61. Jahrgang, Nr. 4 – Mai/Juni 2010 – NADOAW 61 (4) 217-272 – ISSN 1434-4653 – D 11721

Information Wissenschaft & Praxis

4/2010

MIRA
CONSULTING GMBH

WISSENSMANAGEMENT
LANGZEITARCHIVIERUNG
TEXTERKENNUNG
INFORMATIONSBESCHAFFUNG
APPLICATION SERVICE PROVIDING
EMAIL-MANAGEMENT
SCAN-SERVICE
WISSENSNETZ
INHALTSERSCHLIESSUNG
WISSENSBEWAHRUNG

Wissensmanagement by MIRA Consulting

Die kaum mehr zu bewältigende Informationsflut in die richtigen Bahnen zu lenken ist ein vordringliches Problem vieler Unternehmen. Um die täglich neuen und umfangreichen Mengen an Daten, Fakten und Neuerungen dem Unternehmen strukturiert verfügbar zu machen und somit auch Wissens- und Wettbewerbsvorteile realisieren zu können, bedarf es eines gezielten Wissensmanagements, welches ganz auf die Belange des Unternehmens, seiner Mitarbeiter und der Unternehmenskultur zugeschnitten ist.

Wir unterstützen Sie dabei !

VORSPRUNG DURCH
WISSEN

www.mira-consulting.net

Management Information Retrieval Application

Indexieren von Emotionen bei Videos

Geistiges Eigentum in der Wissensgesellschaft

Kostenpflichtige und kostenfreie Patentdatenbanken im Vergleich

Entwicklungen bei elektronischen Medien in der Patentinformation

Oberhofer Kolloquium 2010

Potsdamer I-Science-Tage 2010

iwp Herausgeber: Deutsche Gesellschaft für Informationswissenschaft und Informationspraxis e. V. DGI

Mehr als Information!

Sie als Informationsprofi wissen, worauf es ankommt. Natürlich auf die Inhalte, aber eben auch auf Mehr: GENIOS steht für Qualität, Schnelligkeit, Zuverlässigkeit und Kompetenz. Wir sprechen nicht nur Ihre Sprache, sondern wir verstehen, worauf es Ihnen ankommt. Denn seit mehr als 30 Jahren generieren wir für Sie Informationslösungen mit mehr Wert – maßgeschneidert, kreativ und mit viel Empathie.

Mehr über „Mehr als Information“ unter www.genios.de !

GENIOS

German Business Information

GBI-Genios Deutsche Wirtschaftsdatenbank GmbH
Ein Unternehmen der Frankfurter Allgemeine Zeitung GmbH
und der Verlagsgruppe Handelsblatt GmbH

Editorial

Blick zurück, Blicke nach vorn: Barleben, Paris und Frankfurt

Drei Tagungen prägen in diesem Jahr meine Tätigkeit als DGI-Präsident in besonderer Weise. Drei Tagungen, die so verschieden sind, wie ihre Veranstaltungsorte. Und doch haben sie etwas gemeinsam: darum vor allem soll es am Ende dieses Editorials gehen. Vorab aber ein Blick zurück und zwei Blicke nach vorn.

Barleben: vom 22. bis zum 24. April fand dort das 25. Oberhofer Kolloquium statt, eine inhaltlich und zwischenmenschlich sehr ‚dichte' und intensive Veranstaltung. Überschaubar (wie auch der Veranstaltungsort), aber nicht eigentlich ‚klein': um die 100 Teilnehmer durften an diesen drei Tagen unter einem Dach eine Vielzahl zum Teil wirklich anregender Beiträge aufnehmen. Stellvertretend für viele andere möchte ich dabei den Vortrag von Frau Lorenz hervorheben, die über eine Fraunhofer-weite Mitarbeiterbefragung zur Evaluation von Arbeitsweisen und Informationsbedürfnissen sprach und durch ihren frischen und direkten Vortragsstil und zudem den Verzicht auf Powerpoint als Präsentationssoftware wirklich beispielhaft und mitreißend war.
Thematisch war das WWW allgegenwärtig: ob über „Verifikation von Webquellen", über „Suchmaschinenoptimierung" oder über „Real Time Suche" - durchgehend war das WWW mit seinen neuen Möglichkeiten und Risiken besonders in seiner Ausformung als „soziales Netz" präsent. In einem der letzten Vorträge zum Thema „Ein-Mehr-Thesauri-Szenario auf Basis von SKOS und Crosskonkordanzen" kam dann schließlich schon das ‚Semantische' Web der Linked Open Data zur Sprache, das – so meine Behauptung – die Fachdiskussion in unserem Berufsfeld in den kommenden Jahren massiv prägen wird.
Jedenfalls hat ‚Oberhof' eindrucksvoll belegt, wie weitgehend die Mitglieder unserer Gesellschaft schon in der Realität des WWW angekommen sind - viel weiter nämlich, als es die Darstellung in unserer Web-Präsenz mit dem dort immer noch vorhandenen historischen Positionspapier aus dem Jahr 1998 vermuten lässt! In diesem Zusammenhang steht dann noch eine zweite positive Reminis-

Stefan Gradmann

zenz: nachdem es um unseren Entwurf eines neuen Positionspapiers in den ersten Monaten sehr ruhig geblieben war, kam in Barleben endlich doch eine ernsthafte Diskussion des Entwurfs zustande! Dabei bekam das Autorenkollektiv ‚Vorstand' neben Lob der kompetenten und umfassenden Darstellung auch Kritik zu hören, derzufolge der Entwurf etwa allzu ambitioniert sei – bis hin zu der Aussage, ein solch wortreiches Papier sei einfach gar nicht mehr erforderlich, gefragt sei vielmehr die Kurzform eines ‚mission statement'. Ich habe mich über diese endlich stattgefundene Diskussion sehr gefreut und bedanke mich noch einmal ausdrücklich bei allen Beiträgern für Zustimmung und Kritik: beides wird uns helfen, im Juni eine verbesserte Form des Positionspapiers vorzulegen.

Paris: als Ort wohl der maximale Kontrast zu Barleben, in vieler Hinsicht - so eben auch mit der iExpo, die dort am 9. und 10. Juni stattfinden wird, und auf der ich selbst gleich mit zwei Beiträgen gefordert bin.
Mitveranstalter der iExpo ist unser französischer Schwesterverband GFII, und ich war schon anlässlich der iExpo im letzten Jahr erstaunt über die geballte Professionalität und den Schwung dieser Veranstaltung. Unverkennbar und bestimmend ist bei der iExpo 2010 ein begrifflicher Dualismus wie in der Formulierung des Kongressuntertitels mit „le marché de l'*information* et des *connaissances*" manifest: es geht nicht nur (und vielleicht noch nicht einmal mehr primär) um ‚Information', sondern mindestens ebenso sehr um ‚Wissen'. Ein Blocktitel wie „De la gestion de l'information à la gestion des connaissances" macht diesen Übergang explizit - und dementsprechend sind denn auch im Programm der iExpo die semantischen Technologien omnipräsent. Dabei geht es um Graphen, Netzmodelle von Wissen, um semantische Visualisierung und natürlich auch hier wieder um die Konvergenz des ‚sozialen' und des ‚semantischen' Netzes.
Ich freue mich auch deshalb auf diese Veranstaltung, weil sie sich in einem Punkt entscheidend von unserer Verbandsrealität unterscheidet: hier ist die Informationswirtschaft unter sich, es fehlt also weitgehend der für unsere DGI kennzeichnende Dualismus von wirtschaftsgeprägter und öffentlich-rechtlicher Denkweise. Auch ‚Wissenschaft' ist in Paris nicht wirklich maßgeblich - unter diesen Bedingungen ist es natürlich vergleichsweise einfach, ein sehr geschlossenes und attraktives Programm ‚aus einem Guss' zu erstellen, wie dasjenige der iExpo 2010 ...

Frankfurt: wieder im Rahmen der Buchmesse findet dort vom 7. bis zum 9. Oktober unsere 62. Jahrestagung statt - diesmal mit neuem Titel als 1. DGI-Konferenz. Ihr Titel ist programmatisch und passt zu meinen Beobachtungen in Barleben und Paris: „Semantic Web & Linked Data" als prägende „Elemente zukünftiger Informationsinfrastrukturen" sind thematisch bestimmend.
Dazu passend ist es uns gelungen, mit Ivan Herman den „Semantic Web Activity Lead" des World Wide Web Consortiums (W3C) als Keynote-Speaker zu gewinnen. Ein hochwertiges Programm mit den vier großen Blöcken Forschung, Wirtschaft, Verwaltung und Bildung ist in Vorbereitung, wobei zumindest die ersten beiden Blöcke massiv von „semantischen" Inhalten geprägt sind. Da geht es um Themen wie RDF-basierte Standards/

Techniken, semantisch basierte Suchmaschinen, Open Research Data, Darstellung von Wissen und Ontologien, aber auch etwa um „Semantisch basierte Geschäftsmodelle“ oder um „Weiterverwendung von Informationen des öffentlichen Sektors“ als Linked Data (wobei letztere Thematik von einem französischen Kollegen aus der GFII moderiert wird, denn in diesem Feld ist Frankreich sehr weit und besetzt im europäischen Kontext eine sehr exponierte Position).

Unsere Konferenz ist aber auch Gelegenheit zum persönlichen Kennenlernen oder Wiedersehen: darauf freue ich mich auch in diesem Jahr wieder ganz besonders!
Was nun ist allen drei so ausgesprochen unterschiedlichen Veranstaltungen gemeinsam? Zwei Aspekte (mindestens): da ist zum einen der stetig wachsende Bezug unserer Profession zu den Informations- und Wissensarchitekturen des WWW und insbesondere des ‚Semantic Web'. Das WWW entwickelt sich gerade aus einer gigantischen Hypertextanwendung mit verlinkten Web-‚Dokumenten' zu einem Instrument der Modellierung von Wissen über beliebige Ausschnitte von ‚Welt'.
Solcherart, als ‚Web der Dinge' oder als ‚Linked Data', wird es für unsere Profession nicht nur zu einem zunehmend dominanten Bezugsrahmen, sondern wir können diese Entwicklung auch als Chance begreifen: An die Stelle des reinen Akkumulierens und Zugänglichmachens von Information wie im ‚traditionellen' Fachinformations-Paradigma tritt nun als Primärziel die urdokumentarische Fähigkeit, Information - nun aber mit den Mitteln und unter den Bedingungen des WWW - sinnvoll zu ordnen, zu strukturieren, zu selektieren, zu neuen Einheiten zu aggregieren und aufzubereiten und diese inhaltlich so zu erschließen bzw. mit Kontextbeziehungen zu versehen, dass daraus Wissen entstehen kann (nebenbei gefragt: wo haben Sie diese Formulierung schon gelesen?)
Und gerade auch das Oberhofer Kolloquium war für mich ein schöner Beleg dafür, dass die Mitglieder der DGI diesen Aspekt entweder schon verinnerlicht haben oder doch zumindest zu verstehen beginnen.
Der zweite Aspekt ist absolut trivial, ist aber fast noch wichtiger: allen drei Veranstaltungen ist gemeinsam, dass es sich nicht um ‚virtuelle' Konstellationen mit viel vermittelnder Technik handelt, sondern um wirkliche Treffen wirklicher Menschen - und das ist heutzutage gar nicht mehr so selbstverständlich! Dieser direkte Kontakt der Menschen untereinander, wie ihn auch unsere Veranstaltungen herstellen, ist und bleibt unverzichtbar, wenn wir wirklich etwas erreichen und die dafür erforderlichen belastbaren Kommunikations- und Vertrauensbeziehungen herstellen wollen. Auf dieser Basis ist dann viel in virtuellen, netzbasierten Szenarien realisierbar – aber diese Basis des Bandes zwischen Menschen bleibt essentiell.
Und in diesem letzteren Sinne ist vielleicht das Oberhofer Kolloquium sogar die effektivste Veranstaltung: ich wünsche mir somit, dass wir die Nähe und zum Teil auch Herzlichkeit aus Barleben ein Stück weit in unsere Frankfurter Konferenz hineintragen können.
Vor allem aber wünsche ich mir, möglichst viele von Ihnen in Frankfurt begrüßen zu können.

Ihr Stefan Gradmann

FIZ Technik in die Insolvenz gezwungen

Staat und Wirtschaft waren sich seit 1972 einig, in Deutschland ein wissenschaftliches Informationssystem der Technik aufzubauen und zu betreiben. Sie sahen in der Verfügbarkeit von aktuellen Informationen eine Schlüsselrolle für Innovationen in Wirtschaft, Wissenschaft und Staat. Damit haben sie auch gleichzeitig den allgemeinen Kulturauftrag erfüllt, das von jeder Generation erarbeitete Wissen für die Gegenwart und für die nachfolgenden Generationen präsent zu halten.
Dieser Konsens wurde mit der abrupten Streichung der unverzichtbaren öffentlichen Fördermittel durch das Bundesministerium für Wirtschaft und Technologie zum 30. Juni 2010 aufgegeben. FIZ Technik war mit einem öffentlichen Förderanteil von ca. 40 Prozent daraufhin gezwungen, Insolvenz anzumelden. Es ist für jeden Kenner des Fachinformationsbereichs eine feststehende Tatsache, dass sich derartige Systeme mit einer speziellen Ausrichtung auf den deutschen Wissenschaftsmarkt nicht alleine finanzieren können. Staat und Industrie haben deshalb 38 Jahre lang das außerordentlich effizient arbeitende FIZ-Technik-System gemeinsam durch jährliche Zuwendungen unterstützt.
FIZ Technik hat bis heute den Auftrag, die technisch-wissenschaftliche Information der Technik zur Förderung der Forschung und Entwicklung sowohl für die Einrichtungen der Lehre und Forschung wie auch für die Industrie bereit zu stellen, in hervorragender Weise erfüllt. Fast alle Universitäten, Technischen Hochschulen, Fachhochschulen sowie große Forschungseinrichtungen mit ingenieurwissenschaftlichen Fachrichtungen nutzen insbesondere die FIZ-Technik-eigenen Datenbanken mit den inhaltlichen Schwerpunkten Maschinenbau, Elektrotechnik, Werkstoffe, Fahrzeugbau, Textiltechnik, Umwelttechnik, chemische Technik und weiteren Fachgebieten. Das umfassende Angebot ist ferner insbesondere für große Unternehmen aber auch für viele mittelständische Unternehmen eine unverzichtbare Quelle für die Beobachtung der internationalen Fachliteratur. FIZ Technik weist in seinen Datenbanken vor allem die deutschsprachige, angelsächsische und ostasiatische Fachliteratur nach.
Welche Gründe zu dem außerordentlichen und absolut unüblichen Schritt des BMWi geführt haben, sind nicht zu erkennen. FIZ Technik hat in den vergangenen Jahrzehnten in wesentlichen Bereichen mit anderen Partnern kooperiert. Die längste und erfolgreichste Kooperation im Bereich der inhaltlichen Produktion der Datenbanken besteht seit 1973 bis heute mit der Technischen Informationsbibliothek (TIB). Aber auch Kooperationen im Bereich des Betriebs von Hostrechnern mit FIZ Karlsruhe in den frühen achtziger Jahren und mit Data Star über 13 Jahre sowie die Nutzung des Datenbanksystems der GBI seit zwölf Jahren oder auch die aktuelle Zusammenarbeit mit Forschungseinrichtungen im Bereich des linguistischen Indexierens und Suchens, sind Belege für den sparsamen Umgang mit den Fördermitteln. Weiterhin ist FIZ Technik mit den technisch-wissenschaftlichen Vereinen und den Wirtschaftsverbänden im Bereich der Technik stark vernetzt.
Es scheitert an einem Betrag von jährlich zwei Millionen Euro, die für das Industrieland Deutschland erforderlich sind, um ein eigenes leistungsfähiges Informationssystem für die in der Industrie und Forschung aktiven und an den Hochschulen in Ausbildung befindlichen Ingenieure zu betreiben. Die Politiker sprechen von der Wissensgesellschaft und der Digitalisierung der Information. Aber gleichzeitig sind zwei Millionen Euro zu viel, um ein leistungsfähiges und anerkanntes System mit zu finanzieren.

Der Vorstand der DGI

Inhalt 4/2010

217 **EDITORIAL**

Stefan *Gradmann*
Blick zurück, Blicke nach vorn: Barleben, Paris und Frankfurt

218 Vorstand der DGI
FIZ Technik in die Insolvenz gezwungen

INFORMATIONSWISSENSCHAFT

221 Kathrin *Knautz*, Evelyn *Dröge*, Susanne *Finkelmeyer*, Daniel *Guschauski*, Kerstin *Juchem*, Cynthia *Krzmyk*, Daniel *Miskovic*, Jan *Schiefer*, Evrim *Sen*, Julia *Verbina*, Nils *Werner* und Wolfgang G. *Stock*
Indexieren von Emotionen bei Videos

INFORMATIONSRECHT

237 Volker *Gross*
Aspekte geistigen Eigentums in der Wissensgesellschaft

PATENTINFORMATION

245 Markus *Birkner*
Patentdatenbanken: kostenpflichtige mit Mehrwert oder kostenfreie?

259 Dieter *Geiß*
Aus der Praxis der Patentinformation. Entwicklungen der elektronischen Medien und Dienstleistungen bei den Patentbehörden und Internetprovidern

TAGUNGSBERICHTE

267 Marlies *Ockenfeld*
25. Oberhofer Kolloquium gibt neue Impulse

269 Lars *Müller* und Ute *Rusnak*
Die Datenflut im Forschungsprozess erfordert neue Informationsstrategien

PERSONALIE

258 **Harald Haendler** †

266 **IMPRESSUM**

272 **MEDIAINFO**

U3 **TERMINKALENDER**

Erratum

In der letzten Druckausgabe der IWP 3/2010 war unter dem Editorial und im Impressum auf Seite 154 sowie neben der Abbildung auf Seite 216 der Vorname der Gastherausgeberin Ursula Georgy falsch als Ulrike Georgy abgedruckt. Die Redaktion bittet dieses Versehen zu entschuldigen. In der Online-Version der IWP wurde der Fehler sofort korrigiert.

Missing Link

Westerstrasse 114-116 | 28199 Bremen

Tel.: [0421] 504348 | Fax: [0421] 504316

Erwerbungspartner, mit denen Sie rechnen können

Flexibel

Erfahren

Innovativ

Konditionsstark

Serviceorientiert

Engagiert

Klar

www.missing-link.de

info@missing-link.de

Internationale Versandbuchhandlung

Indexieren von Emotionen bei Videos

Kathrin Knautz, Evelyn Dröge, Susanne Finkelmeyer, Daniel Guschauski, Kerstin Juchem, Cynthia Krzmyk, Daniel Miskovic, Jan Schiefer, Evrim Sen, Julia Verbina, Nils Werner und Wolfgang G. Stock, Düsseldorf

Gegenstand der empirischen Forschungsarbeit sind dargestellte wie empfundene Gefühle bei Videos. Sind Nutzer in der Lage, solche Gefühle derart konsistent zu erschließen, dass man deren Angaben für ein emotionales Videoretrieval gebrauchen kann? Wir arbeiten mit einem kontrollierten Vokabular für neun tionen (Liebe, Freude, Spaß, Überraschung, Sehnsucht, Trauer, Ärger, Ekel und Angst), einem Schieberegler zur Einstellung der jeweiligen Intensität des Gefühls und mit dem Ansatz der broad Folksonomy, lassen also unterschiedliche Nutzer die Videos taggen. Versuchspersonen bekamen insgesamt 20 Videos (bearbeitete Filme aus YouTube) vorgelegt, deren Emotionen sie indexieren sollten. Wir erhielten Angaben von 776 Probanden und entsprechend 279.360 Schiebereglereinstellungen. Die Konsistenz der Nutzervoten ist sehr hoch; die Tags führen zu stabilen Verteilungen der Emotionen für die einzelnen Videos. Die endgültige Form der Verteilungen wird schon bei relativ wenigen Nutzern (unter 100) erreicht. Es ist möglich, im Sinne der Power Tags die jeweils für ein Dokument zentralen Gefühle (soweit überhaupt vorhanden) zu separieren und für das emotionale Information Retrieval (EmIR) aufzubereiten.

Indexing of Emotions in Videos

Object of our empirical research study are depicted and provoked emotions in videos. Are users able to index such emotions consistently? Are the users' votes usable for emotional video retrieval? We worked with a controlled vocabulary for nine basic emotions (love, happiness, enjoyment, surprise, desire, sadness, anger, disgust, and fear), a scroll bar for adjusting the emotions' intensities, and the approach of broad folksonomies. Different users tagged the same videos. The test persons had the task to index the emotions of 20 videos (reprocessed movies from YouTube). We got data from 776 participants and could analyze all in all 279,360 scroll bar values. The consistency of the users' votes is very high; there are stable tag distributions for the emotions of the particular videos. The final shape of the distributions will be reached by the tagging activities of only few users (less than 100). Applying the approach of power tags it is possible to separate the pivotal emotions of every document – if there is any feeling at all. Those document-specific emotions establish the basis of an emotional information retrieval (EmIR) system.

Einleitung

„Ich habe Angst. Ich habe Angst, Dave (…)" entgegnet HAL in Stanley Kubricks 1968er Meisterwerk *2001: Odyssee im Weltraum* dem Astronauten Dave Bowman. HAL ist ein mit einem künstlichen Wesen versehener Supercomputer der Serie 9000, der das Raumschiff „Discovery" autark lenkt. HAL ist zu einer Gefahr für die Besatzung geworden, und Dave Bowman hat sich Zugang zu dem zentralen Computerraum verschafft und beginnt damit, den Computer schrittweise abzuschalten. Computer(systeme), die die Emotionen ihres menschlichen Gegenübers nicht nur wahrnehmen, sondern diese ferner verarbeiten sowie interpretieren – Rosalind W. Picard bündelt dies unter dem Begriff „Affective Computing": „(…) computing that relates to, arises from, or influences emotions" (Picard, 1995, 1) – und darüber hinaus auch auf einer emotionalen Ebene interagieren, sind längst keine bloße Illusion mehr: „Computers are beginning to acquire the ability to express and recognize affect, and many soon be given the ability to ‚have emotions'" (Picard, 1995, 1).

Wir beziehen den Ansatz des Affective Computing auf den Bereich Human-Computer-Interaction (HCI) (Boehner et al., 2007) bzw. auf die Forschungen zum Information Retrieval (Stock, 2007a). Bei unserer Forschungsarbeit geht es nicht um emotional reagierende Computer, sondern darum, ob wir in Retrievalsystemen nach Gefühlen recherchieren können, die in Dokumenten entweder ausgedrückt werden oder die Gefühle beim Betrachter hervorrufen. Wie können wir solche Gefühlsreaktionen überhaupt suchbar speichern und abfragbar machen? Welche Indexierungsverfahren sind einsetzbar? Sind Gefühlseinschätzungen derart konsistent, dass sie auch eindeutig Dokumenten zugeordnet werden können? Dies sind Fragen des jungen informationswissenschaftlichen Forschungsgebietes des Emotionalen Information Retrieval (EmIR) (Schmidt und Stock, 2009). Emotionen schwingen in diversen Dokumenttypen mit. „Gefühlsbeladene" Dokumente finden wir bei

- Texten (z.B. Romanen oder Gedichten),
- Musik,
- Bildern,
- Videos,
- Multimedia-Dokumenten (z.B. Weblogs).

Wir berichten in diesem Artikel über das Indexieren von Basisemotionen bei Videos, nachdem für Musik (Lee und Neal, 2007) sowie für Bilder (Schmidt und Stock, 2009) bereits Ergebnisse vorliegen. Wir versuchen, folgende Forschungsfragen (Ff) einer ersten Beantwortung nahezubringen:

- (Ff 1): Ist zufriedenstellendes Retrieval für Filme im Web derzeit praktisch einsetzbar?
- (Ff 2): Was sind Basisemotionen?
- (Ff 3): Wie können wir Basisemotionen indexieren, um sie im emotionalen Retrieval praktisch einzusetzen?
- (Ff 4): Wie konsistent sind nutzerorientierte Verfahren beim Taggen von Emotionen in Videos?
- (Ff 5): Wie viele unterschiedliche Nutzer braucht man, bis sich eine stabile Verteilung der Gefühls-Tags ergibt?

Gibt es überhaupt stabile Verteilungen bei Gefühls-Tags?
- (Ff 6): Können wir aus den Tagverteilungen Power Tags separieren, die den dargestellten bzw. empfundenen Gefühlen entsprechen?
- (Ff 7): Korrelieren gewisse Emotionen miteinander?

Als Methode setzen wir eine Befragung von Versuchspersonen ein. Insgesamt 776 Probanden haben im Frühsommer 2009 jeweils 20 Filme (bearbeitete YouTube-Dokumente) mittels Schieberegler hinsichtlich neun verschiedener Basisemotionen getaggt.

Videoretrieval (Forschungsfrage 1)

Wie lassen sich Videos überhaupt für Retrievalsysteme zugänglich machen? Die Retrievalforschung unterscheidet grob zwei Ansätze:
- Concept-based Information Retrieval (über Begriffe bzw. Wissensordnungen),
- Content-based Information Retrieval (nur über Inhalte des Dokuments).

Ein zeit- und personalunabhängiger Ansatz ist die automatische, contentbasierte Erschließung von Videos (Aigrain, Zhang, und Petkovic, 1996; Colombo, Del Bimbo und Pala, 1999; Del Bimbo, 1999; Enser, 2008a; Enser, 2008b; Gupta und Jain, 1997). Um contentbasiertes Retrieval (Stock, 2007a, Kap. 31) zu ermöglichen, muss zunächst eine Analyse erfolgen, um das Video für weitere Schritte vorzubereiten. Ein Film besteht aus vielen verschiedenen Szenen, die sich in Kameraeinstellungen, sog. Shots, einteilen lassen. Die Shots wiederum bestehen aus einzelnen Frames (Bildern), welche die kleinsten Videoeinheiten darstellen (Smeaton, 2007).
Dies sind visuelle Komponenten, welche unter den Low-Level Features zusammengefasst werden. Allgemein bezeichnen Low-Level Features Metadaten (Smeaton et al., 2008; Auld und Petrelli, 2008) und elementare Videoattribute, die sich aus visuellen, auditiven und textuellen Komponenten zusammensetzen (Smeaton, 2007). Folgende Low-Level Features kommen für eine automatische Indexierung infrage:
- Dimension (Einzel-)Bild
  - Farbe
  - Textur
  - Gestalt
- Dimension Ton
  - gesprochener Text
  - Musik (Tonhöhe, Rhythmus, Harmonie, Klangfarbe)
  - Geräusche
- Dimension Video
  - Segmentierung: Szenen und Shots
- Bewegung: Kamerabewegung (Schwenken oder Zoomen) und aufgenommene Bewegung.

Die Informationen, die sich aus diesen Low-Level Features erschließen lassen, sind jedoch begrenzt und deshalb nur bedingt für eine Analyse des bedeutungsträchtigen und emotionalen Contents geeignet. Doch gibt es bereits Ansätze, welche versuchen, gefühlsbeschreibenden Content aus den Low-Level Features automatisch zu extrahieren (Chan und Jones, 2005; Chen et al., 2007; Salway und Graham, 2003; Soleymani et al., 2008; Xu et al., 2005). Colombo, Del Bimbo und Pala (1999, 41) geben am Beispiel der Farbdimension zu bedenken: „The mapping of low-level color primitives into emotions is quite complex."
Wir wollen an einem Beispiel verdeutlichen, wie contentbasiertes emotionales Videoretrieval ablaufen könnte. Hanjalic und Xu (2004) versuchen, von Low-Level Features auf den emotionalen Gehalt von Videos zu schließen. Sie verwenden ein zweidimensionales Modell, um Emotionen zu erfassen. Eine erste Dimension ist die Intensität, die andere die Valenz. Die Intensität gibt den Grad der Emotionen an, und deren Skala bewegt sich von energiegeladen, aufgeregt und wachsam bis hin zu ruhig, schläfrig und friedlich. Die Dimension der Valenz oder Wertigkeit zeigt auf, ob die dargestellten Emotionen erfreulich bzw. positiv oder unerfreulich bzw. negativ sind. Auf der Basis dieser beiden Dimensionen soll nun der emotionale Content von den Low-Level Features abgeleitet werden. Zum Erschließen der Wertigkeit der Emotionen werden drei Komponenten berücksichtigt. Diese sind zum einen die Komponente der Bewegung, welche anhand von aufeinanderfolgenden Frames gemessen wird, zum anderen die Komponente des Rhythmus, welche sich aus der Frequenz von einzelnen Schnitten erschließen lässt, sowie die Komponente der Energie des Soundspektrums. Die Dimension der Valenz wiederum wird nur anhand von einer Komponente gemessen. Hier wird lediglich die durchschnittliche Tonhöhe berücksichtigt, wobei eine hohe durchschnittliche Tonhöhe Freude und eine niedrige durchschnittliche Tonhöhe Traurigkeit ausdrücken soll. Das vorgestellte Modell misst diese beiden Dimensionen letztendlich anhand von automatisch extrahierten Zahlenwerten, weshalb eine differenzierte emotionale Indexierung der Videos somit nicht gegeben sein kann. Hierdurch kann lediglich eine Basis geschaffen werden, um besonders positive oder negative emotionale Sequenzen zu bestimmen. Die Autoren geben zu, dass die bis dato erforschten Relationen zwischen den Low-Level Features und der emotionalen Ebene vage sind, weshalb sich eine Abbildung des emotionalen Gehalts eines Videos auf Basis der Low-Level Features als schwierig gestaltet. Contentbasiertes Videoretrieval steckt noch immer in der Entwicklung, obwohl schon einige Fortschritte in Hinblick auf die automatische Extraktion der Low-Level Features zu bemerken sind (Chan und Jones, 2005; Xu, Chia, und Jin, 2005). Jedoch müssen nach Smeaton (2007) weitere Aufgabenschwerpunkte gesetzt werden. Das Ermitteln einzelner Features muss verbessert werden, denn noch sind die Ergebnisse, besonders im automatischen Herausfiltern von semantischen Merkmalen, viel zu ungenau. Bisherige Ansätze im Videoretrieval basieren zudem hauptsächlich auf Keyframe- und Bildretrieval (Auld und Petrelli, 2008; Kim und Kim, 2009), jedoch entspricht dieser statische Ansatz nicht dem gewünschten Videoretrieval, da hier der temporale Aspekt, die Bewegung der Objekte und der Kamera fehlt. Auf Grund dessen eignen sich Keyframes nur bedingt als Kurzfassung von Videos, denn die Auswahlverfahren für diese setzen meist einfach in der Mitte eines Shots an und nicht dort, wo sich der Inhalt des Videos am besten repräsentieren lässt (Kim und Kim, 2009). Des Weiteren bemängelt Smeaton, dass momentan nur einzelne Video-Shots ausschlaggebend sind, was den User bei seiner Videosuche nur bedingt zufriedenstellen kann. Denn anhand eines Shots können kaum semantische Inhalte, wie die Stimmung und Botschaft eines Videos, übermittelt werden. Deshalb sollte stattdessen eine Szene aus bedeutungsträchtigen Shots zusammengestellt werden, die den „wahren" Content widerspiegeln, um auf die Bedürfnisse der Nutzer einzugehen. Smeaton (2007) hebt einen personalisierten Videozusammenschnitt (Smeaton et al., 2008), vergleichbar mit Videotrailern, hervor. Dieser könnte dem User helfen, seine Suchergebnisse besser zu bewerten. Doch momentan beschränkt sich Videoretrieval nur auf einfaches Browsen und einer Keyframe-Darstellung der Ergebnisse.
Das contentbasierte Videoretrieval im Allgemeinen und das Retrieval nach Emotionen im Besonderen bringen also noch einige Herausforderungen mit sich. Wichtig wäre zu bestimmen, welcher konkrete Content (sagen wir: ein lächelndes Gesicht; Bailenson et al., 2008) bestimmte Emotionen beim Betrachter (hier: Freude) auslöst. Forschungen zu solchen „emotionalen Referenzpunkten" (im Anschluss an die „cognitive reference points" von Eleanor Rosch (1975)) stehen noch aus.
Wenden wir uns nunmehr dem conceptbasierten Videoretrieval zu! Snoek und Worring (2009) haben bezüglich des Retrievals von Videos über 300 Literaturquellen ausgewertet und festgestellt, dass in diesen zumeist conceptbasierte Lösungen als vielversprechende Alter-

nativen gesehen werden. Im conceptbasierten Videoretrieval arbeiten wir mit Begriffen, die zwar prinzipiell auch dem Content (vor allem dem gesprochenen Text) automatisch entnommen werden könnten, aber derzeit vorwiegend intellektuell zugeordnet werden (Schweins, 1997). Ein erster Weg führt über Wissensordnungen (beispielsweise Thesauri) und professioneller Inhaltserschließung. Diesen Weg wählen beispielsweise die deutschen und österreichischen Fernsehanstalten bei der Inhaltserschließung von Fernsehproduktionen. Es kommen – bei intellektueller Indexierung – sowohl Klassifikationen als auch ein Thesaurus zum Einsatz (ARD/ORF/ZDF, 2008), Gefühle werden jedoch nicht beschrieben. Beim Vorliegen einer Begriffsordnung für Emotionen wäre es zumindest prinzipiell möglich, den Videos kontrolliertes Vokabular zu Gefühlen intellektuell zuzuordnen (Kim und Park, 2003). Diese Methode ist jedoch stark abhängig von der Person des jeweiligen Indexers. Ein zentrales Problem allen Concept-based Videoretrievals dürfte die Indexierungskonsistenz sein, die bei Bildern – und wohl auch bei Filmen – sehr schwach ausgeprägt ist (Markey, 1984). „The output of the indexing process seemed to be quite inconsistent" (Markkula und Sormunen, 2000, 273). Wir lernen von Goodrum (2000, 64), „manual indexing suffers from low term agreement across indexers (…), and between indexers and user queries." „There is evidence that current systems for image access often fail the user", erinnert Jörgensen (1998, 162). Es gibt bisher keinerlei Erfahrungen mit der intellektuellen Indexierung von Emotionen. „The influence of the photograph's emotional tone on categorization has not been discussed much in previous studies", berichten Laine-Hernandez and Westman (2006). Letztlich steht solch einem Unterfangen im World Wide Web ein praktisches Problem entgegen. Es ist schlicht unmöglich (weil unbezahlbar), die Millionen von verfügbaren Videos durch professionelle Indexer auswerten zu lassen, denn „manual assignment of textual attributes is both time consuming and costly" (Goodrum, 2000, 64). Etwas platt ausgedrückt: So geht es nicht (Enser, Sandom, Hare, und Lewis, 2007).
Der geringere Zeit- und Personalaufwand würde für usergenerierte Tags – also Social Tagging oder kooperatives Indexieren (Jörgensen, 2007) – sprechen, wie sie beispielsweise auf Video-Plattformen wie YouTube zu finden sind (Geisler und Burns, 2007; Melenhorst et al., 2008; van Velsen und Melenhorst, 2009). Folksonomies (Peters, 2009; Smith, 2008; Weller, Peters und Stock, 2010) sind entweder „narrow" (nur der Urheber des Dokuments darf Tags vergeben; Beispiel: YouTube), „extended narrow" (außer dem Urheber dürfen auch seine Freunde Tags verteilen, jeden Tag aber nur einmal; Beispiel: Flickr) oder „broad" (jeder Nutzer des Informationsdienstes darf Tags vergeben, die demnach auch mehrfach einem Dokument zugeordnet werden können; Beispiel: Del.icio.us). Bei narrow und extended narrow Folksonomies stellt sich das Problem der Indexierungskonsistenz – hier sogar besonders pointiert, da keine Information Professionals, sondern Laien indexieren (Peters und Stock, 2007; Peters und Stock, 2008). Bei broad Folksonomies und einer ausreichend großen Zahl an taggenden Nutzern sollten sich bei den hauptsächlich vergebenen Tags keinerlei Konsistenzprobleme ergeben, allerdings entstammen die Tags keinem normierten Vokabular, sondern der Sprache der Nutzer. Da YouTube eine narrow Folksonomy nutzt, können wir die dort durchgeführte Indexierung nicht für unsere Zwecke gebrauchen. So endet die Forschungsfrage 1 mit einem negativen Ergebnis: Weder content- noch conceptbasiertes Videoretrieval führt derzeit zu praktisch brauchbaren Ergebnissen für Filme im WWW.

Basisemotionen (Forschungsfrage 2)

Zur Klärung dessen, was „Emotionen" sind, greifen wir auf Ergebnisse und Diskussionen aus der Psychologie zurück. Die Emotionsforschung kann auf eine lange Tradition zurückschauen, und dennoch war es bisher nicht möglich, eine einheitliche exakte Definition für das Konzept der Emotion vorzulegen. Eine Definition bereitet insofern Schwierigkeiten, da zwar jeder weiß, was eine Emotion ist, die Formulierung einer konkreten Begriffsbestimmung jedoch schwierig ist (Fehr und Russell, 1984). Auch Schmidt-Atzert stellt fest, dass bisher „kein Konsens festzustellen [ist], was man unter einer Emotion zu verstehen hat" (Schmidt-Atzert, 1996, 18). Nach Izard (1994, 20) muss eine vollständige Definition von Emotion drei Aspekte beinhalten:

- das Erleben oder das bewusste Empfinden des Gefühls,
- die Prozesse, die sich im Gehirn und im Nervensystem abspielen und
- das beobachtbare Ausdrucksgebaren, besonders im Gesicht.

