfnis besteht, jedoch der Dokumentbestand sich beständig erweitert (z. B. das Durchsuchen von E-Mail-Listen).

Als die klassische Aufgabe des IR wird die Suche nach Volltextdokumenten verstanden. Im Kontext von Multimedia werden die Prinzipien des klassischen Information-Retrieval auf die Verarbeitung multimedialer Inhalte angewandt. Seien es Videos, Audiodateien, wissenschaftliche Analysedaten oder anderes mehr. In diesem Fall spricht man dann von Multimedia Information-Retrieval.

Selbsttestaufgabe 1.9:

Welche der Aussagen trifft für das Ad Hoc Retrieval zu?

- Es gibt eine Suchanfrage.
- Die Suche läuft unter Zuhilfenahme von Interaktionen von Benutzenden ab.
- Die Suche läuft automatisch ab.
- Es wird ein einzelnes Ergebnis zurückgegeben.
- Es wird eine Menge von Treffern zurückgegeben.

1.4 Multimedia-Information-Retrieval

Historisch gesehen baut das klassische „Information-Retrieval“ auf der Verarbeitung von textuellen Informationen auf. Dies lag schlicht an der Tatsache, dass zu frühen Zeiten der Informationsverarbeitung die Behandlung von multimedialen Inhalten noch nicht digital möglich war. Je mehr Multimedia-Content-Objekte im Laufe der Zeit entstanden und je mehr Standards verabschiedet wurden, desto größer wurde auch die Notwendigkeit, Multimedia in bestehenden Datenbanken oder IR-Systemen zu verwalten. Die Grundlagen dafür wurden u. a. von V.S. Subrahmanian gelegt, der in seinem Buch „Principles of Multimedia Database Systems“ [38] wichtige Transferleistungen von reiner Dokument-basierter Informationsverarbeitung hin zu multimedialen Inhalten geschaffen hat.

Da prinzipiell jedes MMIR-Merkmal auch eine textuelle Repräsentation besitzen kann, lassen sich auf diesem Weg bereits sämtliche Techniken des klassischen IR auch auf Multimedia anwenden. Jedoch gibt es in letzter Zeit mehr und mehr die Betrachtungsweise, dass textbasiertes IR eine Teilmenge des MMIR ist, da Texte nichts anderes als Multimedia-Content-Objekte sind, insbesondere da sie in zunehmendem Maße auch durch andere MM-Typen angereichert werden. Beispielsweise wird anhand einer textuellen Anfrage eine Audio-Datei gesucht oder anhand einer textuellen Eingabe nach Bildern recherchiert. Im Vergleich zu Textdokumenten zeichnen sich Multimedia-daten darüber hinaus durch ein großes Datenvolumen, Heterogenität der Medientypen und Speicherformate, Komplexität von dynamischen Multimedia-Typen, die insbesondere auch noch quasi beliebig kombiniert auftreten können – wie bei Filmen mit Ton – aus. Multimediale Anfragen können an ein MMIR auf verschiedene Art und Weise gestellt werden. Nach [39] können solche Anfragen zunächst auf ganz unterschiedliche Medien als Ergebnistypen abzielen. Eine geschriebene Anfrage kann zum Beispiel auf Textdokumente, auf gesprochenen Text, auf Bilder und auf Filmsequenzen zielen. In [39] liefert beispielsweise die Anfrage „John F. Kennedy in Berlin“ als Ergebnis zum einen ein Textdokument mit der Rede im Originalwortlaut, darüber hinaus aber

auch Bilder von dem Ereignis sowie Filmsequenzen, die Kennedy an verschiedenen Stationen seiner Berlin-Visite zeigen. Weitere gängige Typen von Beispielen sind z. B. die Suche nach Aufnahmen mit einem Sonnenuntergang oder die Suche nach Abbildungen von Kirchen in einer bestimmten Stadt. Die in relationalen Datenbanken verwendeten Anfrageparadigmen auf der Grundlage von formalsprachlichen Methoden wie „*Select... From... Where...*“-Ausdrücke sind auf das Multimedia-Information-Retrieval aufgrund der fehlenden Attributwerte jedoch nicht ohne weiteres übertragbar.

Die im vorigen Abschnitt genannten semantischen Konzepte spielen insbesondere im Bereich des MMIR eine sehr große Rolle, da die textuelle Repräsentation von MM-Merkmalen häufig nur dann erfolgen kann, wenn die Merkmale in einen korrekten semantischen Kontext gesetzt werden können. Die Analyse, ob die textuelle Merkmalsbeschreibung „Jaguar“ zu einem Auto oder einem Tier gehört, wirft bei Bildern, Videos und Audiodateien dieselben Fragestellungen auf, wie bei klassischen Textdokumenten. Nur, dass hier zusätzlich noch weitere MM-Merkmale genutzt werden können, um eine semantische Verortung zu erreichen. Derartige Verortungen können häufig im Multimedia-Umfeld sogar einfacher vorgenommen werden, da die zur Verfügung stehenden Informationen hierfür meist aus mehreren Medien abgegriffen werden können, wie der Social-Media-Bereich eindrucksvoll demonstriert. Hier können Texte, Bilder, Videos, Sprachnachrichten, Kommentare, Interaktionen (wie z. B. Likes) genutzt werden, um den semantischen Kontext von Informationen zu bestimmen.

Im Rahmen dieser Lehrveranstaltung wird daher keine Abgrenzung zwischen „klassischem“ Information-Retrieval und Multimedia-Information-Retrieval vorgenommen. Stattdessen werden die Techniken des „klassischen“ IR im größeren Kontext des MMIR angewandt und ggfs. adaptiert.

1.4.1 Multimedia-Informationssysteme

Die in den vorherigen Abschnitten vorgestellten Konzepte aus dem Bereich des MMIR können nun in den größeren Kontext der Multimedia-Informationssysteme eingegliedert werden. Hierbei spielen auch Themen, wie Archivierung, Historisierung, Zugriff, Verteilung, Speicherung und Bereitstellung von MM-Content-Objekten eine wichtige Rolle. Nicht nur die Größe und Heterogenität großer multimedialer Informationsmengen, sondern insbesondere auch die Vielfalt der unterschiedlichen Bedürfnisse der Benutzenden im Bereich der effektiven und effizienten Nutzung von Multimedia-Kollektionen stellen auch heute noch sehr große Herausforderungen dar und sind nach wie vor aktueller Gegenstand intensiver Forschungs- und Entwicklungsaktivitäten.

Im Vordergrund stehen dabei Fragestellungen wie: „Welche Erwartungen haben die Benutzenden an die Inhaltserfassungs-, Anfrageunterstützungs- und Suchfunktionen einer multimedialen Datenbank? Wie gut erfüllen die von einem multimedialen Informationssystem gelieferten Ergebnisse die Ziele der Recherche bzw. das zugrundeliegende Informationsbedürfnis der jeweiligen Benutzenden und damit der tatsächlich gesuchten Information in ihrer jeweiligen multimedialen Ausprägung?“. Es ist offensichtlich, dass manuell erstellte Beschreibungen der multimedialen Dateninhalte vom Aufwand her, und damit in den allermeisten Anwendungsfällen aus kostenorientierten Gründen nicht praktikabel sind. Es werden daher derzeit insbesondere neue Verfahren zur

Erschließung, Verwaltung und Suche von Multimedia-Daten benötigt, die möglichst weitgehend automatisch ablaufen können. Multimedia-Informationssysteme helfen, diese Aufgaben zu lösen.