Bereits 1981 untersuchten Kleinginna & Kleinginna 100 Aussagen und Definitionen aus einschlägigen Spezialwerken, Wörterbüchern und Einführungstexten und versuchten, eine Arbeitsdefinition zu erstellen: Emotion ist „ein komplexes Muster von Veränderungen, das physiologische Erregung, Gefühle, kognitive Prozesse und Verhaltensweisen einschließt, die in Reaktion auf eine Situation auftreten, welche ein Individuum als persönlich bedeutsam wahrgenommen hat" (Kleinginna und Kleinginna, 1981). Auch Meyer, Reisenzein und Schützwohl (2001) betreiben ihre Forschung auf Basis einer Arbeitsdefinition, da die Definition selbst eine zentrale Fragestellung innerhalb der Emotionsforschung abbildet und damit vielmehr „nicht die Voraussetzung der wissenschaftlichen Emotionsforschung ist, sondern deren Resultat" (Meyer, Reisenzein und Schützwohl, 2001, 50). Dennoch lassen sich Emotionen nach Meyer, Reisenzein und Schützwohl (2001, 24) durch verschiedene Merkmale etwas genauer charakterisieren. So sind Emotionen aktuelle psychische Zustände von Menschen und damit abgrenzbar von anderen Konzepten wie Dispositionen und Stimmungen. Emotionen sind zudem auf ein bestimmtes, nicht unbedingt real existierendes Objekt bzw. auf Objekte gerichtet, welche diese Emotionen auslösen. Als weiteres Merkmal nennen die Autoren, dass sich Emotionen in der so genannten Reaktionstrias aus subjektiven, behavioralen und physiologischen Aspekt zeigen. Auch haben Emotionen eine bestimmte Qualität, Intensität und Dauer. Die Qualität fungiert als Gruppierungsmerkmal, so dass man bei dem Vorkommen einer Emotion in einer Situation wie beispielsweise in *„Person A freut sich"* von einer Zuordnung des Qualitätstyps *Freude* spricht. Die konkreten Realisierungen der Qualitätstypen zeichnen sich, neben einer zeitlichen Begrenzung, zudem durch unterschiedliche Intensitäten aus. Die Stärke einer Emotion lässt sich nach Meyer, Reisenzein und Schützwohl auf einer Skala von niedriger über mittlerer bis hin zu einer starken Intensität festlegen (2001, 29-30).
Wie Emotionen entstehen, versuchen verschiedene Theorien der Emotionsforschung mit unterschiedlichen Ausrichtungen zu klären. Behavioristische Ansätze betonen den Verhaltensaspekt und die auslösenden Bedingungen. Emotionen werden entweder als angeboren angesehen (Furcht, Wut und Liebe) oder entstehen als konditionierte Reaktionsmuster auf bestimmte Reize (Watson, 1930). Weiterhin gibt es evolutionspsychologische Ansätze, welche die evolutionär bedingten adaptiven Funktionen von Emotionen, wie beispielsweise ihre Überlebensfunktion, betonen. Sie gehen auf Darwin (1872) zurück, dessen Hauptanliegen es war, die phylogenetische Entwicklung von Emotionen nachzuweisen, und wurden von diversen Emotionsforschern (McDougall, 1926; Plutchik, 1980; Izard 1971, 1977, 1991; Ekman et al. 1982; Tomkins 1962, 1963) weitergeführt. Dass Emotionen durch die Wechselwirkung zwischen physiologischen Veränderungen bestimmt werden, wird in kognitiv-physiologische Theorien postuliert. Durch die (in-)direkte Wahrnehmung dieser Veränderung entwickelt

und adaptiert der Mensch seine Emotionen (James, 1884; Panksepp, 1982; Frijda, 1986). Attributionstheorien bzw. attributionale Theorien beschreiben, wie der Mensch mittels Ursachenzuschreibung versucht, seine Umwelt zu verstehen und zu kontrollieren. Emotionen sind nach diesen Ansätzen damit Reaktionen auf Handlungsergebnisse (Arnold, 1960; Weiner und Graham, 1984; Lazarus, 1960).
Im Zuge unserer Forschung stellt sich die Frage, welche Emotionen für das Projekt ausgewählt und damit in Hinblick auf das emotionale Videoretrieval herangezogen werden sollen. In der Psychologie findet sich hinsichtlich einer Kategorisierung von Emotionen der Ansatz, Emotionen auf eine kleine, aber relativ fixe Anzahl zu reduzieren. Diese werden als fundamentale Emotionen bzw. als Basisemotionen bezeichnet (Ortony und Turner, 1990). Über ihre Anzahl herrscht jedoch auf Grund der verschiedenen Theorieausrichtungen der Emotionsforscher Uneinigkeit, so dass die unterschiedlichen Basisemotionen in Tabelle 1 aufgezeigt werden.

Sowohl das emotionale Musikretrieval (Lee und Neal, 2007) als auch das emotionale Bildretrieval (Schmidt und Stock, 2009) arbeiten mit fünf Basisemotionen: Traurigkeit, Glück, Ärger, Angst und Ekel. Angesichts der psychologischen Fachliteratur haben wir die Menge der fundamentalen Emotionen angereichert. Als grundlegende Emotionen werden folgende für das Projekt ausgewählt: Trauer, Ärger, Angst, Ekel, Überraschung, Sehnsucht, Freude und Liebe. Unabhängig von diesen in der Emotionspsychologie postulierten Basisemotionen wird zusätzlich in Form von Humor bzw. Witz ein weiterer Aspekt mit aufgenommen, um auch dieser medialen Komponente gerecht zu werden. Unter *Trauer* kann man „die mehr oder weniger stereotypisierten Reaktionen auf Verluste [sehen]. Der Verlust betrifft in der Regel eine geliebte Person, aber auch der Wegfall von materiellen Gütern oder der Verlust von sozialen Werten wie Ehre kann zum Anlass für eine Trauerreaktion werden“ (Hoffmann, 1983, 183). Nach Izard dient Trauer (bzw. Kummer) „äußerst nützlichen Aufgaben, indem er einem selbst und anderen mitteilt, dass nicht alles in Ordnung ist, indem er den Menschen motiviert, das Notwendige zu tun, um den Kummer zu verringern. Kummer macht einen aufgeschlossen für die eigenen Probleme und die Probleme der Welt“ (Izard, 1994, 109-110). *Ärger*-Emotionen werden in allen Kulturen gleich ausgedrückt und besitzen eine gleich Erlebnisqualität (Ekman & Friesen, 1971). „Ärger dient zur Regulierung interpersoneller Beziehungen. Er bringt die Androhung zum Ausdruck, die den zielgerichteten Handlungsablauf einer Person störende Barriere beseitigen oder zerstören zu wollen“ (Bornewasser und Mummendey, 1983, 156). *Angst* und Furcht werden in der Emotionspsychologie teilweise unterschieden (Izard, 1994), teilweise aber auch synonym verwendet. Nach McDougall hat Furcht die biologische Funktion des Vermeidens von Verletzungen oder Tod (1908/1960). Nach Schwarzer ist Angst „die Besorgtheit und Aufgeregtheit angesichts von Situationen, die subjektiv als bedrohlich und ungewiss eingeschätzt werden“ (Schwarzer, 1983, 147). McDougall definiert *Ekel* als ein Abstoßungsinstinkt, welcher den Handlungsimpuls des Zurückweisens bzw. Zurückweichens hervorruft, um Schädigung, Krankheit oder Tod zu vermeiden (McDougall, 1908/1960). Als Ursache kann beispielsweise „[m]aterieller oder psychologischer Verfall (‚alles Verdorbene‘) angesehen werden“ (Izard, 1994, 111). Ekel ist eine der fünf Zivilisationsschranken und „ist eine Provokation, weil wir uns gegen die Nähe wehren und das Ekelhafte gleichzeitig […] unser Interesse weckt“ (Vendrell Ferran, 2008, 233 f.). *Überraschung* als Emotion ist ein vorübergehender Zustand und „resultiert aus einem starken Anstieg in neuraler Stimulierung, typischerweise herbeigeführt durch ein plötzliches, unerwartetes Ereignis. Überraschung dient dem sehr nützlichen Zweck, das Nervensystem von gerade bestehender Emotion und gerade bestehende Kognition zu befreien, so dass das Individuum angemessen auf die Reizsituation und die plötzlicher Veränderung reagieren kann, die es erlebt hat“ (Izard, 1994, 109). *Sehnsüchte* sind „intensive längerfristige und gegenwärtig schwer – oder gar nicht – erfüllbare Wünsche nach Personen, Dingen, Ereignissen oder Erfahrungen aus ihrer Vergangenheit, Gegenwart oder Zukunft“ (Wilhelm, 2008). *Freude* ist „gekennzeichnet durch ein Gefühl von Selbstvertrauen und Bedeutsamkeit, ein Gefühl geliebt zu werden und liebenswert zu sein. […] Freude ist nicht nur gekennzeichnet durch Akzeptierung des Selbst, sondern auch von der Akzeptierung von und gute Beziehungen mit anderen Menschen und der Welt im Allgemeinen, insbesondere mit dem Gegenstand der Freude. Freude an anderen führt zu Vertrauen in andere“ (Izard, 1994, 272). Die romantische *Liebe* „versetzt den gesamten Körper und Geist in Hochstimmung und macht den Menschen offen für Erregung und Ekstase und anfällig für affektiv-kognitive Orientierungen“ (Izard, 1994, 117 f.). Damit stellt Liebe eine besondere Beziehung zwischen den Menschen dar. Das Konzept von *Spaß, Humor und Witz* ist eng verwandt mit dem der Freude und ist fester Bestandteil der sozialen Interaktion. Humor im Sinne des emotionalen Retrieval bei medialen Dokumentationseinheiten ist eine Mischung aus dem standardisierten Humor (wie z.B. in Witzen oder Komödien) und spontanem Humor.

Tabelle 1: Basisemotionen und Vertreter (verändert und erweitert nach Ortony und Turner, 1990, 316).

Fundamental Emotion	Basis for Inclusion	Reference
rage and terror, anxiety, joy	Hardwired	Gray (1982)
expectancy, fear, rage, panic	Hardwired	Panksepp (1982)
pain, pleasure	Density of neural firing	Mowrer (1960)
anger, interest, contempt, disgust, distress, fear, joy, shame, surprise	Unlearned emotional states	Tomkins (1984)
acceptance, anger, anticipation, disgust, joy, fear, sadness, surprise	Relation to adaptive biological processes	Plutchik (1980)
anger, aversion, courage, dejection, desire, despair, fear, hate, hope, love, sadness	Relation to action tendencies	Arnold (1960)
anger, disgust, anxiety, happiness, sadness	Do not require propositional content	Oatley & Johnson-Laird (1987)
anger, disgust, fear, joy, sadness, surprise	Universal facial expressions	Ekman, Friesen & Ellsworth (1982)
anger, disgust, elation, fear, subjection, tender-emotion, wonder	Relation to instincts	McDougall (1908/1960; 1926)
anger, contempt, disgust, distress, fear, guilt, interest, joy, shame, surprise	Hardwired	Izard (1971)
desire, happiness, interest, surprise, wonder, sorrow	Forms of action readiness	Frijda (1986)
fear, grief, love, rage	Bodily involvement	James (1884)
fear, love, rage	Hardwired	Watson (1930)
happiness, sadness	Attribution independent	Weiner and Graham (1984)

Abbildung 1: Schieberegler-Tagging von Emotionen.

„Humor durchbricht unsere Routine auf positive Weise, er unterrichtet uns über soziale Interaktion, er lobt und bestraft für soziales Verhalten in fitnessrelevanten Themen [im Sinne Darwins, Anm. d. Autoren] und er schafft Bindungen und Abgrenzungen" (Schwender, 2001, 220).

Gefühls-Tagging mithilfe von Schiebereglern (Forschungsfrage 3)

Die Beantwortung von Forschungsfrage 1 führt – in positiver Formulierung – zur Erkenntnis, dass die sinnvollste Methode, Videos im WWW zu indexieren, im Einsatz einer broad Folksonomy liegt, dass also viele unterschiedliche Personen dieselben Dokumente indexieren. Forschungsfrage 2 leitet uns dazu, neun Basisemotionen zu akzeptieren. Des Weiteren wissen wir, dass Emotionen über Intensitäten verfügen. Auf diese drei Ergebnisse aufbauend, arbeiten wir beim Indexieren von Gefühlsmomenten

- mit einem kontrollierten Vokabular für unsere neun Basisemotionen,
- mit Schiebereglern zur Einstellung der Intensität pro Emotion (Skala: 0 [keine Ausprägung] bis 10 [starke Ausprägung]),
- mit dem Ansatz der broad Folksonomy, lassen also unterschiedliche Nutzer die Dokumente durch Schiebereglereinstellungen taggen.

Wir folgen damit dem Ansatz von Lee und Neal (2007) für emotionales Musikretrieval und von Schmidt und Stock (2009) für emotionales Bilderretrieval. Einer Idee in Schmidt und Stock (2009, 865) folgend, unterscheiden wir bei den Gefühlen in dargestellte Emotionen und in beim Betrachter ausgelöste Emotionen. Diese müssen schließlich nicht übereinstimmend sein. Stellen wir uns beispielsweise vor: Ein breit lächelnder Hooligan (Lächeln ist Ausdruck von Freude) schlägt auf andere Fussballfans ein, dann wird dies nicht bei allen Betrachtern ebenso Freude auslösen, sondern eher Ärger oder Ekel. Die Nutzer haben ebenso stets die Möglichkeit festzustellen, dass im Video gar keine Emotion ausgedrückt wird bzw. dass sie selbst beim Betrachten nichts empfinden. Zusätzlich baten wir die Nutzer (in einer Art Kommentarfeld) anzugeben, welche Gründe für die Emotionen bzw. deren Intensitäten für sie in Betracht kommen. Abbildung 1 zeigt das von uns eingesetzte Schieberegler-Tagging von Basisemotionen.

Methode: Online-Befragung zu Bearbeitungen von YouTube-Videos

Die methodische Bearbeitung des Forschungsprojektes durchläuft folgende fünf Schritte:

(1) Auswahl und Bearbeitung von Videos als Versuchsdokumente
(2) Pretest des Versuchs
(3) Auswahl und Anschreiben der Probanden
(4) Speichern der Antworten der Versuchspersonen
(5) Auswerten der Rohdaten

(1) Insgesamt wurden aus YouTube 20 Videos ausgewählt: je zwei, von denen die Versuchsleiter meinten, dass sie einer der neun Basisemotionen entsprächen und zwei gefühlsneutrale Filme. Pro Fundamentalgefühl hat ein Film eine musikalische Untermalung und der andere nicht. Hiermit wollten wir uns den Weg offen halten zu analysieren, ob die Musik Einfluss auf die zugeschriebenen bzw. beobachteten Gefühle hat. Diese Daten harren noch der Auswertung. (Ein weiterer Projektschritt, über den wir hier allerdings auch nichts berichten, hat 20 weitere Filme – diesmal jedoch in schwarz/weiß – als Forschungsgegenstand. Hieraus erhoffen wir uns Ergebnisse über das Gefühlstagging bei Farb- bzw. Schwarz-Weiß-Filmen).

Mit einer derart großen Anzahl an Videos, die die Versuchspersonen ansehen und taggen sollen, spielt die Videolänge bei der Selektion und Bearbeitung eine zentrale Rolle. Es galt eine Videodauer zu finden, die einerseits ausreicht, um die gezeigten Emotionen zu transportieren und andererseits so kurz wie möglich ausfällt, um die zumutbare Aufmerksamkeitsspanne der Probanden nicht zu überschreiten. Bei der Einschätzung der zumutbaren Aufmerksamkeitsdauer stützten wir uns auf die takeGas-Studie von DoubleClick & Tomorrow Focus AG (Rhode und Lustig, 2007) und die Studie von Tubemogul (Burch, 2008). Beide Studien untersuchten die durchschnittliche Betrachtungsdauer von Online-Videos. Die Ergebnisse der Studien weichen recht deutlich voneinander ab. Während die takeGas-Studie eine Ausstiegsrate von 45 Prozent ab einer Videolänge von 20 Sekunden verzeichnete, zeigte sich in der Tubemogul-Studie ein vergleichbarer Wert erst ab einer Länge von 60 Sekunden. Da sich beide Studien sowohl in Kontext als auch bezüglich der gezeigten Videoinhalte unterscheiden, können wir die Werte lediglich als Orientierungspunkt verwenden. Die Entscheidung fiel auf eine maximale Länge von 25 Sekunden pro Video. Beim Videoschnitt kam das Videobearbeitungsprogramm Jahshaka zum Einsatz, mit dessen Hilfe die Videos auf die gewünschte Länge zugeschnitten und in ein geeignetes Videoformat konvertiert werden konnten.

(2) In einem dem Versuch vorgeschalteten Pretest konnten die Probanden, 14 Studierende der Informationswissenschaft, auf Unstimmigkeiten sowohl formaler als auch inhaltlicher Natur hinweisen. Nach wenigen Korrekturen ist der Versuchsablauf optimiert worden.

(3) Die Ermittlung der Teilnehmer an der Studie erfolgte mit Hilfe des Email-Verteilers des Rechenzentrums der Heinrich-Heine-Universität Düsseldorf. Die dort aufgeführten Studenten erhielten eine Email, in der auf das Forschungsprojekt hingewiesen und um Teilnahme an der Umfrage gebeten wurde. Aus urheberrechtlichen Gründen wurde die Umfrage durch ein in der Rundmail mitgeschicktes Passwort vor Zugriffen durch universitätsfremde Nutzer geschützt.

(4) Um die Speicherung der Umfragedaten und damit die Auswertung zu gewährleisten, stellte die Heinrich-Heine-Universität Speicherplatz auf einem Webserver mit PHP-Unterstützung und eine MySQL Datenbank zur Verfügung. Daraufhin wurde eine PHP-Anwendung entwickelt, welche die HTML-Seiten der Umfrage dynamisch erzeugte und die Umfragedaten der Nutzer in der Datenbank speicherte. Innerhalb der PHP-Applikation waren die grafischen und funktionalen Eigenschaften der Umfrage hinterlegt. Die einzelnen Umfrageseiten basierten auf einer generischen Vorlage, so wurde jeweils nur die Information des abzuspielenden Videos verändert. Die Anzahl der Videos konnte somit beliebig verändert werden, ohne in den eigentlichen Quellcode der Anwendung eingreifen zu müssen. Jedem Studenten wurde zu Anfang der Umfrage eine eindeutige SessionID zugewiesen, welche über ein Cookie im Browser des Teilnehmers gespeichert wird. Diese SessionID ermöglichte es der PHP-Anwendung, den Nutzer auf jeder Seite der Umfrage zu identifizieren. Eine Löschung des Cookies erfolgte nach der Schließung des Browsers.

Die Hauptbestandteile der Datenbank bildeten drei Tabellen. In einer ersten Tabelle, *emo_user* (vgl. Abbildung 2), wurden die Angaben der Teilnehmer bezüglich ihrer Fakultät und Geschlechts verzeichnet. Auch eine Speicherung der jeweiligen SessionID und des Anfangszeitpunktes erfolgte in diesem Teil der Datenbank. Die eingegeben Video-Bewertungen mittels Schieberegler wurden in den Tabellen *emo_data* und *emo_data_complete* abgelegt. Erfolgte eine vollständige Bewertung aller 20 Videos, zog dies eine Speicherung in der Tabelle *emo_data_complete* nach sich. In der Tabelle *emo_data* hingegen wurden auch die bereits abgegebenen Bewertungen bei abgebrochenen Durchläufen gespeichert. Dadurch wären wir zusätzlich in der Lage gewesen, diese Daten bei der Auswertung zu berücksichtigen. Um den Zugriff auf die vollständigen Datensätze (vgl. Abbildung 3) und damit die Auswertung zu gewährleisten, erfolgte eine Übertragung der Daten auf eine unabhängige, Passwort-gesicherte Seite mit Exportfunktion.

776 Probanden haben alle 20 Videos getaggt. Hiervon sind 36 Prozent männlich und 64 Prozent weiblich. Wir hatten also insgesamt 279.360 Schiebereglereinstellungen (776 Personen * 20 Filme * 18 Emotionen) als Basis für weitere Berechnungen zur Verfügung. Wir verzichteten auf die Berechnung von Signifikanzwerten, da bei solch großen Zahlen (nahezu) alle Resultate zu statistisch signifikanten Ergebnissen führen. Für unsere Zwecke reichen Werte der deskriptiven Statistik vollständig aus, da wir ja ausschließlich

Abbildung 2: Attribute der Datenbank emo_user.

Abbildung 3: Datensatzbeispiel.

Tabelle 2: Konsistenz der emotionalen Voten bei den dargestellten Emotionen.

dargestellte Basisemotion	*Standardabweichung*	*Frauen*	*Männer*
Liebe	0,88	0,94	0,82
Ekel	0,92	0,92	0,93
Trauer	1,11	1,09	1,12
Ärger	1,21	1,11	1,30
Spaß	1,46	1,48	1,44
Glück	1,54	1,56	1,52
Sehnsucht	1,55	1,69	1,42
Angst	1,61	1,69	1,54
Überraschung	1,94	2,15	1,74

Anmerkung: N = 776 Versuchspersonen; 279.360 Schiebereglereinstellungen

Tabelle 3: Konsistenz der emotionalen Voten bei den empfundenen Emotionen.

empfundene Basisemotion	*Standardabweichung*	*Frauen*	*Männer*
Liebe	0,66	0,72	0,63
Trauer	0,91	0,90	0,93
Ärger	1,15	1,16	1,14
Sehnsucht	1,15	1,23	1,07
Ekel	1,24	1,28	1,19
Angst	1,36	1,37	1,34
Glück	1,56	1,66	1,47
Überraschung	1,71	1,77	1,65
Spaß	1,75	2,01	1,63

Anmerkung: N = 776 Versuchspersonen; 279.360 Schiebereglereinstellungen

eruieren möchten, ob konsistentes Indexieren von Emotionen überhaupt möglich ist. Wir sind uns dessen bewusst, dass – mit den Studierenden einer einzigen Universität – weder eine Zufallsstichprobe noch eine repräsentative zielgruppenspezifische Stichprobe vorliegt. Durch die große Zahl an Werten hoffen wir jedoch, zumindest die Tendenzen emotionalen Indexierens von Videos aufspüren zu können.
(5) Zur Auswertung der Rohdaten konnte sowohl auf Excel als auch auf SPSS zurückgegriffen werden, da beide Anwendungen die Analyse der gewonnen Daten in Bezug auf die formulierten Forschungsfragen ermöglichten.

Konsistenz der Gefühlsindexierung (Forschungsfrage 4)

Die Standardabweichung ist eine der wichtigsten Streuungsparameter einer diskreten Zufallsvariable. Sie ist in Analogie zu den entsprechenden empirischen Maßzahlen definiert und setzt metrisch skalierte Zufallsvariablen voraus. Wir haben für alle (dargestellten wie empfundenen) Basisemotionen pro Video die Standardabweichung der Voten berechnet und danach pro Gefühl das arithmetische Mittel der Standardabweichungen bestimmt. Grundsätzlich gilt: Je geringer der Wert, desto konsistenter haben die Probanden votiert.
Insgesamt zeigen sich sowohl bei den dargestellten Emotionen (Tabelle 2) als auch bei den empfundenen Gefühlen (Tabelle 3) recht kleine Werte der Standardabweichung und damit hohe Werte der Konsistenz der Einschätzungen der Versuchspersonen. Insbesondere beim Gefühl der Liebe gibt es (bei einer mittleren Standardabweichung von 0,88 bei der dargestellten Emotion und von 0,66 bei der Empfindung) eine sehr hohe Konsistenz, aber auch Ekel, Trauer und Ärger sind offenbar stark konsensfähig. Die Überraschung sorgt für eine Überraschung, denn diese Emotion ist (mit einem Wert von 1,94 beim dargestellten Gefühl) die umstrittenste im gesamten Bewertungsprozess. Auch beim empfundenen Spaß gibt es große Auffassungsunterschiede (die mittlere Standardabweichung beträgt 1,75).
Die kleinen Zahlen der Standardabweichung bedürfen einer Erläuterung. Häufig dominieren bei den Videos nur wenige Emotionen, meist gar nur ein einziges Gefühl. Dies liegt in der Auswahl der Filme begründet, die ja bewusst nach emotionalen Kriterien selektiert worden sind. Damit bekommen jeweils wenige Basisemotionen einen hohen Intensitätswert und die anderen einen Wert nahe Null. Da die Versuchspersonen konsistent mit etwa Null votiert haben, liegt nicht nur das arithmetische Mittel, sondern auch die Standardabweichung bei einem Wert nahe Null.
Ein Nebenergebnis soll hier nicht verschwiegen werden. Eine Trennung der Probanden nach Geschlecht weist interessante Resultate auf: Bei der dargestellten Emotion waren die Frauen fast durchweg uneiniger als die Männer – lediglich bei den Emotionen Trauer, Ärger und Ekel sind sie in ihren Voten konsistenter. Große Unterschiede bemerken wir bei Sehnsucht und Überraschung, hier sind die Männer weitaus konsistenter. Differenzen liegen auch beim Ärger vor, den die weiblichen Versuchspersonen konsistenter den Videos zuschreiben. Auch bei der empfundenen Emotion gibt es geschlechtsspezifische Unterschiede: insbesondere bei Sehnsucht, Glück und Spaß ergibt sich bei den Männern ein weitaus konsistenteres Gefühlserlebnis. Bei der Untersuchung des Gefühls-Tagging bei Bildern von Schmidt und Stock (2009, 871) ergab sich ein ähnliches Bild: Dort indexierten Frauen bei allen Emotionen weniger konsistent als Männer. Eine Erklärung für diese Beobachtung steht noch aus.
Die Forschungsfrage 4 können wir positiv beantworten: In der Tat beurteilen unsere Versuchspersonen Emotionen in Videos konsistent. Damit ist eine zentral wichtige Hürde in Hinblick auf emotionales Retrieval überwunden. Wir können davon ausgehen, dass Nutzer (zumindest bestimmte abgrenzbare

easy Check
library technologies

LÖSUNGEN – STRATEGISCH & INDIVIDUELL MIT EASYCHECK!

RFID- UND EM-TECHNIK FÜR BIBLIOTHEKEN IN JEDER ART UND GRÖSSE

EasyCheck GmbH & Co. KG
Steinbeisstraße 12
73037 Göppingen
DEUTSCHLAND
Fon +49 (0)7161 808600-0
Fax +49 (0)7161 808600-22
mail@easycheck.org

www.easycheck.org

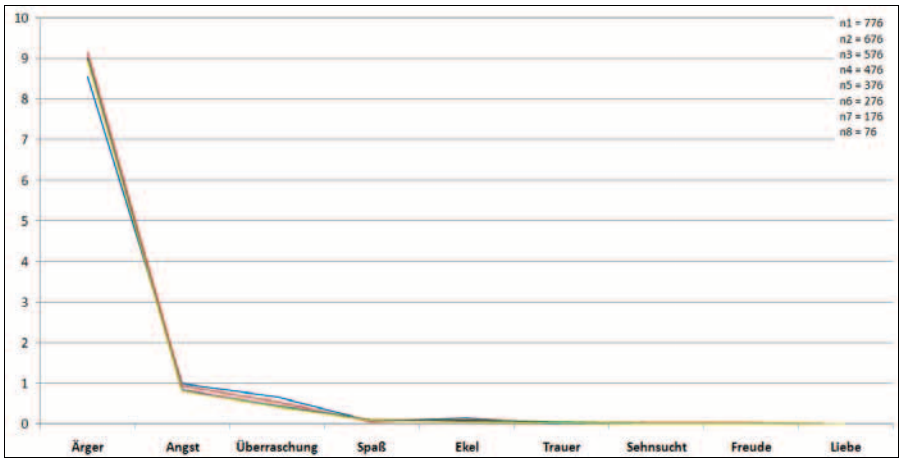


Abbildung 4: Verteilungen der Gefühlsvoten (dargestellte Emotionen) für Video Nr. 20 bei wachsenden Indexiererzahlen.

Tabelle 4: Durchschnittliche Schiebereglereinstellung für Video Nr. 20 (dargestellte Emotionen) bei wachsenden Indexerzahlen.

Block	*Teilnehmerzahl (n)*	*durchschnittliche Schiebereglereinstellung*	*Basisemotion*
1	n = 76	6,3	Ärger / dargestellte Emotionen
2	n = 176	6,6	Ärger / dargestellte Emotionen
3	n = 276	6,8	Ärger / dargestellte Emotionen
4	n = 376	6,7	Ärger / dargestellte Emotionen
5	n = 476	6,7	Ärger / dargestellte Emotionen
6	n = 576	6,7	Ärger / dargestellte Emotionen
7	n = 676	6,6	Ärger / dargestellte Emotionen
8	n = 776	6,6	Ärger / dargestellte Emotionen
1	n = 76	0,7	Angst / dargestellte Emotionen
2	n =176	1,1	Angst / dargestellte Emotionen
3	n = 276	1,0	Angst / dargestellte Emotionen
4	n = 376	0,9	Angst / dargestellte Emotionen
			etc.

Nutzergruppen) den Dokumenten mehr oder minder dieselben Basisemotionen mit zumindest ähnlicher Intensität zuordnen. Ob dieser Sachverhalt global gilt, ist offen. Es ist beispielsweise möglich, dass Nutzer unterschiedlicher Kulturkreise zwar wiederum untereinander konsistent voten, aber – kulturübergreifend – andere Gefühle nennen.

Stabile Verteilung der Gefühls-Tags (Forschungsfrage 5)

Halpin, Robu und Shepherd (2007) haben in ihrer Untersuchung die Dynamik der Tag-Verteilungen in Folksonomies gemessen. Hierbei wurde versucht festzustellen, ab welcher Zeit sich eine stabile Verteilung von Tags in Folksonomies etabliert. Als Maßeinheit wurden jeweils Monate gewählt. Bezogen auf das kollaborative Tagging ist diese Art der Untersuchung durchaus sinnvoll, da in Folksonomies Nutzer in den Social-Media-Plattformen kontinuierlich Dokumente taggen. Bei unserem Forschungsprojekt haben die Probanden dagegen nicht zeitverteilt über Monate hinweg, sondern innerhalb einer kurzen Zeitspanne alle zur Verfügung gestellten Videos angesehen und indexiert. Es gilt herauszufinden, ab welcher Anzahl von Nutzern sich eine stabile Verteilungsform herauskristallisiert. Danach ändert sich wohl die Anzahl der Schiebereglereinstellungen, die aber keine Auswirkung auf den Verlauf der Verteilungskurve hat.

Um nun herauszufinden, ab wie viel Teilnehmern sich eine Stabilität anstellt, werden die Teilnehmerzahlen der jeweiligen Videos in acht Blöcke mit n Teilnehmern unterteilt: 1. Block: n = 76, 2. Block: n = 176 usw. bis zum 8. Block: n = 776. Für jeden dieser Blöcke wurde die durchschnittliche Schiebereglereinstellung pro Emotion für jedes Video ausgerechnet. Wir wollen das Vorgehen am konkreten Beispiel des Videos Nr. 20 verdeutlichen (Tabelle 4). Die durchschnittliche Schiebereglereinstellung für Ärger (als dargestellte Basisemotion) beträgt für die ersten (willkürlich selektierten) 76 Teilnehmer 6,3, für 176 Personen 6,6 und so weiter bis zum endgültigen Wert bei 776 Probanden von 6,6. Das gleiche Verfahren wiederholt sich für die übrigen Gefühle. Alle Verteilungen der dargestellten wie der empfundenen Emotionen für die 20 Videos (für alle Blöcke an Teilnehmern) sind im Anhang beigefügt.

Bei den Auswertungen wird deutlich, dass eine stabile Verteilung bereits im ersten Block erreicht wird. Größere Ausreißer, die sich vom arithmetischen Mittelwert wegbewegen, werden aufgrund der Menge der Nutzer wieder „korrigiert". Das heißt: In der Praxis des emotionalen Indexierens benötigen wir keinesfalls sehr großen Mengen an taggenden Nutzern; ganz im Gegenteil, die Verteilung der Basisemotionen „steht" bereits bei wenigen Dutzenden von Usern. Es ergeben sich drei unterschiedliche Verteilungsformen: *erstens* finden die User keine vorherrschenden Gefühle, so dass sich gar keine Verteilung ergibt (Beispiele: Videos Nr. 7 und 15); *zweitens* zeigen sich Power Law-Verteilungen, bei der ein einziges Gefühl alle anderen dominiert (Beispiele: Videos Nr. 11, 19 und 20, jeweils bei der dargestellten Emotion), drittens ergeben sich invers-logistische Verteilungen, bei denen mehrere Items stark vertreten sind (Beispiele: Video Nr. 5 bei der dargestellten Emotion, Video Nr. 3 bei der empfundenen Emotion). Sowohl das Power Law als auch die invers-logistische Verteilung sind typische Kurvenverläufe bei Rangordnungen informationswissenschaftlicher Gegenstände (Stock, 2006).

Power Tags bei Gefühlen (Forschungsfrage 6)

Können wir aus den Verteilungen diejenigen Basisemotionen eindeutig separieren, die das jeweilige Video am besten emotional beschreiben? Zur Lösung dieses Problems machen wir vom Konzept der „Power Tags" Gebrauch (Peters, 2009, 363-372; Peters und Stock, 2007; Peters und Stock, 2008; Stock, 2007b; Stock und Stock, 2008, 172). „Power tags are tags that best describe the resource's content, or the platform's focal point of interest, according to Collective Intelligence (…), since they reflect the implicit consensus of the user community" definiert Peters (2009, 363). Uns interessieren hier nur die Power Tags auf der Ressourcenebene. Peters (2009, 370) betont, dass die Beschränkung einer Suche auf Power Tags die Präzision dieser Recherche erhöht, da der dokumentspezifische „lange Schwanz" an Tags nicht mehr durchsucht wird.

In Übereinstimmung mit Schmidt und Stock (2009, 871) arbeiten wir mit einem

Tabelle 5: Power Tags bei den dargestellten Basisemotionen.

Video Nr.	Dargestellte Emotion	Mittelwert	Standard-abweichung	Verteilungsform	Distanz
1	Überraschung	4,5	3,60	invers-logistisch	1,0
2	keine				
3	Überraschung Spaß Freude	7,7 5,8 4,0	3,26 2,81 1,15	invers-logistisch	1,9 1,8
4	Angst	5,1	2,25	invers-logistisch	2,2
5	Spaß Überraschung	5,7 5,4	3,40 3,55	invers-logistisch	0,3
6	Ekel	5,0	3,97	Power Law	3,1
7	keine				-
8	keine				
9	Liebe	8,1	4,04	Power Law	5,5
10	keine				
11	Ärger	8,9	3,94	Power Law	8,1
12	keine				
13	Trauer	7,4	3,17	Power Law	4,4
14	Spaß	4,7	4,02	Power Law	2,5
15	keine				
16	Liebe	7,6	4,22	Power Law	4,7
17	Sehnsucht	6,2	4,75	Power Law	2,8
18	Freude	6,2	3,83	Power Law	2,5
19	Angst	8,9	4,05	Power Law	5,3
20	Ärger	6,6	3,35	Power Law	5,8

Anmerkung: Alle Emotionen mit einer Intensität ≥ 4. Distanz: Abstand zur Intensität der nächsten Emotion. N = 776 Versuchspersonen (für jedes Video).

Tabelle 6: Power Tags bei den empfundenen Basisemotionen.

Video Nr.	Empfundene Emotion	Mittelwert	Standard-abweichung	Verteilungsform	Distanz
1	Spaß	5,4	3,39	Power Law	2,7
2	Ekel	6,4	3,59	Power Law	4,8
3	Überraschung Spaß	5,7 5,2	3,77 4,96	invers-logistisch	0,5
4	keine				
5	Spaß Überraschung	5,0 2,9	5,00 3,39	invers-logistisch	2,1
6	Ekel	7,5	4,95	Power Law	6,3
7	keine				
8	keine				
9	keine				
10	keine				
11	keine				
12	Angst	4,5	4,05	Power Law	2,4
13	Trauer	5,0	3,76	Power Law	4,2
14	Spaß	4,5	4,07	Power Law	2,5
15	keine				
16	keine				
17	keine				
18	keine				
19	Angst	5,9	4,14	Power Law	3,4
20	keine				

Anmerkung: Alle Emotionen mit einer Intensität ≥ 4. Distanz: Abstand zur Intensität der nächsten Emotion. N = 776 Versuchspersonen (für jedes Video).

Schwellenwert für die Aufnahme einer Basisemotion als Kandidatin für einen Power Tag von größer oder gleich 4. Für alle Videos, bei denen nicht mindestens ein Gefühl eine Schiebereglereinstellung von mindestens 4 erreicht, wird „keine Emotion" notiert. Unabhängig von der Verteilungsform wird eine Gefühlsbenennung nur dann notiert, wenn die Basisemotion eine Intensität von mindestens 4 vorweisen kann. Bei den dargestellten Emotionen wird bei sechs Videos kein herausragendes Gefühl getaggt, bei den Empfindungen passen die Versuchspersonen bei elf Videos.

Da wir nur über insgesamt neun Basisemotionen verfügen, erscheint es sinnvoll, beim Vorliegen einer Power Law-Verteilung (wie in Abbildung 4) alle Terme ab dem zweiten Rangplatz als „langen Schwanz" anzusehen und dementsprechend abzuschneiden. Eine solche Verteilung folgt der Formel $f(x) = C / x^a$, wobei x der Rangplatz, C eine Konstante und a ein Wert zwischen ca. 1 und ca. 2 ist. Wir nehmen zur Sicherheit einen kleinen Wert für a an (derzeit rechnen wir mit 1). Eine Kurve ist genau dann ein Power Law, wenn der Wert des Items auf Rangplatz 2 gleich oder weniger als die Hälfte des Wertes des Items auf Rang 1 ist. Für alle anderen Fälle wird als Verteilungsform „invers-logistisch" notiert (dies entspricht nicht der Lehrmeinung in der Literatur – die den Wendepunkt der Verteilung als Schwellenwert definiert (Peters, 2009, 369) –, ist aber leicht anzuwenden). Bei den invers-logistischen Verteilungen werden alle Gefühle notiert, deren Intensitätswert größer/gleich 4 ist. So kann es durchaus passieren (wie bei den dargestellten Basisemotionen von Video 3), dass drei Gefühle indexiert werden. Alle in den Tabellen 5 und 6 genannten Gefühle werden in den Datenfeldern für dargestellte bzw. empfundene Basisemotionen in den Datensätzen der Videodokumente gespeichert und stehen so für eine Recherche nach Gefühlen zur Verfügung.

Die Standardabweichungen bei der Bewertung sowohl der dargestellten als auch der empfundenen Top-Emotionen sind ziemlich hoch – zumindest im Vergleich zu den Durchschnittswerten in den Tabellen 2 und 3. So ein Ergebnis deutet darauf hin, dass die Teilnehmer sich nicht besonders einig darüber sind, wie hoch jedes dominante Gefühl tatsächlich bewertet werden soll, obwohl prinzipiell Konsens darüber herrscht, welche Emotion vorherrschend ist (und nur diese Beobachtung ist für unsere Zwecke wichtig). Ist für einen Nutzer ein Wert von, sagen wir, 6 Ausdruck einer starken Intensität, so schiebt ein anderer den Regler auf 10, um das Gleiche auszusagen.

Korrelationen zwischen den Basisemotionen (Forschungsfrage 7)

Welche Basisemotionen treten gemeinsam auf? Gibt es typische Gefühlskonstellationen? Zur Beantwortung dieser Fragen errechnen wir die Korrelationskoeffizienten (nach Pearson, zweiseitig) für alle untersuchten Gefühle.

Sind Gefühle ansteckend (Wild, Erb und Bartels, 2001)? Wenn ja, müssten die in den Videos dargestellten Emotionen auch von den Versuchspersonen empfunden werden. Bei den Power Tags gibt es in der Tat breite Übereinstimmungen zwischen dargestellten und empfundenen Gefühlen: Bei Video 1 ist es der Spaß, bei Videos 3 und 5 sind es Überraschung und Spaß, bei 6 Ekel, bei 13 Trauer und bei Video Nr. 19 die Angst. Die Korrelationen zwischen den gleichen Gefühlen, also zwischen als „dargestellt" und als „empfunden" eingeschätzten gleichen Basisemotionen, zeigt Tabelle 7. Insbesondere bei Trauer und Liebe scheinen diese Gefühle in der Tat ansteckend zu sein: Dargestellte Trauer bzw. dargestellte Liebe korrelieren bei vielen Betrachtern mit erlebter Trauer bzw. mit Liebesgefühl. Dies gilt aber auch (und dies wegen der diversen Null-Bewertungen) für den Fall, dass ein nicht dargestelltes Gefühl mit demselben nicht empfundenen Gefühl korreliert. Wir sollten die Ergebnisse in diesem Abschnitt also nicht überbewerten.

Tabelle 7: Sind Gefühle „ansteckend"?

Basisemotion	*Korrelation (dargestellt – empfunden)*
Trauer	+0,47
Liebe	+0,47
Spaß	+0,39
Ekel	+0,39
Angst	+0,38
Überraschung	+0,38
Ärger	+0,32
Sehnsucht	+0,24
Freude	+0,18

Von den vielen signifikanten Korrelationen (alle mit einer Irrtumswahrscheinlichkeit von kleiner als 0,01) sei hier nur auf die auffälligsten eingegangen (Tabelle 8). Bei der Betrachtung aller Datensätze treten die dargestellte Freude und der dargestellte Spaß, sowie die empfundene Freude und der empfundene Spaß häufig gemeinsam auf: Die Korrelationskoeffizienten liegen hier bei +0,476 und bei +0,386. Beide Emotionen scheinen bei den Umfrageteilnehmern oft gleich verwandt zu werden, weshalb die Schieberegler hier sehr ähnlich bewegt wurden.

Negative Korrelationen bestehen vorwiegend zwischen positiven und negativen Gefühlen. Dargestellte Angst steht demnach im Gegensatz zu dargestellter Freude (-0,16) und zu dargestelltem Spaß (-0,16), empfundene Angst zu dargestellter Freude (-0,15) und dargestelltem Spaß (-0,15). Analoge Gegensatzpaare liegen bei Freude und Trauer sowie bei Spaß und Trauer vor.

Die Basisemotionen Sehnsucht und Überraschung zeigen keine durchgehend klaren Korrelationen zu positiven bzw. negativen Gefühlen. Überraschung korreliert (sogar stark) positiv mit dargestelltem (+0,35) und empfundenem Spaß (+0,32), aber negativ beispielsweise mit dargestellter Liebe (-0,12) und dargestellter Trauer (-0,08). Sehnsucht korreliert positiv z.B. mit empfundener Liebe, aber negativ etwa mit (dargestelltem wie empfundenem) Ärger.

Betrachtet man nur die Werte der männlichen Versuchsteilnehmer, so finden sich die gleichen Auffälligkeiten wie zuvor, wobei diese in diesem Fall zum Teil allerdings noch stärker ausgeprägt sind. Die Korrelation zwischen dargestellter Freude und dargestelltem Spaß beträgt hierbei sogar +0,513. Bei den weiblichen Teilnehmern sind die Korrelationen im Schnitt weniger hoch.

Positive Emotionen korrelieren oft positiv mit anderen positiven Emotionen und negativ mit den negativen Emotionen. Das gleiche gilt umgekehrt auch für negative Emotionen. Allerdings gibt es auch Emotionen wie Sehnsucht und Überraschung, die stark mit positiven, aber auch mit negativen Emotionen korrelieren. Hier scheint es auf den Kontext anzukommen, der bewirkt, dass diese Emotionen eher positiv oder eher negativ eingeschätzt werden.

Fazit und Ausblick

Wir wollen abschließend unsere Forschungsfragen beantworten.

(Ff 1): Ist zufriedenstellendes Retrieval für Filme im Web derzeit praktisch einsetzbar? Contentbasiertes Videoretrieval befindet sich nach wie vor in den Kinderschuhen und ist in der Praxis derzeit nicht nutzbringend einsetzbar. Es gibt zwar Experimente, content-based Emotionen in Videos zu erkennen, von zufriedenstellenden Lösungen sind wir noch weit entfernt. Begriffsorientiertes Videoretrieval arbeitet entweder mit kontrolliertem Vokabular oder mit nutzergenerierten Tags. Erschließt ein einziger Indexer ein Video oder liegen narrow bzw. extended narrow Folksonomies vor, stellt sich das – hier sehr massiv auftretende – Problem der Indexierungsinkonsistenz. Für die Bewältigung der Massen an Videos im Web kommt ausschließlich der Einsatz von Folksonomies infrage. Der optimale Weg wäre demnach die Anwendung von broad Folksonomies. Diese werden aber von derzeit betriebenen Videoservices (wie YouTube) nicht verwendet. Ein zufriedenstellendes Videoretrieval im Web ist somit heutzutage nicht anzutreffen. In der Praxis brauchbares emotionales Videoretrieval gibt es überhaupt nicht.

Tabelle 8: Korrelationen zwischen den Basisemotionen.

		Alle Korrelationen																	
		Dargestellte Emotionen									Empfundene Emotionen								
		Liebe	Freude	Spaß	Überr.	Sehns.	Trauer	Ärger	Ekel	Angst	Liebe	Freude	Spaß	Überr.	Sehns.	Trauer	Ärger	Ekel	Angst
Dargestellte Emotionen	Liebe	1																	
	Freude	0,18	1																
	Spaß	-0,1	**0,48**	1															
	Überraschung	0,01	**0,23**	**0,35**	1														
	Sehnsucht	**0,24**	-0,02	-0,12	-0,06	1													
	Trauer	0,04	-0,14	-0,15	-0,12	**0,34**	1												
	Ärger	-0,12	-0,11	-0,1	-0,07	-0,12	-0,09	1											
	Ekel	-0,1	-0,06	-0,05	0,18	-0,1	-0,1	-0,07	1										
	Angst	-0,09	-0,16	-0,16	0,14	0,01	0,08	-0,04	**0,23**	1									
Empfundene Emotionen	Liebe	**0,47**	0,08	-0,07	-0,04	0,13	0,01	-0,08	-0,07	-0,07	1								
	Freude	0,15	0,18	0,2	0,12	-0,02	-0,1	-0,07	-0,06	-0,11	0,15	1							
	Spaß	-0,11	0,17	**0,39**	**0,3**	-0,12	-0,13	0	-0,05	-0,08	-0,06	**0,39**	1						
	Überraschung	-0,12	0,08	**0,23**	**0,33**	-0,1	-0,08	-0,05	0,09	0,11	-0,07	0,16	**0,32**	1					
	Sehnsucht	0,2	0	-0,1	-0,08	**0,24**	0,13	-0,09	-0,08	-0,07	**0,29**	0,03	-0,09	-0,09	1				
	Trauer	0,01	-0,12	-0,13	-0,1	**0,23**	**0,47**	-0,07	-0,09	0,03	0,06	-0,09	-0,12	-0,09	0,17	1			
	Ärger	-0,07	-0,04	-0,02	-0,03	-0,06	-0,05	**0,32**	-0,04	-0,02	-0,05	-0,07	-0,07	-0,03	-0,04	-0,02	1		
	Ekel	-0,12	-0,12	-0,11	0,02	-0,13	-0,13	-0,07	**0,39**	0,08	-0,08	-0,1	-0,11	0,08	-0,1	-0,1	0,02	1	
	Angst	-0,1	-0,15	-0,14	0,05	-0,1	-0,05	0,02	0,18	**0,38**	-0,06	-0,11	-0,11	0,12	-0,06	-0,01	0,03	0,24	1

N=15520 Wertungen
776 Teilnehmer
Pearson, zwei-seitig

Anmerkung: N = 776 Versuchspersonen; 279.360 Schiebereglereinstellungen; Korrelationen nach Pearson (zweiseitig).

(Ff 2): Was sind Basisemotionen? Voruntersuchungen zum emotionalen Retrieval (Lee und Neal, 2007, sowie Schmidt und Stock, 2009) sowie das psychologische Schrifttum zu Emotionen legen nahe, dass es mindestens fünf (Freude, Trauer, Angst, Ärger, Ekel), wahrscheinlich aber eher neun (zusätzlich: Überraschung, Sehnsucht, Liebe, Spaß) fundamentale Gefühle gibt. Emotionen haben eine Intensität, können also als schwach oder stark wahrgenommen werden.

(Ff 3): Wie können wir Basisemotionen indexieren, um sie im emotionalen Retrieval praktisch einzusetzen? Zur Abhilfe dieser misslichen Lage im Videoretrieval schlagen wir den Einsatz von kontrolliertem Vokabular (Bezeichnungen der Basisemotionen) und Schiebereglern (zur Abschätzung der Intensität der Emotion) im Rahmen einer broad Folksonomy zur Indexierung des emotionalen Gehalts von Videos vor.

(Ff 4): *Wie konsistent sind nutzerorientierte Verfahren beim Taggen von Emotionen in Videos?* Die Konsistenz der Nutzervoten, gemessen an der Standardabweichung vom Mittelwert, ist derart hoch (ca. 1 bis 2 bei einer Skala von 0 bis zehn), dass wir von einer genügenden Übereinstimmung der Indexer ausgehen können. Einige Gefühle – allem voran Liebe – sind stark konsensfähig.

(Ff 5): *Wie viele unterschiedliche Nutzer braucht man, bis sich eine stabile Verteilung der Gefühls-Tags ergibt? Gibt es überhaupt stabile Verteilungen bei Gefühls-Tags?* Es gibt in der Tat stabile Verteilungen der Gefühle pro Video. Diese Verteilungen zeigen sich bereits bei geringen Nutzerzahlen (< 100) und bleiben in der Form stabil, egal, wie viele weitere User das Video taggen. Wir können drei typische Verteilungsformen separieren: keine Emotion (und damit keine Verteilung) vorhanden, Power Law-Verteilung und invers-logistische Verteilung.

(Ff 6): Können wir aus den Tagverteilungen Power Tags separieren, die den dargestellten bzw. empfundenen Gefühlen entsprechen? Bei der Verteilung der Basisemotionen nach Power Law ergibt sich genau ein Gefühl, das dem Dokument als emotionaler Zugangspunkt zugeordnet wird. Bei invers-logistischen Verteilungen sind es mehrere Gefühle (zwei oder drei), die dem Video zukommen. Wir müssen zwischen dargestellten und empfundenen Emotionen unterscheiden, da die Nutzer diese auch unterschiedlich beurteilen.

(Ff 7): Korrelieren gewisse Emotionen miteinander? Bei den Korrelationen unterscheiden wir zwischen der Analyse gleicher Emotionen bei Darstellung und Empfindung sowie der Analyse der Zusammenhänge zwischen unterschiedlichen Gefühlen. „Ansteckende" Basisemotionen (also solche mit hohen Korrelationswerten zwischen dargestellt und empfunden) sind vor allem Trauer und Spaß. Gefühlsübergreifend positiv korrelieren besonders ausgeprägt Freude und Spaß, negativ beispielsweise Angst mit Freude und mit Spaß.

Ausblick. Wir können anhand der Resultate unseres groß angelegten Indexierungsexperiments mit hoher Wahrscheinlichkeit annehmen, dass Nutzer in der Lage sind, (vorgegebene) Basisemotionen bei Videos konsistent zu erkennen. Um stabile Verteilungen der Gefühle pro Video zu erhalten, benötigen wir nur recht kleine Zahlen (kleiner hundert) an beteiligten Usern. Diese Ergebnisse stehen im Einklang mit den Resultaten beim emotionalen Bilderindexieren (Schmidt und Stock, 2009). Dies ermutigt uns, als nächsten Schritt im EmIR-Forschungsprogramm die Entwicklung eines Prototyps eines Retrievalsystems für emotionale Dokumente – als „Media Emotion Search (Memose)" – in Angriff zu nehmen.

Danksagung. Alexander Kochanek, Natalia Mamerow und Michael Mamerow halfen bei der Vorbereitung und Durchführung der empirischen Untersuchung.

Literatur

Aigrain, P., Zhang, H., & Petkovic, D. (1996). Content-Based Representation and Retrieval of Visual Media: A State-of-the-Art Review. In: Multimedia Tools and Applications, 3(4), 179-202.

ARD/ORF/ZDF (2008). Regelwerk Mediendokumentation: Fernsehen. Richtlinien für die Formalbeschreibung und Inhaltserschließung von Fernsehproduktionen (Version 1.0, Stand: 9. April 2008).

Arnold, M.B. (1960). Emotion and Personality. New York: Columbia University Press.

Bailenson, J.N., Pontikakis, E.D., Mauss, I.B., Gross, J.J., Jabon, M.E., Hutcherson, C.A., Nass, C., & John, O. (2008). Real-time Classification of Evoked Emotions Using Facial Feature Tracking. In: International Journal Human-Computer Studies, 66(5), 303-317.

Boehner, K., DePaula, R., Dourish, P., & Sengers, P. (2007). How Emotion is Made and Measured. In: International Journal of Human-Computer Studies, 65(4), 275-291.

Bornewasser, M., & Mummendey, A. (1983). Ärger. In: Euler, H., & Mandl, H. (Eds.): Emotionspsychologie. Ein Handbuch in Schlüsselbegriffen (S. 156-162). München: Urban & Schwarzenberg.

Burch, D. (2008). How Much of a Typical Video Online Is Actually Watched? www.tubemogul.com/research/report/18 [10.5.2009]

Chan, C.H., & Jones, G.J.F. (2005). Affect-Based Indexing and Retrieval of Films. In: Proceedings of the 13th Annual ACM International Conference on Multimedia (S. 427-430). New York: ACM.

Chen, L., Chen, G.C., Xu, C.Z., March, J., & Benford, S. (2008). EmoPlayer: A Media Player for Video Clips with Affective Annotations. In: Interacting with Computers, 20(1), 17-28.

Colombo, C., Del Bimbo, A., & Pala, P. (1999). Semantics in Visual Information Retrieval. In: IEEE Multimedia, 6(3), 38-53.

Darwin, C. (1872). The Expression of the Emotions in Man and Animals. London: John Murray.

Del Bimbo, A. (1999). Visual Information Retrieval. London: Academic Press.

Ekman, P., & Friesen, W.V. (1971). Constants Across Cultures in the Face and Emotion. In: Journal of Personality and Social Psychology, 17, 124-129.

Ekman, P., Friesen, W.V., & Ellsworth, P. (1982). What Emotion Categories or Dimensions Can Observers Judge from Facial Behavior? In: Ekman, P. (Ed.): Emotion in the Human Face (S. 39-55). New York: Cambridge University Press.

Enser, P. (2008a). Visual Image Retrieval. In: Annual Review of Information Science and Technology, 42, 3-42.

Enser, P. (2008b). The Evolution of Visual Information Retrieval. In: Journal of Information Science, 34(4), 531-546.

Enser, P., Sandom, C., Hare, J., & Lewis, P. (2007). Facing the Reality of Semantic Image Retrieval. In: Journal of Documentation, 63(4), 465-481.

Fehr, B., & Russell, A. (1984). Concept of Emotion Viewed from a Prototype Perspective. In: Journal of Experimental Psychology: General, 113, 464-486.

Frijda, N, H. (1986). The Emotions. New York: Cambridge University Press.

Geisler, G., & Burns, S. (2007). Tagging Videos: Conventions and Strategies of the YouTube Community. In: Proceedings of the 7th ACM/IEEE-CS Joint Conference on Digital Libraries (S. 480). New York: ACM.

Goodrum, A.A. (2000). Image Information Retrieval: An Overview of Current Research. In: Informing Science, 3(2), 63–67.

Gupta, A., & Jain, R. (1997). Visual Information Retrieval. In: Communications of the ACM, 40(5), 71-79.

Halpin, H., Robu, V., & Shepherd, H. (2007). The Complex Dynamics of Collaborative Tagging. In: Proceedings of the 16th International World Wide Web Conference (WWW'07), Banff, Canada, 2007 (S. 211-220). New York: ACM.

Hanjalic, A., & Xu, L. (2005). Affective Video Content Representation and Modeling. In: IEEE Transactions on Multimedia, 7(1), 143-154.

Hoffmann, N. (1983). Trauer. In: Euler, H., & Mandl, H. (Eds.): Emotionspsychologie. Ein Handbuch in Schlüsselbegriffen (S. 183-188). München: Urban & Schwarzenberg.

Izard, C, E. (1994). Die Emotionen des Menschen. Eine Einführung in die Grundlagen der Emotionspsychologie. Weinheim: Psychologie Verlags Union.

Izard, C, E. (1977). Human Emotions. New York: Plenum Press.

Izard, C, E. (1971). The Face of Emotion. New York: Appleton-Century-Crofts.

James, W. (1884). What is an Emotion? In: Mind, 9, 188-205.

Jörgensen, C. (1998). Attributes of Images in Describing Tasks. In: Information Processing & Management, 34(2/3), 161–174.

Jörgensen, C. (2007). Image Access, the Semantic Gap, and Social Tagging as a Paradigm Shift. In: Proceeding of the 18th Workshop of the American Society for Information Science and Technology Special Interest Group in Classification Research, Milwaukee, WI.

Kim, H., & Kim, Y. (2009). A Two-Step Model for Video Key-Frame Determination. In: Proceedings of the 72nd ASIS&T Annual Meeting. Thriving on Diversity – Information Opportunities in a Pluralistic World.

Kim, M.G., & Park, Y.T. (2003). Information Retrieval Agents Using Emotional Features. In: Palau Salvador, C.E. (Ed.): Proceedings of the Second IASTED International Conference on Communication Systems and Networks. Anaheim: Acta Press.

Kleinginna, P.R., & Kleinginna, A.M. (1981). A Categorized List of Emotion Definitions, with Suggesti-

ons for a Consensual Definition. In: Motivation and Emotion, 5(4), 345-379.

Laine-Hernandez, M., & Westman, S. (2006). Image Semantics in the Description and Categorization of Journalistic Photographs. In: Proceedings of the 69th Annual Meeting of the American Society for Information and Technology. Information Realities: Shaping the Digital Future of All.

Lee, H.J., & Neal, D. (2007). Toward Web2.0 Music Information Retrieval: Utilizing Emotion-Based, User-Assigned Descriptors. In: Proceedings of the 70th ASIS&T Annual Meeting. Joining Research and Practice: Social Computing and Information Science.

Markey, K. (1984). Interindexer Consistency Tests: A Literature Review and Report of a Test of Consistency in Indexing Visual Materials. In: Library & Information Science Research, 6, 155–177.

Markkula, M., & Sormunen, E. (2000). End-User Searching Challenges Indexing Practices in the Digital Newspaper Photo Archive. In: Information Retrieval, 1(4), 259–285.

McDougall, W. (1908/1960). An Introduction to Social Psychology. Boston, Mass.: Methuen.

McDougall, W. (1926). An Outline of Abnormal Psychology. Boston, Mass.: Luce.

Melenhorst, M., Grootveld, M., van Setten, M., & Veenstra, M. (2008). Tag-Based Information Retrieval of Video Content. In: Proceedings of the 1st International Conference on Designing Interactive User Experiences for TV and Video (S. 31-40). New York: ACM.

Meyer, W.U., Reisenzein, R., & Schützwohl, A. (2001). Einführung in die Emotionspsychologie. Band I: Die Emotionstheorien von Watson, James und Schachter. Bern: Verlag Hans Huber.

Ortony, A., & Turner, T.J. (1990). What's Basic About Basic Emotions? In: Psychological Review, 97, 315-331.

Panksepp, J. (1982). Toward a General Psychobiological Theory of Emotions. In: The Behavioral and Brain Sciences, 5, 407-467.

Peters, I. (2009). Folksonomies. Indexing and Retrieval in Web 2.0. Berlin: De Gruyter Saur. (Knowledge and Information. Studies in Information Science).

Peters, I., & Stock, W.G. (2007). Folksonomy and Information Retrieval. In: Proceedings of the 70th ASIS&T Annual Meeting. Joining Research and Practice: Social Computing and Information Science.

Peters, I., & Stock, W.G. (2008). Folksonomies in Wissensrepräsentation und Information Retrieval. In: Information – Wissenschaft und Praxis, 59(2), 77-90.

Petrelli, D., & Auld, D. (2008). An Examination of Automatic Videoretrieval Technology on Access to the Contents of an Historical Video Archive. In: Program: Electronic Library and Information Systems, 42(2), 115-136.

Picard, R.W. (1995). Affective Computing. M.I.T Media Laboratory Perceptual Computing Section Technical Report, 321. Cambridge, Mass.: MIT.

Plutchik, R. (1980). A General Psychoevolutionary Theory of Emotion. In: Emotion: Theory, Research, and Experience, Vol. 1: Theories of Emotion (S. 3-33). New York: Academic.

Rohde, U., & Lustig, S. (2007). Die takeGas-Studie von DoubleClick und der TOMORROW FOCUS AG. DoubleClick: Digital Advertising: http://emea.doubleclick.com/DE/downloads/pdfs/070827_Berichtsband%20takeGAS-Studie_final.pdf [10.5.2009]

Rosch, E. (1975). Cognitive Reference Points. In: Cognitive Psychology, 7(4), 532-547.

Salway, A., & Graham, M. (2003). Extracting Information about Emotions in Films. In: Proceedings of the 11th ACM International Conference on Multimedia (pp. 299-302). New York: ACM.

Schmidt, S., & Stock, W.G. (2009). Collective Indexing of Emotions in Images. A Study in Emotional Information Retrieval. In: Journal of the American Society for Information Science and Technology, 60(5), 863–876.

Schmidt-Atzert, L. (1996). Lehrbuch der Emotionspsychologie. Stuttgart: Kohlhammer.

Schwarzer, R. (1983). Angst und Furcht. In: Euler, H., & Mandl, H. (Eds.), Emotionspsychologie. Ein Handbuch in Schlüsselbegriffen (S. 147-154). München: Urban & Schwarzenberg.

Schweins, K. (1997). Methoden zur Erschließung von Filmsequenzen. Köln: FH Köln. (Kölner Arbeitspapiere zur Bibliotheks- und Informationswissenschaft; 5).

Schwender, C. (2001). Medien und Emotionen. Wiesbaden: Deutscher Universitäts-Verlag.

Smeaton, A.F. (2007). Techniques Used and Open Challenges to the Analysis, Indexing and Retrieval of Digital Video. In: Information Systems, 32(4), 545-559.

Smeaton, A.F., Wilkins, P., Worring, M, de Rooij, O., Chua, T.S., & Luan, H. (2008). Content-Based Videoretrieval: Three Example Systems from TRECVid. In: International Journal of Imaging Science and Technology, 18(2-3), 195-201.

Smith, G. (2008). Tagging. People-Powered Metadata for the Social Web. Berkeley, Cal.: New Riders.

Snoek, C.G.M., & Worring, M. (2009). Concept-Based Videoretrieval. In: Foundations and Trends in Information Retrieval, 4(2), 215-322.

Sokolowski, K. (2002). Emotion. In: Müsseler, J. & Prinz, W. (Eds.): Allgemeine Psychologie (S. 337-384). Heidelberg, Berlin: Spektrum Akademischer Verlag.

Soleymani, M., Chanel, G., Kierkels, J.J.M., & Pun, T. (2008). Affective Ranking of Movie Scenes Using Physiological Signals and Content Analysis. In: Proceedings of the 2nd ACM Workshop on Multimedia Semantics (pp. 32-39). New York: ACM.

Stock, W.G. (2006). On Relevance Distributions. In: Journal of the American Society for Information Science and Technology, 57(8), 1126-1129.

Stock, W.G. (2007a). Information Retrieval. Informationen suchen und finden. München, Wien: Oldenbourg.

Stock, W.G. (2007b). Folksonomies and Science Communication. A Mash-Up of Professional Science Databases and Web 2.0 Services. In: Information Services & Use, 27, 97-103.

Stock, W.G., & Stock, M. (2008). Wissensrepräsentation. Informationen auswerten und bereitstellen. München: Oldenbourg.

Tomkins, S.S. (1962). Affect, Imagery, Consciousness, Vol. I: The Positive Affects. New York: Springer Publishing.

Tomkins, S.S. (1963). Affect, Imagery, Consciousness, Vol. II: The Negative Affects. New York: Springer Publishing.

Van Velsen, L., & Melenhorst, M. (2009). Incorporating User Motivations to Design for Video Tagging. In: Interacting with Computers, 21(3), 221-232.

Vendrell Ferran, I. (2008). Die Emotionen. Gefühle in der realistischen Phänomenologie. Berlin: Akademie Verlag.

Watson, J. B. (1930). Behaviorism. Chicago, Ill.: University of Chicago Press.

Wild, B., Erb, M., & Bartels, M. (2001). Are Emotions Contagious? Evoked Emotions While Viewing Emotionally Expressive Faces: Quality, Quantity, Time Course and Gender Differences. In: Psychiatry Research, 102(2), 109-124.

Weller, K., Peters, I., & Stock, W.G. (2010). Folksonomy. The Collaborative Knowledge Organization System. In: Dumova, T., & Fiordo, R. (Eds.): Handbook of Research on Social Interaction Technologies and Collaborative Software: Concepts and Trends (S. 132-146). Hershey, PA: Information Science Reference.

Weiner, B., & Graham, S. (1984). An Attributional Approach to Emotional Development. In: Izard, E., Kagan, J., & Zajonc, R. (Eds.): Emotions, Cognition, and Behavior (S. 167-191). New York: Cambridge University Press.

Wilhelm, K. (2008). Zwischenstation Sehnsucht. In: MaxPlanckForschung, Nr. 3, 66-72.

Xu, M., Chia, L., & Jin, J. (2005). Affective Content Analysis in Comedy and Horror Videos by Audio Emotional Event Detection. In: IEEE International Conference on Multimedia and Expo (ICME 2005).

Information Retrieval, Film, inhaltliche Erschließung, empirische Untersuchung, Folksonomy, Schieberegler, emotionales Retrieval, EmIR, Basisemotion, Gefühl, Tagging, Tag-Verteilung, Power Tag

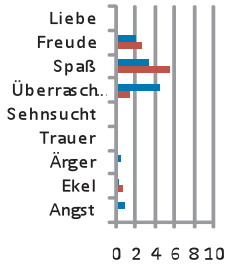
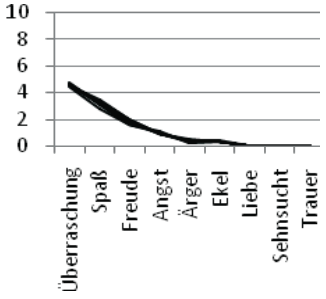
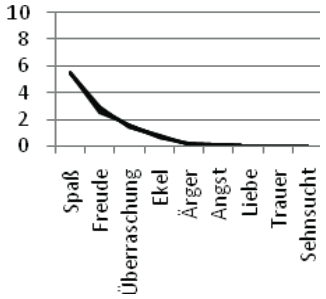
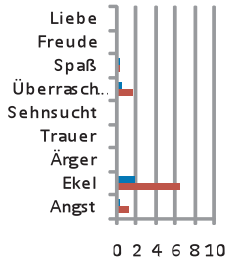
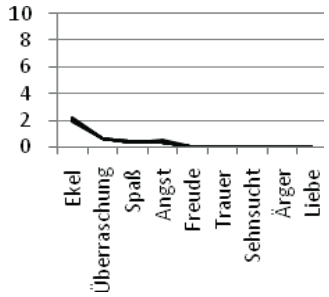
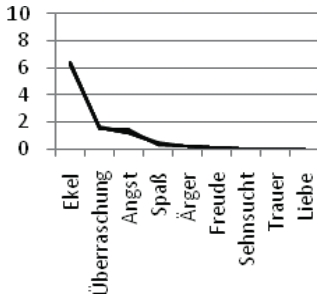
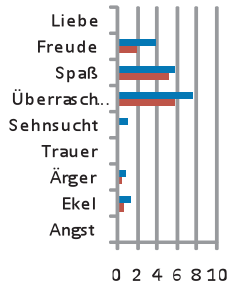
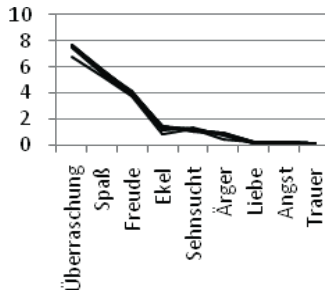
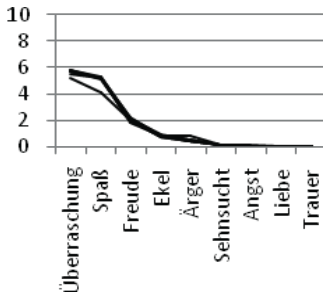
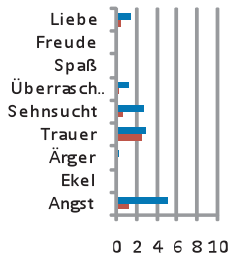
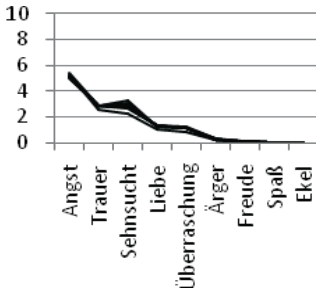
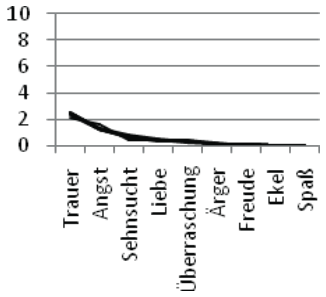
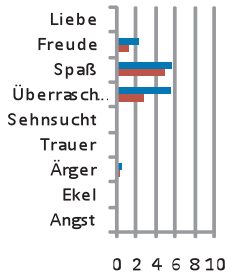
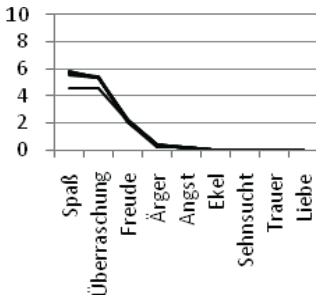
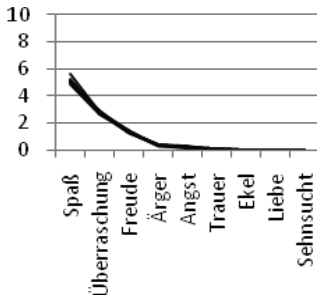
Das Autorenteam

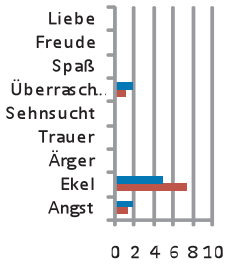
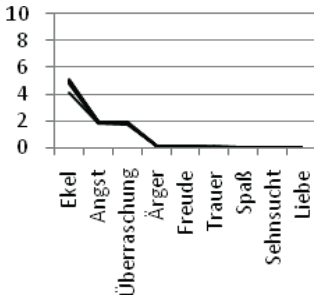
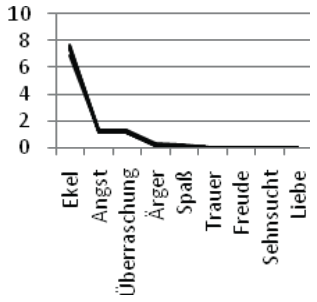
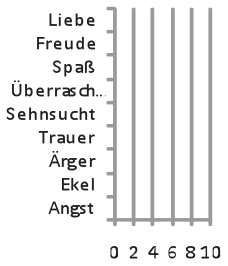
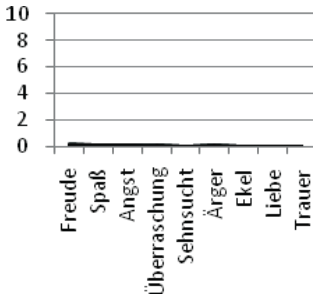
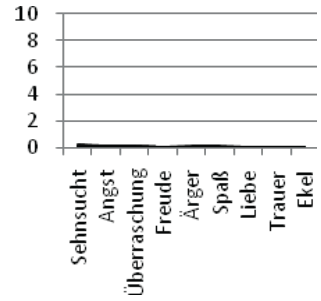
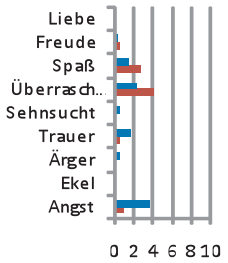
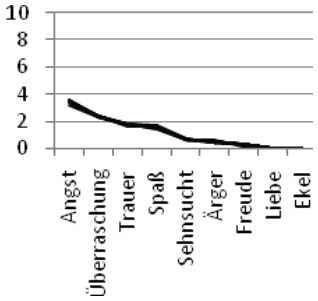
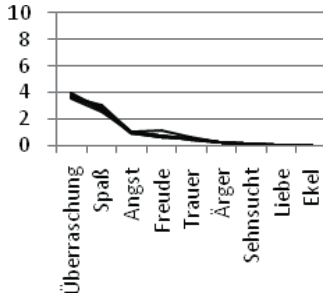
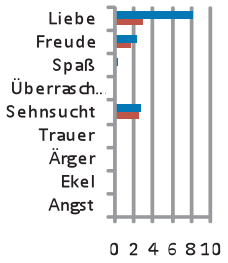
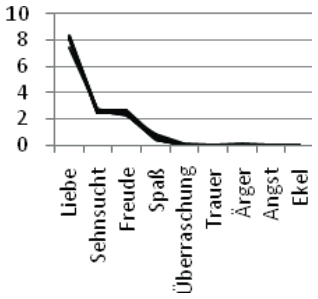
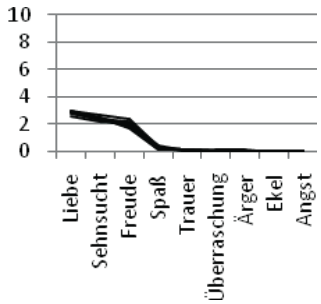
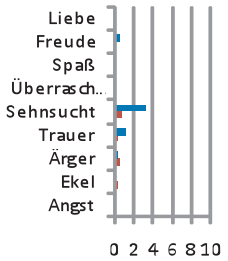
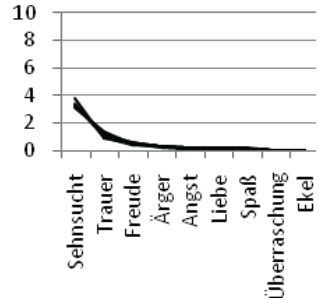
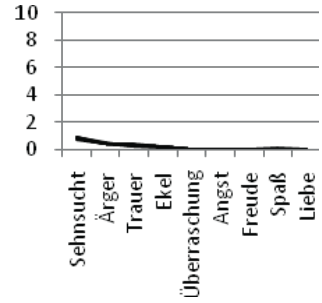
Evelyn Dröge, Susanne Finkelmeyer, Daniel Guschauski, Kerstin Juchem, Kathrin Knautz, Cynthia Krzmyk, Daniel Miskovic, Jan Schiefer, Evrim Sen, Julia Verbina und Nils Werner sind Master- oder Promotionsstudierende der Informationswissenschaft an der Heinrich-Heine-Universität Düsseldorf.

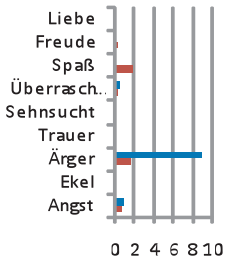
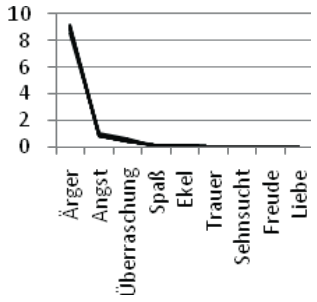
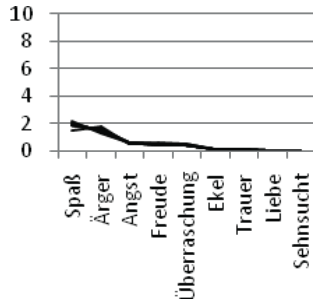
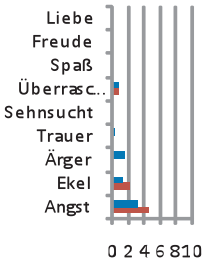
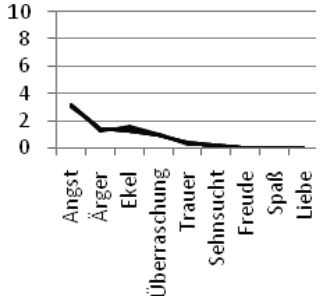
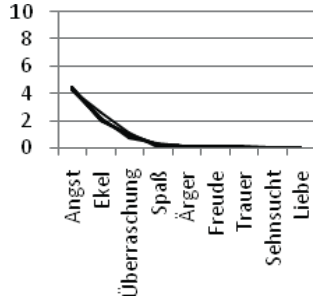
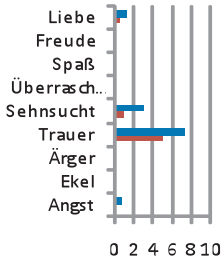
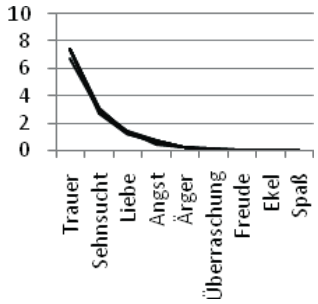
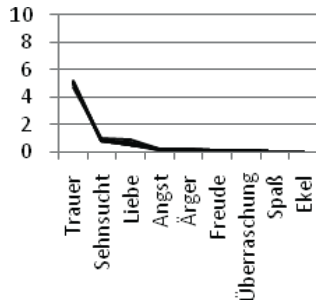
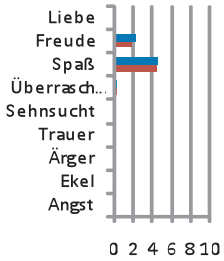
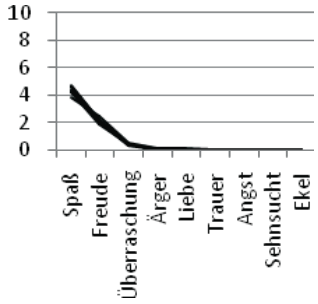
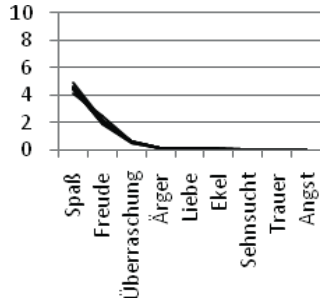
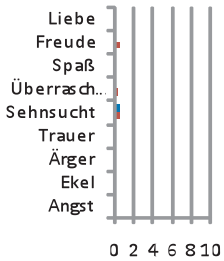
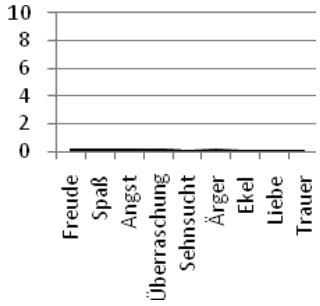
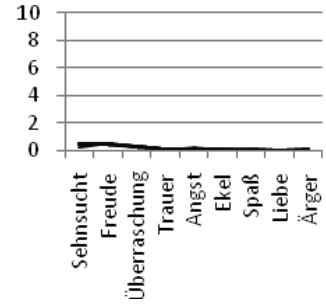
Wolfgang G. Stock ist Leiter der Abteilung für Informationswissenschaft in Düsseldorf.