- Ein Multimedia-Informationssystem verfügt typischerweise über folgende Komponenten:
- Repräsentationskomponente der jeweiligen Multimedia-Content-Objekte, meist mit einer formalen Beschreibung der im jeweiligen Content Objekttyp enthaltenen Informationen.
- Anfragerepräsentationskomponente zur Unterstützung der formalen Beschreibung des Informationsbedürfnisses eines Benutzenden.
- Suchkomponente mittels der nach bestimmten Regeln mithilfe von Pattern-Matching-Prozeduren MM-Content-Objekte gefunden werden, die für die Anfrage der Benutzenden relevant sind.
- Speicher- und Verteilungskomponenten, die die unterschiedlichen MM-Content-Objekte physikalisch speichern, verteilen, konvertieren oder archivieren.
- Prozessrepräsentationskomponente, in der die unterschiedlichen MMIR-Prozesse abgebildet werden.
- Datenrepräsentationskomponente, die das Datenmodell der im MMIS gespeicherten Informationen beschreibt.

Die Struktur der über ein MMIS zugreifbaren Daten wird analog zur klassischen Datenbanktheorie über ein Datenmodell beschrieben. Kemper et al. definieren dies in dieser Weise: *„Das Datenmodell legt die Modellierungskonstrukte fest, mittels der man ein computerisiertes Informationsabbild der realen (bzw. des relevanten Ausschnitts) der Welt generieren kann[...] es legt die generischen Strukturen und Operatoren fest [!], die man zur Modellierung einer bestimmten Anwendung ausnutzen kann.“* [40]

Ein MMIS lässt sich damit insofern als ein System beschreiben, das auf Basis von Daten, die auf der Spezifikation eines Datenmodells beruhen, Information für den Zugriff der Benutzenden bereitstellt (siehe auch **Fehler! Verweisquelle konnte nicht gefunden werden.**). Aus Sicht der Benutzenden lassen sich drei zentrale Komponenten identifizieren: die Systemeingabe, die Verwaltung und die Systemausgabe.

Die Eingabe in das MMIS kann in drei Bereiche unterteilt werden. Der erste Bereich ist der, welcher die Query der User/innen beschreibt. Eine gängige Form der Übermittlung einer Query wird über die Angabe von Zeichenketten vorgenommen, wobei die Zeichenketten Bestandteil einer formalen Sprache (u. a. SQL [41] oder SPARQL [42]) sein können oder umgangssprachlich. Dies könnten aber z. B. auch im Falle des Musik-Retrievals Melodien sein, die vorgesummt werden, oder ein Bild im Falle der Suche nach ähnlichen Bildern.

Passend zur Query muss das MMIS auf eine Dokument-/Medienobjektsammlung zugreifen können, gegen die es die Query sinnhaft abbilden kann. Dieser Teil der Lehrveranstaltung ist zunächst

vorwiegend auf die Verarbeitung von Dokumenten mit Volltexten beschränkt. Potenziell könnten die Objekte der Sammlung aber, analog zur Verarbeitung von Text, im MMIR jeglichen multimedialen Inhaltstyp haben.

Neben der Query und den zu durchsuchenden Elementen wird zusätzlich die Berücksichtigung externer Wissensquellen benannt. Diese externen Wissensquellen können bestimmte Datenbanken, aber auch menschliche Expertise sein. Das über diese Quellen verfügbare Wissen wird u. a. genutzt, um Querys zu optimieren (z. B., indem die Bedeutung eines Wortes einer umgangssprachlichen Query festgelegt wird, z. B. „Jaguar“, Automarke oder Tier?). Auch wird dieses Wissen genutzt, um den glatten Text eines Dokuments der Dokumentsammlung in eine für das Retrieval verwertbare Form zu überführen. Dieser Vorgang wird als Indexieren bezeichnet.

Die Verarbeitungskomponente bekommt die Eingabe und verarbeitet diese für unterschiedliche Zwecke. Bevor eine Query überhaupt verarbeitet werden kann, müssen jedoch erst alle Elemente der Dokumentsammlung in eine für die Retrievalfunktion verständliche Repräsentation überführt worden sein. Der sog. Indexierungsvorgang untersucht dafür den Inhalt der Elemente der betrachteten Dokumentsammlung und überführt diesen, unter Umständen unter Zuhilfenahme externen Wissens, in eine entsprechende Repräsentation (den Index). Der Index wird in einem für das MMIS zugänglichen Speichermedium abgelegt.

Auch die Query muss gegebenenfalls um externes Wissen angereichert werden und dann in eine für die Retrievalfunktion verständliche Repräsentation überführt werden. Die Retrievalfunktion bekommt Repräsentationen der Query und der Elemente von Dokumenten und vergleicht diese hinsichtlich ihrer Relevanz. Gegebenenfalls wird das berechnete Ergebnis noch sortiert und danach an die Systemausgabe weitergegeben.

Die Systemausgabe visualisiert das Ergebnis der Verarbeitungskomponente für die User/innen und reichert dieses gegebenenfalls noch um weitere Informationen an.

Im Folgenden werden nun überblicksartig formale Modelle beschrieben, die in der Regel den unterschiedlichen Realisierungen der verschiedenen Komponenten als gemeinsames Beschreibungsmittel zugrunde liegen. Diese unterscheiden sich unter anderem dadurch, ob sie Dokumente nach dem Ansatz des Exact Match, also der genauen Übereinstimmung, oder nach dem Prinzip des Best Match, also der bestmöglichen Übereinstimmung, als relevant selektieren. Die bekanntesten Vertreter sind für die erste Methode das boole'sche Information-Retrieval-Modell, für die letztere das Vektorraum- sowie das probabilistische Information-Retrieval-Modell. Zwischen den genannten Modellen existiert jedoch auch eine Reihe von Mischformen.

Selbsttestaufgabe 1.10:

Welche Komponenten eines MMIS kennen Sie?

1.5 *Übersicht der zentralen MMIR Modelle*

Im Themengebiet Information-Retrieval (bzw. Multimedia-Information-Retrieval) haben sich eine Reihe von Modellen für die Berechnung oder Durchführung verschiedener Aufgaben entwickelt. In diesem Abschnitt werden die zentralen Modelle eingeführt und vorgestellt.

1.5.1 *Exact-Match-Modelle*

Das boole'sche Information-Retrieval ist das älteste und am weitesten verbreitete Information-Retrieval-Modell [43]. Es basiert auf exakter Erfüllung einer Anfrage durch ein Dokument/Medienobjekt. Anfragen werden mithilfe der Standardoperatoren boole'sche Logik formuliert. Spezifikationen der Suchanfrage werden mit den Inhaltsrepräsentationen der Datenbasis verglichen und nur jene Dokumente bzw. deren Stellvertreter als Treffer präsentiert, welche diesen Spezifikationen genau entsprechen. Eine darüber hinausgehende Quantifizierung der Relevanzwerte der einzelnen Dokumente/Medienobjekte der Ergebnismenge wird nicht durchgeführt. Die Datenbasis wird bezüglich einer Anfrage in zwei Untermengen unterteilt: in die Menge der als voll zutreffend zugelassenen und in die Menge der als nicht zutreffend verworfenen Dokumente/Medienobjekte.

Eine Anfrage besteht dabei aus Termen. Im Text-Retrieval sind dies in der Regel Anagramme, Suchworte, Phrasen, oder wiederum Dokumentfragmente anderer Granularität, die durch boole'sche Operatoren verknüpft werden. Ein Dokument erfüllt die Anfrage, wenn es die Terme in der geforderten prädikatenlogischen Konstellation enthält. Im Falle von anderen Multimedia-Content-Typen wird dieses Prinzip häufig auf die textuelle Repräsentation der MM-Merkmale angewendet.