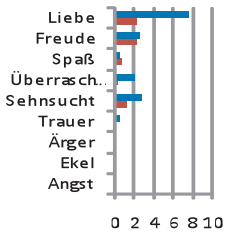
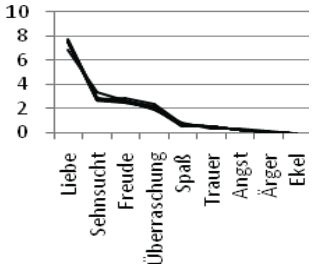
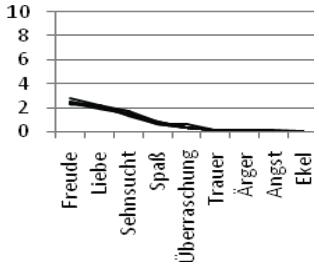
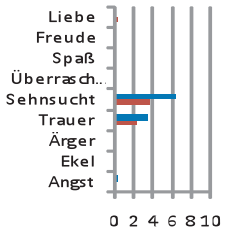
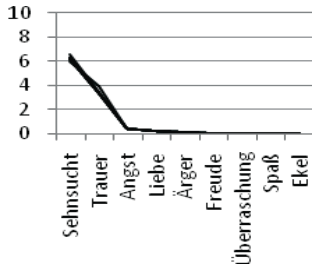
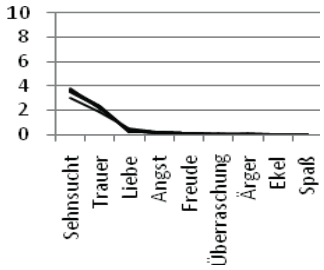
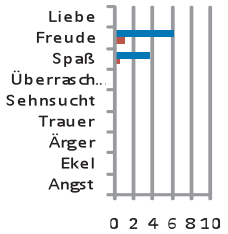
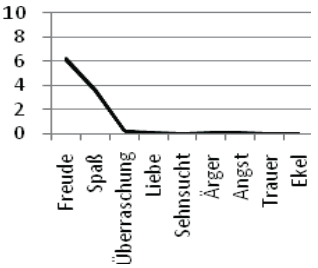
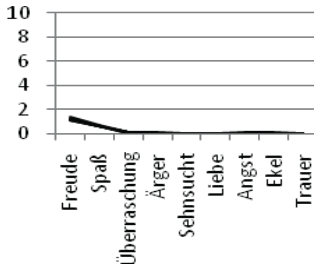
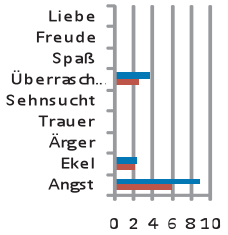
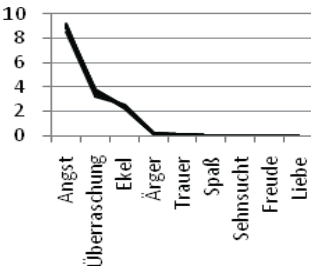
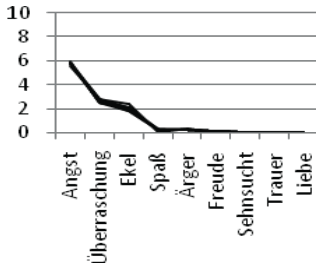
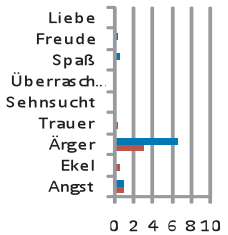
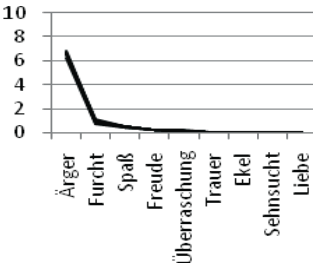
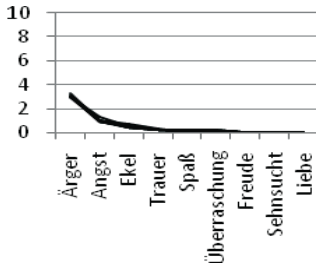
Kontakt:
Kathrin.Knautz@uni-duesseldorf.de

Anhang

Video-Details	Gesamt-Ergebnis der Umfrage	Dargestellte Emotionen	Empfundene Emotionen
Video 1 Länge: 23 Sek. Musik: ja Quelle: Youtube youtube.com/watch?v=6HyDgu4UNTk 			
Video 2 Länge: 17 Sek. Musik: nein Quelle: Youtube youtube.com/watch?v=FCTMJwKI4xU 			
Video 3 Länge: 24 Sek. Musik: nein Quelle: Youtube youtube.com/watch?v=Zb_F9Xg2X7s 			
Video 4 Länge: 23 Sek. Musik: ja Quelle: Youtube youtube.com/watch?v=iF_J9rLXuTA 			
Video 5 Länge: 9 Sek. Musik: nein Quelle: Youtube youtube.com/watch?v=YTfY5NG67aQ 			

Video-Details	Gesamt-Ergebnis der Umfrage	Dargestellte Emotionen	Empfundene Emotionen
Video 6 Länge: 16 Sek. Musik: ja Quelle: Youtube youtube.com/watch?v=tDVs-wxw7Tc 			
Video 7 Länge: 8 Sek. Musik: nein Quelle: Youtube youtube.com/watch?v=fEnxxIYc_0Y 			
Video 8 Länge: 20 Sek. Musik: ja Quelle: Youtube youtube.com/watch?v=0GgYumfE2GY 			
Video 9 Länge: 21 Sek. Musik: ja Quelle: Youtube youtube.com/watch?v=YqU8QLxQFL0 			
Video 10 Länge: 11 Sek. Musik: nein Quelle: Youtube youtube.com/watch?v=q2PZLwbicQQ 			

Video-Details	Gesamt-Ergebnis der Umfrage	Dargestellte Emotionen	Empfundene Emotionen
Video 11 Länge: 21 Sek. Musik: nein Quelle: Youtube youtube.com/watch?v=jdhkUL-5FI4 			
Video 12 Länge: 21 Sek. Musik: nein Quelle: Youtube youtube.com/watch?v=jMjnH8H5tJ8 			
Video 13 Länge: 17 Sek. Musik: nein Quelle: Youtube youtube.com/watch?v=yPoJ6tYw_pE 			
Video 14 Länge: 21 Sek. Musik: nein Quelle: Youtube youtube.com/watch?v=ZLIf6N7LA4s 			
Video 15 Länge: 16 Sek. Musik: ja Quelle: Youtube youtube.com/watch?v=u8U32srGqQc 			

Video-Details	Gesamt-Ergebnis der Umfrage	Dargestellte Emotionen	Empfundene Emotionen
Video 16 Länge: 21 Sek. Musik: nein Quelle: Youtube youtube.com/watch?v=MujfX70w42w 			
Video 17 Länge: 22 Sek. Musik: ja Quelle: Youtube youtube.com/watch?v=kZa7OYWPFPc 			
Video 18 Länge: 5 Sek. Musik: ja Quelle: Youtube youtube.com/watch?v=aCDgdbywRZI 			
Video 19 Länge: 21 Sek. Musik: ja Quelle: Youtube youtube.com/watch?v=1tEtA28T39M 			
Video 20 Länge: 21 Sek. Musik: ja Quelle: Youtube youtube.com/watch?v=sBkLhQ1FMb4 			

Aspekte geistigen Eigentums in der Wissensgesellschaft

Volker Gross, Potsdam

Vor dem Hintergrund der Entwicklung der Informations- und Kommunikationstechnologien hat sich eine so genannte digitale Ökonomie konstituiert. Die damit einhergehende ‚Entmaterialisierung' der Ökonomie und die Entstehung eines digitalen Warenuniversums erfährt allerdings verschiedene Interpretationen. Von den einen als Bedrohung traditioneller Geschäftsmodelle begriffen, ist es den anderen Anknüpfungspunkt in erster Linie technikinduzierter Emanzipationsvorstellungen. Das geistige Eigentum wird in diesem Zusammenhang zu einem Konfliktfeld zwischen restriktiver Sicherung exklusiver Eigentumsrechte und Verteidigung der ‚Wissens-Allmende'. Beide Positionen werden summarisch vorgestellt und als zwei Seiten eines dem geistigen Eigentum grundsätzlich inhärenten Widerspruchs identifiziert. Die ‚Wissensgesellschaft' bezeichnet hierbei den politökonomischen Rahmen dieser Auseinandersetzung um Verwertungsmöglichkeit des Wissens und seiner Verfügbarkeit als entscheidendem Produktionsfaktor.

Aspects of Intellectual Property in the Knowledge Society

In the context of the development of the information and communication technology a digital economy has emerged. This is accompanied by a 'dematerialisation' of the economy and the evolution of a universe of digital goods which both provoke different interpretations. On the one hand this is considered as a threat to traditional business models. On the other hand it allows to tie in with ideas of technology based emancipation. Intellectual property becomes the point at issue in the dispute between restrictive measures to provide exclusive ownership and the defence of the 'commons of the mind'. Both positions get to be presented in summary and identified as two sides of a contradiction inherent to intellectual property. The 'knowledge society' here names the polit-economic context of this conflict about the commercial exploitability of an intellectual good and its availability as resource crucial to production.

Der Begriff der Wissensgesellschaft prägt seit der Mitte der 1990er Jahre die Diskussion in Wissenschaft und Politik. Im Anschluss an Daniel Bells erstmals 1973 unter dem Titel ‚The Coming of the Post-Industrial Society' veröffentlichten Studie, können als Merkmale der Wissensgesellschaft zunächst vor allem die Rede von der ‚Tertiarisierung', d.h. einer zunehmenden „sektoralen Verschiebung von Produktions- zu Dienstleistungstätigkeiten"[1] sowie der Aufstieg des Wissens als eigenständigem, wenn nicht entscheidendem Produktionsfaktor neben Arbeit und Kapital bestimmt werden. ‚Klassische' für die Industriegesellschaft typische, produktive Arbeiten würden zusehends abgelöst durch ‚Wissensarbeit' und ‚Wissensmanagement', als sowohl autonomeren als auch kreativeren und damit schließlich, so Bell, humaneren Formen der Arbeit. Zugleich erhöhe die allgemein verfügbare Information die Partizipationsmöglichkeiten und zu guter Letzt würde auch das Nord-Süd-Gefälle im Zuge der Globalisierung von Information und Kommunikation verringert.

Als Hintergrund dieser wohl nicht zu Unrecht als optimistisch zu bezeichnenden Prognose erweist sich eine Entwicklung in den modernen Industriegesellschaften, die in ihren Anfängen schon angelegt, in den 1970er Jahren als ein parallel verlaufender Prozess der Restrukturierung des westlichen Kapitalismus und der so genannten ‚mikroelektronischen Revolution' im Zuge der Krise des Fordismus einen neuen Schub erfährt. Die Mikroelektronik erlaubt es, die technische Unbeweglichkeit der ‚assembly Line' zu überwinden. Die mehr oder minder eindimensional zweckgerichtete, mechanische Maschine wird durch die elektronische ‚Universalmaschine' abgelöst. Deren Funktionalität wird durch Software bestimmt, Materie und Energie, so Andreas Wiebe, werden durch Information ersetzt[2]. Während die Fertigungsprozesse damit erheblich flexibilisiert werden können, erlaubt die Fortentwicklung der Informations- und, seit den 1980er Jahren, auch der Kommunikationstechnologien die Einführung neuer Organisationsformen von Märkten und Unternehmen. Diesen Veränderungen vorausgesetzt ist die Technisierung des Wissens. Die schon das Rückgrat des Taylorismus darstellende Trennung von Disposition und Ausführung sowie die Formalisierung der sich daraus ergebenden Kontroll- und Steuerungsprozesse werden nun als „'Code' in Computern" umgesetzt[3]. Diese Vereinheitlichung der Informationsform ermöglicht schließlich die Integration bislang räumlich, zeitlich und personell unverbundener Informationssysteme zu einem, so Baukrowitz et al., globalen sozio-technischen System[4].

Mit der Entwicklung des Internets und damit einhergehend nicht zuletzt auch durch die allgemeine Verbreitung von PCs in Privathaushalten vollzieht sich im Laufe der 1990er Jahre eine ausgreifende Durchdringung und Anbindung sämtlicher gesellschaftlicher Bereiche an das Internet. Daraus ergibt sich neben den innerbetrieblich-organisatorischen Rationalisierungsmöglichkeiten im Weiteren eine Dynamisierung der Außenbeziehungen von Unternehmen und die Entstehung einer ‚digitalen Ökonomie'. Damit sind allerdings nicht nur die Trans-

1 Vgl. Daniel Bell: Die Nachindustrielle Gesellschaft, S. 3. Frankfurt a.M./New York,1976.

2 Vgl. Andreas Wiebe: Patentschutz und Softwareentwicklung – ein unüberbrückbarer Gegensatz?. In: Gehring/Lutterbeck (Hg.): Open Source Jahrbuch, S. 278. Berlin, 2004.

3 Vgl. Christoph Engemann: Electronic Government – vom User zum Bürger, S. 129. Bielefeld, 2003.

4 Vgl. Andrea Baukrowitz/Andreas Boes /Rudi Schmiede: Die Entwicklung der Arbeit aus der Perspektive ihrer Informatisierung. 2000, S. 6. www.ifs.tu-darmstadt.de/fileadmin/soziologie/Schmiede_Publikationen/chemnitz_010223.pdf

formationen im Bereich der Produktion sowie der Zirkulation und Konsumtion bezeichnet, sondern auch die Vorstellung eine ‚De-' oder ‚Entmaterialisierung' der Ökonomie und die Entstehung eines digitalen Warenuniversums. Diese Entmaterialisierung erweist sich dabei nicht zuletzt als Anknüpfungspunkt vor allem technikinduzierter Emanzipationsvorstellungen, die den Rahmen der Ökonomie überschreiten. Gerade das Internet hat in diesen Visionen einen prominenten Platz. Während es seinen Charakter als Wissenschaftsnetz im Laufe der 1990er Jahre und spätestens mit dem Rückzug der National Science Foundation (NSF) aus der Entwicklung und Finanzierung zusehends verlor, setzte mit der Übernahme durch private Internet-Provider seine Kommerzialisierung ein und es drängten schließlich immer mehr Unternehmen auf die digitalen Marktplätze. Diese Kommerzialisierung des Internets, d.h. die Entstehung eines digitalen Massenmarktes, verbindet sich neben einer zunehmenden, für geschäftliche Transaktionen notwendigen Kontrolle, wie der Nutzer-Identifizierung und ihres Verhaltens im Sinne präziser Marktbeobachtung und direkter Kundenansprache, auch mit der Frage nach der Institutionalisierung digitaler Eigentumsrechte. Demgegenüber erschien und erscheint bis heute das Internet manchen als Medium eines weltweiten, demokratischen Zugangs zu Wissen jenseits von Bürokratien und Hierarchien. Die ‚Freiheit der Rede' werde schließlich mit der Vernetzung und der ‚Entstofflichung' des gesellschaftlichen Wissensflusses eine „kumulative Wirkung des geteilten Wissensvorrats" entfalten können[5]. Schon früh finden sich in diesem Zusammenhang Beschreibungen, die das Internet im Jargon postmoderner Philosophie als universelles, nicht aber totalisierbares Netz nicht nur am adäquatesten bezeichnet sehen, sondern es nachgerade als ihre Verwirklichung begreifen: „Das Web, weit davon entfernt, eine gestaltlose Masse zu sein, verbindet eine offene Vielzahl von Gesichtspunkten, doch diese Verbindung wird transversal verwirklicht, als Rhizom, ohne göttlichen Gesichtspunkt, ohne überragende Vereinheitlichung"[6].

Wo die technikinduzierten Hoffnungen allerdings mit der, die digitale ‚Wissens-Allmende' bedrohenden Kommerzialisierung konfrontiert werden, sehen deren Anhänger eine Wiederholung der Geschichte am Werk: Die „Einzäunung des nicht greifbaren Allgemeinguts des Geistes"[7] in Analogie zu den ‚Enclosures' im England des 18. Jahrhunderts, der ‚so genannten ursprünglichen Akkumulation', nach Karl Marx. Entgegen der mit der grundsätzlich durch die Ablösung von einem materiellen Träger gegebenen Ubiquität, also der räumlichen und zeitlichen Unabhängigkeit der digitalen ‚Immaterialgüter', beschworenen Befreiung des Wissens erweist sich gerade dieser Umstand als Ausgangspunkt einer (erneut) an Bedeutung gewinnenden Auseinandersetzung um die Eigentumsrechte an geistigen ‚Gegenständen'. Denn tatsächlich hat das vollständige Auseinanderfallen von kognitivem Inhalt und materieller Form, bzw. die ‚Entmaterialisierung' der Ausdrucksform zunehmend diesen kognitiven Inhalt mindestens praktisch selbst zum Eigentumsgegenstand geraten lassen und es scheint damit schließlich wirklich eine Parzellierung des Wissens ‚an sich' stattzufinden.

Die durch die Digitalisierung ermöglichte Ablösung geistiger ‚Gegenstände' von einem materiellen Träger und die daraus resultierende unendliche Vervielfältigbarkeit ohne Qualitätsverlust oder nennenswerte Kosten sowie im Zusammenhang der allgemeinen Vernetzung scheinbar unkontrollierbare Verbreitung lässt einen Widerspruch zu Tage treten, der dem geistigen Eigentum von Beginn an eingeschrieben ist und im Urheberrecht unter der Bedingung traditioneller, d.h. analoger, Medien weitgehend unwidersprochen eine Form der Regulierung gefunden hatte. Parallel zu den unter das Urheberrecht fallenden wissenschaftlichen und kulturellen Schöpfungen (*propriété littéraire et artistique*) lässt sich dieser Konflikt zunächst an durch das Patentrecht geregelten und als solchen auf die Produktionssphäre bezogenen (im weiteren Sinne ‚technischen') Erfindungen (*propriété industrielle*) in Gestalt einer immanenten Kritik seiner ökonomischen Zweckmäßigkeit aufzeigen. So stellte der Ökonom Fritz Machlup 1958 in seinem im Auftrag des *Subcommittee on Patents, Trademarks and Copyrights* verfassten Bericht über das Patentsystem fest, dass aus dem mit dem Patent verliehenen exklusiven Recht des Inhabers „to the commercial exploitation of inventions, secure to theses owners profits" möglicherweise „monopolistic restrictions which hold down production and hinder he most efficient utilization of resources" resultierten. Derweil bestehe vor allem über die Notwendigkeit des erst durch die Verwertungsmöglichkeit eingeräumten Anreizes zur Erfindung Uneinigkeit[8]. Der Widerspruch lässt sich folglich als einer zwischen der Förderung von Innovation (Steigerung der Effizienz der Produktion) und der des Anreizes zur Wissensproduktion (Verwertungsmöglichkeit des Produkts) oder allgemeiner als der zwischen Produktions- und Zirkulationssphäre beschreiben. Entlang dieser zwei Pole entfalten sich schließlich auch die Konfliktlinien in der aktuellen Auseinandersetzung um das geistige oder digitale Eigentum.

Während auf der einen Seite die klassischen Geschäftsmodelle der Content-Industrie – worunter in diesem Zusammenhang zuerst vor allem Musik- und Filmindustrie zu zählen waren, nunmehr aber zusehends, wenn auch mit Verzögerung und mit der sukzessiven Etablierung von E-Books genauso die Buchindustrie fällt[9] – durch die Nichtrivalität und Nichtausschließbarkeit geistiger ‚Gegenstände' grundsätzlich bedroht werden, insofern sie auf der unter analogen Bedingungen mit der Bindung an einen materiellen Träger noch selbstverständliche Exemplarität aufsetzen, finden auf der anderen Seite die heterogenen Interessen von Internet-Nutzern und -Aktivisten sowie öffentlichen Institutionen im Namen der Informationsfreiheit zusammen. Gegenüber letzteren insistiert die im Sinne der Content-Industrie argumentierende Property Rights Theorie auf der Notwendigkeit exklusiver Eigentumsrechte an geistigen ‚Gegenständen', um deren Marktfähigkeit gerade auch unter den Bedingungen des Trends zur Ubiquität und Pervasivität des Wissens gewährleisten zu können. Die Knappheit von ‚Gütern', d.h. die Exklusivität von Eigentum stellt dabei in der materiellen ebenso wie in der immateriellen Welt die „Grundvoraussetzung für eine effiziente Allokation über Märkte" dar[10]. Knappheit bildet eines der zentralen Axiome der Property Rights Theorie, insofern nur so die Warenförmigkeit von Gütern garantiert werden und auf dieser Grundlage der Anreiz zu ihrer Produktion aufrechterhalten werden kann. Anreiz zur Produktion aber ist unter der Bedingung individueller Nutzenmaximierung handelnder Wirtschaftssubjekte an die Verwertungsmöglichkeit des jeweils geschaffenen Produkts und damit notwendig an seine Eigentumsform gebunden.

5 Vgl. Jeannette Hofmann: Digitale Unterwanderungen. Der Wandel im Innern des Wissens. In: Aus Politik und Zeitgeschichte 36(2001), S. 5f.

6 Vgl. Pierre Lévy: Cyberkultur. Universalität ohne Totalität. In: Bollmann/Heilbach: Kursbuch Internet, S. 62. Hamburg, 1996.

7 Vgl. James Boyle: The Second Enclosure Movement and the Construction of the Public Domain. 2003, S. 37. www.ifs.tu-darmstadt.de/fileadmin/soziologie/Schmiede_Publikationen/chemnitz_010223.pdf.

8 Vgl. Fritz Machlup: An Economic Review of the Patent System. Study of the Subcommittee on Patents, Trademarks and Copyrights of the Committee on the Judiciary United States Senate, S. 44. Washington, 1958.

9 Vgl. Skillnet 2009. www.skillset.org/uploads/pdf/asset_13980.pdf?1; vgl. auch Catherine Neilan: Digital Skills gap now 'critical' for publishers. 2009. www.thebookseller.com/news/94322-digital-skills-gap-now-critical-for-publishers.html.

10 Vgl. Heike Walterscheid: Exklusives Eigentum an geistigen Werken, das Recht auf Kopie und effiziente Allokation. Eine ökonomische Analyse. In: Fechner (Hg.): Die Privatkopie. Juristische, ökonomische und technische Betrachtungen, S. 97. Ilmenau, 2007.

Prominent ist in diesem Zusammenhang Garrett Hardins Aufsatz über die *Tragedy of the Commons*, demnach „[r]uin is the destination toward which all men rush, each pursuing his own best interest in a society that believes in the freedom of the commons. Freedom in the commons brings ruin to all"[11]. Erst unter der Voraussetzung privater Aneignung und gesicherter Verfügungsrechte über Privateigentum kann folglich effizientes Wirtschaften ermöglicht werden. Und erst damit realisiert sich die harmonisierende Vermittlung des Marktes als optimale Verteilung knapper Güter garantiert durch den Preis. Da das Handeln der Wirtschaftssubjekte sich unweigerlich an der Maximierung ihres individuellen Nutzens orientiert, sind gesicherte Eigentumsrechte eine notwendige Bedingung für eine effiziente Produktion, einerseits um die Verwertung der Produkte zu garantieren, andererseits um gerade im Bereich der Informationsprodukte die Kosten der Nachahmung nicht unter das Niveau der eigenen Produktion sinken zu lassen, damit der Anreiz zur Investition in Kreativität und Innovation nicht erlischt. Der „Strom der Kreativität auf Autoren- und Künstlerseite" würde, so schwant es manchem, buchstäblich „zum Versiegen" gebracht und anstelle dessen eine „Recyclingkultur" die „Entropie des Internets" einleiten[12].

Die Sicherung der Eigentumsrechte fällt dem Staat zu, der somit durch Senkung der Transaktionskosten, d.h. der Kosten im Zusammenhang der inhaltlichen Gestaltung und personellen Zuordnung sowie zur Durchsetzung und Kontrolle von Verfügungsrechten, steuernd auf das Verhalten der Wirtschaftssubjekte einwirken kann. Daneben allerdings spielen auch so genannte „externe Effekte" eine Rolle, die derart nur bedingt aufgelöst werden können. Mit der Möglichkeit die Marktmechanismen zu hintergehen besteht dementsprechend zum einen ein ‚Internalisierungserfordernis', d.h. das Erfordernis der Etablierung eines Unrechtbewusstseins[13], zum anderen sollen Kontroll-Techniken diese externen Effekte auffangen.

Während in der analogen Welt die private Aneignung geistigen Eigentums noch relativ problemlos zu gewährleisten war, zeigt sich schnell, dass die Verbreitung des Internet das herkömmliche, auf Knappheit von Informationsprodukten beruhende Distributionsmodell unterminiert. Angesichts des Versagens des ‚alten' Urheberrechts unter den Bedingungen der Digitalisierung und Vernetzung hat gerade die Musik- und die Filmindustrie in den mit neuen Technologien gegeben Möglichkeiten weniger das Potential neuer Nutzungsarten (bspw. On-Demand-Angebote, Abonnieren statt Kaufen, Personalisierung, etc.) als vielmehr eine Bedrohung der traditionellen Geschäftsmodelle u.a. durch ‚Piraterie' zu entdecken vermocht und versucht, dem Problem auf zwei Ebenen zu begegnen, einerseits durch die Forderung einer entsprechenden Rechtsentwicklung, andererseits über die Entwicklung einer zur Wahrung ihrer Interesse geeigneten Technologie. Im Bereich der Rechtsentwicklung zeigt sich dabei, dass für das nationalstaatlich verfasste Recht die Sicherung von digitalen Eigentumsrechten vor dem Hintergrund der Globalität des Internet ein nachgerade unlösbares Problem darstellt, dem sich auch über die nationalen Umsetzungen der von der WIPO verabschiedeten Richtlinien im *Digital Millenium Copyright Act* und der *EU-Richtlinie 2001/29/EG* nur bedingt im Sinne eines vereinheitlichten, globalisierten Rechts entsprechen lässt. Auf der Ebene der Technologieentwicklung wird dementsprechend versucht vor allem dem Problem der ‚Piraterie' mittels unterschiedlicher Ansätze beizukommen, die zusammengefasst unter dem Oberbegriff der Systeme zum so genannten *Digital Rights Management* (Kurz: DRM) eine ‚Re-Individualisierung' digitaler Produkte zur Sicherung ihrer Warenform ermöglichen sollen[14]. Es handelt sich hierbei allerdings nicht um zwei je alternative Herangehensweisen, sondern Recht und Technologie stellen einander ergänzende Maßnahmen dar, insofern das Umgehen des Kopierschutzes, ebenso wie Werkzeuge zu diesem Zweck einzuführen, zu verkaufen, verbreiten oder vermieten, zur Straftat erklärt wird[15].

Die Durchsetzung und Kontrolle von Urheber- und verwandten Schutzrechten mithilfe technologischer Maßnahmen zielt in erster Linie auf eine fortgesetzte ‚De-Anonymisierung' des Internets, d.h. einerseits die Ermöglichung einer rechtsgültigen Authentifizierung von Internet-Nutzern sowie andererseits die Herstellung einer ‚Quasi-Exemplarität' digitaler Produkte[16]. Erst unter diesen Bedingungen ist eine tatsächliche Regulierung des wirtschaftlichen Verkehrs qua Vertragsrecht und damit eine personelle Zuweisung von Verfügungsrechten ebenso

11 Garrett Hardin: The Tragedy of the Commons. 1968. www.sciencemag.org/cgi/content/full/162/3859/1243.

12 Vgl. Michael Roesler-Graichen: Copyright und Rechtemanagement im Netz. In: Aus Politik und Zeitgeschichte (2009) 42/43, S. 18.

13 Vgl. Walterscheid, a.a.O.: 99f.

14 Vgl. Tanja Lieckweg: Cyberlaw. Ein Überblick. 2002, S. 21. http://notesweb.uniwh.de/wg/wiwi/wgwiwi,nsf/contentbykey/TLIG-5C4CSD-DE-pdf.

15 Vgl. Volker Grassmuck: Wissenskontrolle durch DRM: von Überfluss zu Mangel. In: Hofmann (Hg.): Wissen und Eigentum, S. 168. Bonn, 2006.

16 Vgl. Engemann, a.a,O., S. 43.

Gesamtlösungen für Digitalisierungsprojekte

ImageWare

MyBib eDoc® Workflow-System
BCS-2® Scansoftware
C-3 Periodikaerschließung
MyBib eRoom Elektronischer Lesesaal
SCANROBOT® Bookeye® Buchscanner

ImageWare Components GmbH | Am Hofgarten 20 | 53113 Bonn | Germany | www.imageware.de

wie mit den Vereinzelungsmechanismen durch DRM das Begreifen eines digitalen Produkts als ‚Quasi-Sache' im Sinne bürgerlichen Rechts umsetzbar. Die somit technologisch erzeugte Knappheit erweist sich dabei allerdings als problematisch, insofern nicht mehr das Informationsprodukt selbst erworben werden kann, sondern bloß beschränkter Zugang zu ihm gewährt wird. Damit gerät jedoch vor allem die, sowohl in technischer als auch in rechtlicher Hinsicht, tendenziell sich anbahnende Unterminierung der Begründung des Urheber- und Patentrechts, Innovation durch die Ermöglichung von Erneuerung bzw. Variation bestehenden Wissens zu fördern, in die Kritik[17]. Die Dichotomie von Form und Inhalt, des vom Urheberrecht geschützten Werks und der nicht unter diesen Schutz fallenden in ihm enthaltenen Ideen, wird durch die technischen Sicherungsmaßnahmen praktisch aufgehoben. Im Unterschied zu den analogen Bedingungen zeigt sich hier, dass zunehmend kein Eigentum im eigentlichen Sinne mehr an den Manifestationen von Informationsprodukten erworben werden kann. Anstelle dessen reguliert ein durch Lizenzen, deren Einhaltung technisch erzwungen wird, streng geregelter Zugang, der neben der Art der Verwendung auch die Anzahl der Verwendungen ebenso wie die Medien, auf die eine Übertragung stattfinden darf, die Nutzung. Als Beispiele seien hier das *End User License Agreement* (EULA) sowie das *Next Generation Secure Computing Base* (NGSCB) der Firma Microsoft genannt. Die Konsequenz sei, so Christian Schmidt, nicht eine Ablösung des Eigentums durch Zugang, wie dies Jeremy Rifkin postulierte, sondern die Durchsetzung des Eigentumsparadigmas und Begründung von Monopolen im Bereich der geistigen Formen[18].

Die geistigen Formen, begriffen als Produktionsformen und Kulturgut der Gesellschaft, gelten zunächst als Bereich, auf den Zugriff und dessen Nutzung gerade im Horizont seiner Weiterentwicklung und der Innovationsschöpfung durch Urheber- und Patentrecht im Sinne einer Balance zwischen öffentlichem Nutzen und privater Verfügung reguliert werden. Im Zuge der Digitalisierung und allgemeinen Vernetzung erfährt diese Balance jedoch eine zunehmende Schlagseite zugunsten einer effizienten Allokation. Im Rahmen des Urheberrechts eingeräumte Ausnahmeregelungen wie die Privatkopie, d.h. die Verwendung urheberrechtlich geschützter Materialien zu Zwecken öffentlicher Bildung und der Anregung wissenschaftlicher Forschung, werden technisch tendenziell ausgesetzt und, wo diese Maßnahmen rechtlich durch das Verbot der Umgehung des Kopierschutzes legitimiert werden, schließlich auch faktisch durch einzig von den Rechteinhabern bestimmten Lizenzen abgelöst. Das Recht auf Privatkopie, so Rüdiger Theiselmann in seiner Untersuchung über die *rechtlichen Implikationen der digitalen Werkverwertung*, sei jedoch abhängig zu machen von der Erwägungsgrundlage des Gesetzgebers und demnach sei schon unter der Bedingung analoger Kopiertechnik diese allenfalls „zulässig", nicht aber verpflichtend einzuräumen. Gerade mit den Vervielfältigungsmöglichkeiten der Digitaltechnik seien die Urheberrechte ungleich stärker gefährdet und somit ein Schutz des Kopierschutzes legitim. Die Informationsfreiheit so Theiselmann weiter, stehe durch das Verbot der Privatkopie in digitaler Form nicht in Gefahr, da die analoge Kopie von diesem Verbot schließlich nicht betroffen sei[19]. Und Heike Walterscheid verweist darauf, dass das Recht auf unentgeltliche Privatkopie für Dritte eine Verringerung des Verfügungsrechts des Eigentümers bedeute. Es handle sich um eine Teilenteignung des Urhebers, zwar im Namen der Steigerung gesellschaftlicher Wohlfahrt, allerdings im Widerspruch zum Persönlichkeitsrecht und stelle „die wichtige Bedeutung exklusiven Eigentums für eine effiziente Allokation in Frage" [20].

Die Ersetzung des Rechts durch Lizenzen, d.h. zwischen Nutzern und Rechteinhabern geschlossenen Verträgen über die Nutzung, zunächst vor allem im Bereich der Software verbreitet, gewinnt bald auch für andere digitale Werkarten Bedeutung. Damit wird mit den durch das Urheberrecht noch vorgesehenen Ausnahmeregelungen zugunsten der Interessen der Allgemeinheit am Zugang zum Wissen zuletzt das Balance-Prinzip in Gänze aufgehoben. Darüber hinaus verliert auch die zeitliche Beschränkung des Urheberschutzes seine Gültigkeit. Bestimmte die Schutzdauer noch den Zeitraum, bevor ein Werk ‚gemeinfrei' wurde, also in den Besitz der Öffentlichkeit überging, so können Lizenzverträge eine zeitlich unbegrenzte Geltung für sich in Anspruch nehmen. Entgegen des Eindrucks, dass im Internet Urheberrechte, oder vielmehr Eigentumsrechte, nur schlecht geschützt werden könnten, zeigt sich, dass mithilfe von Technologie, kryptographischen Programmen zur Nutzerauthentifizierung und Personalisierung digitaler Produkte, dieser Schutz als Kontrolle der Nutzer gerade auf der Grundlage der Möglichkeiten des Internets perfektioniert wird. In diesem Zusammenhang ist es bisweilen nicht ganz ersichtlich, ob das Argument, jede Schutzmaßnahme sei schließlich umgehbar oder könne ‚gehackt' werden, ein effektiver Schutz folglich im Internet nicht durchsetzbar, nicht am ehesten von den Gegnern von DRM ins Feld geführt wird. Lawrence Lessig allerdings schrieb schon 1999: *„We have every reason to believe that cyberspace, left to itself, will not fulfil the promise of freedom. Left to itself, cyberspace will become a perfect tool of control. […] [T]he invisible hand of cyberspace is building an architecture that is quite the opposite of what it was at cyberspace's birth. The invisible hand, through commerce, is constructing an architecture that perfects control – an architecture that makes possible highly efficient regulation"*[21].

Die Content-Industrie habe sich das Internet als ein interaktives 500 Kanäle Fernsehen vorgestellt, als ein „promotion medium and mail-order machine"[22], stattdessen aber bestehe es, so die Verteidiger der Informationsfreiheit, aus 50 Millionen Nutzern, durch die das Wissen seine vollendet dialogische Form gewönne[23]. Das Wissen, dessen eigentümlicher Charakter es sei, desto größeren Nutzen zu erzeugen, je mehr es sich verbreite, finde im Internet sein adäquates Medium, insofern hier seine freie Zirkulation und kooperative Fortentwicklung befördert würden. Entgegen aller privaten Eigentumsansprüche sei es ein Gemeingut und die Maßnahmen zur ‚Einzäunung' dieser ‚Wissens-Allmende', dem ‚Rohmaterial' aller künftigen Entwicklungen, unternommen durch die Content-Industrie, von fataler Auswirkung auf den „Fortschritt des Wissens" und auf die freie Rede[24]. Eine Übernutzung, eine *Tragedy of the Commons*, wie sie im Anschluss an Garrett Hardin beschworen werde, sei in Bezug auf das Wissen, das sich in der Nutzung nicht verbrauche, gar nicht möglich. Im Gegenteil könne Neues nur produziert werden, wo auf schon veröffentlichtes Wissen zugegriffen, „aus dem großen kulturellen Pool kollektiver Kreativität" geschöpft werde, um in diesen wieder einzugehen[25]. Das Internet, organisiert als Allmende, jenseits hierarchischer Strukturen nach dem Prinzip

17 Vgl. Christian Schmidt: Die zwei Paradoxien des geistigen Eigentums. In: Deutsche Zeitschrift für Philosophie 52(2004)5, S. 767; Volker Grassmuck: Freie Software. Zwischen Privat- und Gemeineigentum, S. 122-123. Bonn, 2004.

18 Vgl. Schmidt: a.a.O., S. 762.

19 Vgl. Rüdiger Theiselmann: Geistiges Eigentum in der Informationsgesellschaft. Rechtliche Implikationen der digitalen Werkverwertung, S. 97ff., 103. München, 2004.

20 Vgl. Walterscheid, a.a.O.: 118f.

21 Lawrence Lessig: Code and other Laws of Cyberspace, .5-6. New York, 1999.

22 Jessica Litman: Digital Copyright and the Progress of Science. 2002. http://portal.acm.org/citation.cfm?id=792558&dl=GUIDE&coll=GUIDE&CFID=62054969&CFTOKEN=38696395.

23 Vgl. Volker Grassmuck: Die Wissens-Allmende. 2000, unter: http://waste.informatik.hu-berlin.de/Grassmuck/Texts/wissens-almende.html.

24 Vgl. Rainer Kuhlen: Wissen als Eigentum? 2000, S. 5. www.wissensgesellschaft.org/themen/publicdomain/wisseneigentum.pdf.

25 Vgl. Grassmuck, a.a.O. 2000.

DABIS.eu
Gesellschaft für Datenbank-InformationsSysteme mbH

Ihr Partner für Archiv-, Bibliotheks- und DokumentationsSysteme

BIS-C 2000

Archiv- und Bibliotheks-InformationsSystem

DABIS.com · alle Aufgaben · ein Team

Synergien: Qualität und Kompetenz
Software: Innovation und Optimierung
Web · SSL · Warenkorb und Benutzeraccount
Lokalsystem zu Aleph-Verbünden

Software · State of the art · Open Source

Leistung	Sicherheit
Standards	Offenheit
Stabilität	Verläßlichkeit
Generierung	Adaptierung
Service	Erfahrenheit
Outsourcing	Support
Dienstleistungen	Zufriedenheit

GUI · Web · Wap · XML · Z 39.50

Archiv — Bibliothek

singleUser	System	multiUser
Lokalsystem		Verbund
multiDatenbank		multiServer
multiProcessing		multiThreading
skalierbar		stufenlos
Unicode		multiLingual
Normdaten		redundanzfrei
multiMedia		Integration

Portale

http://www.vthk.eu
http://www.landesbibliotheken.eu
http://www.bmlf.at
http://www.volksliedwerk.org
http://www.dabis.org:18093

DABIS.com
Heiligenstädter Straße 213
1190 · Wien, Austria
Tel.: +43-1-318 9 777-10
Fax: +43-1-318 9 777-15
eMail: office@dabis.com
http://www.dabis.com

DABIS.de
Herrgasse 24
79294 · Sölden/Freiburg, Germany
Tel.: +49-761-40983-21
Fax: +49-761-40983-29
eMail: office@dabis.de
http://www.dabis.de

Zweigstellen: 61350 · Bad Homburg vdH, Germany / 1147 · Budapest, Hungary / 39042 · Brixen, Italy

der Kooperation funktionierend, ermutige Nutzer grundsätzlich zu innovativem Schaffen und werde damit zum „Modell der Wissensgesellschaft insgesamt“[26]. Die Regeln, die für die ‚old' Economy gelten, sind im virtuellen Raum nicht länger von Bedeutung: *„Your legal concepts of property, expression, identity, movement, and context do not apply to us. They are all based on matter, and there is no matter here“*[27].

Angesichts der im Gegensatz zu den Eigenschaften der Dinge der materiellen Welt stehenden Nichtausschließbarkeit und Nichtrivalität von Wissen, dem ‚Mangel an Mangel', der im virtuellen Raum herrsche und die Warenförmigkeit der Informationsprodukte unterlaufe, seien Verknappungen mithilfe von Kopierschutzmaßnahmen bloß künstliche und perpetuierten nicht nur die Spaltung zwischen ‚Information-Haves' und ‚Have-Nots', sondern weiteten sie aus[28]. Die Durchsetzung exklusiver Eigentumsrechte im Internet wirke sich vor allem als Blockierung einer sinnvollen Nutzung – der prozesshaften Verküpfung – des Wissens aus, indem seine Privatisierung die Transaktionskosten ins Uferlose steigen lasse, wo Beschaffung des nötigen Wissens durch Ausschließlichkeitsrechte erschwert würde. Karl-Heinz Ladeur spricht in diesem Zusammenhang in Anlehnung an Heller und Eisenberg[29] von einer *Tragedy of the Anti-Commons*[30]. Wo mit dem Urheber- und Patentrecht der Schutz der Eigentumsrechte eines Autors oder einer Erfinderin geltend gemacht würden, entspreche dies zwar dem Konzept John Lockes von der Begründung des Rechts auf Eigentum durch Arbeit, das von Kant, Hegel und Fichte auf den Bereich der geistigen Produktion übertragen wurde, sei aber doch bloß ein ‚Nebenziel'. Die entscheidendere Absicht, das ‚Hauptziel' der Institution dieser Rechte sei dagegen die Förderung des Interesses der Gesellschaft an Innovation und in deren Folge des allgemeinen Wohlstandes gewesen, zu dem solcher Schutz nur den (pekuniären) Anreiz zu kreativem Schaffen hatte bieten sollen[31]. Heute aber schlage dieses Konzept nicht zuletzt mit der Aussetzung des Rechts auf Privatkopie vollständig um und verabsolutiere den individuellen Anspruch, demnach geistige Tätigkeit Eigentum schaffe, gegenüber dem gesellschaftlichen Interesse. Dabei handle es sich hier immer weniger um den tatsächlichen Schutz materieller und ideeller Interessen der Urheber, sondern um das Verwertungsinteresse der Content-Industrie. Die Wissensproduzenten selbst würden in diesem Zusammenhang bloß instrumentalisiert und als moralisches Argument ins Feld geführt, das mehr Aussicht auf Zustimmung verspreche als die Beschwörung der Bedrohung kommerzieller Verwertungsinteressen durch ‚Piraterie'. Gerade das Persönlichkeitsrecht als Begründung des Urheberrechts – des europäischen *droit d'auteur* – habe sich in der modernen Wirtschaft für die notwendige Verkehrsfähigkeit der Eigentumsrechte als unzureichend erwiesen und sei dementsprechend nicht nur durch die internationalen Rechtsanpassungen zunehmend in die Bedeutungslosigkeit gedrängt worden, sondern werde auch faktisch durch nicht verhandelbare Abtretungsverträge zu Ungunsten der Urheber ausgehebelt[32]. Verstärkt werde dies noch durch die Konzentration und medienübergreifende Integration der Content-Industrie, die gerade unter den Bedingungen der Digitalisierung die Kontrolle der gesamten Wertschöpfungskette anstrebe. Damit würde die Herausbildung von „Wissensmonopolen“ vorangetrieben, wofür gerade die Firma Microsoft ein sprechendes Beispiel gebe[33].

Neben den Tauschbörsen im Internet sehen viele Vertreter der Informationsfreiheit ihre Vorstellungen im Bereich der Softwareproduktion Wirklichkeit werden. Die 1989 entwickelte GNU General Public License (GPL) polt sozusagen das exklusive Verfügungsrecht proprietärer Software in ein „inkludierendes ‚gedeihliches'“[34] Verfügungsrecht um und macht die Veränderbarkeit sowie die Möglichkeit der Weiterverbreitung des ‚offenen' gegenüber dem ‚geschlossenen' Quellcode von Software zur lizenzrechtlichen Bedingung. GPL soll so einen ‚Allmende'-ähnlichen Raum offenen Wissens schaffen, innerhalb dessen weltweit miteinander vernetzte und kooperierende Nutzer und Programmierer an Software arbeiten, sie verbessern, weiterentwickeln und – wenn auch nicht unbedingt kostenlos – nutzen können. Damit würden neue Formen der Produktion, gerade auch solche jenseits des klassischen Anreiz-Modells und frei vom Konkurrenzdruck entstehen[35]. Aber nicht nur idealistische, sondern vor allem pragmatische Zielsetzungen prägen die Entwicklung der Open Source Software. Insofern wird durchaus mit der höheren Effizienz der Entwicklung und der besseren Qualität der Produkte argumentiert und man ist einer Kommerzialisierung gegenüber aufgeschlossen. So entstehen Lizenzen nach dem Vorbild der GPL, die eine spätere Verwertung, bspw. in Gestalt von ‚Mischformen' sowohl offenen als auch geschlossenem Codes zulassen, und damit nicht zuletzt in zunehmendem Maße Open Source Software auch für die Industrie interessant machen[36]. In diesem Rahmen schließlich werden Geschäftsmodelle entwickelt, die Schule machen sollen. Zwar kann Open Source Software, da selbst kein ‚knappes Gut' aufgrund lizenzrechtlicher Bestimmungen, nicht verkauft werden, jedoch haben sich bald alternative Verwertungsstrategien entwickelt, in deren Zusammenhang sie den entgeltfreien Kern des Geschäfts mit Zusatzleistungen darstellt. Auf der Grundlage einer solchen „Gabenwirtschaft“ als „vorkommerziellem Stadium“ werden, wie Ladeur schreibt, „Programme verschenkt […], damit an anderer Stelle ein Lock-in-Effekt entsteht, der dann durch ein Angebot von Spezialprogrammen oder Dienstleistungen für einzelne Nutzer durchaus zu einem ökonomischen Erfolg führen kann“[37]. Einerseits als Gegengewicht zum Quasi-Monopol der Firma Microsoft, andererseits als kostenloses Produktionsmittel – in der Verwendung von proprietärer Software nicht mehr zu unterscheiden –, als Tool zur Herstellung kommerzieller Software oder als ihr weiterhin offener Bestandteil erweist sich Open Source Software als integrierbar in den Verwertungszusammenhang. Die Entwicklung kann allerdings auch als ein Austesten neuer Produktionsformen und Flexibilisierungsstrategien begriffen werden. Gerade die hohe Motivation der Entwickler in ihrer Freizeit unbezahlt, für die Reputation innerhalb der *Community*, an Software zu arbeiten, sich dazu im globalen Ausmaß zu vernetzen und miteinander zu kooperieren, lässt sich leicht übersetzen in ein vorbildliches Modell der Selbstaktivierung und -ökonomisierung der Humanressourcen, die unter flexiblen, global dezentralen Produktionsbedingungen bei ungeregelten Arbeitszeiten Innovation durch ‚Netzwer-

26 Bernd Lutterbeck: Die Zukunft der Wissensgesellschaft. In: Hofmann: Wissen und Eigentum, S. 320, 331. Bonn, 2006.

27 John Perry Barlow: The Declaration of Independence of Cyberspace. 1996. http://homes.eff.org/~barlow/Declaration-Final.html.

28 Vgl. Sabine Nuss: Download ist Diebstahl? Eigentum in einer digitalen Welt. In: PROKLA 126(2002)1, S. 13.

29 Vgl. Michael A. Heller; Rebecca S. Eisenberg: Can Patents Deter Innovation? The Anticommons in Biomedical Research. 1998. www.sciencemag.org/cgi/content/full/280/5364/698.

30 Vgl. Karl-Heinz Ladeur: Die Dynamik des Internet als Herausforderung der Stabilität des Rechts. 2002, S. 19-20. http://www2.jura.uni-hamburg.de/ceri/publ/download02.pdf.

31 Vgl. Kuhlen, a.a.O. 2000, S. 8.

32 Vgl. Rainer Kuhlen: Erfolgreiches Scheitern. Eine Götterdämmerung des Urheberrechts, S. 88. Boizenburg, 2008; vgl. auch Cyrill Rigamonti: Geistiges Eigentum als Begriff und Theorie des Urheberrechts, S. 90ff. Baden Baden, 2001.

33 Vgl. Grassmuck: a.a.O. 2004, S. 83ff.

34 Vgl. Petra Haarmann: Copyright und Copyleft. Vermittlung im Falschen oder falsche Unmittelbarkeit. 2004, S. 195. www.exit-online.org/pdf/Exit0108PHCopyleft.pdf.

35 Vgl. Sabine Nuss; Michael Heinrich: Freie Software und Kapitalismus. 2002. www.roter-ruhr-uni.com/cms/IMG/pdf/504Nuss-Heinrich.pdf.

36 Vgl. Sabine Nuss: Zur Verwertung allgemeinen Wissens. Ein kapitalistisches Geschäftsmodell mit freier Software. 2002a. http://wbk.in-berlin.de/wp_nuss/wp-content/uploads/2006/10/argumentfertig.pdf.

37 Ladeur, a.a.O., S. 17.

kintelligenz' schaffen. Gerade insofern Software ein „grundlegender Baustein der Wissensgesellschaft"[38] ist, fragt Sabine Nuss, ob vor diesem Hintergrund die Open Source Bewegung nicht als „Avantgarde einer postindustriellen Produktionsweise" zu bezeichnen sei[39].

Die Verwertungsmodelle aus dem Umfeld der Open Source Software, die über Services und Zusatzleistungen die mit dem Kernprodukt nicht zu erzielenden Gewinne zu realisieren versuchen, stehen nun gewissermaßen Modell für ähnliche, zwischen Innovation und Anreiz vermittelnden Ansätzen im Feld der kulturellen und wissenschaftlichen Produktion. Das 2001 von Lawrence Lessig in expliziter Anlehnung an die GPL gegründete Projekt *Creative Commons* (CC) gibt dafür ein mittlerweile recht prominentes Beispiel[40]. Ähnlich wie im Bereich der Software wird ein ‚sekundäres Einkommen' über den Verkauf von Umfeldprodukten und zusätzliche Serviceleistungen erwirtschaftet. Im Internet, so wird in diesem Zusammenhang bisweilen gerne festgestellt, herrsche nicht eine „Ökonomie des Kapitals, sondern eine der Aufmerksamkeit", die Filter- und Verteilungsfunktion der Verlage werde nun folglich von „anderen Instanzen", wie sozialen Netzwerken, Blogs und Suchmaschinen übernommen[41]. Dieser im Rahmen digitaler Distributionsweisen immer wieder angeführte Autonomiezugewinn der Autoren/Künstlerinnen gegenüber Verlagen findet sich allerdings relativiert, wo die verschärfte Bedeutung öffentlicher Sichtbarkeit und Reputation mitunter Folge der erhöhten Konkurrenz ist und nur bedingt erstrebenswert sein mag sowie im Hinblick auf eine Schwächung bei der Durchsetzung von Copyrightansprüchen[42]. Für die Verlage wie für die Musikindustrie gilt derweil, so will es unter anderem die Prognose Tim Renners, dass sich in Zukunft ohnehin immer weniger mit dem direkten Verkauf von *Content* verdienen lasse und stattdessen zunehmend auf *Convenience* gesetzt werden müsse. Eine solche Entwicklung von einer „Bestands- zu einer Leistungsökonomie" sagte die IT-Journalistin Esther Dyson allerdings schon im Dezember 1994 voraus: „In short, intellectual *processes and services appreciate; intellectual assets depreciate*"[43].

Die aktuelle Diskussion um Art und Umfang des geistigen Eigentums vor dem Hintergrund der Digitalisierung involviert neben den zwei besprochenen Positionen schließlich auch den Staat in seiner Funktion die gesellschaftlichen Reproduktionsbedingungen zu garantieren. Sofern dem Wissen in der so genannten Wissensgesellschaft eine zentrale Rolle als Produktionsfaktor zuerkannt wird, gewinnt die Regulierung von Innovationsförderung und Sicherung des Anreizes unter veränderten Bedingungen für die Gesetzgebung Bedeutung. Die Durchsetzung digitaler Eigentumsrechte und damit einhergehend die Bestrebungen der De-Anonymisierung, d.h. rechtsgültiger Nutzerauthentifizierung im Internet, so zeigt sich in diesem Zusammenhang, liegt nicht allein im Interesse der Content-Industrie, sondern verbindet sich damit auch mit einem Anliegen des Staates selbst. Dies ist nicht zuletzt im Horizont der fortschreitenden Verschärfung der internationalen Konkurrenz im Zuge der Krise des fordistischen Akkumulationsregimes und der Diskussion um die ‚Globalisierung' seit Beginn der 1990er Jahre und verstärkt betonten Standortpolitik zu betrachten. Wissen wird als ein entscheidender Faktor des Wachstums und der Wettbewerbsfähigkeit nationaler Ökonomien begriffen. So heißt es bspw. in der EU-Urheberrechtsrichtlinie: *„Ein harmonisierter Rechtsrahmen zum Schutz des Urheberrechts und der verwandten Schutzrechte wird durch erhöhte Rechtssicherheit und durch die Wahrung eines hohen Schutzniveaus im Bereich des geistigen Eigentums substantielle Investitionen in Kreativität und Innovation einschließlich der Netzinfrastruktur fördern und somit zu Wachstum und erhöhter Wettbewerbsfähigkeit der europäischen Industrie beitragen, und zwar sowohl bei den Inhalten und der Informationstechnologie als auch allgemeiner in weiten Teilen der Industrie und des Kultursektors. Auf diese Weise können Arbeitsplätze erhalten und neue Arbeitsplätze geschaffen werden"*[44]. Insofern die Informations- und

38 Lutterbeck, a.a.O., S. 329.
39 Nuss, a.a.O. 2002, S. 22.
40 Vgl. Michael Spielkamp: Creative Commons. Andere Zeiten, andere Lizenzen. 2005. http://www.irights.info/index.php?id=150.
41 Vgl. Albrecht Hausmann: Zukunft der Gutenberg-Galaxis. In: Aus Politik und Zeitgeschichte (2009)42/43, S 39; vgl. auch Tobias Moorstedt: Zum Mitnehmen. 2009. www.fluter.de/de/83/heft/8014/?tpl=162.
42 Vgl. Hans Geser: Copyright und Copy left. Prekäre immaterielle Eigentumsverhältnisse im Cyberspace. 1999. http://socio.ch/intcom/t_hgeser08.htm.
43 Esther Dyson: Intellectual Property on the Net. 1994. http://w2.eff.org/IP/ip_on_the_net.html ; vgl. auch Esther Dyson: Intellectual Value. 1995. http://www.wired.com/wired/archive/3.07/dyson.html.
44 Richtlinie 2001: §4.

Kommunikationstechnologien zu einem wesentlichen Moment der Rationalisierung der Produktion sowie der intra- und interorganisatorischen Vernetzung geworden sind und der damit einhergehende Aufstieg der digitalen Ökonomie eine neue Art stoffloser Produkte hervorgebracht hat, können rechtliche Anpassungen um restriktive Eigentumsrechte zu garantieren als Maßnahme zur Sicherung der Reproduktionsbedingungen der kapitalistischen Ökonomie begriffen werden. Für den Staat ist dies nicht zuletzt im Hinblick auf die Besteuerung der damit entstehenden Mehrproduktanteile durchaus von Interesse. Derweil legitimiert sich staatliche Politik auch immer über ihre Nützlichkeit bezüglich des ‚Allgemeinwohls', worunter gerade das wirtschaftliche Wachstum des nationalen Standorts fällt, u.a. als Bedingung des Erhalts und der Schaffung von Arbeitsplätzen, sowie mittelbar über die dadurch entstehenden staatlichen Einnahmen. Das Allgemeinwohl ist somit wiederum an die Reproduktionsbedingungen des Kapitals, d.h. an die die nationale Wettbewerbsfähigkeit garantierenden Verwertungsbedingungen, geknüpft.

Die Ausweitung der Eigentumsordnung, die ‚Einzäunung' des ‚öffentlichen Guts' Wissen, wie sich momentan beobachten lässt, vollzieht sich dabei nicht als eine im Rahmen des Übergangs zur Wissensgesellschaft immerzu konstatierte ‚sektorale Verschiebung', sondern kann vielmehr begriffen werden als Subsumtion der auf dem Internet als ihrem zentralen Medium beruhenden ‚digitalen Ökonomie' über die Verrechtlichung seiner Dynamik. Etabliert wird die Exklusivität digitalen Eigentums und damit wird eine rechtliche Absicherung der Austauschverhältnisse in der Zirkulationssphäre als Bedingung der „Monetarisierung von Netzwerkbeziehungen" geschaffen[45]. Die Kritik an diesen Vorgängen, die sich auf die Nichtausschließbarkeit des Wissens, auf den ‚Mangel an Mangel' als dessen ‚natürlichem' Charakteristikum beruft, geht insofern fehl als sie verkennt, dass es sich hierbei um eine soziale und rechtliche Konstruktion, d.h. um den gesellschaftlich bestimmten Charakter und die damit implizierte Funktion des Wissens handelt.

In diesem Prozess der Verrechtlichung zeigt sich allerdings bald, dass mit der Verwertungsmöglichkeit digitaler Produkte unter der Voraussetzung der Herstellung exklusiver Verfügungsrechte im Internet eine innovative Selbstblockade einhergeht, wo horrende Transaktionskosten aufgrund des aufwendigen Erwerbs erforderlicher Lizenzen für wissensintensive Fortentwicklungen in der Produktion Ineffizienzen entstehen lassen. Dieser Widerspruch ist verschiedentlich als „digitales Dilemma" angesprochen worden[46]. Während der Gesetzgeber durch das Urheberrecht (wie durch das Patentrecht) noch versuchte eine Balance zwischen öffentlichem, d.h. allgemeinem – im Sinne der Aufrechterhaltung des kapitalistischen Reproduktionsprozesses – Interesse und privater Aneignung zu schaffen, muss diese Balance nun durch die unter den Bedingungen der Digitalisierung erschwerten Verwertungsmöglichkeiten neu austariert werden. Diese Integration technischer Neuerungen bzw. deren Effekte ins Recht und deren diskontinuierliches Moment werden dabei, so scheint es, ignoriert oder heruntergespielt. Unter Bezug auf das von John Locke geprägte und auf den Bereich der Wissensproduktion übertragene Dogma, dass Arbeit ein Recht auf Eigentum schaffe, werden dessen eigene Voraussetzungen tendenziell unterminiert. Diese Voraussetzung, die strukturellen Bedingungen der Schaffung neuen Wissens, von Locke jenseits des sozialen Zusammenhangs als ‚natürlich' den Individuen zukommende Fähigkeiten unterstellt, gerät durch deren Subsumtion unter das vermittels des Rechts restriktiv geregelte geistige Eigentum in eine prekäre Lage. Je mehr das geistige Eigentum die Kreation der Form gegenüber ihrer Interpretation, d.h. ihrer variierenden Wiederholung abdichtet, desto eher laufen, so Schmidt, der Prozess des Erwerbs von Fähigkeiten und ihre Ausübung allmählich leer. So entstehe der paradoxe Effekt, dass bei gleichzeitiger Forderung nach immer mehr Kreativität und Innovation deren Möglichkeiten durch die „Privatisierung der Formen des Produktionsprozesses" zunehmend beschnitten würden[47].

Die Problematik der Analogisierung von dinglichem und geistigem Eigentum erweist sich somit erst im Zusammenhang seiner Produktion. Wird geistiges Eigentum demnach grundsätzlich als Produktionsform begriffen (sowohl im technischen Sinne als auch in Gestalt kulturellen Wissens als Voraussetzung literarischer oder künstlerischer Produktion), ist mit seiner restriktiven Privatisierung, d.h. dem Verbot der Variation etc., die Voraussetzung der Aneignung gefährdet. Auf der Ebene der Zirkulation, die durch eine vorausgegangene Aneignung immer schon bedingt ist, lässt sich dagegen eine der dinglichen Eigentumsordnung tatsächlich analoge Sphäre auch für das geistige Eigentum konstatieren. Gerade insofern Eigentum in der Zirkulationssphäre vor allem als bloß juristischer Ausdruck erscheint, mag es nicht verwundern, dass rechtstheoretische Untersuchungen zu dem Ergebnis kommen, dass ungeachtet des Unterschieds zwischen Sacheigentum und geistigem Eigentum in Bezug auf ihre jeweilige Beschaffenheit, die Differenzen geringer seien, als sie zunächst erschienen[48]. Die Transaktionskosten, die auf der einen Seite durch die Absicherung exklusiver Verfügungsrechte an geistigem Eigentum gesenkt werden sollten, fallen schließlich aber an anderer Stelle als Steigerung der Produktionskosten an. Dies ist u.a. einerseits der Fall im Zusammenhang des Wissenserwerbs durch Unternehmen, also im Rahmen wissensintensiver Produktion, andererseits auf Seiten derjenigen, die sich, um die Erfolgsaussichten auf den Verkauf ihrer Arbeitskraft zu steigern, in zunehmendem Maße gezwungen sehen, nicht unerhebliche Investitionen in Bildung zu tätigen. Die aktuelle Auseinandersetzung stellt sich so als eine zweier Positionen dar, die vor dem Hintergrund einer allgemeinen ökonomischen Dynamik je eine Seite eines grundlegenden Widerspruchs zwischen Einschluss und Offenheit des Wissens, d.h. zwischen der für die Verwertung geistigen Eigentums notwendigen Warenform und der Garantie effizienter Produktionsbedingungen, repräsentieren. Es handelt sich hierbei folglich nicht um eine auf das Feld des geistigen Eigentums beschränkbare Diskussion, sondern die Auseinandersetzung ordnet sich um Funktion und Gestalt des geistigen Eigentums in einem gesamtökonomischen Zusammenhang ein, dessen aktuelle Gestalt die so genannte Wissensgesellschaft ist.

45 Lars Meyer: Theorie und Politik des Eigentums in der Wissensgesellschaft, S. 93. Bremen, 2004.

46 Vgl. Lieckweg, a.a.O., S. 2f.; vgl. Ladeur, a.a.O., S. 19f.

47 Vgl. Schmidt, a.a.O., S. 763, 767f.

48 Vgl. Volker Jänich: Geistiges Eigentum. Eine Komplementärerscheinung zum Sacheigentum?, S. 229. Tübingen, 2002.

Urheberrecht, Gesellschaft, Information, Informationspolitik, Informationswirtschaft, Arbeit, Innovation

DER AUTOR

Volker Gross

Jahrgang 1979. Absolvent des Diplom-Studiengangs Bibliothekswesen an der Fachhochschule Potsdam. Erhielt 2009 den Peis für die beste Diplomarbeit im Fachbereich Bibliotheksmanagement an der Fachhochschule Potsdam, die diesem Artikel zugrunde liegt.

grosz.v@gmail.com

Patentdatenbanken: kostenpflichtige mit Mehrwert oder kostenfreie?

Markus Birkner, Stuttgart

Seit einigen Jahren sind Patentinformationen auch zunehmend durch Datenbanken von Patentämtern sowie Patentsuchmaschinen kostenfrei im Web verfügbar. Diese Angebote stellen eine wachsende Konkurrenz für die kommerziellen Datenbankhersteller und -anbieter dar. Oft ist jedoch nicht bekannt, welche umfangreichen Möglichkeiten die kostenpflichtigen Datenbanken bieten und welche erheblichen Einschränkungen bei den kostenfreien Quellen zu beachten sind. In diesem Artikel werden daher der Leistungsumfang von kostenfreien Patentdatenbanken und die Mehrwerte von kostenpflichtigen Patentdatenbanken im Überblick vorgestellt. Für effektive und effiziente Recherchen eignen sich Retrievalsprachen in einer standardisierten Host-Umgebung. Hochwertige Mehrwertdatenbanken zeichnen sich vor allem durch qualitativ verbesserte Daten und zusätzliche Indexierungen aus. Wenn sie bei Patentrecherchen durch Primärdatenbanken ergänzt werden und beide eine höchstmögliche Abdeckung aufweisen, sind verlässliche und vollständige Trefferlisten möglich. Auch für den weiteren Bearbeitungsprozess stehen bei kommerziellen Anbietern zahlreiche unterstützende Funktionen, Programme und eine Kundenbetreuung zur Verfügung.

Patent databases: fee required but value-added or free of charge?

Since several years patent information has been increasingly available through databases by patent offices as well as patent search engines which are freely accessible on the Web. These offers represent a growing competition for the commercial database producers and providers. However it is often little known which extensive possibilities the fee required databases offer and that there are considerable restrictions when using free of cost sources. This article therefore shows the scope of services of free of charge patent databases and the additional benefits of commercial patent databases in an overview. Retrieval languages in a standardized host environment are suitable for effective and efficient patent searches. Value-added databases are characterized through qualitative improved data and additional indexing. If they are used to supplement patent search with primary databases and if both show the highest possible coverage reliable and complete hit lists are possible. Another advantage for the users of databases by commercial suppliers is the availability of the numerous supporting functions and programs, as well as customer service.

1 Einleitung

Seit den 1990er-Jahren sind zunehmend kostengünstige und vor allem kostenfreie Patentinformationen - verstärkt im Volltext - im Internet für jeden verfügbar. So stellen insbesondere Patentämter und andere Einrichtungen ihre eigenen oder extra produzierten Datenbanken frei ins Web und erweitern ständig deren Leistungsumfang. Durch die steigende Anzahl dieser Angebote haben die kommerziellen Datenbankhersteller und Hosts in den letzten Jahren viel Konkurrenz erhalten.

Auch die Nutzerseite gerät zunehmend in Bedrängnis: der allgemeine Kostendruck steigt weiter und die Finanz- und Wirtschaftskrise führt zu nie gekannten Kosteneinsparungen. Doch gerade in wirtschaftlich sehr schweren Zeiten steckt in Patentinformationen viel ungenutztes Potenzial. Ihr Charakter als Produktionsfaktor wird nur selten wahrgenommen. Die Kosten verursachenden Stellen in den Einrichtungen haben große Schwierigkeiten ihre Kosten und Leistungen dem Unterhaltsträger gegenüber zu rechtfertigen und ihre internen oder externen Kunden sind oft nicht mehr bereit für die Beschaffung von Informationen entsprechend zu bezahlen, da es anscheinend inzwischen alles bei Google & Co. kostenlos gibt. Diese Einstellung kann in Hinblick auf Doppelforschung, immer kürzer werdenden Produktentwicklungszeiten sowie Investitionsrisiken verheerende Auswirkungen haben. Allgemein betrachtet wird die Beschaffung von Patentinformationen oft nur noch als Kostenfaktor angesehen. All dies führt allmählich zu einem veränderten Nutzerverhalten: hin zu kostenfrei verfügbaren Patentinformationen im Web.

Den Herstellern kostenpflichtiger Patentdatenbanken war schon immer der Stellenwert ihrer qualitativ hochwertigen Informationen und Mehrwerte bewusst. Sie haben den eben geschilderten Trend seit langem erkannt und versuchen daher mit entsprechenden neuen Mehrwertangeboten ihre Daseinsberechtigung aufrecht zu erhalten und ihre Kunden zu binden. Aus Interessenten- und Nutzersicht ist jedoch oft nicht bekannt oder transparent ersichtlich, welche Vorteile und Mehrwerte kostenpflichtige Patentdatenbanken konkret bieten.

Dieser Beitrag erhebt nicht den Anspruch, die Vorteile aller kostenpflichtigen Patentdatenbanken vollständig und im Detail abzubilden. Stattdessen werden exemplarische Beispiele aus dem Markt der Patentdatenbanken ausgewählt. Der Schwerpunkt liegt dabei beim Datenbankangebot des Hosts STN International, weil in Deutschland vor allem dessen Datenbanken genutzt werden. Damit soll aufgezeigt werden, in

Dieser Artikel besteht überwiegend aus Auszügen der Bachelorarbeit des Autors: Birkner, Markus (2009): Mehrwerte kostenpflichtiger Patentdatenbanken im Vergleich zum Leistungsumfang kostenfreier Patentdatenbanken. Stuttgart: Hochschule der Medien. http://opus.bsz-bw.de/hdms/volltexte/2009/672/pdf/Birkner_Bachelorarbeit.pdf. Sämtliche Literaturangaben sind dort zu finden. Bei den Angaben und Betrachtungen handelt es sich um Momentaufnahmen während des Bearbeitungszeitraumes dieser Abschlussarbeit von April bis Juli 2009.

welchen Bereichen Mehrwertangebote bereits existieren, potentiell möglich sind und wie andere Datenbankhersteller ihr Angebot verbessern oder erweitern könnten. Diese Übersicht kann ebenso auch institutionellen Nutzern als Argumentationsgrundlage gegenüber ihren Unterhaltsträgern dienen: Warum wird es, trotz kostenfrei verfügbaren Patentinformationen, in den meisten Fällen auch in Zukunft notwendig und sinnvoll sein, kommerzielle Patentdatenbanken mit den entsprechenden Mehrwerten zu nutzen?
Jede Datenbank weist zwangsläufig auch Nachteile auf und ist nicht für jede Fragestellung geeignet. Auf diese Nachteile und Anwendungsmöglichkeiten kann hier nur vereinzelt eingegangen werden. Im Folgenden bleiben Inhouse-Patentdatenbanken unberücksichtigt, da die Zugänglichkeit und Vergleichbarkeit dieser besonderen Datenbankart nicht gegeben ist. Somit ist mit den folgenden Ausführungen keine Abwägung und Hilfestellung zur Auswahl der „richtigen" Datenbank möglich.

2 Leistungsumfang von kostenfreien Patentdatenbanken

Prinzipiell lassen sich kostenfrei verfügbare Patentinformationen im Web unterteilen in kostenfrei zugängliche Patentdatenbanken der Patentämter und kostenfreie Patentsuchmaschinen. Im Folgenden werden beide Typen unter der Bezeichnung kostenfreie Patentdatenbanken zusammengefasst. Patentinformationen die einen wesentlichen Mehrwertcharakter aufweisen sind in diesen Angeboten nicht zu finden.
Kostenfreie Patentdatenbanken zeichnen sich vor allem durch einfach auszuführende Recherchen aus, die überwiegend menügeführt erfolgen. Meist ist es nicht möglich, komplexe Suchanfragen, bestehend aus mehreren thematischen Aspekten, zu generieren. Des Weiteren sind die Funktionen von Retrievalsprachen so gut wie nicht vorhanden.
Bei der einfachsten Rechercheart steht nur ein einziges Eingabefeld ohne Einschränkung auf spezifische Felder zur Verfügung. Nach Eingabe einzelner Wörter oder Phrasen wird i.d.R. im Volltext danach gesucht. Diese Art der Suche ist insbesondere bei der Nutzung von allgemeinen Web-Suchmaschinen weit verbreitet - mit den sich daraus ergebenden und bekannten Nachteilen.
Für feldspezifische Suchen stehen Suchformulare zur Verfügung. Je nach Indexerstellung werden unterschiedlich viele Suchfelder angeboten. Die verschiedenen Suchfelder können meist nur mit dem AND-Operator miteinander kombiniert werden. In einigen Fällen ist bspw. auch die Anzahl der Begriffe, die pro Suchfeld maximal eingegeben werden können, stark begrenzt. In den seltensten Fällen sind Suchanfragen, die auf einer Syntax basieren, möglich. Werden diese angeboten, sind die formulierten Recherchean fragen unübersichtlich und in ihrer Anwendung nicht sonderlich praktikabel. Durch einen fehlenden allgemeinen Standard sind die Suchfunktionen, z.B. die Symbole für Trunkierungen, nicht einheitlich anwendbar.
Im Gegensatz zu den kostenpflichtigen Angeboten ist bei den kostenfreien Patentdatenbanken die Verfügbarkeit von Volltexten recht oft gegeben. Sowohl bei den Patentämtern als auch bei den Patentsuchmaschinen ist häufig die Anzeige und der Download der Dokumente im Format PDF möglich; daneben bietet bspw. das US-amerikanische Patentamt USPTO ihre Volltexte im hochauflösenden Format TIFF an. Bei den Suchmaschinen wird in der Trefferliste jedoch oft direkt auf die entsprechenden Volltexte bei den Patentämtern verlinkt.
Durch diese häufig verfügbaren Volltexte wird vielfach auch eine Volltextsuche angeboten. Um dies zu ermöglichen, werden die Inhalte der Patentschriften mit OCR-Schrifterkennung eingelesen. Jedoch sind nicht immer alle Bestände bereits retrospektiv bearbeitet. Außerdem können bei diesem automatisierten Verfahren Fehler aufgrund nicht korrekter Interpretationen der Schriftzeichen durch die Software auftreten. Eine anschließende Überprüfung und Korrektur der Daten zur Qualitätsverbesserung findet aber nur selten statt. Zudem sind nicht alle Dokumente für die OCR-Schrifterkennung tauglich. Diese Problematik spiegelt sich zum Teil auch bei kostenpflichtigen Patentdatenbanken wider.
Durch die Berücksichtigung aller Feldinhalte der Dokumente bei einer Volltextsuche wird zwar die höchstmögliche Vollständigkeit der Suchergebnisse, bezogen auf die jeweilige Datenbank, erreicht. Dieser hohe Recall-Wert führt jedoch zu einer geringen Precision, d.h. zu einem erheblichen Anteil nicht oder wenig relevanter Treffer. Denn dabei werden auch jene Dokumente berücksichtigt, in denen die Suchwörter nicht im Fokus der Erfindung stehen, sondern bspw. nur in einer allgemeinen Aufzählung erscheinen, als Produktname mit anderer Bedeutung genannt werden oder nur einen sehr geringen Bezug zum erwähnten Patent haben. Die Relevanzprüfung sowie Aussortierung dieser Treffer kann sehr zeitraubend und durch die hohen Personalkosten auch sehr kostenintensiv sein.
In einer Fallstudie aus dem Bereich Pharmazie wurden u.a. die Ergebnisse einer Volltextsuche auf ihre Relevanz hin untersucht: über 30 Prozent der Patentdokumente waren nicht relevant. Allerdings waren unter den Treffern der Volltextsuche auch hoch relevante Patente dabei, die von den Mehrwertdatenbanken nicht gefunden wurden. Dies ist hauptsächlich auf die Zeitverzögerung zurückzuführen. Während die Volltextdatenbanken für gewöhnlich sehr zeitnah nach der Veröffentlichung der Patente ihren Datenbestand aktualisiert haben, dauert bspw. die vollständige Generierung der Mehrwerte bei der Datenbank CAplus bis zu 27 Tage. Ein weiterer Grund liegt z. T. auch an der intellektuellen Erschließung. Dabei werden naturgemäß Prioritäten gesetzt und es wird eine Auswahl der zu indexierenden Inhalte vorgenommen. Nur eine Volltextsuche kann diese aussortierten Elemente erfassen und damit Inhalte suchbar machen, die vom Indexierer zunächst nicht für relevant erachtet wurden.
Die Suche nach asiatischen Patentdokumenten bereitet, insbesondere bei den kostenfreien Quellen, z. T. erhebliche Probleme. Sie sind i.d.R. mit asiatischen Schriftzeichen verfasst und oft fehlen Übersetzungen sowie zusätzliche Indexierungen. Außerdem bieten bspw. nicht alle asiatischen Patentämter eine englischsprachige Suche an oder es sind einige Suchfunktionen teilweise nur in der Landessprache und mit den dazugehörigen Schriftzeichen möglich.
Im Laufe der Zeit haben einige Patentämter mit ihren Datenbankangeboten enorm aufgeholt und Boden zu den kommerziellen Angeboten gut gemacht. So stellen sie bspw. deutlich mehr Daten zur Verfügung und bieten z. T. weiterentwickelte Recherche- und Analysewerkzeuge an. Stellvertretend für die große Vielzahl und Vielfalt an Online-Informationsressourcen der Patentämter werden im Folgenden einige Merkmale der Angebote des Deutschen Patent- und Markenamtes, des Europäischen Patentamtes und der WIPO vorgestellt.

Deutsches Patent- und Markenamt – DEPATISnet:

- Datenbestand: ca. 66 Mio. Patentveröffentlichungen aus über 70 Ländern
- für deutsche Schriften: Volltextrecherche, Verfügbarkeit zum Publikationstag
- Subdokumente und Volldokumente im PDF-Format
- Recherche mit Suchbegriffen oder Querverweisen des IPC-Verzeichnis
- Anzeige der Familienmitglieder
- IKOFAX-Recherchemodus: komplexe Suchanfragen, Syntax ähnlich Messenger, zahlreiche Suchfelder
- Assistentenrecherche: Erstunterstützung durch Experten aus Patentinformationszentren
- interaktiver Info-Lotse: Online-Service, interaktive Erstellung einer Suchstrategie, Nutzung des Remote-Desktop-Verfahrens durch Rechercheexperte und Kunde

- DPMApublikationen: Plattform zur Veröffentlichung der amtlichen Publikationen, geeignet für Überwachungs- und Verletzungsrecherchen
- DPMAkurier: auf DPMApublikationen basierender SDI-Dienst, begrenzte Überwachungskapazitäten
- DPINFO: amtliches elektronisches Register für die vom DPMA vergebenen Schutzrechte, Auskunft über aktuelle Rechts- und Verfahrensstände

Europäisches Patentamt – esp@cenet:

- Datenbestand: über 60 Mio. Dokumente, 5 Mio. Zusammenfassungen
- bestehend aus drei Teildatenbanken:
  - Worldwide: Anmeldungen aus mehr als 90 Ländern
  - EP – esp@cenet: Anmeldungen beim Europäischen Patentamt der letzten 24 Monate
  - WIPO – esp@cenet: Anmeldungen bei WIPO der letzten 24 Monate
- Anzeige der INPADOC-Familie sowie Rechtsstandsdaten
- Integration einer maschinellen Übersetzung
- ECLA-Klassifikationsrecherche, Nummernsuche
- Helpdesk, Nutzerforum
- Register Plus: sämtliche Informationen über europäische Anmeldungen, Verfahrens- und Rechtsstandsdaten, z.B. auch Schriftverkehr, ergänzt um Benachrichtigungsdienst WebReg MT
- Europäischer Publikationsserver: wöchentliche Bereitstellung der EPO-Veröffentlichungen in den Formaten PDF, XML oder ZIP
- IPscore: Software zur Analyse und Verwaltung des Patentportfolio

WIPO – PATENTSCOPE:

- Datenbestand: ca. 1,5 Mio. internationale PCT-Anmeldungen
- Erweiterte Suche mit Syntaxfunktionen
- ansprechende und informative Gestaltung der Trefferliste sowie der einzelnen Treffer
- Optionen für Trefferliste: Sortierung, Listenlänge, angezeigte Felder
- hilfreiche Links zu anderen Dokumenten (Statusberichte, Prüfberichte, etc.)
- RSS-Feeds für individuelle Suchkriterien
- vier automatisch generierte Graphiken zur Trefferauswertung

Die Betreiber mehrerer Datenbanken und Patentsuchmaschinen haben ihre Inhalte für die Volltext-Indexierung durch allgemeine Web-Suchmaschinen zugänglich gemacht. Dadurch wird hauptsächlich die Sichtbarkeit der freien Patentinformationen erhöht.

Vielen Nutzern von Patentsuchmaschinen sind deren Nachteile gar nicht bewusst. Neben den spärlich vorhandenen Suchfunktionen ist einer der wesentlichen Nachteile die mangelnde oder durchweg fehlende Transparenz. Häufig fehlen vor allen Dingen detaillierte Angaben zur inhaltlichen, geographischen sowie zeitlichen Abdeckung. Und wenn nicht bekannt ist, was den Datenbestand umfasst, können auch keine Aussagen getroffen werden, inwieweit die gefundenen Informationen vollständig sind. Darüber hinaus ist, wie bei den allgemeinen Websuchmaschinen, insbesondere das Ranking naturgemäß überhaupt nicht transparent. Da oft keine Möglichkeiten zur Sortierung der Treffer vorhanden sind, gestaltet sich die Relevanzprüfung nach den eigenen Kriterien schwierig und zeitraubend. Bei nahezu allen Patentsuchmaschinen ist es nicht möglich, im Index von ausgewählten Feldern zu blättern. Ebenso können nur selten gezielte Faktenrecherchen durchgeführt werden.