Gehen wir zum Beispiel von den Operatoren AND, OR und NOT aus, und nehmen an, dass a, b, g und d Terme sind. Die Anfrage $(a \text{ OR } b \text{ OR } g) \text{ AND } (\text{NOT } d)$ wird von allen Dokumenten erfüllt, die a enthalten und b oder g enthalten und d nicht enthalten. Eine Unterscheidung der Dokumente/Medien-Objekte in der Ergebnismenge findet nicht statt, d. h. es wird keinerlei Unschärfe der Beschreibung bei der Relevanzbewertung der Dokumente/Medienobjekte innerhalb oder außerhalb der Ergebnismenge berücksichtigt. Das führt dann beispielsweise dazu, dass für die Benutzenden nicht mehr zu erkennen ist, ob ein Dokument/Medienobjekt außerhalb der Ergebnismenge nur einen Term oder alle Terme nicht erfüllt. Aufgrund der Analogien zu Mengen und ihren Operatoren können boole'sche Anfragen als Mengenoperationen auf Termmengen angesehen werden.

In einem boole'schen IR-System wird jedes Dokument/Medienobjekt durch einen Satz von inhaltlichen Indizes repräsentiert, der sich aus Termen des für die Indexierung gewählten Vokabulars zusammensetzt. In der Regel erfolgt auch die Indexierung ungewichtet, d. h. ein Term wird entweder als in einem Dokument/Medienobjekt vorhanden oder nicht vorhanden vermerkt. Normalerweise sind die Indizes von häufig vorkommenden Stopwörtern bereinigt. Haben Benutzende eine Anfrage eingegeben, wird sie mit den Repräsentationen der Dokumente/Medienobjekte in der Datenbasis verglichen, und die berechnete Ergebnismenge wird präsentiert. Dieser Prozess kann auch wiederholt unter Veränderung der Anfrage ablaufen. Häufig wird die Formulierung der

Anfrage durch funktionale Erweiterungen der Anfragesprache, wie etwa die Ausweitung der Anfrage durch Suchterm-Trunkierung oder Einführung von Bereichs- oder Größenoperatoren, unterstützt. Nicht selten ist auch eine Erweiterung oder Einengung der Suche durch die Möglichkeit der Einbindung eines Thesaurus oder strukturierten Wörterbuches sowie durch zusätzliche Operatoren oder Makro-Operatoren wie $a \text{ XOR } b$, das ja bekanntermaßen eine Vereinfachung von $(a \text{ AND } (\text{NOT } b)) \text{ OR } (b \text{ AND } (\text{NOT } a))$ darstellt.

Trotz seiner weiten Verbreitung besitzt dieses Information-Retrieval-Modell einige bedeutsame Nachteile, die sich in größeren Datenbasen sowie bei steigenden Ansprüchen und sinkenden Kenntnissen der Benutzenden bemerkbar machen. Boole'sche Systeme bieten Benutzenden, die sich mit den verwendeten Indexierungstermen auskennen, bzw. die bereits früher im gleichen Sachgebiet recherchiert haben, ein effektives Retrieval. Die Effizienz sinkt jedoch rapide mit dem steigenden Maß der Unvertrautheit der Benutzenden mit einem Sachgebiet. Das Modell erweist sich darüber hinaus hinsichtlich der Information-Retrieval-Effizienz in bestimmten Fällen anderen Systemen unterlegen. Nachteilig wirkt sich beispielsweise die „Alles-oder-Nichts“-Eigenschaft der boole'schen Suche aus, da sich die maximale Anzahl der als relevant angezeigten Dokumente/Medienobjekte nicht je nach Menge des insgesamt vorhandenen Materials variieren lässt. So kann diese fehlende Kontrolle über die maximale Größe des Ergebnisses dazu führen, dass bei einer zu allgemeinen Suchformulierung eine zu große Menge von Dokumenten/Medienobjekten aus einer entsprechend großen Datenbasis als für die Suchanfrage zutreffend herausgefunden werden. In solchen Fällen muss über Umformulierungen und weitere Einschränkungen die Suche eingegrenzt werden. Schwierig hierbei ist es insbesondere für Benutzende ohne Expertise bezüglich der Anwendung und Steuerung der Anfragekonstruktion die Auswirkungen der logischen Operatoren in allen Konsequenzen zu verstehen.

Die einfache Klasse boole'scher Anfragen besteht aus einem einzigen Suchterm. Eine solche Anfrage ist wahr für Dokumente/Medienobjekte, die mit diesem Term indexiert wurden. Aus einer abstrahierenden Sicht gesehen besteht in diesem Fall zwischen der Anfrage und dem Ergebnis eine ungewichtete und ungeordnete 1:N-Relation, die von den Benutzenden kognitiv zu erfassen ist. Komplexere boole'sche Anfragen entstehen durch die Kombination mehrerer Terme mittels der oben genannten Operatoren, deren Wahrheitsgehalt kombiniert überprüft wird. In diesem Fall stehen Anfrage und Ergebnis also in einer ungewichteten und ungeordneten M:N-Beziehung. Im boole'schen Modell wird implizit von der Annahme ausgegangen, dass die Benutzenden in der Lage sind, ihr Informationsbedürfnis vollständig und präzise als logischen Ausdruck, der kompatibel zur Indexierungspraxis des Systems ist, zu formulieren. Darüber hinaus wird davon ausgegangen, dass die Benutzenden die MxN-Relation zwischen den Komponenten einer komplexen Anfrage und denen einer Ergebnismenge erfassen und verstehen können. Dies ist bei naiven Benutzenden jedoch nicht unbedingt der Fall. Eine gewisse Vertrautheit mit dem Vokabular der in der Datenbasis gespeicherten Dokumente/Medienobjekte ist zudem unverzichtbar, da eine sehr exakte Formulierung der Anfrage unter Einbeziehung möglicher Synonyme oder Wortstämme über Erfolg oder Misserfolg entscheiden kann. Eine Vertrautheit mit dem Vokabular ist besonders wichtig, wenn z. B. in Volltextdatenbanken ein komplexer Sachverhalt auf einige wenige Suchwörter abgebildet werden muss.

Ein weiterer Nachteil ist die Schwierigkeit, nur partiell relevante Dokumente/Medienobjekte aufzufinden. Dieser lässt sich nur über nacheinander eingegebene verschachtelte Anfragen umgehen. Jede neue Formulierung teilt wiederum die Menge der Dokumente/Medienobjekte in eine Untermenge, auf welche die Suchbedingung zutrifft und in eine Untermenge, für die das nicht der Fall ist. Da zudem jeder der Komponenten der Suche das gleiche Gewicht zugemessen wird, ist die Ausgabe der aufgefundenen Dokumente in Form einer geordneten Liste, welche die zutreffendsten Dokumente an oberster Stelle präsentiert, nicht möglich. Einige neuere Systeme versuchen, dem Abhilfe zu schaffen, indem bestimmten Termen der Suchanfrage Gewichte zugeordnet werden können, anhand derer die Ergebnisse geordnet und in der Reihenfolge ihrer Relevanz ausgegeben werden. Bei Extended-Boolean-Retrieval-Systemen bilden geordnete 1xN- bzw. MxN-Relationen die Grundlage der Beziehung zwischen Anfrage und Ergebnis. Es gab auch Versuche, die Ergebnisse einer boole'schen Anfrage anhand des Überschneidungsgrades zwischen Dokument-/Medienobjektinhalt und den Termen der Suchformulierung einem Gewichtungsprozess zu unterwerfen. In solchen Systemen werden die Dokumente/Medienobjekte erst nach boole'schen Kriterien extrahiert und dann in einem zweiten Durchlauf gewichtet.

1.5.2 Best-Match-Modelle

Bei den auf bestmöglicher Erfüllung der Anfrage beruhenden Best-Match-Modellen werden die Vektorraum-, die Wahrscheinlichkeitstheorie und die Theorie der Fuzzy-Logik bemüht, um Modelle zu entwickeln, welche die Nachteile des boole'schen Information-Retrieval-Modells vermeiden. Hierbei sind zwar die Information-Retrieval-Funktionen und die Anfrageformulierungen die gleichen wie beim herkömmlichen boole'schen Retrieval, jedoch werden auch gewichtete Dokument-/Medienobjektbeschreibungen und Anfrageformulierungen zugelassen. Dahinter steht der Gedanke, dass es tatsächlich häufig schwierig ist, zu entscheiden, ob ein Dokument/Medienobjekt mit einem bestimmten Term indexiert werden sollte oder nicht.

Der Indexierer muss bei Implementierungen des Best-Match-Modelles nach einem auf unscharfer Logik (engl. Fuzzy logic) basierenden Algorithmus entscheiden, inwieweit ein Term ein Dokument charakterisiert und dies durch einen Wert zwischen 0 und 1 ausdrücken. Zentrale Frage dieses Modells ist also die Bewertung der Relevanz, die ein Term für ein Dokument/Medienobjekt hat. Mittels einer solchen Information-Retrieval-Funktion kann für jedes Dokument ein separater Wert berechnet und eine Rangordnung der Ausgabedokumente/-medienobjekte erstellt werden.

Dieser Vorteil wird jedoch dadurch relativiert, dass viele Nachteile des boole'schen Information-Retrievals beibehalten werden, bei dem eine Gewichtung der Terme nicht vorgesehen ist. Nach wie vor ist die Formulierung der Anfrage für den naiven Benutzer schwierig. Weiterhin ist es nachteilig, dass die Termgewichte willkürlich zugeordnet werden und somit abhängig von der subjektiven Einstufung eines Indexierers sind - was umso schwerer ins Gewicht fällt, als sie nicht dynamisch veränderbar sind.

Im Best-Match-Retrieval wird die Forderung nach exakter Erfüllung der Anfrage aufgegeben und dafür eine Sortierung der Dokumente nach dem Grad ihrer Erfüllung der Anfrage eingeführt. Durch die Beispielanfrage (a OR b OR g) AND (NOT d) könnte durchaus ein Dokument/Medienob-

jekt gefunden werden, in dem der Term d , trotz der Forderung $\text{NOT}(d)$, vorkommt. Dies ist sinnvoll, wenn d als wenig relevant eingestuft und der Rest der Anfrage mit hoher Übereinstimmung erfüllt wird.

1.5.3 **Ranking-Modelle**

Bei den Modellen, die sich unter dem Begriff der Best-Match-Modelle einordnen lassen, wird im Gegensatz zum boole'schen Modell nicht länger gefordert, dass die Spezifikationen einer Anfrage exakt erfüllt werden, sondern es wird eine geordnete Auswahl von Dokumenten/Medienobjekten aus der Datenbasis berechnet, wobei die Ergebnisdokumente in der Reihenfolge präsentiert werden, in der sie als für die Anfrage relevant berechnet wurden. Die derart geordneten Dokumente/Medienobjekte können dabei auf zwei Arten geordnet werden: beim **Weak Ordering** können zwei Dokumente/Medienobjekte den gleichen Wert bzw. Rang einnehmen, beim **Linear Ordering** müssen Unterschiede vorhanden sein – üblich ist das Weak Ordering.

Ein typisches Maß für Ähnlichkeit ist die Anzahl der sowohl in der Anfrage als auch im jeweiligen Dokument/Medienobjekt vorkommenden Terme oder textuellen Repräsentationen von MM-Merkmalen. Bei einem so implementierten System ist das Ergebnis eine geordnete Liste, in der diejenigen Dokumente/Medienobjekte, welche die meisten Terme mit der Anfrage gemeinsam haben, an erster Stelle stehen und so dem Benutzenden als erste präsentiert werden. Je nach Ansatz wird dabei gegebenenfalls noch eine Gewichtung der indizierten Terme berücksichtigt. Diese Gewichtung erfolgt in vielen Fällen anhand von statistischen Häufigkeitswerten und beruht beispielsweise darauf, wie oft ein Term in der gesamten Datenbasis vorkommt, möglicherweise im Verhältnis zu allen dort gespeicherten Termen. Ein wichtiger statistischer Wert ist dabei die inverse Dokumenthäufigkeit. Ein hoher Wert ergibt sich daraus, dass ein Term sehr oft in einem Dokument/Medienobjekt, jedoch eher selten in der Datenbasis vorkommt. Ein weiterer wichtiger statistischer Wert ist der Term-Diskriminationswert. Dieser ist ein Maß für die Eigenschaft eines Terms, ein Dokument/Medienobjekt von anderen unterscheidbar zu machen. Je seltener also ein Term in der Datenbasis vorkommt, umso höher wird dieser Wert liegen. Die Implementierung eines solchen Algorithmus setzt voraus, dass die Größe der Dokumenten-/Medienobjektkollektion sowie die Anzahl der Dokumente/Medienobjekte, die von jedem Term indiziert werden, bekannt ist. Gerade in einer interaktiven Umgebung, in der die zuvor gefundenen Objekte häufig der Umformulierung der ursprünglichen Anfrage dienen, ist es besonders wichtig, mithilfe von Ranking-Verfahren Unterscheidungen zwischen den aufgefundenen Dokumenten/Medienobjekten zu treffen. Falls die Anfrageformulierung stark bedeutungstragende Begriffe enthält, ist die Wahrscheinlichkeit recht hoch, dass an oberster Stelle der Ergebnisliste Dokumente/Medienobjekte hoher Relevanz stehen. Vorteil dieser Methode ist, dass die Suche mit einer ungeordneten Liste von Schlüsselwörtern durchgeführt werden kann. Es ist nicht nötig, boole'sche Operatoren gezielt anzuwenden oder über die logischen Konsequenzen von Formulierungen nachzudenken. Auch die Kontrolle der Größe des Ergebnisses stellt kein Problem dar, da es den Benutzenden freigestellt ist, wie viele der gefundenen Dokumente/Medienobjekte sie sich anzeigen lassen wollen. Es ist außerdem nicht länger erforderlich, alle Dokumente/Medienobjekte mit allen angeführten Kriterien anzuzeigen, denn es kann ein Schwellwert festgelegt werden.

In einigen Systemen, die das Retrieval auf diese Art und Weise durchführen, wird auch eine boole'sche Suchmethode angeboten, um geübten Benutzenden die Formulierung logischer und präziser, eventuell mehrfach verschachtelter Anfragen zu ermöglichen. Zusätzlich können Rückkopplungsmechanismen realisiert sein, die bei einem zweiten Suchlauf auf der Grundlage von Relevanzbewertungen des Benutzenden zu den Trefferdokumenten die Termgewichte der Anfrage neu berechnen und so zu einem veränderten Ergebnis führen. In Versuchen hat sich gezeigt, dass die Einbeziehung zusätzlicher Suchterme anhand ihres Vorkommens in bereits als relevant beurteilten Dokumenten/Medienobjekte zu einer Verbesserung der Information-Retrieval-Ergebnisse führt. Bei diesem Verfahren erwies sich die Gewichtung nach Relevanz derjenigen überlegen, die lediglich anhand statistischer Gesichtspunkte wie der Vorkommenshäufigkeit des entsprechenden Terms in der Dokumenten-/Medienobjektsammlung vollzogen wurde. Diese war jedoch wiederum der gänzlich ungewichteten Suche überlegen. Die Anzahl und Qualität der nach Relevanz gewichteten Suchterme hatte einen deutlichen Einfluss auf die Systemperformanz.