Abgesehen von diesen gravierenden Nachteilen steckt in einigen von ihnen aber auch ein Potential, das oft nicht komplett ausgeschöpft wird. Die folgende Auflistung bietet einen Einblick in den Leistungsumfang einiger Patentsuchmaschinen:

FreePatentsOnline:

- Datenbestand: US-Patente, europäische Patente, japanische Patente und WIPO (PCT)
- Erweiterte Suche mit Syntaxfunktionen
- informative Hilfefunktionen sowie benutzerfreundliche Gestaltung der optisch ansprechenden Website, insbesondere im Bereich Suchformulare
- RSS-Feeds zur Überwachung aktueller Patente, nur für US-Patente, nicht für individuelle Suchkriterien sondern vordefiniert durch IPC-Klassen
- farbliche Hervorhebung der Suchbegriffe in Dokumentenanzeige, zur schnelleren Trefferbeurteilung und besseren Übersichtlichkeit
- störende Werbung bei Einzeltrefferanzeigen
- zur Nutzung erweiterter Funktionen ist kostenfreies Konto notwendig, bspw. für Anzeige von PDF-Volltexten oder Struktursuche
- SureChem Free Chemical Search, u.a. kostenfreie Struktursuche mit intuitiv bedienbarer graphischer Suchoberfläche, kostenpflichtige Erweiterung SureChem Pro mit zusätzlichen Funktionen

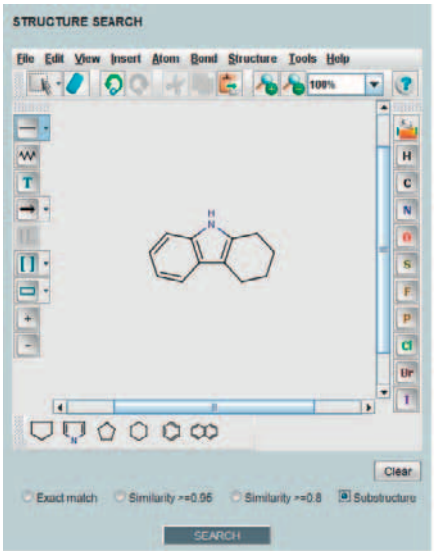


Abbildung 1: Struktursuche SureChem von FreePatentsOnline.

Google Patents:

- Verlinkung zu zitierten und zitierenden Patenten
- Direktzugriff auf PDF-Dateien, keine Verlinkung zum Volltext beim Patentamt
- RSS-Feeds für individuelle Suchkriterien
- Datenbestand: nur USPTO

WikiPatents:

- Einbeziehung der interessierten Öffentlichkeit, Funktionsweise sowie Vor- und Nachteile ähnlich Wikipedia, Nutzer kommentieren und bewerten Patente, für Mitarbeit kostenfreie Registrierung notwendig
- WikiPatents Marketplace integriert, zu Verkaufs- oder Lizenzierungszwecken angebotene Patente, offen einsehbare Werte: geschätzter Wert, prozentualer Lizenzsatz, Marktanteil und geschätzte Marktgröße

Patent Lens:

- von australischer nonprofit Forschungseinrichtung Cambia
- Datenbestand: WIPO, EPA, USPTO, Australisches Patentamt
- Erweiterte Suche mit Syntaxfunktionen
- Suche mit Zeiträumen:
  - Veröffentlichungstag oder Anmeldetag
  - nur für erteilte US-Patente: Zeitpunkt an dem Patentschutz voraussichtlich abläuft oder bereits abgelaufen ist
- strukturierte Anzeige der Familienangaben und Rechtsstandsdaten aus INPADOC
- RSS-Feeds für individuelle Suchkriterien
- Regular BLAST und Mega BLAST: Suche nach Protein- und DNA-Sequenzen

IP Newsflash:

- Portal für gewerbliche Schutzrechte
- bündelt verschiedenartige Informationen aus unterschiedlichen Quellen
- Schwerpunkte: Überwachung von IPC-Klassen, Statistiken, Nachrichten, Rechtsfälle
- gezielte Suche nach Familienmitgliedern und Rechtsstandsdaten einer Pu-

blikationsnummer, basiert auf INPADOC-Daten
- eigentliche Patentsuche erfolgt in FreePatentsOnline

PatentStorm:
- Datenbestand: nur USPTO
- Blättern im alphabetischen Index von Erfindern und Prüfern
- Blättern im chronologischen Index von Patentanträgen und erteilten Patenten
- RSS-Feeds für vorgegebene Kategorien
- Anzeige von PDF-Dokumenten erst nach kostenfreier Registrierung

PatentCluster:
- Datenbestand: im Aufbau, derzeit nur Teil der US-Patente
- Thematische Gruppierung und Visualisierung der Resultate
- Erweiterte Suche mit Syntaxfunktionen und zahlreichen Suchfeldern
- umfangreiche Verlinkungen in den Ergebnissen

Das Unternehmen Intellectual Property Exchange Limited entwickelt und betreibt u.a. die Datenbank SurfIP des Patentamtes von Singapur. Daneben stellt es auch die eigene Patentsuchmaschine IPEXL zur Verfügung. Als Erweiterung für diesen Dienst wurde der sog. *IPEXL Excel Client* für die Trefferweiterverarbeitung und -analyse veröffentlicht. Dabei handelt es sich um eine Excel-Datei, die einmalig heruntergeladen und gespeichert werden muss. In der Regel funktionieren Hilfsprogramme für die Weiterverarbeitung und Analyse von Treffern nach folgendem Muster: Die Recherche erfolgt auf der ursprünglichen Suchoberfläche, danach werden die Treffer in das Hilfsprogramm exportiert und schließlich dort bearbeitet. Bei dieser Lösung wird jedoch eine andere Vorgehensweise gewählt, indem die verschiedenen Bearbeitungsschritte in einer einfach anzuwendenden Excel-Datei vereint werden.
Die Suchfunktionalitäten sind über Verknüpfungen zur entsprechenden Website integriert. Nach einer Suchanfrage werden die Treffer direkt in das Tabellenblatt importiert und können dort nach Belieben z.B. ausgewählt, farblich markiert, sortiert oder mit Hilfe von Diagrammen analysiert werden. Somit erfolgt die Recherche, Trefferbearbeitung und Analyse direkt in dem Auswertungstool. Diese alternative Herangehensweise bietet zahlreiche Vorzüge: Es wird nur noch ein Programm für die anstehenden Arbeitsschritte benötigt, durch die Verwendung von Excel muss kein zusätzliches Programm installiert werden, sämtliche bekannten und gewohnten Excel-Funktionen können individuell nach eigenen Bedürfnissen verwendet werden, asiatische Schriftzeichen werden unterstützt und durch das XLS-Format ist eine Weiterleitung oder Integration der Daten ohne Weiteres möglich. Im Prinzip sind solche Excel-Dateien auch für andere Patentdatenbanken denkbar und wünschenswert.

Die Verwendung von Metasuchmaschinen ist speziell für das Feld der Patentinformationen ein interessanter und relativ neuer Ansatz. Beispielsweise bietet das Patentamt von Singapur mit ihrer Metasuche SurfIP solch ein Werkzeug für die gleichzeitige Suche in diversen Patentquellen an. Je nach Suchfunktion stehen bis zu 13 individuell selektierbare Datenbanken von Patentämtern zur Verfügung. Darunter befinden sich bspw. auch die Länder China, Korea, Taiwan, Japan und Thailand. Nach Eingabe der Suchparameter und Auswahl der zu durchsuchenden Patentquellen erfolgt die Anzeige der Treffer der verschiedenen Ämter jeweils in eigenen Registerkarten. Jedes aufgeführte Dokument ist jeweils direkt mit dem dazugehörigen Treffer in der Ursprungsdatenbank verlinkt. Zusätzlich werden in einer separaten Registerkarte alle Treffer der jeweils ersten Ergebnisseite jeder Datenbank aggregiert angezeigt. Die Sortierung der Treffer geschieht nach dem von den allgemeinen Suchmaschinen her bekannten Prinzip des Relevance Ranking.

Ein wesentlicher Nachteil dieses Angebotes ist die z.T. sehr lange Wartezeit bis die Ergebnisse der verschiedenen Länder vorliegen. Außerdem ist die Aussagekraft der kombinierten Trefferanzeige kritisch zu bewerten. Dies liegt neben dem Relevance Ranking an unterschiedlich großen Treffermengen der jeweils ersten Ergebnisseite. Im aufgeführten Beispiel (Abb. 3) werden von den vier Datenbanken 100 Treffer angezeigt. Diese setzen sich wie folgt zusammen: 50 von USPTO, 25 von WIPO, 15 von EPO und 10 von UK-IPO. Eine gleichberechtigte Berücksichtigung im Ranking ist somit nicht gegeben. Auffällig dabei ist auch, dass zwar die Hälfte dieser Treffer von USPTO stammt, jedoch keiner davon unter den ersten 30 Treffern zu finden ist. Durch die einfach anzuwendende, länderübergreifende Suche ist mit SurfIP ein guter erster Überblick möglich. Aufgrund der Länderabdeckung und der andersartigen Anordnung der Treffer scheint diese Metasuche als Unterstützung für Verletzungsrecherchen geeignet zu sein.
Auf den ersten Blick machen all diese kostenfreien Angebote den Anschein, als ob auch Laien ohne Vorkenntnisse zum Patentwesen einigermaßen gute Patentrecherchen durchführen könnten. Unerfahrenen Patentrechercheuren sei allerdings geraten, sich Unterstützung z.B. in Patentinformationszentren zu holen. Dies ist bspw. auch notwendig um die Recherchen korrekt aufbauen zu können und die Ergebnisse richtig zu interpretieren. Aber selbst erfahrenen Patentrechercheuren ist davon abzuraten, ausschließlich kostenfreie Patentdatenbanken zu nutzen. Was diesen Angeboten fehlt und welche Nutzungseinschränkungen dies bedeutet, wird deutlich beim Betrachten der kostenpflichtigen Patentdatenbanken, insbesondere bei den Mehrwertdatenbanken.

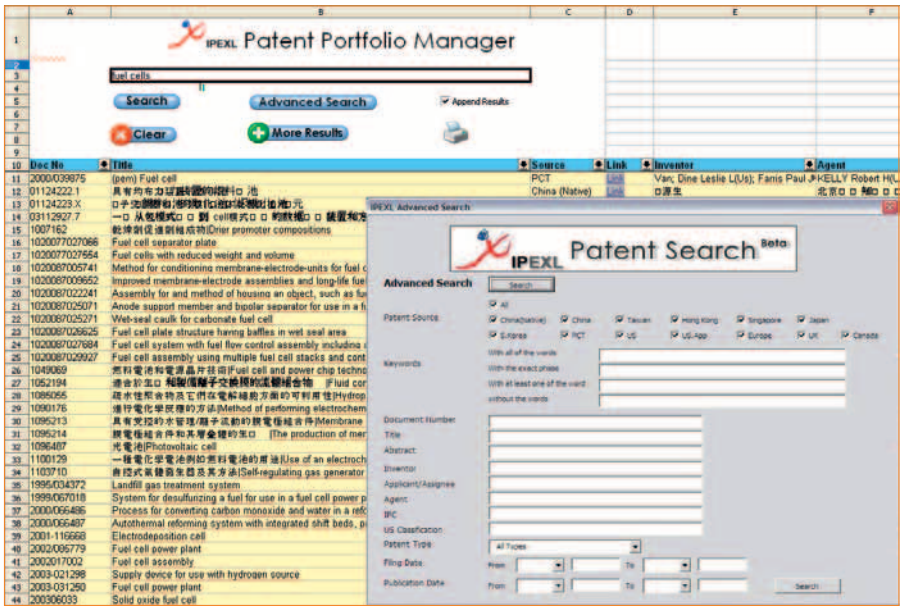
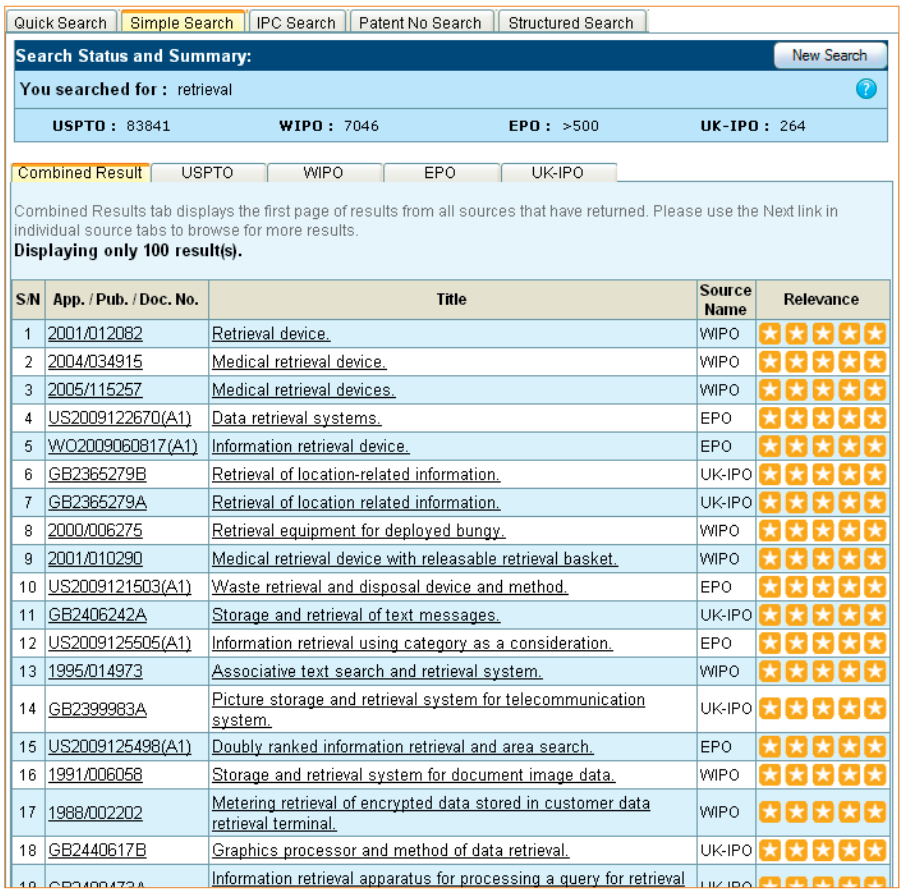


Abbildung 2: IPEXL: Recherche, Trefferbearbeitung und Analyse integriert in einer Excel-Datei.

3 Potentielle Mehrwerte im Überblick

3.1 Abdeckung

Eine der schwierigsten Aufgaben bei Patentrecherchen ist die richtige Auswahl der relevanten Datenbanken. Da keine Patentdatenbank alle Patente enthält und demnach nur für gewisse Fragestellungen geeignet sein kann, sind detaillierte und aktuelle Kenntnisse über die



Quick Search | Simple Search | IPC Search | Patent No Search | Structured Search

Search Status and Summary: New Search

You searched for : retrieval

USPTO : 83841 **WIPO :** 7046 **EPO :** >500 **UK-IPO :** 264

Combined Result | USPTO | WIPO | EPO | UK-IPO

Combined Results tab displays the first page of results from all sources that have returned. Please use the Next link in individual source tabs to browse for more results.

Displaying only 100 result(s).

S/N	App. / Pub. / Doc. No.	Title	Source Name	Relevance
1	2001/012082	Retrieval device.	WIPO	
2	2004/034915	Medical retrieval device.	WIPO	
3	2005/115257	Medical retrieval devices.	WIPO	
4	US2009122670(A1)	Data retrieval systems.	EPO	
5	WO2009060817(A1)	Information retrieval device.	EPO	
6	GB2365279B	Retrieval of location-related information.	UK-IPO	
7	GB2365279A	Retrieval of location related information.	UK-IPO	
8	2000/006275	Retrieval equipment for deployed bungy.	WIPO	
9	2001/010290	Medical retrieval device with releasable retrieval basket.	WIPO	
10	US2009121503(A1)	Waste retrieval and disposal device and method.	EPO	
11	GB2406242A	Storage and retrieval of text messages.	UK-IPO	
12	US2009125505(A1)	Information retrieval using category as a consideration.	EPO	
13	1995/014973	Associative text search and retrieval system.	WIPO	
14	GB2399983A	Picture storage and retrieval system for telecommunication system.	UK-IPO	
15	US2009125498(A1)	Doubly ranked information retrieval and area search.	EPO	
16	1991/006058	Storage and retrieval system for document image data.	WIPO	
17	1988/002202	Metering retrieval of encrypted data stored in customer data retrieval terminal.	WIPO	
18	GB2440617B	Graphics processor and method of data retrieval.	UK-IPO	
19	GB2489473A	Information retrieval apparatus for processing a query for retrieval	UK-IPO	

Abbildung 3: Kombinierte Trefferanzeige mehrerer Datenbanken in SurfIP.

inhaltliche, geographische und zeitliche Abdeckung unerlässlich. In vielen Fällen überschneiden sich die Inhalte der verschiedenen Angebote in unterschiedlichem Umfang. Jedoch kommt es auch vor, dass manche Daten nur in einer Quelle recherchierbar sind. So wurde bspw. in einer Fallstudie festgestellt, dass mehr als 30 Prozent der Patente nur in einer der ausgewählten Mehrwertdatenbanken zu finden waren. In diesem repräsentativen Fall bedeutet es: jede Datenbank liefert zusätzliche Treffer und für eine höchstmögliche Vollständigkeit sind alle relevanten Datenbanken notwendig. Dieses Problem besteht insbesondere bei Patenten kleinerer Patentämter, die keine weiteren Familienmitglieder besitzen. Sie sind oft nur im Deckungsumfang einzelner Mehrwertangebote, wie z.B. CAplus oder DWPI, enthalten.

Generell lassen sich die Angaben zur Abdeckung einzelner Datenbanken nur schwer miteinander vergleichen. Selbst innerhalb einer Datenbank unterscheiden sich bspw. die abgedeckten Zeiträume der verschiedenen Patentämter, Dokumentenarten oder Fachbereiche z. T. erheblich, sodass eine pauschale Jahreszahl für den Beginn der zeitlichen Abdeckung stets kritisch hinterfragt werden sollte.

Als Beispiel hierfür seien die Jahreszahlen 1907 bei CAplus, 1836 bei INPADOCDB und 1963 bei WPINDEX (DWPI) im Vergleich zu den spezifischen Abdekkungszeiträumen in dieser Zusammenstellung[1] genannt. Die Unterschiede sowohl innerhalb einer Datenbank als auch im Vergleich zu den anderen sind z. T. erheblich. Auffallend ist z.B., dass CAplus bei den meisten Ländern zwar am frühesten mit dem Nachweis beginnt, bei manchen Ländern (bspw. Hongkong, Neuseeland, Portugal, Singapur) erfolgt die Dokumentation aber deutlich später als bei beiden Kontrahenten. Außerdem sind einige Länder (bspw. Kuba, Indonesien, Island, Marokko, Ukraine) ausschließlich von INPADOCDB erfasst und es müssen die vielen Besonderheiten bei den Publikationsarten beachtet werden.

Neben solchen Übersichten liefern vor allem die Datenbankbeschreibungen und Handbücher der jeweiligen Hersteller detaillierte und relativ aktuelle Informationen zum Inhalt.

Aus diesen geschilderten Gründen sind professionelle Patentrechercheure auf die gleichzeitige Nutzung mehrerer, sich ergänzender Datenbanken angewiesen. In Frage kommen dafür hauptsächlich die Angebote der drei bedeutendsten Hosts für Patentinformationen: Dialog, Questel und STN.

1 vgl. STN International (2008): Dokumentart-Codes in STN-Patentdatenbanken. www.stn-international.de/fileadmin/be_user/STN/pdf/search_materials/patents/pkcodes_compl.pdf.

Da sich ein Kunde im Regelfall für einen dieser Hosts entscheidet, ist für ihn nicht nur die Abdeckung der einzelnen Datenbanken von Interesse, sondern insbesondere die Zusammensetzung oder Exklusivität des Gesamtangebotes. Beispielsweise benötigen Kunden aus den Bereichen Chemie und Pharmazie verstärkt die Datenbanken DWPI, CAplus, REGISTRY sowie INPADOCDB in ihrem Portfolio. Einziger Anbieter all dieser Mehrwertangebote mit einer standardisierten Hostumgebung ist STN.

Im Folgenden werden exemplarisch einzelne Datenbanken aufgelistet, die sich in *Teilbereichen* entweder durch besondere, sehr spezielle oder sogar exklusive Inhalte oder eine sehr umfangreiche Abdeckung auszeichnen. Dies soll vor allem die mögliche Bandbreite und Abdeckung von Patentinformationen demonstrieren.

***Inhaltliche* Abdeckung:**

- CAplus:
  - 31 Mio. Dokumente
  - Chemie, Biochemie, chemische Verfahrenstechnik, etc.
  - intellektuell erweiterte bibliographische Daten
  - Substanzinformationen aus REGISTRY
  - eigentlich eine Fachdatenbank aber sehr gut für Patentrecherchen im Bereich Chemie geeignet
  - alle Arten von Fachpublikationen nachgewiesen
- Chinese Patent Abstracts in English:
  - alle in China veröffentlichte Patente
  - in Englisch übersetzte Titel und Abstracts
- DGENE:
  - Sequenz-Datenbank mit 8,1 Mio. Nukleinsäure-Sequenzen und 3,4 Mio. Proteinsequenzen
  - mehr als die Hälfte der Sequenzdaten in keiner anderen Sequenz-Datenbank verfügbar
  - viele Indexierungsdaten von DWPI
- DWPI:
  - 37,2 Mio. Dokumente in 17,4 Mio. Patentfamilien
  - eine Patentfamilie je Dokument
  - weltweit die Datenbank mit den aufwändigsten intellektuellen Erweiterungen der bibliographischen Daten
- INPADOCDB:
  - größte bibliographische Patentdatenbank weltweit
  - 65 Mio. Dokumente in 36 Mio. Patentfamilien, 83 Mio. Zitierungen
  - umfangreichste Daten zu Rechts- und Verfahrensständen sowie Familienzugehörigkeiten
  - gleicher Datenbestand in INPAFAMDB, dort repräsentiert jeder Datensatz eine Patentfamilie
  - z.B. auch asiatische Gebrauchsmuster

– viele zusätzliche Abstracts; Europäische Klassifikationen (ECLA, ICO, IDT)
- LITALERT:
  – Informationen zu Gerichtsverfahren in den USA: Verstöße gegen Patent- und Markenrecht
  – z.T. bisher noch unveröffentlichte Informationen
- MARPAT:
  – 748.000 Markush-Strukturen für Patente in CAplus
- PATDPA:
  – Offenlegungs-, Auslege- und Patentschriften
  – auch Rechtsstände und Gebrauchsmuster
- PatentWeb:
  – 50 Mio. Dokumente im Volltext
- PCI:
  – Zitierungen aus den Recherchebericht en der Prüfer: Patente und Literatur → Verwendung wie herkömmliche Bibliographien
  – 8 Mio. Patentfamilien mit 98 Mio. Patentzitierungen und 11 Mio. Literaturzitierungen
  – Nutzung der gleichen Familienstruktur wie DWPI
- PlusPat:
  – 54,9 Mio. Dokumente
  – in FamPat repräsentiert jeder Datensatz eine Patentfamilie
- RDISCLOSURE:
  – defensiv publizierte Veröffentlichungen von Erfindungen: durch Offenlegung können sie nicht mehr als Patent angemeldet werden
  – in den meisten Fällen anonyme Veröffentlichungen zur Geheimhaltung der eigenen Forschungsaktivitäten
  – wichtig für Recherchen zum Stand der Technik
  – Volltexte; Quelle: Zeitschrift ‚Research Disclosure'
- REGISTRY:
  – weltweit größte Substanz-Datenbank
  – 86 Mio. Dokumente
  – 33 Mio. organische und anorganische Substanzen, 59,4 Mio. Sequenzen
  – jede Substanz mit eindeutiger CAS REGISTRY Number
- USPATFULL:
  – 5,4 Mio. Volltexte des Patentamtes der USA
  – auch Gebrauchsmuster

***Geographische* Abdeckung:**
- CAplus:
  – 59 Patentbehörden, auch kleinere Patentämter
  – gute Abdeckung asiatischer Patente
- DWPI:
  – 41 Patentbehörden, auch kleinere Patentämter
  – gute Abdeckung asiatischer Patente und Gebrauchsmuster
- INPADOCDB:
  – 80 Länder, auch kleinere Patentämter
  – allgemein sehr weite, geographische Abdeckung → wertvoll für technologie-orientierte Suchanfragen
- PCTFULL:
  – PCT-Anmeldungen der 118 PCT-Mitgliedsstaaten im Volltext
- PlusPat:
  – 85 Patentbehörden, auch kleinere Patentämter
- TotalPatent:
  – bibliographische Daten von 96 Patentbehörden, Volltexte von 22 Patentbehörden

***Zeitliche* Abdeckung:**
- CAplus:
  – ab 1907, vereinzelt auch früher
- DWPI First View (WPIFV):
  – in Verbindung mit DWPI für aktuellen Datenbestand
- INPADOCDB:
  – ab 1836
- USPAT2:
  – in Verbindung mit USPATFULL für aktuellen Datenbestand
- USPATOLD:
  – von 1790 bis 1975

Was nützen Datenbanken mit hohem Abdeckungsgrad, aufwändigen Indexierungen, vielfältigsten Retrievalfunktionen oder hilfreichen Analysewerkzeugen, wenn der Datenbestand nicht aktuell ist?

3.2 Aktualität der Inhalte

Speziell bei Recherchen zum Stand der Technik, Neuheitsrecherchen oder zur Wahrung von Einspruchsfristen ist es zwingend notwendig, dass die Ergebnisse vollständig und auf dem neuesten Stand sind. Ansonsten verfehlen die Recherchen ihre Zweckbestimmung. Im Idealfall gelangen die Datensätze neuer Dokumente ohne Zeitverzögerung direkt nach deren Veröffentlichung in die Datenbank. In der Praxis können die neuesten Informationen in manchen Datenbanken allerdings mehrere Monate alt sein.
Im Vergleich zu den meisten kostenfreien Angeboten und den Datenbanken mit First-Level-Daten ist das Kriterium der Aktualität für Patentdatenbanken mit Mehrwerten eine wirkliche Herausforderung. Während viele der Primärdatenbanken sogar tagesaktuell sind, nimmt die aufwändige Generierung von Mehrwerten erhebliche Zeit in Anspruch. Beispielsweise dauerte es Anfang des Jahres 2003 im Durchschnitt 125 Tage bis eine Erstveröffentlichung des amerikanischen Patentamtes in DWPI erschien. Für die dreimonatige Widerspruchsfrist in Deutschland war dies eindeutig zu lange. Thomson Reuters hat die Aktualität der Daten daraufhin verbessert. Im Juli 2008 betrug die durchschnittliche Bearbeitungszeit für DWPI 49 Tage. Angestrebt ist eine vollständige Mehrwertgenerierung innerhalb von 30 Tagen nach der Veröffentlichung. Es ist prinzipiell davon auszugehen, dass den Dokumenten aus größeren Ländern eine höhere Priorität zuteil wird, sodass diese schneller bearbeitet werden.
Datenbankproduzenten oder Hosts versuchen eine hohe Aktualität ihrer Datenbanken zu gewährleisten, indem sie die (Meta)Daten direkt von den Informationsproduzenten elektronisch erhalten, für bestimmte Dokumente oder Dokumenttypen Prioritäten im Bearbeitungsprozess setzen, den Ablauf mit Statistiken der Durchlaufzeiten überwachen und Updates in kurzen Zeitabständen durchführen.
Abgesehen von der Bearbeitungszeit bei Mehrwertdatenbanken spielt bei allen Datenbanken auch die Updatefrequenz eine Rolle. Wünschenswert wäre im Prinzip eine möglichst häufige Aktualisierung, welche allerdings mit einem höheren Aufwand bei den Datenbankanbietern verbunden ist. Der Turnus reicht von monatlich wie z.B. bei JAPIO, über wöchentlich bei PATDPA und alle 3 bis 4 Tage bei DWPI bis hin zu täglich bei REGISTRY.

3.3 Fehlerbereinigung und Datenqualität

Zur Erhöhung der Datenqualität stehen bei der eigenen Datenbankproduktion hauptsächlich folgende Maßnahmen zur Verfügung:
- „Verwendung von Regelwerken, Authority Listen (Zeitschriften, Autoren, Fachbegriffe etc.) und Standards (ISO Codes, Nummernformate, Maßeinheiten)
- Automatisierte Prüfungen wie Rechtschreibeprüfungen [sic] und Dublettenerkennung sowie Plausibilitäts- und Konsistenztests
- Automatische Berechnungen von ableitbaren numerischen Daten
- Datensichtung anhand von Prüfausdrucken und Registerausgaben
- Kontrolle von Fehlerprotokollen der verschiedenen Datenprozessierungsprogramme
- …strukturelle Einheitlichkeit aller Dokumente – auch über Regelwerksänderungen hinweg oder bei Einsatz neuer Formate, Standards und Methoden…"

Hingegen können Hosts, als Anbieter von Datenbanken anderer Produzenten, die Daten aus rechtlichen Gründen nicht beliebig editieren. Sie müssen die Einhaltung der geistigen Eigentumsrechte wahren. „Wo ein Copyright zu beachten ist (Sekundärdatenbanken), kann Qualitätsverbesserung nur ‚on top' geschehen.

Das heißt über aufgesetzte Maßnahmen wie Standardisierungen, zusätzliche Klassifizierungen oder auch Vernetzung von Inhalten."
Je größer die Datenbestände sind, desto wichtiger ist die Verwendung von Standards und Regelwerken. Folgende Aspekte werden erst durch einheitlich aufgebaute Daten möglich:

- strukturierte, effektive und effiziente Recherchen
- hohe Recall-Werte
- Erkennen von zusammengehörigen Informationen
- Datenaustausch
- Data Mining
- automatisierte Analysen und Visualisierungen.

Bei den bibliographischen Angaben bestehen zwar Standards, wie bspw. INID-Codes oder Ländercodes für eine sprachunabhängige Erschließung, allerdings gibt es in diversen Bereichen länderspezifische Eigenheiten. Besonders deutlich wird dies am Beispiel der diversen Nummern in Patentdokumenten. Die Anmelde-, Prioritäts- und Publikationsnummern können aus Ländercodes, Jahresangaben, Provinzkennungen und spezifischen Nummern für das jeweilige Patent bestehen. Es gibt keine weltweit gültige Syntaxregel, die die Anzahl, Reihenfolge und Länge der Bestandteile einer Nummer einheitlich regelt. Somit geht jedes Patentamt beim Aufbau seiner Nummern unterschiedlich vor.
Für die Rechercheure würde dies bedeuten, dass sie keine Multifile-Suche mit Hilfe von Patentnummern - mit der sonst gewohnten hohen Vollständigkeit - und auch keinen anschließenden Dublettenabgleich durchführen können. Durch das Fehlen einer Standardisierung in diesem Bereich muss nachgearbeitet werden. Als Folge stellt bspw. STN für all seine Patentdatenbanken nachträglich einen eigenen Standard der diversen Patentnummern her. Dazu werden zunächst die angelieferten Nummernformate analysiert. Anschließend werden für jedes Land intellektuell Umsetzungsregeln erstellt, die stets aktualisiert und erweitert werden müssen. Erst dadurch ist das Retrievalsystem in der Lage alle Datensätze datenbankübergreifend zu identifizieren, sodass mit einer Recherche alle in Frage kommenden Datenbanken gleichzeitig durchsucht und die Dubletten eliminiert werden können.
Inhaltliche Korrekturen zur Erhöhung der Genauigkeit können vorgenommen werden, wenn die Primärdaten der Datenlieferanten kein Copyright besitzen. Diese Rohdaten stammen meist von Patentämtern. Um eine spätere Nachvollziehbarkeit gewährleisten zu können, empfiehlt es sich die originalen Daten beizubehalten und die korrigierten Informationen in Ergänzung einzupflegen. Doch welche Fehler sollten bereinigt werden? Als Beispiel hierfür wird im Folgenden die korrigierende Zusammenführung von Patentfamilien in der Datenbank INPADOCDB von STN vorgestellt.
Bekanntermaßen ist es das Ziel von Patentfamilien, weltweit alle Schutzrechtsdaten der Patentdokumente, die zu einer Erfindung gehören, zu gruppieren und recherchierbar zu machen. Gerade in diesem Punkt spielt die Vollständigkeit die entscheidende Rolle. Die Bildung von Patentfamilien geschieht durch den Abgleich der Aktenzeichen der prioritätsbegründenden Anmeldungen. Alle Nachanmeldungen, die eine bestimmte Priorität in Anspruch nehmen, werden dem Ursprungsaktenzeichen zugeordnet. Sind jedoch diese Prioritätsnummern nicht absolut einwandfrei, können die Dokumente auch nicht mit anderen Familienmitgliedern zusammengeführt werden. Fehler in den Nummern können durch die Angabe einer falschen Prioritätsnummer auf der Originalanmeldung, durch Schreibfehler in den Nachanmeldungen sowie durch Fehlformatierungen beim Datenimport oder -export entstehen. Dementsprechend erfolgt die Mehrwertgenerierung durch die exakte Plausibilitätsprüfung der Nummern und Korrektur der fehlerhaften Daten.
Dazu legt das FIZ Karlsruhe zunächst in einer Prüftabelle alle Standards der Nummernformate an und aktualisiert diese regelmäßig. Bei der anschließenden automatisierten formalen Qualitätsprüfung werden die vom Europäischen Patentamt angelieferten Rohdaten mit den hinterlegten Standards verglichen. Entstehende Abweichungen werden schließlich redaktionell untersucht und bereinigt. Durch diese Prüfroutinen können die bisher fehlenden oder falschen Familienmitglieder identifiziert und korrigiert werden, sodass die neuen Familieninformationen zuverlässig und vollständig sind.

All diese Korrekturen sind erst dann sinnvoll, wenn sie auch auf Dauer erhalten bleiben. Je nach Verlauf von Patentverfahren ändern die Patentämter entsprechend die Inhalte und schicken die neuen kompletten Datensätze mit den Ergänzungen an die Datenbankhersteller oder Hosts. Die alten Sätze dürfen aber nicht ungeprüft mit den neuen Informationen überschrieben werden, da sich in ihnen eventuell die bereits erkannten Fehler erneut befinden. Um sicherzustellen, dass die schon durchgeführten Korrekturen nicht verloren gehen und nochmals intellektuell bearbeitet werden müssen, führt bspw. auch das FIZ Karlsruhe bei Neulieferungen von bereits korrigierten Datensätzen komplexe Abgleiche, Selektionen und ggf. Zusammenführungen durch. Nur so kann eine Konsistenz und Kontinuität in der Datenhaltung gewährleistet werden.
Diese Veredlung von Primär- und Sekundärdaten ist sehr aufwändig und für Außenstehende oft nicht ersichtlich. Aber gerade diese Mehrwerte sind es, die professionelle Datenbanken ausmachen.

3.4 Indexierung

Die Aktivitäten der Datenbankhersteller in den Bereichen Fehlerbereinigung und Indexierung gehen oft direkt ineinander über. Ein Beispiel hierfür ist die Datenbank PCTGEN, mit der Recherchen in Sequenzen durchgeführt werden können. Die Originaldaten stammen von der WIPO und werden vom FIZ Karlsruhe aufbereitet. Dazu werden die Dateien zunächst mit einer OCR-Schrifterkennung eingelesen und in mehreren Bearbeitungsschritten in TXT-Dateien umgewandelt. Anschließend erfolgt eine manuelle Korrektur dieser Daten, gefolgt von automatisierten Plausibilitätsprüfungen. Als Ergebnis steht pro Sequenz ein durchsuchbarer Datensatz in hoher Qualität zur Verfügung.

Die Internationalität des Patentwesens zeigt sich eindrucksvoll in der Vielzahl der Sprachen und Schriftzeichen, in denen Patentdokumente verfasst werden. Dies erschwert die sprachenübergreifende Recherche mit Stich- oder Schlagwörtern in erheblichem Ausmaß. Die IPC als international gültiges Klassifikationssystem versucht die sprachliche Unabhängigkeit bei Recherchen zwar herzustellen, aber insbesondere bei Recherchen, bei denen es auf Vollständigkeit ankommt, wie z.B. Recherchen zum Stand der Technik oder Neuheitsrecherchen, kann sich nicht alleinig auf die IPC verlassen werden.
Für die Indexierung und spätere Auswertung von Patentinformationen gilt eine wesentliche Besonderheit und Herausforderung des Patentwesens: Die Antragsteller möchten zwar einen möglichst umfassenden Schutz gewährt bekommen, wollen aber nur so wenig Informationen wie möglich transparent offenlegen. Daraus resultierend ist die Patentsprache schwer verständlich, umständlich und sehr allgemein formuliert. Um nun doch weltweite Recherchen mit Suchwörtern in einer einheitlichen Sprache durchführen zu können und verständliche Erfindungsbeschreibungen sowie Ansprüche zu erhalten, gibt es vereinzelte Datenbanken mit Übersetzungen oder eigens verfassten Inhalten in englischer Sprache. Die Herstellung dieser Inhalte ist äußerst aufwändig. Viele Rechercheure sehen darin jedoch einen der bedeutendsten Mehrwerte der generiert werden kann. Durch diese erweiterten Textinformationen ergibt sich bspw. der wesentliche Vorteil einer zügigeren und zuverlässigeren Relevanzprüfung, da sie

dem Nutzer ein besseres Verständnis für die neue Erfindung bieten.
Bei der Erstellung eigener Abstracts oder sonstiger Indexierungselemente, wie z.B. Schlagwörter, ist an dieser Stelle auf eine weitere Besonderheit im Patentwesen hinzuweisen. Der technische Inhalt eines Patentes kann innerhalb einer Patentfamilie zwischen den verschiedenen Dokumenten der Familienmitglieder durchaus variieren. Wenn nun zwei Datenbankhersteller für die Indexierung eines Patentes jeweils ein anderes Dokument der Patentfamilie auswählen, können sich dadurch Unterschiede bei den resultierenden Mehrwerten ergeben.

Der bedeutendste Vertreter für intellektuelle Erweiterungen der bibliographischen Metadaten und tiefe sachliche Erschließungen ist DWPI mit folgenden besonderen Merkmalen:

- neu geschriebene, informative Titel und Abstracts in klarer englischer Sprache für alle Datensätze durch Fachleute - unabhängig von der Originalsprache des Patentes; dadurch werden auch fremdsprachige Patente, z.B. aus Asien, suchbar, die kein englischsprachiges Dokument in ihrer Patentfamilie besitzen
- Einordnung in ein eigenes Klassifizierungs- und Indexierungssystem, den sog. Manual Codes und Derwent Classes; im Fokus steht dabei die Technologiebetrachtung aus Sicht der Industrie und nicht wie bei der IPC die Abdeckung aller Ansprüche im Detail
- zusätzliches Indexierungssystem für chemische Strukturen, den sog. Chemical Fragmentation Codes und Polymer Codes
- spezieller Index mit standardisierten Namen und zusammenführenden Codes für Patentanmelder, den sog. Derwent Patentee Codes
- intellektuelle Zusammenführung von verwandten Patenten anderer Länder mit dem Basisdokument zu Patentfamilien
- standardisierte Formate für Veröffentlichungsnummern und Aktenzeichen
- weitere Standardisierungen und Fehlerkorrekturen.