1.5.4 Vektorraum-Modelle

Auf der Grundlage des Best-Match-Ansatzes wurden einige weitere Modelle entwickelt. Als Überbegriffe sind hier die Modelle des Vektorraum- und des probabilistischen Information-Retrieval zu nennen. Vektorbasierte Modelle sind im Bereich des Best-Match-Retrieval weit verbreitet. Grundlage dieses Modells ist die Idee, dass ein Dokument als Vektor in einem mathematischen Raum, dem Inhaltsraum der Dokument-/Medienobjektmenge, aufgefasst werden kann. Die inhaltliche Verwandtschaft von Dokumenten/Medienobjekten untereinander, oder bezüglich einer Anfrage, kann durch den numerischen Abstand von Vektoren zum Ausdruck gebracht werden. Zentrale Eigenschaft dieses Modells ist die Berücksichtigung der Relevanz eines Begriffes bezüglich eines Dokumentes/Medienobjektes und der Relevanz eines Suchbegriffs bezüglich der Anfrage. Dem Vektorraummodell liegt die Annahme zugrunde, dass der euklidische Begriffsraum existiert, in dem die Inhalte von Dokumenten/Medienobjekten und Anfragen als einzelne Punkte lokalisiert sind. Das System antwortet auf eine Anfrage, indem es die Dokumente/Medienobjekte präsentiert, die räumliche Nähe zur Anfrage aufweisen. Eine Möglichkeit der Implementierung solcher Systeme ist die Repräsentation eines jeden Dokumentes/Medienobjektes und jeder Anfrage durch einen Indexvektor, der jeweils die Dokument-/Medienobjektterme mit dem ihnen zugemessenen Gewicht enthält. Verschiedene Entfernungsberechnungen sind möglich, wobei häufig das Kosinus-Maß verwendet wird. Ein Vorteil dieser Entfernungsberechnung ist, dass die gewonnenen Werte gleichzeitig für ein gewichtetes Ranking herangezogen werden können. Hierbei wird von der Annahme ausgegangen, dass das dem Anfragevektor räumlich am nächsten stehende Dokument/Medienobjekt zugleich das relevanteste für die jeweilige Anfrage ist. Diese Systemeigenschaft, die für den Benutzenden eine Hilfe bei der Beurteilung der gefundenen Dokumente/Medienobjekte darstellt, kann auch in Systemen eingesetzt werden, die auf dem boole'schen Information-Retrieval-Modell aufbauen. Um den Dokumenten-/Medienobjektraum tatsächlich räumlich interpretieren zu können, wird er als orthogonal angenommen, d. h. alle Termvektoren werden als linear unabhängig betrachtet und sind normiert. Für den Benutzenden ist der vieldimensionale Raum, der von den Dokument-/Medienobjekttermen aufgespannt wird, jedoch nur schwer vorstellbar, sobald er über die dritte Dimension hinausgeht. Ein Vektorenvergleichssystem führt Vergleiche zwischen Anfrage- und Dokument-/Medienobjektvektoren durch und präsentiert

ein geordnetes Ergebnis in absteigender Reihenfolge der berechneten Ähnlichkeitswerte. Dies bietet nicht nur die Möglichkeit einer Kontrolle der Ergebnisgröße, sondern es werden auf diese Weise auch iterative Maßnahmen der Anfrageverfeinerung ermöglicht.

Für die Generierung der Vektoren stehen mehrere Möglichkeiten zur Auswahl. So kann zum einen lediglich auf binäre Art und Weise entschieden werden, ob ein Term in einem Dokument/Medienobjekt vorkommt oder nicht. Nach dieser Methode besitzen alle Vektoren einer Datenbasis die gleiche Dimensionalität, und es wird lediglich vermerkt, ob ein Term in einem Dokument/Medienobjekt vorhanden ist oder nicht. Die Alternative besteht in der Berechnung von Termgewichten. In der einfachsten Form wäre dies eine Zählung der Vorkommenshäufigkeit der einzelnen Terme im jeweiligen Dokument/Medienobjekt. Auf der Vorkommenshäufigkeit der einzelnen Terme beruhen auch ausgefeiltere Berechnungen von Termgewichten, die im einfachsten Fall lediglich eine Normalisierung der Werte des oben beschriebenen Vektors erbringen, meist jedoch auf komplexeren Algorithmen beruhen und beispielsweise die Vorkommenshäufigkeit eines Terms in der gesamten Datenbasis mit einbeziehen. Je näher in einem solchen Fall das Gewicht des Terms bei 1 liegt, desto bedeutungsschwerer ist er. Aufgrund der räumlichen Interpretation der Dokument-/Medienobjektrepräsentation als Vektoren können als Information-Retrieval-Funktion Ähnlichkeitsberechnungen zwischen Dokument-/Medienobjekt- und Anfragevektoren vorgenommen werden. Schmitt [48] nennt verschiedene Ähnlichkeitsmaße, die zur Berechnung der Anfrage-Dokument/Medienobjekt-Ähnlichkeit in der Literatur vorgeschlagen wurden.

Vorteilhaft erweist sich die Einfachheit der Anfrageformulierung, bei der selbst eine einfache Aneinanderreihung von Termen zu Ergebnissen führt. Als problematisch beurteilt wird dagegen die Gleichbehandlung aller Texte ohne Abklärung der Frage, inwieweit beispielsweise das Gewicht eines Terms in einem Referat zu einem bestimmten Thema anders bestimmt werden sollte als dasjenige des gleichen Terms in einem längeren Volltext. Eine nachträgliche Umgewichtung ist kaum möglich, was ebenfalls einen Nachteil darstellt. Darüber hinaus kann es vorkommen, dass zwei völlig verschiedene Dokumente/Medienobjekte, die zufällig die gleiche Entfernung zur Anfrage aufweisen, den gleichen Rang haben, obwohl sie im Dokumenten-/Medienobjektraum völlig unterschiedliche Positionen einnehmen. Für das Retrieval bedeutet dies, dass einander ähnliche Dokumente/Medienobjekte mit großer Wahrscheinlichkeit für dieselbe Anfrage relevant sind und dass nach dem Auffinden eines relevanten Dokumentes/Medienobjektes ähnliche Dokumente gesucht und so weitere Treffer erzielt werden können. Im Sinne des Vektorraummodells lässt sich hierbei inhaltliche Ähnlichkeit als räumliche Nähe interpretieren und mittels der Dokument-/Medienobjektvektoren berechnen und numerisch ausdrücken.

1.5.5 Weitere Modelle

Neben den hier vorgestellten Modellen gibt es eine Reihe von weiteren Modellen, die jedoch im Rahmen dieser Lehrveranstaltung nicht vertieft werden. Beim **Clustering-Modell** werden beispielsweise Häufungen von Dokumenten/Medienobjekten genutzt, um auf Basis eines Ähnlichkeitsmaßes den Abstand eines Dokuments von einem Cluster zu ermitteln und somit eine Vorausswahl bei umfangreicherem Retrieval zu treffen. **Probabilistische Modelle** schätzen mit Mitteln der Wahrscheinlichkeitsrechnung ab, wie gut ein Ergebnis das Interesse der Benutzenden erfüllt und weichen damit grundlegend vom boole'schen Modell ab. In probabilistischen Modellen wird

angenommen, dass sowohl in den Repräsentationen der Anfrage als auch der Dokumente/Medienobjekte Unsicherheiten liegen, die sich auf die zwischen ihnen bestehende Relevanzbeziehung auswirken und deren Berechnung ungenau werden lassen. **Inferenz-Modelle** beziehen zusätzlich zu Wahrscheinlichkeiten auch noch Einzeleigenschaften der Dokumente/Medienobjekte in die Suche mit ein. Somit können sowohl transitive Eigenschaften als auch maschinelles Lernen besser unterstützt und abgebildet werden.

Selbsttestaufgabe 1.11:

Welche Eigenschaften werden beim „Exact-Match“-Modell gefordert?