Auch CAplus weist bei der Indexierung einige Besonderheiten auf. Beispielsweise nutzen CAplus und zahlreiche weitere Datenbanken von STN die sog. CAS Registry Number, um damit die bibliographischen Daten anzureichern. Sobald eine chemische Substanz in die Datenbank REGISTRY von CAS aufgenommen wird, erhält sie eine einmalige CAS Registry Number, mit der sie eindeutig identifizierbar ist. Jeder bekannte chemische Stoff kann dadurch ohne die aufwändige Verwendung von spezifischem Fachvokabular oder Summenformeln ohne Verwechslung benannt, zitiert und recherchiert werden. Dieses Nummernsystem wurde zu einem international anerkannten Standard für die zweifelsfreie Bezeichnung jeglicher Substanzen.

Des Weiteren unternimmt CAS zahlreiche Anstrengungen zur erweiterten inhaltlichen Erschließung aller Einträge in CAplus. Dazu zählen z.B. englischsprachige Abstracts und speziell auf den Chemiebereich zugeschnittene Thesauri.
Was eine gute Indexierung im Zusammenspiel mit einer hohen Datenqualität und umfangreichen Retrievalfunktionen bewirken können, zeigt folgendes Beispiel eindrücklich. In einer hier bereits erwähnten Fallstudie aus dem Bereich Chemie wurde in den Datenbanken REGISTRY, CAplus, DWPI und MARPAT nach Patenten mit dem Arzneistoff Pantoprazol mit Hilfe von Strukturformeln und Stich-/Schlagwörtern recherchiert. Die Trefferanzahl betrug nach Entfernung der Dubletten 587. Beachtlich dabei ist die sehr hohe Treffergenauigkeit: Lediglich ein Prozent der Treffer war nicht relevant.

Als Kernelement der Indexierung professioneller Datenbanken gilt eine hohe Granularität der Informationen. Darunter versteht man die Bildung möglichst kleiner Informationseinheiten in fein untergliederte Felder mit jeweils spezifischen Bedeutungen. Dabei werden die Metadaten eines Teilbereiches, die sonst oft in einem Feld zusammengefasst sind, in ihre Bestandteile zerlegt und nach deren Eigenschaften und Zugehörigkeiten unterschieden. Das Ergebnis ist ein neues Datenformat. „Je feinkörniger die Angaben, desto eindeutiger die Indexgenerierung innerhalb der Datenbank und somit die Recherchemöglichkeiten." Dadurch wird ein gezieltes Retrieval überhaupt erst möglich, welches die Treffergenauigkeit und damit auch die Kundenzufriedenheit erheblich erhöht.

3.5 Retrievalfunktionen

Die von Hosts zur Verfügung gestellten Retrievalsprachen, wie z.B. Dialog von Dialog oder Messenger von STN, stellen durch ihre zahllosen Möglichkeiten bedeutende Suchwerkzeuge dar. Ihre kompetente Anwendung für Patentrecherchen verlangt eine sorgfältige Einarbeitung, umfangreiche Kenntnisse über das Patentwesen und über den detaillierten Aufbau von Patentdokumenten, regelmäßige Rechercheпraxis sowie das laufende Studium der Neuerungen. Für laienhafte oder semiprofessionelle Recherchen sind sie daher kaum geeignet. Wenn ihr voller Funktionsumfang ausgeschöpft wird, werden sie vermutlich auch in Zukunft den Recherchefunktionen kostenfreier Dienste bei weitem überlegen sein.

Für eine effektive und effiziente Arbeit mit Retrievalsprachen ist spezielle Software, teilweise noch Kommunikationssoftware genannt, sehr nützlich bzw. notwendig. Damit können die Recherchen kostenlos offline vorbereitet werden, indem bspw. die Suchbegriffe, Suchlogik und Verknüpfungen der Suchschritte in einer Kommandodatei gespeichert werden. Auch spezielle Suchen nach chemischen Strukturen oder Biosequenzen lassen sich so ohne Kostendruck vorbereiten. Dies reduziert die Anschaltzeit-Gebühren auf ein Minimum und dient der Erleichterung regelmäßiger Abfragen. Auf Wunsch protokollieren die Programme sämtliche Aktionen und Ergebnisse einer Recherche. Des Weiteren unterstützen sie die Arbeit durch zahlreiche weitere Funktionen, wie bspw. Analyse sowie Weiterverarbeitung der Resultate, und Schnittstellen zu anderen Programmen.

Ein essentieller Vorteil von Retrievalsprachen ist die Nutzung der Multifile-Suche. Hosts bereiten die von ihnen angebotenen Datenbanken so auf, dass die Suchfelder der verschiedenen Datenbanken eine einheitliche Struktur aufweisen und mit denselben Kommandos durchsuchbar sind. Dadurch wird es möglich, mit einer Abfrage in mehreren Datenbanken gleichzeitig zu recherchieren und anschließend die Dubletten zu entfernen. Meistens werden von Seiten der Anbieter aus, thematisch zusammengehörige Datenbanken bereits in einer vordefinierten Kategorie unter einem Gruppenname subsumiert. So bietet bspw. STN seine derzeit 39 patentrelevanten Datenbanken unter dem Cluster *Patents* an. Der Nutzer kann sich aber auch jederzeit sein eigenes Portfolio der zu durchsuchenden Datenbanken individuell zusammenstellen.

Insbesondere im Bereich der Patentrecherchen ist diese datenbankübergreifende Suche äußerst hilfreich. Hier ist i.d.R. eine vollständige Abdeckung aller relevanten Dokumente gefragt. Durch die Verteilung der relevanten Informationen auf zahlreiche Datenbanken wären für eine Fragestellung ansonsten diverse Recherchen mit unterschiedlichen Feldbezeichnungen notwendig. Jedoch weisen nicht alle Datenbanken eines Hosts einen identischen Umfang ihrer Suchfelder auf. Es gibt sogar teilweise Feldinhalte, die nur in einer Datenbank enthalten sind und somit ein Alleinstellungsmerkmal darstellen. Daher ist es trotz komfortabler Multifile-Suche notwendig, die detaillierten Datenbankbeschreibungen zu Rate zu ziehen. Sie enthalten neben Angaben zur Abdeckung, den Anzeigeformaten und einem Beispieldokument insbesondere eine detaillierte Aufstellung und Beschreibung sämtlicher Suchfelder mit Beispielen.

Die folgende Auflistung zeigt eine Auswahl an bedeutenden Funktionen der Retrievalsprachen am Beispiel von Messenger:

- Multifile-Suche, Crossfile-Suche
- Boolsche Operatoren und Nachbarschaftsoperatoren
- feldbezogene Suche in beliebig vielen Feldern
- sämtliche Maskierungsarten (links, innen, Anzahl Zeichen, etc.)
- zeilenweise Eingabe von Suchschritten, die einzelnen Zeilen können beliebig miteinander kombiniert werden
- Auflisten der Feldinhalte
- Suche von Thesaurusbegriffen
- Patentfamilien suchen und sortieren
- Aussortieren von gekennzeichneten Dokumenten zur Generierung einer präzisen Trefferliste
- Entfernung von Dubletten
- Trefferanzeige genau definierbar, tabellarische Anzeige
- farbliche Hervorhebung der Suchbegriffe in Trefferumgebung
- feldspezifische Sortierung der Antworten

BOEH	BOEHRINGER INGELHEIM AGROVET
BOEH	BOEHRINGER INGELHEIM DIAGNOST
BOEH	BOEHRINGER INGELHEIM GMBH
BOEH	BOEHRINGER INGELHEIM INT GMBH
BOEH	BOEHRINGER INGELHEIM INT KG
BOEH	BOEHRINGER INGELHEIM KG
BOEH	BOEHRINGER INGELHEIM LAB
BOEH	BOEHRINGER INGELHEIM LTD
BOEH	BOEHRINGER INGELHEIM PHARMA
BOEH	BOEHRINGER INGELHEIM VETMEDICA
BOEH	BOEHRINGER SOHN C H
BOEH	BOEHRINGER SOHN CHEM FAB C H
BOEH	BOEHRINGER SOHN GMBH C H
BOEH	BOEHRINGER SOHN INGELHEIM C H
BOEH	BOEHRINGER SOHN INT GMBH
BOEH	CELAMERCK GMBH & CO KG
BOEH	GEBR BOEHRINGER GMBH

Abbildung 4: Beispiel eines Derwent Patentee Code mit allen dazugehörigen Einträgen.

- Selektierung und Übertragung der Suchergebnisse
- Auswertung der Suchergebnisse mit Statistikkommandos (z.B. für Erfinder-, Anmelder-, Länder-, Fachgebiets- oder Zitierungsstatistiken)
- Einrichtung von SDI-Profilen
- Datenspeicherung.

Ein weiterer sehr wichtiger Mehrwert von Retrievalsprachen ist das Blättern in Indexen. Gerade bei Namensfeldern in den bibliographischen Daten sind verschiedene Ansetzungen keine Seltenheit. Durch diese Funktion können unterschiedliche Schreibweisen bspw. in den Feldern Patentanmelder oder Erfinder festgestellt werden, um damit die Recherche auszuweiten und so dem Ideal einer vollständigen Recherche näher zu kommen.

Als sehr hilfreiche Besonderheit in diesem Bereich sind die sog. Derwent Patentee Codes in der Datenbank Derwent World Patents Index zu erwähnen. Dies ist ein zusätzlicher, manuell erstellter Index. Bei jedem Patent in der Datenbank wird dem jeweiligen Patentinhaber ein standardisierter Name und ein Code aus vier Buchstaben zugeordnet. Die Codebezeichnung lehnt sich i.d.R. an den Namen des Patentinhabers an. Alle zusammengehörigen Firmenbezeichnungen oder Tochterfirmen werden unter einem einheitlichen Code geführt. Durch die Recherche mit diesem Code können gleichzeitig alle dazugehörigen Namens-

NEUERSCHEINUNG Juni 2010

Patentinformetrie:
Analyse und Verdichtung von technischen Schutzrechtsinformationen

Jasmin Schmitz

Patentinformationen spielen eine entscheidende Rolle bei der Vorbereitung der eigenen Forschungs- und Entwicklungsaktivitäten und ermöglichen eine Beobachtung von Mitbewerbern und Technikgebieten. Patentschriften sind nicht nur Rechtsdokumente, sie enthalten vielmehr technisches Wissen, welches häufig nicht in Fachzeitschriften publiziert wird. Darüber hinaus geben sie beispielsweise Auskunft über die wirtschaftliche Situation eines Unternehmens, sind also auch Träger von Wirtschaftsinformationen. Diese Informationen können mittels Patentrecherche und anschließender Analyse zutage gefördert werden. Eine Herausforderung bei der Recherche in Patentdatenbanken ist die immense Treffermenge, die – kombiniert mit der Komplexität der Dokumente und einer hohen Informationsdichte – einen erheblichen Zeitaufwand bei der Weiterverarbeitung bedeutet. Hier leistet die Informetrie als informationswissenschaftlicher „Methodenkasten" zur Analyse von Datenbankinhalten gute Dienste. Mit ihren Verdichtungsmöglichkeiten bietet sie Instrumente, um verborgene Informationen in Dokumentmengen aufzuspüren und dabei auch die Informationsmenge zu verkleinern.
Die Autorin geht der Frage nach, welche erprobten informetrischen Methoden auf den Patentbereich übertragbar sind und welche Erkenntnismöglichkeiten und Anwendungsfelder sich daraus ableiten lassen. Außerdem untersucht sie, wie mit bestehenden Patentanalysesystemen Patentinformetrie betrieben werden kann. Auf der Basis eines ausführlichen Marktüberblicks und ausgehend von Detailanalysen ausgewählter Produkte erarbeitet sie schließlich Verbesserungsvorschläge für solche Systeme.

Leserkreis
Patentrechercheure, Verantwortliche in F&E-Abteilungen, Informationswissenschaftler, Wissensmanager, Information Broker, Informationsanbieter, Bibliothekare, Studierende

Jasmin Schmitz
Patentinformetrie: Analyse und Verdichtung von technischen Schutzrechtsinformationen
Mit einem Geleitwort von Wolfgang G. Stock.
Frankfurt am Main 2010, 328 Seiten, DGI Schrift (Informationswissenschaft – 12)
ISBN 978-3-925474-69-9, EUR 29,80 (für DGI-Mitglieder EUR 23,90)
Bestellungen an die DGI-Geschäftsstelle:
Hanauer Landstraße 151-153, 60314 Frankfurt am Main, Fax 069 4909096, E-Mail: publikation@dgi-info.de

ansetzungen, Namensänderungen sowie Firmenzukäufe abgedeckt werden.
Ein weiteres Beispiel für Online-Thesauri ist der IPC-Thesaurus von STN. Er ist in allen STN-Patentdatenbanken enthalten, in denen die aktuelle IPC-Version 8 verfügbar ist. In ihm sind alle Codes der International Patent Classification von der ersten bis zur aktuellen Version integriert. Die Ansicht ist hierarchisch aufgebaut und beinhaltet die Definitionen der IPC-Codes. Dieser Thesaurus ist vor allem nützlich für die Suche in IPC-Bereichen. Entsprechend dem neuen dynamischen IPC-System wird er laufend aktualisiert.

Eine besondere und recht spezielle Art der Recherche stellt die Struktursuche in der Datenbank MARPAT von CAS dar. MARPAT beinhaltet die Patente aus CAplus mit sog. Markush-Strukturen. Dies sind generische Strukturformeln mit denen in Patenten viele ähnliche, chemische Verbindungen abgedeckt werden können. Durch die Recherche strukturähnlicher Verbindungen mit Hilfe von Summenformeln ist es möglich, Patentdokumente zu finden, die bei herkömmlichen Suchen fehlen würden. Zur Vorbereitung dieser Recherchen können die Strukturen komfortabel und offline in der Kommunikationssoftware STN Express gezeichnet werden.

3.6 SDI-Dienste

Die Zielgruppe von Überwachungsrecherchen sind Rechercheure, die gleiche Suchabfragen in regelmäßigen Abständen wiederholen. SDI-Dienste tragen wesentlich zur Reduzierung der Informationsflut bei. Im Patentwesen werden neue Anmeldungen hauptsächlich auf Fachgebiete, Anmelder und Erfinder überprüft. Ein konkreter Anwendungsfall kann z.B. das rechtzeitige Erkennen von Schutzrechtsverletzungen oder eine kontinuierliche Marktbeobachtung sein. Vor allem bei den Hosts stehen umfangreiche Optionen zur Einrichtung von individuellen Rechercheprofilen und zur Übermittlung der Ergebnisse zur Verfügung. Des Weiteren lassen sich je nach Datenbank unterschiedliche Überwachungsfrequenzen festlegen. In der Regel kann der Suchlauf automatisch bei jedem Update einer Datenbank durchgeführt werden.
Interessenten für SDI-Dienste müssen nicht zwangsläufig auch Kunden eines Hosts sein. Beispielsweise bietet Questel mit dem PatLegal Monitor einen Profildienst für patentrechtliche Fragestellungen an, bei dem die Nutzer keinen Vertrag mit Questel abschließen müssen sondern eine Rechnung erhalten.
Da der Zeitfaktor bei diesem Neuerscheinungsdienst die wesentliche Rolle spielt, ist dieser Service nur dann sinnvoll, wenn sich auch der Datenbankinhalt immer auf dem neuesten Stand befindet und regelmäßig durch Updates aktualisiert wird.

3.7 Trefferanzeige und -weiterverarbeitung

Ein weiterer Mehrwert der meisten kostenpflichtigen Patentdatenbanken stellt die individuell konfigurierbare Trefferanzeige dar. Dafür stehen bspw. folgende Funktionen zur Verfügung: Auswahl der anzuzeigenden Felder, tabellarische Anzeige, farbliche Hervorhebung der Suchbegriffe in der Trefferumgebung oder die Sortierung nach eigenen Kriterien.
Der Rechercheprozess endet jedoch keineswegs mit einer Trefferliste. Die Resultate müssen je nach Arbeitsauftrag auf Relevanz geprüft, inhaltlich bewertet, aufbereitet und in eine präsentationsfähige Zusammenfassung gebracht werden. Insbesondere Hosts bieten dafür leistungsfähige Module in Ergänzung zu ihren Retrievalprogrammen an. Entsprechende Such- und Analysewerkzeuge, Sortier- und Auswahlfunktionen, mehrfarbiges Highlighting von Begriffen, Annotationen, etc. können automatisiert diese aufwändigen Nacharbeiten unterstützen.

Neben den Programmen für die direkte Weiterverarbeitung der Treffer gibt es zahlreiche weitere Anwendungsfelder und Zielgruppen für unterstützende Systeme. Sie können z.B. dem Portfolio-Management, der Verwaltung und Überwachung eigener Patente oder der Zitationsanalyse dienen. Bei der Auswahl zwischen all diesen Modulen und separaten Programmen gilt es vor allem einen wichtigen Punkt zu beachten: Die Rechercheergebnisse müssen verlustfrei und anwenderfreundlich von der Rechercheoberfläche in die Weiterverarbeitungssoftware exportiert werden können.

Nach der Recherche in bibliographischen Datenbanken sind oft Volltexte ausgewählter Patente gewünscht. Mehrere Datenbankanbieter stellen daher einen eigenen Lieferdienst für Volltexte von Patenten oder andere Dokumentenarten bereit. Beispiele dafür sind: DigiPat von Questel, FIZ AutoDoc sowie Thomson Patent Store.

3.8 Auswertungs-, Analyse- und Visualisierungswerkzeuge

Schon *einzelne* Patentdokumente enthalten wertvolle Informationen. Doch nicht auf alle Fragestellungen geben überschaubare Treffermengen die passenden Antworten. Sobald eine Vielzahl an Patenten das Ergebnis einer korrekten Recherche ist, sind Auswertungs-, Analyse- und Visualisierungswerkzeuge äußerst hilfreich, wenn nicht sogar manchmal die einzig sinnvolle Lösung.

Damit bspw. Unternehmen in Zeiten stetig wachsenden Konkurrenzdrucks, kurzen Produktlebenszyklen, Finanz- und Wirtschaftskrisen, Globalisierung usw. überleben können, sind sie mehr denn je auf handlungsorientiertes *aggregiertes* Wissen über ihre Markt- und Wettbewerbssituation – Competitive Intelligence – angewiesen.

„Unter Competitive Intelligence (CI) versteht man die kontinuierliche Recherche, Sammlung, Analyse und Aufbereitung von Wettbewerbsinformationen. Der Einsatz von Wettbewerbsbeobachtung und -analyse in Großunternehmen dient als Grundlage für Planungen und Entscheidungen in der Unternehmensstrategie, in Forschung und Entwicklung, bei Investitionen und im Marketing. Konkurrenzbeobachtung ermöglicht eine schnelle Reaktion auf Marktveränderungen zur Erhaltung der Wettbewerbsfähigkeit.“ Ein wichtiger Anteil der dafür notwendigen Informationen lässt sich aus den weltweiten Patentanmeldungen ablesen.
Bei klassischen Patentrecherchen erfolgt jedoch meist eine Suche nach einzelnen speziellen Patenten und ggf. die Nutzung der Analysefunktionen der Retrievalsprachen. Dies kann folgende Nachteile bedeuten:

- viele Fragestellungen können nicht beantwortet werden
- manche Fragestellungen können beantwortet werden, aber nur mit erhöhtem finanziellen und zeitlichen Aufwand
- für Analysen sind mehrere Rechercherchevorgänge notwendig
- die ausgegebenen Informationen sind i.d.R. nur für Rechercheure selbsterklärend, für Entscheidungsträger ist oft eine zeitraubende Aufbereitung notwendig.

Durch die Auswertung *vieler* Patentdokumente lassen sich zahlreiche weitere Informationen gewinnen. So wird u.a. die Beantwortung dieser Fragestellungen möglich:

- Markttendenzen, neue Technologiebereiche, Innovationen
- Wettbewerbssituation und Konkurrenzbeobachtung
- führende Erfinder/Forscher in Fachgebieten
- geographische Schwerpunkte der Anmeldungen je Technologie.

Als exemplarisches Beispiel für ein Analyse- und Visualisierungswerkzeug wird im Folgenden die Software STN AnaVist 2.0 für Nutzer von STN-Datenbanken vorgestellt.

Die wesentlichen Zielsetzungen dieses Werkzeugs bestehen in der Lieferung von aggregierten Entscheidungsgrundlagen für komplexe strategische Fragestellungen sowie dem identifizieren und sichtbar machen von ansonsten verborgenen Informationen und Zusammenhängen aus einer umfangreichen Treffermenge. Hierbei handelt es sich z.B. um Konkurrenten oder Marktführer und deren Patentstrategien bzw. -portfolio, Key Researcher oder auch Markttrends und Innovationen durch die Tendenz der Patentanmeldungen.
Die bedeutendsten Ansätze und Merkmale von STN AnaVist sind:

- professionelle Datengrundlage durch STN-Datenbanken, z.B. CAplus, DWPI, PCTFULL
- strukturierte Daten: Anwendung von Datamining, unstrukturierte Daten: Anwendung von Textmining
- umfangreiche Treffermenge einer Recherche kann nach individuell differenzierbaren Fragestellungen hin untersucht werden
- graphische Oberfläche: Erstellung und *kontextabhängige* Veränderung von Diagrammen und Graphiken per Mausklick → ermöglicht quasi interaktive Analyse und Visualisierung
- i.d.R. ist weitere Aufbereitung der Informationen nicht mehr nötig
- Visualisierungssets können bei den Kunden von mehreren Nutzern bearbeitet werden (sog. Shared Projects) → ergibt unterschiedliche Versionen mit ggf. weiteren Informationen, Zusammenhängen und Sichtweisen
- Details der Patentdokumente können im gleichen Programm angezeigt werden.

Bei der Datenanalyse erfolgt die Beantwortung der individuellen Fragestellungen quasi mit Hilfe einer interaktiven ‚Recherche durch Markieren'. Durch das Markieren von Datenfeldern/Balken in Diagrammen werden die Auswahlkriterien bzw. Bedingungen festgelegt. Als Ergebnis werden gleichzeitig all die Teilbereiche der Datenfelder anderer Diagramme mitmarkiert, die der Bedingung entsprechen. Jedes Markieren bedeutet im Prinzip einen separaten Schritt einer klassischen Patentrecherche. Es können bis zu acht unterschiedliche Bedingungen gleichzeitig festgelegt werden, z.B. für den Vergleich von Unternehmen oder Jahren.

Durch diese simultanen und zügig hintereinander durchführbaren Treffermarkierungen lassen sich in kurzer Zeit – und ohne Zusatzkosten – viele Fragestellungen ad hoc beantworten. Da für diese Abfragen eine graphische Oberfläche zur Verfügung steht, sind innerhalb eines vorbereiteten Visualisierungssets keine Retrievalkenntnisse notwendig. Dies erhöht den Nutzerkreis fertig erstellter Visualisierungssets erheblich.
Des Weiteren lassen sich mehrere Diagramme gleichzeitig anzeigen und verschiedene Auswahlkriterien mit unterschiedlichen Farben belegen. Somit werden Zusammenhänge zwischen sonst isoliert betrachteten Informationseinheiten – insbesondere im Zeitablauf – auf einen Blick sichtbar. Auch besteht die Möglichkeit auf Besonderheiten zu stoßen, die bei einem gezielten Retrieval nicht aufgefallen wären.
Die graphische Darstellung bietet außerdem den Vorteil, dass erst dadurch gewisse Fragestellungen beantwortet werden können. So lässt sich bspw. in einer Übersicht darstellen, welcher Erfinder für welches Unternehmen wie viele Patente einer IPC-Untergruppe angemeldet hat. Dabei fallen insbesondere die Erfinder auf, die im Zeitablauf oder parallel für unterschiedliche Unternehmen gearbeitet haben.

Ein Nachteil von STN AnaVist ist, dass sich die Aussagen zu Patentaktivitäten ausschließlich auf quantitative Auswertungen stützen. Aussagen zur Bedeutsamkeit oder Qualität der betrachteten Patente sind nicht möglich. Als Lösung wäre eine Ergänzung um Patentkennzahlen denkbar, die bspw. im System WISSPAT eingesetzt werden.
Des Weiteren gilt es bei der Beurteilung der Zahlen immer die allgemeinen Besonderheiten im Patentwesen zu berücksichtigen. Beispielsweise unterscheidet und verändert sich das Veröffentlichungsverhalten erheblich oder es werden nicht bei jedem Patent die Erfinder genannt.

3.9 Nutzerschulungen

Ein oft übersehener Mehrwert sind professionelle Nutzerschulungen. Kein Patentinformationssystem ist so benutzerfreundlich gestaltet und intuitiv bedienbar, als dass es ohne Erklärungen und Training auskommen könnte. Regelmäßige, umfangreiche, detaillierte und praxisnahe Schulungen können die Kundenzufriedenheit deutlich steigern, zu einer höheren Nutzungsfrequenz beitragen und die Kundenbindung stärken. Dabei sollte unbedingt darauf geachtet werden, dass sie von sehr versierten Patentrechercheuren durchgeführt werden. Ansonsten besteht die Gefahr, dass Kunden, die seit Jahren tagtäglich mit einer Datenbank arbeiten, mehr darüber wissen als der Trainer.

Nutzerschulungen bieten viel Potenzial für die Anpassung der Datenbanken und Dienstleistungen an die Bedürfnisse und Wünsche der Kunden. In kaum einer anderen Situation können Hosts und Datenbankhersteller ihren Nutzern bei der

Die richtige Information zu jeder Zeit an jedem Ort - weltweit. Das spart Zeit und Geld, Mitarbeiter und Material. Das ist Verwaltungsmodernisierung. Unsere zertifizierte Software strukturiert, automatisiert und sichert langfristigen Zugriff.

Zum Beispiel mit:

DoRIS
Elektronische Akte und Ad hoc Workflow - von der Registratur bis zum Leistungsbereich
www.registratur.com

HPR
Effiziente Planung und Berechnung von Pensionsrückstellungen und Beihilfen
www.pensionsystem.de

RENTE
Kompakte Vorgangsbearbeitung in Zusatzversorgungskassen und Versorgungswerken
www.rente3.de

HAESSLER
www.haessler.com

HAESSLER Information GmbH
Mühlstraße 58 | 75328 Schömberg
07235 9709 0 | info@haessler.com

Anwendung ihrer Produkte besser über die Schulter schauen und ihre Probleme erkennen.

3.10 Kundenbetreuung

Neben regelmäßigen Nutzerschulungen bieten gerade betreuende – und nicht akquirierende - Kundenkontakte eine gute Möglichkeit, die aktuellen Kundenwünsche und -bedürfnisse zu erfahren. Zu einer erfolgreichen Kundenbetreuung zählt hauptsächlich ein kompetentes und professionelles Helpdesk mit einer guten Erreichbarkeit sowie zügigen und zufriedenstellenden Antworten. All dies ist bei den kostenfreien Angeboten im Prinzip nicht vorhanden.

Eine gute Kundenbetreuung zeichnet sich auch durch ein umfangreiches und stets aktuelles Informationsangebot auf der Website aus. Dies kann z.B. Newsletter, Handbücher, ausführliche Dokumentationen der Datenbanken und Retrievalfunktionen sowie weitere Trainingsmaterialien beinhalten. Darüber hinaus bietet bspw. das FIZ Karlsruhe einen Recherche-Service in STN-Datenbanken an. Einige der Anlässe für die Nutzung dieses Angebotes können sein:

- kein eigener Zugang zu den Fachdatenbanken
- fehlendes Know-how oder fachfremde Recherchethemen für die eigene Experten fehlen (z.B. Sequenzrecherchen)
- zeitliche Engpässe
- Outsourcing.

An dieser Stelle sei ein kritischer Nebenaspekt erwähnt, der alle kostenfreien Patentdatenbanken mit unverschlüsselter HTTP-Verbindung betrifft – vermutlich also die Mehrheit. Aus den Suchwörtern, IPC-Klassen und sonstigen Angaben, die von den Nutzern in die Suchformulare eingetragen werden, lassen sich oft Rückschlüsse zu deren Forschungsaktivitäten oder sonstigen Interessen ziehen. Dies können durchaus wertvolle Informationen für Wettbewerber sein. Werden die Angaben nicht über verschlüsselte HTTPS-Verbindungen übertragen, wäre es Konkurrenten mit genügend krimineller Energie prinzipiell möglich, zu verfolgen wonach recherchiert wird.

Im Gegensatz dazu spielt Datenschutz und absolute Vertraulichkeit bei kommerziellen Anbietern i.d.R. eine große Rolle. Hosts wie bspw. Questel und STN bieten dementsprechend verschlüsselte HTTPS-Verbindungen für ihre Webzugänge oder eigene Kommunikationssoftware mit speziellen Verschlüsselungen an. Für den ebenso sehr sensiblen Bereich des Recherche-Service gelten strenge Selbstverpflichtungen. Auch dürfen die Suchanfragen weder gespeichert noch ausgewertet werden.

3.11 individuelle/flexible Preismodelle

So gut alle Mehrwerte auch sein mögen, letztendlich spielt der Preis die ausschlaggebende Rolle bei der Entscheidung für oder gegen einen Anbieter. Wichtig dabei ist, dass die Preismodelle zu den individuellen Nutzungsgewohnheiten passen. Wird ein Interessent wirklich zu einem Kunden, wenn bspw. eine Datenbank nur über ein jährliches Abonnement zu beziehen ist, der Interessent sie aber vermutlich nur vereinzelt nutzen wird und somit sehr hohe Durchschnittskosten je Recherche bezahlen müsste? Auf der anderen Seite könnten z.B. größere Einrichtungen mit mehreren Nutzern durchaus an pauschalen Gebühren interessiert sein, um somit eine bessere Kostenkontrolle und -planung zu erreichen.

Einige Datenbanken werden je nach Leistungsumfang in unterschiedlich teuren Versionen angeboten. Dadurch kann der Kunde wählen, wie viel Mehrwerte er benötigt und dementsprechend zahlen muss. Ein Beispiel hierfür ist der Derwent World Patents Index bei STN:

- WPINDEX: ohne Abonnement
- WPIDS: mit Abonnement
- WPIX: mit Abonnement und erweiterten Abstracts
- WPIFV: neu veröffentlichte Dokumente
- LWPI: nicht aktualisierte, kostengünstige Lerndatenbank.

In seltenen Fällen haben Hosts Datenbanken im Angebot, deren Inhalte zwar identisch sind, aber eine Wahl unter verschiedenen Preismodellen erlauben. So bietet STN bspw. die Datenbank CAplus mit zwei zusätzlichen Varianten an:

- CAplus: geringe Anschaltzeit-Gebühren plus Gebühren für Suchbegriffe; entspricht der Standard-Gebührenberechnung
- HCAplus: höhere Anschaltzeit-Gebühren aber keine Gebühren für Suchbegriffe; geeignet für Recherchen mit hoher Anzahl an Suchbegriffen
- ZCAplus: keine Anschaltzeit-Gebühren aber höhere Gebühren für Suchbegriffe; geeignet für zeitintensive Recherchen und Trefferbewertungen.

Eine ähnliche Flexibilität bieten auch die bei Hosts verwendeten Anzeige- und Druckformate. Je nach Umfang der ausgegebenen Informationen werden keine oder unterschiedlich hohe Gebühren erhoben.

In diesem entscheidenden Bereich ist von Kundenseite aus selbstverständlich eine höchstmögliche Kostentransparenz gewünscht. Hier sollte unterschieden werden zwischen Hosts und sonstigen Anbietern. Viele sonstige Anbieter von kostenpflichtigen Angeboten nennen keine definitiven Gebühren. Da es sich i.d.R. dabei um Abonnements handelt, werden Interessenten oft auf Kontaktformulare verwiesen, damit ihnen ein individuelles Angebot unterbreitet werden kann. Bei den Hosts ist im Regelfall die Preisgestaltung zwar deutlich differenzierter, aber dafür klar ersichtlich. So bieten bspw. Questel und STN auf ihren Webseiten detaillierte Preislisten für ihre Dienste an.

Fazit

Der Zugang zu Patentinformationen ist immer leichter geworden, aber diese Informationen zu verstehen und richtig zu interpretieren ist immer schwieriger geworden. Wissen die Nutzer, was sie nicht gefunden haben bzw. was ihnen fehlt?
Die Verwendung von Patentsuchmaschinen ist nur eingeschränkt zu empfehlen. Da die Abdeckung dieser Angebote im Regelfall recht gering und oft auf US-Patente fokussiert ist, kann bei weitem keine Vollständigkeit der Recherche erzielt werden. Ebenso sind die Recherchemöglichkeiten i.d.R. sehr eingeschränkt. Es fehlen u.a. zahlreiche Suchfelder, Trunkierungsoptionen, das Blättern in Indexen, Verknüpfungen von Suchschritten, usw. Eine gezielte Suche mit übersichtlichen Treffermengen ist daher nicht möglich. Das Ranking der Ergebnisse ist schwer nachvollziehbar und eine Sortierungsfunktion fehlt meistens. Sie bieten sich allerdings an, wenn erste Vorrecherchen benötigt werden, Teilangaben von Patenten bekannt sind oder Volltexte gefragt sind.

Die im Web frei verfügbaren Datenbanken der Patentämter haben sich i.d.R. zu wertvollen Informationsressourcen, mit z. T. umfangreichem oder vereinzelt auch einmaligem Datenbestand, entwickelt. Jedoch sind die Unterschiede in Bezug auf Leistungsumfang, Qualität und Benutzerfreundlichkeit teilweise erheblich. Ihnen fehlt es hauptsächlich an erweiterten Suchfunktionalitäten, Mehrwertdaten sowie Analysewerkzeugen. Die fortschreitenden Aktivitäten der Patentämter dürfen den Markt der Patentinformationen aber nicht aus dem Gleichgewicht bringen. Es müssen noch genügend Möglichkeiten für Datenbankhersteller bestehen ihre Mehrwertangebote fortzusetzen. Daher sind Mehrwertangebote wie bspw. Analyse- und Visualisierungstools immer notwendiger und eine Chance für kommerzielle Anbieter.
Diese müssen aber unter allen Umständen vermeiden, ihre Gebühren weiter anzuheben. Ansonsten schädigen sie

sich früher oder später selber. Denn tendenziell gesehen sinkt, insbesondere bei vielen Budgetverantwortlichen, die Einsicht für Informationen überhaupt etwas zu bezahlen. Wenn die Informationen dann auch noch teurer werden, ist der Schritt hin zur Abbestellung der kostenpflichtigen Angebote nicht mehr weit. Vergleichbar mit der Zeitschriftenkrise setzt sich dann die Spirale weiterer Preiserhöhungen in Gang. Ein geringerer Absatz bedeutet für den Anbieter, dass er seine Kosten auf immer weniger Kunden verteilen muss, was zwangsläufig zu einer Erhöhung der Einzelpreise führt. Um diese Entwicklung zu kompensieren, versuchen die Anwender bei nutzungsabhängigen Preismodellen die Datenbanken dementsprechend weniger zu verwenden. Dies führt jedoch außerdem zu einer Verringerung der Qualität ihrer Dienstleistungen, sowohl für die internen als auch für die externen Kunden.

Bei der Kosten- und Preisdiskussion darf ein Punkt nicht unberücksichtigt bleiben: Auch die Herstellung der, für die Anwender frei verfügbaren, Datenbanken der Patentämter verursacht bei den Behörden z. T. enorme Kosten. Beispielsweise ist das Europäische Patentamt in der Lage ausreichende Ressourcen für ihr hochwertiges Angebot esp@cenet aufzubringen. Doch woher kommen diese finanziellen Mittel? Sie stammen von den Gebühren der Anmelder. Würde das EPA stattdessen Nutzungsentgelte für esp@cenet verlangen müssen, wären diese sicherlich mit denen kommerzieller Anbieter vergleichbar.
Vielfach werden bei Kosten-Nutzen-Abwägungen fälschlicherweise nur die Kosten für die Datenbanknutzung berücksichtigt. Korrekterweise sollten aber auch die Personalkosten hinzugezogen werden. Bei den kostenfreien Angeboten ist im Regelfall ein wesentlich höherer Zeitbedarf für die Relevanzprüfung und Weiterverarbeitung der Treffer notwendig. Werden nun diese Zeiten und die damit entstehenden Personalkosten mit in die Rechnung einbezogen, können sich die Gesamtkosten beider Angebotsarten durchaus einander angleichen.
Für die Beantwortung der entscheidenden Frage ob kostenpflichtige Mehrwertdatenbanken notwendig sind, muss zunächst eine Gegenfrage gestellt werden: Wofür werden die Informationen benötigt? Bei der ausschließlichen Recherche in kostenfreien First-Level-Daten besteht vor allem die große Gefahr, dass nicht alle relevanten Patentdokumente gefunden werden. Sobald die benötigten Informationen als Grundlage für geschäftspolitische, patentrechtliche oder forschungsrelevante Entscheidungen mit weit reichenden Konsequenzen dienen, ist ein solches Vorgehen nicht ausreichend. Bei professionellen Recherchen muss jedoch davon abgeraten werden, sich ausschließlich auf die Datenbanken mit Mehrwertdaten zu verlassen. Für entscheidungsrelevante Recherchen mit hoher wirtschaftlicher Bedeutung ist daher folgendes Vorgehen zu empfehlen: Recherche in möglichst allen relevanten Mehrwertdatenbanken und die anschließende Vervollständigung der Treffer mit First-Level-Daten – kostenpflichtige und kostenfreie. Wie erläutert kann es nämlich durchaus vorkommen, dass relevante Patente zum Zeitpunkt der Recherche ausschließlich in kostenfreien Beständen recherchierbar sind.

Kostenfreie Patentdatenbanken eignen sich verstärkt für folgende Anwendungsfälle:

- einfache Recherchen, z.B. wenn Patentnummer oder Erfinder bekannt sind
- erster Einstieg in neues Fachgebiet
- Vorrecherche zu kostenpflichtigen Abfragen
- Vervollständigung von Recherchen in Mehrwertdatenbanken
- Download ausgewählter Dokumente deren Relevanz bereits recherchiert wurde
- Recherchen außerhalb eines Beschäftigungsverhältnisses, da hier keine Personalkosten für den erhöhten Zeitbedarf anfallen.

Mehrwertangebote hingegen bieten im Allgemeinen eine sehr zuverlässige Suche mit vielfältigsten Retrievalfunktionen, einen hohen Abdeckungsgrad, einwandfreie korrekte Daten mit zusätzlichen Indexierungen sowie hilfreiche Werkzeuge zur Weiterverarbeitung und Analyse der Treffer. Sie bieten sich insbesondere dann an, wenn zuverlässige und umfassende Informationen benötigt werden.
Um sich gegenüber den kostenfreien Patentdatenbanken weiter behaupten zu können, müssen die kommerziellen Hersteller die Qualität und Verlässlichkeit sowie den Abdeckungsgrad ihrer Produkte auch weiterhin verbessern. Denn dies sind bedeutende Gründe für Kunden in kostenpflichtige Informationen zu investieren.
Insbesondere im Bereich der Indexierung gibt es z. T. noch Verbesserungspotenzial. Manche der Personen, die die intellektuelle Indexierung vornehmen, fehlt es an Wissen und Erfahrung bei der Patentrecherche, sodass ihre Eingaben nicht immer alle relevanten Informationen, wie z.B. Codes, abdecken. Doch gerade durch die manuelle Indexierung besteht für die kommerziellen Produzenten die Chance, sich erfolgreich von der großen Masse der frei verfügbaren Volltextdatenbanken abzusetzen.
Des Weiteren dürfte es sicher ratsam sein, weiter verstärkt auf die Kunden-

B.I.T. online
NEUERSCHEINUNGEN

Regina Pfeifenberger
Pocket Library
Band 27:
ISBN 978-3-934997-30-1, 2010
Brosch., 112 Seiten € 24,50

Ann Christine Marr
Serious Games
Band 28:
ISBN 978-3-934997-31-8, 2010
Brosch. 164 Seiten, € 24,50

Ursula Georgy
Erfolg durch Innovation
Band 29:
ISBN 978-3-934997-32-5, 2010
Brosch., 170 Seiten, € 24,50

www.b-i-t-online.de

wünsche und -bedürfnisse einzugehen. Somit könnten neue Mehrwerte generiert werden, die in der täglichen Praxis zu einer effizienteren sowie effektiveren Nutzung führen und damit die Kundenzufriedenheit steigern. Zu einer dieser Zukunftsthemen kann die semantische Suche werden - durchaus auch im Bereich der Patentinformationen. So könnte sie bei der allgemein formulierten und schwer verständlichen Patentsprache bspw. Hilfestellungen mit Vorschlägen zu Synonymen geben.