1.5.6 Anfragekonstruktion

Die Formulierung des Informationsbedürfnisses ist für Benutzende häufig eine Herausforderung. Gleiches gilt für die Interpretation der Anfrage eines Benutzenden im MMIR bzw. MMIS. Dies liegt daran, dass an dieser Stelle eine Übersetzung des Informationsbedürfnisses in technische Verfahren, wie z. B. Indexstrukturen, erfolgen muss und der Benutzende einerseits nicht weiß oder wissen kann, welche Strukturen das MMIR-System erfordert, das MMIR-System andererseits aber nicht davon ausgehen kann, dass Benutzende in der Lage sind, speziell optimierte und ggfs. syntaktisch korrekte Anfragen zu formulieren. Daher haben sich eine Reihe von Standard-Verfahren entwickelt:

- **Stichwort-basierte Anfragen:** Die Verwendung einzelner Stichworte, wie z. B. bei einer Internet-Recherche via Google, hat sich als sehr praktikabel aus Sicht der Benutzenden herausgestellt. Die Benutzenden geben lediglich ein oder mehrere textbasierte Stichworte an, um ihr Informationsbedürfnis zu beschreiben und überlassen dem System die Übersetzung auf technische Strukturen.
- **Anfragekonstruktion mit syntaktischen Regeln:** Um Zusammenhänge oder Regeln zwischen den eingegebenen Stichwörtern zu spezifizieren, stellen einige MMIR- oder MMIS-Systeme syntaktische Erweiterungen zur Verfügung. So kann dann bspw. ein Informationsbedürfnis weiter konkretisiert oder eingeschränkt werden, um die Menge der Ergebnisse zu reduzieren. Die Anfrage „*+(+Obama +Barak) -Michelle) +(New York))*“ würde in diesem Fall alle Ergebnisse mit Barack Obama ohne Michelle und einer Verbindung zu dem Begriff New York beschreiben.
- **Anfragesprachen:** Für komplexere Systeme und entsprechend im Umgang ausgebildete Benutzende werden häufig auch sog. Query Languages (QL) bereitgestellt. Der gängigste Vertreter hierfür ist die Structured Query Language (SQL) aus dem Datenbankumfeld und erlaubt in klar definierten syntaktischen Regeln die Spezifikation komplexer Anfragen. Ein weiterer, gängiger Vertreter ist SPARQL, eine Anfragesprache, mit der semantische Beziehungen beschrieben werden können.
- **Query by Example:** Speziell im Multimedia-Kontext gibt es häufig auch das Anfragemuster nach Beispiel. In diesem Fall wird eine ähnlichkeitsbasierte Suche durch Angabe eines Beispiel-Objektes spezifiziert. Bekannt und praktisch ist dieses Muster beispielsweise bei der Suche nach ähnlichen Bildern oder Musikstücken, da hier jeweils ein klar umrissenes Beispiel für die

Anfrage vorliegt. Bei Texten oder Videos ist dieses Anfragemuster aufgrund der Vielzahl möglicher Einzelszenen oder Textpassagen zwar grundsätzlich möglich, in der Praxis jedoch seltener anzutreffen.

- **Anfrageverfeinerung und Relevanz-Feedback:** In vielen Systemen können die Anfragen permanent überarbeitet und verfeinert werden, um das Anfrageergebnis zu optimieren (engl. *Query Refinement*). Häufig wird hierbei auch eine Möglichkeit geschaffen, die Benutzenden gezielt die Markierung bestimmter, besonders relevanter (oder irrelevanter) Treffer ermöglicht, um somit weitere Information über das Informationsbedürfnis eines Benutzenden zu erhalten und in den MMIR-Prozess zu integrieren (engl. *Relevance-Feedback*).

Diese Mechanismen bieten nunmehr Benutzenden die Möglichkeit, Anfragen an MMIR-Systeme zu stellen und Ergebnismengen zur Befriedigung des Informationsbedürfnisses zu erhalten. Die Akzeptanz eines MMIR-Systems hängt dabei natürlich unmittelbar von der Qualität der gelieferten Information ab. Möglichkeiten, diese zu bewerten und messbar zu machen, werden im nächsten Abschnitt vorgestellt.

Selbsttestaufgabe 1.12:

Erklären Sie den Begriff „Query by Example“.

1.6 MMIR-Evaluation

Eine quantifizierbare Bewertung von MMIR-Prozessen und -Ergebnissen ist die Basis für Qualitätsaussagen in Bezug auf MMIR-Prozesse oder MMIS. Hierbei gilt es, diverse Methoden anzuwenden, um die Effektivität („wie gut“) oder die Effizienz („wie schnell“) zu erfassen. Ein zentraler Begriff hierbei ist die „Relevanz“. Allgemeine Definitionen aus bekannten Wörterbüchern definieren die Relevanz folgend: „*relation to the matter at hand*“ oder „*Wichtigkeit in einem bestimmten Zusammenhang*“. Eine kompakte Darstellung von Relevanz in Verbindung mit MMIS aus einem Wörterbuch beschreibt Relevanz als „*the ability (as of an MMIS system) to retrieve material that satisfies the needs of the user*“. In diesem Zusammenhang lassen sich die folgenden, unterschiedlichen Ebenen für eine Relevanzbetrachtung identifizieren (nach Saracevic):

- **Systemische oder algorithmische Relevanz:** Die Auswahl eines Dokuments/Medienobjektes über ein IRS findet über den Vergleich von Repräsentationen der Query und der Dokumente auf Basis eines Algorithmus statt. Der automatisierte Vergleich beruht dabei immer auf gewissen Annahmen, die dem Vergleich zugrunde liegen.
- **Aktualität oder subjektive Relevanz:** Die Verbindung zwischen der Query und einem relevanten Dokument sollte das gleiche inhaltliche Thema sein. Die Aktualität oder subjektive Relevanz ergeben sich aus der Gleichheit des Themas der beiden Repräsentationen.
- **Kognitive Relevanz oder Pertinenz:** Die Verbindung zwischen dem verfügbaren Wissen der User/innen und dem kognitiven Bedarf bestimmt die kognitive Relevanz (z. B. Informationsgehalt, Neuheit oder Qualität der Information).

- **Situative Relevanz oder Nützlichkeit:** Sie besteht in der Verbindung zwischen der Situation, Aufgabe und dem vom IRS gefundenen Dokument. Hier spielen u. a. folgende Faktoren eine Rolle: Nützlichkeit, um eine Entscheidung zu treffen, oder die Reduktion der Unsicherheit.
- **Motivation oder affektive Relevanz:** Sie besteht in der Verbindung zwischen der Intention, den Zielen und der Motivation von User/innen und den gefundenen Dokumenten/Medienobjekten. Zufriedenheit und Erfolg sind u. a. Kriterien, um die Motivation oder affektive Relevanz abzuleiten.

Diese Relevanzbetrachtung führt auch dazu, dass nunmehr ein Gradmesser für die Güte eines MMIR- oder MMIS-Systems berechnet werden kann. Generell wird im MMIR zwischen einer user/userinorientierten Evaluation und einer systemorientierten Evaluation unterschieden. Die user/userinorientierte Evaluation bezieht Benutzende von MMIR-Verfahren in die Evaluation ein, um auf Basis ihrer Rückmeldung einen Schluss auf die Qualität eines IR-Prozesses zu ziehen. Die systemorientierte Evaluation kommt ohne Benutzende aus und setzt den Fokus auf die Bewertung der Systemperformanz (= Effizienz) und Genauigkeit (= Effektivität). Ein Standardvorgehen in der Evaluation der Effektivität eines MMIR-Prozesses ist das sog. Benchmarking, bei dem die Ergebnisse eines MMIR-Algorithmus auf Basis einer reproduzierbaren Aufgabenstellung unter Verwendung einer definierten Datenbasis gemessen und mit anderen Algorithmen verglichen werden. Der verwendeten Datenbasis kommt hierbei speziell im MMIR eine besondere Stellung zu, da die Qualität der Datenbasis und der damit verbundenen Metadaten für automatisiertes Benchmarking eine wichtige Rolle spielen. Soll beispielsweise gemessen werden, wie viele Personen in einer Bilddatenbank erkannt werden, muss für Benchmarking-Verfahren die Anzahl der Personen in der Bilddatenbank bekannt sein. Hierzu werden Testkollektionen erstellt, die dann meist manuell überprüft und annotiert werden und somit einen belegbaren Standard für Evaluierungen darstellen. Diese werden häufig „Goldstandard“ genannt.

In Benchmarking-Verfahren werden bestimmte Metriken auf Grundlage der Testkollektion berechnet. Eine Auswahl dieser Metriken wird im Verlauf dieses Kapitels eingeführt. Die Dokumente der Test-Kollektion lassen sich durch die Unterscheidung in relevant/nicht relevant und gefunden/nicht gefunden für jede konkrete Anfrage, wie in **Fehler! Verweisquelle konnte nicht gefunden werden.** dargestellt in verschiedene Teilmengen zerlegen. Der Begriff „true positive“ beschreibt hierbei die korrekten und gefundenen Treffer, „false positive“ bezeichnet gefundene Treffer, die jedoch nicht relevant sind. Relevante, aber nicht gefundene Objekte werden mit „false negative“ und irrelevante und auch nicht gefundene Objekte mit „true negative“ bezeichnet. Gewünscht ist, dass das MMIS ausschließlich True-Positives und True-Negatives erzeugt, wovon natürlich nur die True-Positives ausgegeben werden. Dies entspricht einer binären Klassifikation.

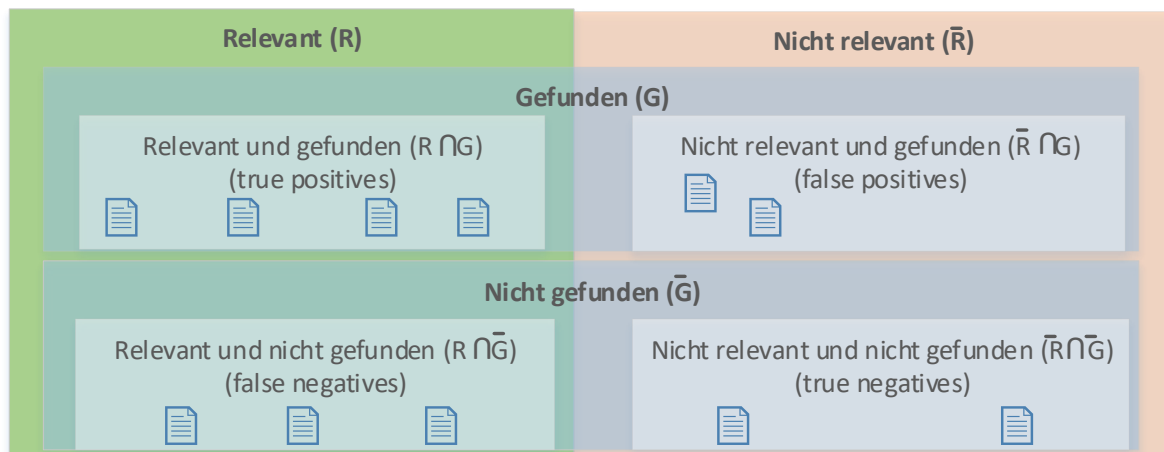


Abbildung 10 - Einteilung der Dokumentenmenge

In diesem Zusammenhang haben sich zwei Begriffe zur Klassifikation durchgesetzt: Precision und Recall, die wie folgt formalisiert werden:

$$Recall = \frac{G \cap R}{R}$$

$$Precision = \frac{G \cap R}{G}$$

Intuitiv bedeutet das, dass der Recall angibt, wie gut der MMIR-Prozess darin ist, alle relevanten Dokumente/Medienobjekte zu finden. Die Precision hingegen gibt an, wie vollständig die Ergebnisse sind, d. h. wie gut der MMIR-Prozess darin ist, alle nicht relevanten Dokumente/Medienobjekte auszuklammern.

Selbsttestaufgabe 1.13:

Erklären Sie den Begriff „Recall“.

1.7 Zusammenfassung

In dieser ersten Lektion wurden nunmehr die zentralen Begriffe, Konzepte und Zusammenhänge vorgestellt. Sie haben einen Überblick über den Übergang von isolierten digitalen Formaten hin zu Multimedia erhalten und die gängigsten Formatvertreter kennen gelernt. Diese Formate sind in der Lage, Informationen zu repräsentieren und können im Bereich Multimedia genutzt werden, um Merkmale auszuprägen. Die Begriffe Information, Wissen und die daraus abgeleiteten Konzepte des Information-Retrieval und der Multimedia-Informationssysteme wurden ebenso beschrieben, wie die zentralen Modelle, um Informationen zu finden und die Qualität der Ergebnisse zu evaluieren.

Auf Basis dieser einführenden Überlegungen werden in den folgenden Lektionen nun die Details zur Repräsentation und Extraktion von Merkmalen, der Speicherung und Verteilung von Multimedia-Content-Objekten und Erstellung von Multimedia-Index-Strukturen beschrieben.

1.8 Lösung der Selbsttestaufgaben

Selbsttestaufgabe 1.1 Text, Bilder, Video, Audio

Selbsttestaufgabe 1.2 HTML muss nicht wohlgeformt oder gültig sein.

Selbsttestaufgabe 1.3 Welche der Eigenschaften trifft auf JSON-Dokumente zu?

- Austausch von Informationen
- Überprüfung der strukturellen Gültigkeit ist möglich
- Dokumente sind kürzer als XML-Dokumente

Selbsttestaufgabe 1.4 Welche Bildformate unterstützen Transparenz?

- TIFF
- PNG
- GIF

Selbsttestaufgabe 1.5 Assets sind Multimedia-Objekte, die für Benutzer einen Wert haben, d. h. wichtig für diese sind.

Selbsttestaufgabe 1.6

- Technische Merkmale
- Allgemeine Merkmale
- Formatspezifische Merkmale
- Inhaltsbasierte Merkmale
- Semantische Merkmale

Selbsttestaufgabe 1.7 Daten > Wissen > Information > Entscheidung / Lösung

Selbsttestaufgabe 1.8 Die klassische Wissenslücke, die das Informationsbedürfnis von Benutzern im Vergleich zur existierenden Wissensbasis beschreibt.

Selbsttestaufgabe 1.9

- Es gibt eine Suchanfrage.
- Die Suche läuft automatisiert ab.
- Es wird eine Menge von Treffern zurückgegeben.

Selbsttestaufgabe 1.10

- Repräsentationskomponente
- Anfragerepräsentationskomponente
- Suchkomponente
- Speicher- und Verteilungskomponente
- Prozessrepräsentationskomponente
- Datenrepräsentationskomponente
- Systemeingabe

Aus Sicht der Nutzenden

- Systemeingabe
- Verwaltungskomponente
- Systemausgabe

Selbsttestaufgabe 1.11

- eine Form des boole'schen IR
- Ziel ist die exakte Erfüllung einer Anfrage durch ein Ergebnis
- XOR-Operationen sind möglich

Selbsttestaufgabe 1.12

Ein Multimediaobjekt (z. B. ein Bild) wird als Referenz für eine Anfrage nach Ähnlichkeit angegeben.

Selbsttestaufgabe 1.13

Recall gibt an, wie gut der MMIR-Prozess die relevanten Dokumente gefunden hat. Precision gibt an, wie vollständig die Ergebnisse sind, d. h. wie gut alle nicht relevanten Ergebnisse ausgeklammert wurden.

Literaturverzeichnis

- [1] U. Library of Congress, „Sustainability of Digital Formats: Planning for Library of Congress Collections,” 2021. [Online]. Available: <https://www.loc.gov/preservation/digital/formats/intro/intro.shtml>.
- [2] Adobe, „PDF Format Specification,” 2022. [Online]. Available: <https://www.pdfa.org/resource/pdf-specification-index/>.
- [3] Microsoft, „Open Office Format Specification,” 2022. [Online]. Available: https://docs.microsoft.com/en-us/openspecs/office_standards/ms-docx/728a7abc-7f55-40dc-90a7-1276ff53c8b2.
- [4] W3C, „HTML Format Specification,” 2011. [Online]. Available: <https://www.w3.org/TR/2011/WD-html5-20110405/>.
- [5] W3C, „XML Format Specification,” 2008. [Online]. Available: <https://www.w3.org/TR/xml/>.
- [6] Json, „JSON Format Specification,” 2018. [Online]. Available: <https://www.json.org/json-en.html>.
- [7] Adobe, „TIF Format Specification,” 2022. [Online]. Available: <https://www.loc.gov/preservation/digital/formats/fdd/fdd000022.shtml>.
- [8] W3C, „PNG Format Specification,” 2003. [Online]. Available: <https://www.w3.org/TR/2003/REC-PNG-20031110/>.
- [9] W3C, „GIF Format Specification,” 1987. [Online]. Available: <https://www.w3.org/Graphics/GIF/spec-gif87.txt>.
- [10] JPEG, „JPEG Format Specification,” 2001. [Online]. Available: <https://jpeg.org/jpeg/>.
- [11] Adobe, „RAW Format Specification,” 2022. [Online]. Available: <https://www.adobe.com/creativecloud/file-types/image/raw.html>.
- [12] Microsoft, 2021. [Online]. Available: <https://docs.microsoft.com/en-us/windows/win32/directshow/avi-riff-file-reference>.

-
- [13 Apple, 2016. [Online]. Available:
] <https://developer.apple.com/library/archive/documentation/QuickTime/QTFF/QTFFPreface/qtffPreface.html>.
- [14 MPEG, 2022. [Online]. Available: <https://mpeg.chiariglione.org/standards/mpeg-4>.
]
- [15 SMPTE, 2019. [Online]. Available: <https://ieeexplore.ieee.org/document/8984681>.
]
- [16 Apple, 2022. [Online]. Available: <http://www-mmmsp.ece.mcgill.ca/Documents/AudioFormats/AIFF/AIFF.html>.
]
- [17 Microsoft, 2022. [Online]. Available: <http://www-mmmsp.ece.mcgill.ca/Documents/AudioFormats/WAVE/WAVE.html>.
]
- [18 LoC, 2022. [Online]. Available:
] <https://www.loc.gov/preservation/digital/formats/fdd/fdd000012.shtml>.
- [19 MIDI_Association, 2022. [Online]. Available: <https://www.midi.org/specifications-old/item/standard-midi-files-smf>.
- [20 T. Vaughan, Multimedia: Making it work, ISBN: 0-07-881869-9: Osborne / McGraw-Hill,
] 1993.
- [21 T. BBC, „Safeguarding the BBC's archive,“ BBC, 2020. [Online]. Available:
] https://www.bbc.co.uk/blogs/bbcinternet/2010/08/safeguarding_the_bbc_archive.html.
- [22 Statista, „Smartphones cause a photography boom,“ Statista / Business Insider, 2020.
] [Online]. Available: <http://www.businessinsider.com/12-trillion-photos-to-be-taken-in-2017-thanks-to-smartphones-chart-2017-8>.
- [23 J. Clement, „Social Media - Statistics & Facts,“ 2021. [Online]. Available:
] <https://www.statista.com/topics/1164/social-networks>.
- [24 Statista, „Value of the global entertainment and media market from 2011 to 2024,“ 2021.
] [Online]. Available: <https://www.statista.com/statistics/237749/value-of-the-global-entertainment-and-media-market/>.
- [25 Sony, „α7R IV 35-mm-Vollformatkamera mit 61,0 MP,“ 2020. [Online]. Available:
] <https://www.sony.de/electronics/wechselobjektivkameras/ilce-7rm4>.

- [26 Xiaomi, „Redmi Note 10 Pro,“ [Online]. Available: <https://www.mi.com/de/product/redmi-note-10-pro/>.
- [27 DicomLookup.com, „Dicom Lookup,“ 2022. [Online]. Available: <http://dicomlookup.com/>.
- [28 M. Webster, „Merriam Webster,“ 2022. [Online]. Available: <https://www.merriam-webster.com/dictionary/information>.
- [29 Duden, „Duden,“ 2019. [Online]. Available: <https://www.duden.de/rechtschreibung/Daten>.
- [30 Rowley, „Representations of the DIKW hierarchy,“ *Journal of Information Science*, 33(2), p. 163–180, 2007.
- [31 R. Kuhlen, „Verarbeitung von Daten, Repräsentation von Wissen, Erarbeitung von Information,“ *Primat der Pragmatik bei informationeller Sprachverarbeitung.*, pp. 1-22, 1985.
- [32 R. & R.-N. B. Baeza-Yates, „Modern information retrieval,“ *Modern information retrieval*, Nr. 463, p. Glossary URL: <http://people.ischool.berkeley.edu/~hearst/irbook/glossary.html> [2019.12.10]., 1999.
- [33 H. M. C. D. & R. P. Schütze, Introduction to information retrieval (Vol. 39, pp. 234-265)., Cambridge: Cambridge University Press., 2008.
- [34 N. J. Belkin, „Anomalous states of knowledge as a basis for information retrieval,“ *Canadian journal of information science*, Bd. 5, Nr. 1, p. 133–143, 1980.
- [35 D. Ellis, „A behavioural model for information retrieval system design,“ *Journal of information science*, Bd. 15, Nr. 4-5, p. 237–247, 1989.
- [36 M. J. Bates, „Toward an integrated model of information seeking and searching,“ *The New Review of Information Behaviour Research*, Bd. 3, Nr. 1, p. 1–15, 2002.
- [37 B. M. L. D. V. K. T. L. & A. S. J. Hemminger, „Information seeking behavior of academic scientists,“ *Journal of the American society for information science and technology*, Bd. 58, Nr. 14, p. 2205–2225, 2007.
- [38 V. Subrahmanian, Principles of Multimedia Database Systems, ISBN: 978-1-558-60466-7: Morgan Kaufmann Publishers, 1998.

-
- [39 W. G. Stock, *Information Retrieval: Informationen suchen und finden*, Oldenbourg, 2007.
]
- [40 K. A., *Datenbanksysteme: Eine Einführung*, ISBN: 978-3486257069: Oldenbourg
] Wissenschaftsverlag, 2001.
- [41 ISO, „ISO/IEC 9075-2:2016, Information technology — Database languages — SQL,“ 2016.
] [Online]. Available: <https://www.iso.org/standard/63556.html>.
- [42 W3C, „SPARQL 1.1 Query Language,“ 2013. [Online]. Available:
] <https://www.w3.org/TR/sparql11-query/>.
- [43 R. Ferber, *Information Retrieval. Suchmodelle und Data-Mining-Verfahren für
] Textsammlungen und das Web*, dpunkt.verlag GmbH, 2003.
- [44 T. Gruber, „A translation approach to portable ontology specifications.,“ *Knowledge
] Acquisition*, Bd. 5, Nr. 2, p. 199–220, 1993.
- [45 M. H. P. H. M. & S. N. Ehrig, „Similarity for Ontologies – a Comprehensive Framework.,“
] *Workshop Enterprise Modelling and Ontology: Ingredients for Interoperability*, at PAKM,
2004.
- [46 D. & S. D. Wilson, *Relevance Theory, Handbook of Pragmatics*, Blackwell., 2002.
]
- [47 T. Saracevic, „Relevance reconsidered,“ *In Proceedings of the second conference on
] conceptions of library and information science (CoLIS 2)*, Bd. New York: ACM, p. 201–218,
1996.
- [48] I. Schmitt, *Ähnlichkeitssuche in Multimedia-Datenbanken: Retrieval, Suchalgorithmen und
Anfragebehandlung: Retrieval, Suchalgorithmen und Anfragebehandlung*

000 000 000 (00/25)

00000-0-00-S1

Alle Rechte vorbehalten
© 2025 FernUniversität in Hagen
Fakultät für Mathematik und Informatik