Erst die Summe und das Zusammenspiel der genannten Mehrwerte zeichnen sehr gute Patentdatenbanken aus. Es ist z.B. von geringem Nutzen, wenn eine Datenbank professionelle Retrievalfunktionalitäten bietet, aber der Datenbestand nicht auf dem aktuellsten Stand ist oder eine zu geringe Abdeckung aufweist.

Das Feld der Patentinformationen ist äußerst heterogen. Keine Datenbank wird je alle patentrelevanten Daten, Volltexte usw. in sich vereinen können. Dies wäre auch nicht erstrebenswert und sinnvoll. Daher wird sich für die Datenbankhersteller wohl auch in Zukunft die Frage stellen: entweder Spezialisierung auf ein Fachgebiet mit spezieller Kernkompetenz als Alleinstellungsmerkmal oder eine möglichst breite Abdeckung um vielen Kundenwünschen und Fragestellungen nachkommen zu können?

Vermutlich kennen die Nutzer kostenfreier Dienste und die Interessenten von kommerziellen Angeboten meist nur einen geringen Teil der möglichen Mehrwerte von kostenpflichtigen Patentdatenbanken. Viele Vorteile sind nicht direkt als solche erkennbar, wie bspw. die Aktivitäten im Bereich der Fehlerbereinigung, oder tauchen selten in einer Datenbankbeschreibung explizit auf. Wenn die Mehrwerte nicht bekannt sind, ist verständlicherweise auch die Bereitschaft, für aufbereitete Patentinformationen entsprechend zu bezahlen, nicht vorhanden. Eine denkbare Lösung wäre daher eine offensivere Produktpräsentation und ein Marketing, bei dem die detaillierte Nennung der konkreten Mehrwerte im Vordergrund steht.

Im schlimmsten Fall führen falsche oder lückenhafte Patentrecherchen zu kostspieligen Erfahrungen oder weit reichenden Konsequenzen. Muss es wirklich erst soweit kommen um die Einsicht zu erlangen, dass kostenpflichtige Patentdatenbanken mit Mehrwerten notwendig sind? Die Nutzer von kostenfreien Patentdatenbanken mögen zwar meist mit ihren Suchergebnissen zufrieden sein. Wer aber erst einmal die Datenbestände, Möglichkeiten und Mehrwerte von professionellen und qualitativ hochwertigen Recherchesystemen kennen gelernt hat, möchte diese in Zukunft nicht mehr missen.

Mehrwertdienst, Patentdatenbank, Suchmaschine, Patentdokumentation, Host, Information Retrieval, Fachgebietsabdeckung, Qualität, Kosten, Indexierungsverfahren, STN International, Derwent

Der Autor

Markus Birkner

ist gelernter Industriekaufmann und staatl. gepr. Betriebswirt. Ergänzend dazu studierte er Bibliotheks- und Informationsmanagement (B.A.) an der Hochschule der Medien in Stuttgart. Seit dem Abschluss im Jahr 2009 ist er beim Fraunhofer-Informationszentrum Raum und Bau IRB im Bereich der Volltextanbindung tätig.
markus-birkner@gmx.de

Personalien

Harald Haendler †

Dr. agr. Harald Haendler, seit 1969 Mitglied der DGI, ist am 18. April 2010 im Alter von 88 Jahren verstorben. In seinem Geburtsort Kiel absolvierte Harald Haendler nach der Schule eine landwirtschaftliche Lehre. Nach Kriegsdienst und Gefangenschaft sowie praktischen landwirtschaftlichen Tätigkeiten besuchte er 1946 die Höhere Landbauschule in Schleswig, die er als staatlich geprüfter Landwirt abschloss. 1947 nahm Harald Haendler ein Studium der Landwirtschaft an der Universität Kiel auf, erwarb 1949 das Diplom und wurde 1952 mit einer Arbeit zu Qualitätsfragen von Gärfutter promoviert. Dort arbeitete er von 1949 bis 1950 im Tierzuchtinstitut und war dann für zwei Jahre bei der Bundesforschungsanstalt für Milchwirtschaft in Kiel beschäftigt. Anschließend nahm Harald Haendler bei verschiedenen Unternehmen in der Industrie leitende Positionen ein.

1961 kam Harald Haendler nach Hohenheim, wo er zunächst das dem Institut für Tierernährungslehre angegliederte Archiv für Futtermittel leitete und das „Hohenheimer Klassifikationssystem für Futtermittel entwickelte. 1963 wurde die Dokumentationsstelle der Universität Hohenheim gegründet, deren Leiter er bis zu seinem Eintritt in den Ruhestand 1986 war. Aufgabe der Dokumentationsstelle war die systematische Literaturdokumentation auf dem Gebiet der Tierischen Produktion sowie der Bereiche Agrargeschichte, Beratung und Ausbildung. Die Geschichte dieser seinerzeit größten und bedeutendsten agrarwissenschaftlichen Dokumentationsstelle Deutschlands beschrieb er abschließend in einem Beitrag in den Mitteilungen der Gesellschaft für Bibliothekswesen und Dokumentation des Landbaus dar, deren Vorsitz er von 1970 bis 1978 innehatte.

Lehraufträge am Lehrinstitut für Dokumentationswesen der DGD, an der Universität Stuttgart und an der Fachhochschule für Bibliothekswesen Stuttgart ermöglichten ihm, sein umfassendes Wissen weiterzugeben. Von seinen umfassenden Spezialkenntnissen zeugen die fast 200 Beiträge zur Agrardokumentation und seine Tätigkeit in verschiedenen internationalen Gremien. So war er ab 1970 im Executive Committee und von 1975 bis 1985 Vizepräsident der International Association of Agricultural Librarians and Documentalists (IAALD), die ihm 1985 die Ehrenmitgliedschaft verlieh. In Würdigung seiner nationalen und internationalen Verdienste um die Agrardokumentation wurde ihm 1987 das Verdienstkreuz am Bande der Bundesrepublik Deutschland verliehen, nachdem er von 1983 bis 1986 im Rahmen des Bund-Länder-Verwaltungsabkommens über die Agrardokumentation Erster Vorsitzender des „Fachbeirats Literatur“ des FIS Ernährung, Landwirtschaft und Forsten war und bereits zuvor beim Planungsprozess im Rahmen des IuD-Programms im Sachverständigenkreis für das FIS mitgewirkt hatte.

Die DGI wird Herrn Dr. Harald Haendler ein ehrendes Andenken bewahren.

Aus der Praxis der Patentinformation

Die Entwicklung der elektronischen Medien und Dienstleistungen bei den Patentbehörden und Internetprovidern

Dieter Geiß, Castrop-Rauxel

Die 37. Arbeitssitzung der AGM (Arbeitsgruppe Elektronische Medien in der Patentinformation) beim Technischen Informationszentrum des Deutschen Patent- und Markenamtes in Berlin am 12. November 2009 stand wie in jedem Jahr unter dem Motto „Fortschritte in der Patentinformation". Die Arbeitsgruppe besteht seit 37 Jahren und hat inzwischen rd. 80 Mitglieder aus Wirtschaft und Behörden, die bei Bedarf regelmäßig informiert werden. In diesem Jahr haben wieder 38 Fachleute den Weg nach Berlin gefunden, um sich mit Fragen und Problemen der Patentinformation auseinanderzusetzen.

Weitere Lücken in den digitalen Datenbeständen wurden geschlossen und jeder Provider trägt Sorge dafür, dass benutzerfreundliche Recherchewerkzeuge angepasst und ergänzt werden. Es ist sehr erstaunlich, aber auch heute ist noch kein Ende dieser Entwicklung abzusehen. Der Informationssuchende ist ohne sich ständig weiterzubilden auf verlorenem Posten und hat ständig gewissenhaft zu prüfen, welche der breit gefächerten Dienstleistungsangebote für ihn und für seine Firma tatsächlich in Betracht kommen und zum wirtschaftlichen Erfolg führen.

An dieser Stelle, wie in jedem Jahr, ein Appell an die Nutzer: zögern Sie nicht und informieren Sie die Ämter bzw. Provider, wenn sie bei Ihrer Arbeit Fehler oder Ungereimtheiten bei den Rechercheergebnissen oder Softwaretools feststellen sollten. Nur durch eine aufmerksame Mitarbeit ist in Zukunft sicherzustellen, dass in der Welt der komplexen Datenstrukturen Fehler beseitigt und benutzerfreundliche Lösungen gefunden und erarbeitet werden können.

Die Organisatoren der AGM Arbeitsgruppe: von links Dr. Wissmann DPMA, Herr Lingscheid BBS, Frau Franke DPMA und Herr Jagstorff BBS (neuer Vorsitzender der AGM).

Patentbehörden

Europäisches Patentamt

Dienstleistungen des EPA

Die Patentinformationsdienste des Europäischen Patentamtes wurden den Bedürfnissen der Nutzer und der technischen Entwicklung angepasst. Die ESPACE Serien wie die CD-ROM bzw. DVD-Serien ESPACE ACCESS, ESPACE ACCESS EPC, ESPACE FIRST und ESPACE WORLD werden nicht mehr produziert.

Als Ersatz wird eine neue ESPACE GPI aufgelegt, die in Verbindung mit anderen Online-Diensten der EPO/WIPO (esp@cenet/PATENTSCOPE) genutzt werden kann. Für die Patentinformationszentren sind Überbrückungsmaßnahmen angekündigt.

Dieser neu aufgelegte GPI Index stellt DOCDB Daten für Fortgeschrittene bereit. Der Dienst enthält bibliographische Daten von über 75 Patentämtern, dazu besser aufbereitete Daten (ca. 100 Suchkriterien), eine einfache Datenextraktion und -weiterverarbeitung, eine einfachere Handhabung, die lokale Nutzung oder Nutzung über Internet und vielfältige Suchmöglichkeiten. Hinzu kommen Zitierungen der Patente und der Nichtpatente, dazu die Kategorie der Zitierungen, Klassifikationsangaben nach IPC1-8, ECLA, ICO und nationale Klassifikationssystemen, einfache Patentfamilien (inkl. Angabe des repräsentativen Familienmitgliedes), Benennungen (PCT, EPC, EPC Erstreckung) und die Suche mit „Näherungsoperatoren" in Titeln und Zusammenfassungen.
Alle DOCDB-Bibliographiedaten sind in der Datenbank indexiert. Der Zugriff kann zur Laufzeit auf OPS(Open Patent Service)-Daten erfolgen, wie auch auf die repräsentativen Abbildungen, die INPADOC-Patentfamilien und den INPADOC Rechtsstand.

In Abbildung 1 sind die Zusammenhänge erkennbar, wie über das GPI Servlet die Verbindung zu dem OPSV2 Server hergestellt werden kann. Dieses Servlet kann auch dazu genutzt werden über Mimosa Flex oder andere Internet basierte Produkte zu recherchieren. Es stellt eine einfache HTTP-Schnittstelle in Mimosa selbst bereit, ohne ein SOAP (Simple Object Access Protocol) nutzen zu müssen.
Über das GPI-Suchfenster kann in drei Indexgruppen (leichte, einfache und detaillierte Suche) in über 100 Einzelindices recherchiert werden. Trefferlisten und Suchprofile können gespeichert und für die Überwachung von Wettbewerbern bereitgestellt werden. Über Links wer-

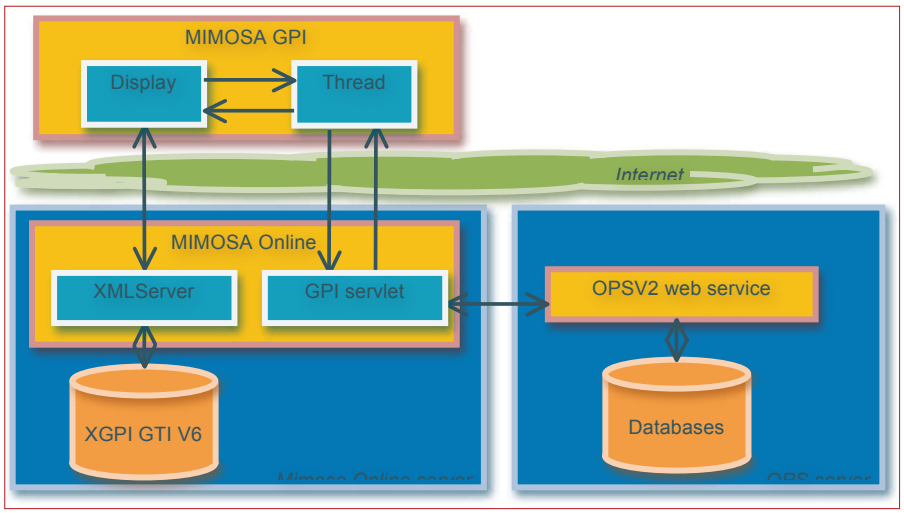


Abbildung 1: EPA Nutzung von OPS V2 in MIMOSA GPL.

den Volltextdokumente in esp@cenet aufgerufen oder über OPS die Familien und Rechtsstandsdaten. Der Dienst ermöglicht die Benutzung eigener Statistikwerkzeuge zur Erstellung maßgeschneiderter Statistiken. Er wird wöchentlich aktualisiert.

Mimosa Flex (Abb. 2) ist die nächste Generation der Benutzeroberfläche des Amtes. Die wichtigsten Attribute sind die vollständig integrierte Internet-Anwendung, der modulare Aufbau und je nach Bedarf die Nachladung benötigter Module aus dem Netz. Dazu lassen sich an Spezialaufgaben angepasste Module laden für z. B. statistische Auswertungen oder für eine Visualisierung der Ergebnisse. Obwohl Mimosa Flex als Internetanwendung konzipiert, ist der Zugriff auf lokale Daten, auf Festplatte, CD-/DVD-ROM oder USB-Gerät möglich. Die Anwendung ist völlig in vorhandene Websites integrierbar. Eine bewusste Installation durch Benutzer ist daher nicht notwendig.
Die Website des Publikationsservers wird noch überarbeitet.

Die weltweite **Rechtsstandsdatenbank** des EPA (INPADOC PRS) umfasst gegenwärtig Daten von 54 Ländern und Organisationen mit Datensätzen über 72 Millionen Rechtstandsereignisse für ungefähr 19 Millionen Patentanmeldungen.
Das monatliche Wachstum beträgt ca. 2,3 Millionen Ereignisse bei annähernd 560.000 zusätzlichen Patentanmeldungen.
Die wichtigsten Entwicklungen bei den Rechtsstandsdaten sind die Erweiterungen der Abdeckung im Jahr 2009, so um die neuen Länder China und Russland. Es erfolgte auch eine Verbesserung der bestehenden Sammlungen wie z.B. bei den US Assignments, von denen ungefähr 3,6 Millionen „Assignment Events" 2009 in die Rechtsstandsdatenbank geladen wurden.
Die chinesischen Daten wurden ab der 18. Woche 2009 als Frontfile in die Datenbank PRS geladen. Die aktuellen Daten werden jeden Monat verarbeitet. Das Laden des gesamten Altbestandes (Backfile) der chinesischen Rechtsstandsdaten wurde in der 35. Woche 2009 abgeschlossen.
Die russischen Rechtsstandsdaten wurden ab der 29. Woche 2009 für russische Patente und Gebrauchsmuster in die Datenbank aufgenommen. Gegenwärtig sind nur aktuelle Daten (Frontfile) verfügbar. Das Laden der Daten begann mit den Mitteilungen des russischen Patentblattes vom 10. April 2009. Daten noch gültiger Patente und Gebrauchsmuster der ehemaligen Sowjetunion wurden auch aufgenommen.
Das EPA bemüht sich verstärkt um Rechtsstandsdaten von allen EPÜ Staaten, die über die derzeit verfügbare „post-grant" Information hinaus zu erhalten sind, sowie von den lateinamerikanischen Staaten, aus Japan und Indien.
Der Umzug von **Register Plus** erfolgte nach https://register.epoline.org. Verbessert wurden der sicherere Zugang (ohne Kennung und SmartCard), die Navigation, die Benutzerstatistiken, die Benutzerunterstützung, die Verwendung von „Bookmarks" für Recherchen und zusätzliche Funktionen, die unter dem Portal bisher nicht integrierbar waren.
Das WebRegMT verbleibt im Portal. Bei der Neugestaltung von Register Plus wurde das Datumsformat „dd.mm.yyyy" geändert. Die Titel sind inzwischen in allen drei Amtssprachen verfügbar. Neue Datenfelder wurden eingeführt für BIOM: „Biologisches Material: In der Anmeldung wird die Hinterlegung von biologischem Material erwähnt, Akteneinsicht für Einzelheiten", für die Felder REJR01, REJR02: „Zurückweisung des Antrags auf Widerruf des Patents" und für das Feld ORAL01: „Absetzung einer mündlichen Verhandlung". Eine SmartSearch ist auch in Register Plus möglich.
Ferner sind geplant: die Einführung eines Formulars zum Melden von Datenfehlern, die Verbesserung der Suchfunktionalität so z. B. die Anwendung von „Fuzzy Logik" in der Nummernsuche, die Verbesserung des Dokumentendruckes und die Verbesserung der Filterfunktion, z. B. durch die Zusatzfelder OPPO1, OPPO2 und andere.

In **esp@cenet** ist zunächst eine verbesserte Navigation in den Hilfetexten vorgesehen. Dazu kommt die Anzeige des Datenumfangs bei Auswahl einer Datenbank, die Anzeige der PCT Veröffentlichungsnummer „Also published as", eine verbesserte Nummernsuche, die Anzeige aller Veröffentlichungsstufen in der Trefferliste, mehr Suchbegriffe je Suchfeld und größere Suchfelder, die Nutzung der ECLA anhand der Bilder auf der ersten Seite, die Anzeige aller Zitate, Einführung von Parsern für die Übertragung von Teilen der Ansprüche auf eine andere Website, die farbliche Hervorhe-

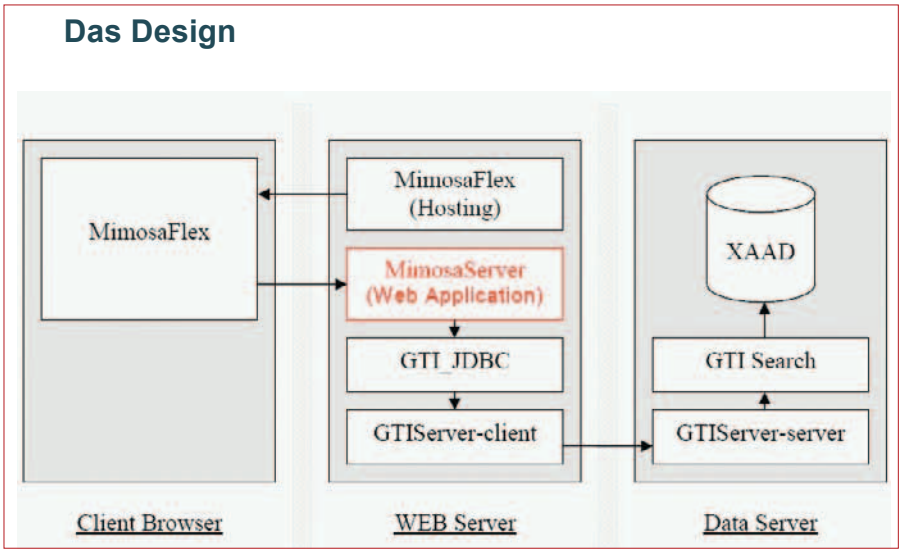


Abbildung 2: EPA Mimosa Flex Design.

bung der Suchbegriffe und Sortierfunktionen, die Volltextsuche und RSS Feeds. Alle diese Verbesserungen stehen seit November 2009 zur Verfügung.

Deutsches Patent- und Markenamt

Schutzrechtsanmeldungen

Die Anmeldung von Schutzrechten kann bekanntlich schon seit längerer Zeit online über DPMAdirekt erfolgen und zwar für Patente (deutsche, europäische und PCT-Anmeldungen), Marken, Gebrauchsmuster, Einspruchs- und Beschwerdeverfahren in Patentsachen, Beschwerden in Markensachen und Einzugsermächtigungen. Geschmacksmusteranmeldungen werden voraussichtlich erst ab April 2010 angenommen, da noch auf die Verabschiedung einer Rechtsverordnung „ERVDPMAV“ gewartet werden muss.
Die Vorteile für Anmelder sind bekannt. Die neue Gebührenordnung gilt seit dem 1. Oktober 2009. Eine elektronische Anmeldung kostet demnach mit bis zu zehn Ansprüchen 40,- EUR, ab dem 11. Anspruch zzgl. je 20,- EUR je Anspruch.
Seit Oktober 2009 gibt es eine neue Benutzeroberfläche mit der Demo-Version einer Geschmacksmusteranmeldung mit dem direkten Aufruf aller Funktionen (Erstellen, Editieren, Signieren und Versenden der Vorgänge sowie das Empfangen der Anmeldebenachrichtigung) und dem Speichern von Adressen in der integrierten Datenbank.
Bekanntlich ist eine qualifizierte, personenbezogene Signaturkarte (gesetzliche Anforderung) notwendig, die bei allen allgemein verbindlichen Rechtsgeschäften zum Einsatz kommt und zu einer sicheren Kommunikation in Deutschland führt. Die Liste aller Anbieter ist unter www.bundesnetzagentur.de zu finden. Die Signaturanbieter müssen eine verbindliche Überprüfung (bis zu 30 Jahre) der Signatur garantieren.
Die DPMAdirekt Statistik zeigt, dass elektronisch 20 Prozent mehr Patente, fünf Prozent mehr Marken und sieben Prozent mehr Gebrauchsmuster als im Vorjahr (hochgerechnet auf das ganze Jahr) angemeldet wurden.

DEPATISnet

Folgende Verbesserungen konnten bei DEPATISnet eingeführt werden. Seit dem 1. Oktober 2009 gibt es keine „Session-ID“ mehr. Es werden eigene URL-Adressen pro Treffer eingeführt, jeweils eine für die bibliographischen Daten und die PDF Anzeige. Damit kann die Trefferliste herunter geladen werden.
Die .TXT-Dateien enthalten nun ebenfalls URL-Adressen. Damit ist ein direkter Aufruf über Hyperlinks zu DEPATISnet möglich. Es erfolgt auch eine Markierung des letzten Treffers in der Trefferliste (nach PDF-Anzeige). Die Trefferlisten wurden ebenfalls neu konfiguriert. Eine neue Recherche beginnt mit der Übernahme der Einstellungen der letzten Recherchesitzung (Cookies). Bei der Nutzung der Prüfstoff-IPC wird die Anzeige in den „Bibliographischen Daten“ mit Link zur IPC vorgenommen und bei der Expertenrecherche wird die Auswahl der verfügbaren Felder systematisch sortiert.
Inzwischen ist der Datenbestand von DEPATISnet auf 67.221.211 Datensätze angewachsen (39.528.196 PDF-Sätze und 27.693.015 bibliographische Datensätze).
Die weiteren Planungen für das Jahr 2010 sehen vor: das Highlighting des Suchbegriffs in einem Treffer, die Ablage neuer DE-Dokumente im DEPATIS-Archiv nach neuem PDF/A-Standard (in zeichencodierter Form), den Ausbau der barrierefreien Volltextanzeige, die Einführung einer dritten Anzeigemöglichkeit im XML-Format zunächst für DE-Dokumente ab 1987, die Umwandlung der XML-Daten „on the fly“ in PDF-Dokumente, lesbar für Screen Reader-Software („Barrierefreiheit“), und die Ausnutzung der Acrobat Reader-Funktionen für einen geringeren Speicherplatz.

DPMAregister Patente

Bekanntlich soll mit der Zusammenführung von DPINFO und DPMApublikationen das DPMAregister geschaffen werden (siehe auch den Zeitplan der Projekte in Abbildung 3). Die Vorzüge liegen auf der Hand. Es gibt eine tägliche Aktualisierung, eine Barrierefreiheit durch entsprechende Styleguides und sehr gute Recherchefunktionalität, wobei die Daten identisch mit den hausinternen „Mastersystemen“ sind.
Bei der Einführung des DPMAregister wird eine Überarbeitung der Recherchemodi vorgenommen. Der Einsteigermodus wurde gestrafft, die komplexe Expertenrecherche modifiziert. Neu eingeführt wurden die Schnellsuche und das Monitoring. Ferner konnte eine übersichtlichere Darstellung der Verfahrenshistorie bereitgestellt werden. Die Gesamttrefferanzahl der Recherche wurde auf 1000 heraufgesetzt, die Auswahl der Treffer 10, 50, 500 oder 1000 wurde vereinfacht. Ein Registerauszug im PDF-Download mit DPMA-Kopf ist verfügbar. Der Lieferrhythmus von DPMAkurier kann „täglich“ erfolgen.

DPMAregister Marken

Das DPMAregister für Marken sieht mit der Recherchierbarkeit von bisher schwer auffindbaren Markentexten (gesperrte Schreibweise, Sonderzeichen etc.) eine Verbesserung der Suchstrategien vor. Die Darstellung der Marke mit Bild erfolgt in der Trefferliste und der Markentext im Tooltip. Die Suche der Leitklasse oder Klasse erfolgt in einem Feld. Die zurückgenommenen oder zurückgewiesenen Markenanmeldungen werden in einer Trefferliste mit dem Zusatz versehen: Aktenzustand „Eintragung nicht möglich“.
Die Registerauskunft zu einem Aktenzeichen enthält den aktuellen Stand des Registers, die Verfahrensdaten und, wenn gewünscht, auch die Detailansicht zu einzelnen Verfahren mit Datum des Eintrags in der Datenbank (EDVT) und der Veröffentlichung im Marken- und Geschmacksmusterblatt unter Auskunft zu einer Markenanmeldung.
Festzuhalten ist, dass bis zu einer gesetzlichen Neuregelung das elektronische „Blatt“ im PDF weiterhin die amtliche Publikation bleibt. Die neue Monitoring-Recherche erfolgt über eine eigene Suchmaske.

DPMAregister Geschmacksmuster

Das DPMAregister Geschmacksmuster wurde neu gestaltet. Es wurde eine musterorientierte Darstellung mit eigenem Aktenzeichen pro Muster gewählt. Die

Zeitplan	04/2009	12/2009	Mitte 2010	Ende 2010	ab Ende 2011/2012
Marke	DPMAregister				
GSM	DPMApublikationen + DPINFO		DPMAregister vsl. ab 28.01.2010		
Patent, Gbm	DPMApublikationen + DPINFO			DPMAregister	
Akteneinsicht, online Patent,	Akteneinsicht konventionell in den Recherchesälen oder über DPMA-Kopierservice erhältlich				DPMA register
EU-Marken EU-GSM	CTM-Online RCD-Online				DPMA register
Int. Marken (DE) Int. GSM (DE)	MadridExpress HagueExpress				DPMA register

** ELSA Marken vsl. ab 2013*

Abbildung 3: DPMA Zeitplan für die Einführung von DPMA Register.

Bildwiedergabe wird in der Trefferliste angezeigt. Die Trefferliste ist konfigurierbar als Bildergalerie mit Pop-Up-Fenster, Matrix und Tabellen.
Die zweite Projektphase sollte am 28. Januar 2010 starten.

World Intellectual Property Organisation

Die Reform der IPC

Die World Intellectual Property Organisation arbeitet ständig an der Verbesserung der internationalen Patentklassifikation (siehe auch den Report CER/41 unter www.wipo.int/meetings/en/details.jsp?meeting_id=17524). Der Core Level wurde inzwischen die Hauptgruppenebene. Die nationalen Patentämter können wählen, ob sie nach der Hauptgruppenebene oder der Unterklassenebene klassifizieren. Die Bezeichnungen „advanced level“ und „core level“ wurden abgeschafft.
Die 5. Version der IPC erlaubt einen Blick in die Zukunft der Klassifikation. Die Arbeitsgruppe ist um China und Südkorea erweitert worden, so dass nicht mehr von einem trilateralen Gremium gesprochen werden kann. Entwickelt wird eine Hybridklassifikation, die in die IPC eingearbeitet wird. Im Wesentlichen soll sich die neue Version an der ECLA ausrichten mit eventueller Übernahme bestimmter F-Terms der japanischen Klassifikation.
Der Arbeitsplan der Arbeitsgruppe sieht vor, dass die IPC-Version am 1. Januar 2011 nur noch eine Klassifikationsebene enthält. Das neue Design der IPC im Internet (siehe auch www.wipo.int/edocs/mdocs/classifications/en/ipc_ce_41/ipc_ce_41_www_118494.ppt) sollte ebenfalls am 1. Januar 2011 vorliegen mit zusätzlichen Reklassifizierungswarnungen, mit Hinweisen, wo die Reklassifikation noch nicht abgeschlossen werden konnte, mit neuen Suchbegriffen und der Möglichkeit der „cross reference search“. Die Warnhinweise werden für die Ämter obligatorisch.

Patentscope

Die Recherchеplattform Patentscope gilt nur für PCT-Veröffentlichungen. Rechercheerleichterungen werden durch das Wortstemming (die automatische Abtrennung der Flexionsendungen), eine machtvolle Retrievalsprache „Lucene“ mit Wildcards und Näherungsoperatoren, einer Rankingliste, der Eliminierung von leeren Feldern, dem Einsatz der Fuzzy-Recherchen und vielen Formatvarianten für die Patentnummern bereitgestellt. Bisher gibt es noch keine Suchmöglichkeiten nach der Dokumentenart oder nach Daten von Zweitpublikationen. Dies wird aber bald nachgeholt.
Es sollen möglichst alle Daten des Übergangs in die nationalen Phasen in Patentscope aufgenommen werden. Die kleineren Ämter werden von der WIPO dabei unterstützt.
Mit ARDI (Access to research for development and innovation) soll den Entwicklungsländern unter die Arme gegriffen werden. Zurzeit versorgen zwölf Verlage mit 50 Zeitschriften 107 Entwicklungsländer über dieses Programm.

WIPO Gremien

Es werden neue Komitees geschaffen und andere aufgelöst. Über ein neues Komitee, das sich mit der weltweiten Infrastruktur des gewerblichen Rechtsschutzes beschäftigen soll, muss noch die Generalversammlung abstimmen. Die Hauptaufgaben werden darin gesehen, die Arbeitsgrundlagen zu verbessern, gemeinsame Werkzeuge zu erarbeiten und eine Annäherung für eine internationale bessere Zusammenarbeit zu schaffen, mit einer Politik, die eine freie Verfügbarkeit der Schutzrechtsinformationen fördert.
Weitere Tätigkeitsfelder sind die Weiterentwicklung der Normen (z.B. ST.10/C) und Empfehlungen für den Gebrauch der URLs im Hinblick auf die Nutzung der „uniform resource identifiers“.
Vorgesehen sind Berichte an das Internationale Büro und das EPA über den Eintritt von Patentdokumenten in die nationale Phase und die Entwicklung von STAD eine Datenbank für Verwaltungsnormen, die weltweit Gültigkeit haben sollen. Geplant ist ein Symposium der Patentbehörden über die weltweite IP-Infrastruktur. Die Zusammenfassungen werden im Internet veröffentlicht. Es geht um die Vermeidung von Doppelarbeit sowie die Bildung von Ländergruppen zur gemeinsamen Nutzung von Recherche- und Prüfungsergebnissen (wie z.B. die Vancouver Group mit AU, CA und GB oder die südamerikanische Gruppe). Eine Road Map für die Vereinheitlichung der Qualitätsmaßstäbe bei der Recherche und Prüfung soll erarbeitet werden. Dies wird insbesondere bei kleineren Ämtern auf große Schwierigkeiten stoßen. Der Austausch von Prioritätsdokumenten (DAS: Digital Access Service) ist in Arbeit.

Internetprovider für Patentinformationen

LexisNexis

Das Informationsgeschäft von LexisNexis gliedert sich in drei Bereiche: den Bereich Reed Business – NEWS Informationen, den Bereich Elsevier – Naturwissenschaftliche Informationen und den Bereich LexisNexis Recht – News und Business Informationen. Für IP-Lösungen sind Recherchen in der Patent- und Nichtpatentliteratur möglich.
So gibt es die TotalPatent Workflow-Lösung für die Online-Patentrecherche und -analyse. Dabei können aus 24 Sprachen Volltexte mit englischer Maschinenübersetzung bereitgestellt werden und dies aus 96 Ländern.
Wichtige Referenzen der Nichtpatentliteratur unter lexis.com sind die „Research and Disclosure“ Daten unter ip.com und die Elsevier Volltextliteratur im wissenschaftlichen Bereich.
Mit dem Dienst IP DataDirect–Corporate können aus 105 Ländern Patentdaten bereitgestellt werden, dabei von 25 Sprachen Volltexte mit Maschinenübersetzung. Hinzugekommen sind im Jahr 2009 Daten aus Russland, China, Indien, Brasilien und weiteren Staaten.

Semantische Recherche

LexisNexis bietet im Bereich der IP-Lösungen schon seit einiger Zeit auch semantische Recherchen in der Patentliteratur und der Nichtpatentliteratur an. Probleme ergeben sich bei der Häufung von Synonymen und der zum Teil ungleichartigen Nomenklatur der Patentsprache.
Die Grundlage der semantischen Recherche ist die Analyse der eingegebenen Konzepte mit semantischer Technologie, um dem Patentrechercheur eine zusätzliche, komplementäre Möglichkeit zur Booleschen Recherche anzubieten. Verbesserungen der Rechercheergebnisse ergeben sich aus einer besseren Transparenz und Übersichtlichkeit, sowie der besseren Kontrolle und Steuerung der Recherche.
Die semantische Recherche ersetzt aber in keinem Fall die Erfahrung des Rechercheurs in der Patentrecherche.
Die semantische Recherche baut auf der Arbeit von PureDiscovery™ auf, mit der LexisNexis eine Partnerschaft abgeschlossen hat. Sie besteht aus einer sehr gut ausgestatteten und ständig aktualisierten Ontologie für den Patentbereich. Hintergrund sind die Patentdatenbank des USPTO und alle Volltexte der wissenschaftlichen Zeitschriften von Elsevier. Alle Verknüpfungen zwischen Worten und Konzepten werden kategorisiert und in unterschiedliche semantische „Gehirne“ gespeichert. Ein dynamischer Index wird erzeugt. Es wird quasi eine künstliche Intelligenz mit einem konstanten Lernprozess geschaffen. Dadurch wird eine dauerhafte Aktualisierung der unterschiedlichen Quellen erreicht. Der Index ist damit nicht statisch. Neue Technologien in der Patentliteratur werden in immer weiter verzweigten semantischen Gehirnen kategorisiert
Insgesamt werden 2700 englischsprachige Quellen in TotalPatent als Ontologiegrundlage verwendet zusammen mit Quellen aus lexis.com.

In TotalPatent gibt es dazu auf der Rechercheoberfläche zu dem Tab für die Boole'sche Recherche jetzt auch einen Tab für die semantische Recherche. Die Recherche kann aus Worten, Sätzen, Paragraphen, Ansprüchen, etc. bestehen.
Als erstes Ergebnis wird in graphischer Darstellung eine Begriffswolke erzeugt. Worte und Phrasen werden durch Inferenz des semantischen Gehirns bestimmt. Die Größe und die Farbe der entstandenen Worte kennzeichnen das Ranking dieser Worte innerhalb der Recherche. Ein Klick auf ein beliebiges Wort gibt dem Nutzer die Möglichkeit, die Wichtigkeit der Zuordnung des Begriffs im Ranking zu erkennen oder auch bestimmte Worte aus der Recherche auszuschließen. Der Rechercheur kann auch zusätzliche Konzepte einfügen. Der untere Teil der Rechercheoberfläche eröffnet, wie bei der Booleschen Recherche, die Möglichkeit, die Recherche mit weiteren Kriterien zu verfeinern, wie z. B. mit den Patentanmeldern, den Erfindern, den Klassifikationsangaben oder auch einem Datum oder mit Ländern.
Die Treffermenge des ersten Suchschrittes kann durch weitere Suchstrategien semantisch weiter verfeinert werden.

Lieferung von Rohdaten über Corporate

Corporate ist der Dienstleistungsbereich von LexisNexis, der Patentrohdaten liefert.
Die Datensammlung verfügt über eine Volltextsammlung mit Backfile und Frontfile von 25 Ländern, so die wichtigsten Quellen der Mindest PCT-Dokumentation wie die US, EP, WO, DE, FR, GB Dokumente und die JP A Daten als Zusammenfassungen, so auch andere wichtige europäische Quellen wie AT, BE, CH, DD, DK, ES, FI, IE, IT, LU, MC, NL, PT, SE dazu kommen weitere wichtige englischsprachigen Quellen wie AU und CA und einzigartige Quellen, die nicht elektronisch zur Verfügung stehen wie GB B Dokumente.

Hinzu kommen 105 bibliographische Sammlungen mit Zusammenfassung und einer wöchentlichen Aktualisierung am Tag der Veröffentlichung.
In der letzten Zeit wurden die Sammlungen mit den Patendaten weiterer Länder ergänzt mit Daten aus Indien, aus Brasilien, aus China, aus Argentinien und aus Russland. Im Jahre 2010 sollen noch weitere Länder hinzukommen.
Neu ist die maschinelle Übersetzung in die englische Sprache, die Rechtsstandinformation, die zitierten und zitierenden Patente, die Patentfamilien, sowohl die Daten zur einer engen wie auch Daten zur erweiterten Patentfamilie, sowie die Suche in PDF-Dateien.

Die Lieferung der Daten erfolgt im XML WIPO ST.36 Format. Weitere Details unter caterina.dauvergne@lexisnexis.com

Minesoft

PatBase

Minesoft stellte **PatBase** (s. Abb. 4) vor. Neben einer besseren Sortierung wurden auch die Darstellung und der Export von Daten überarbeitet. Die Sortierung erfolgt nach Relevanz und nach der Anzahl der Zitierungen. Die Exportoptionen für PDF Formate, für Bilddarstellungen und fünf weitere benutzerdefinierte Formate können gewählt werden für die Wiedergabe von Trefferanzeigen mit der Auswahl verschiedener Länder, für den Titel, das Abstract und weitere Felder.

Neue Suchoptionen wurden implementiert, so z. B. die Suche mit nichtlateinischen Buchstaben in chinesischen, japanischen, koreanischen sowie in russischen Datenbanken. Bei der Suche nach deutschen Texten kann eine Linkstrunkierung erfolgen. Ein Nummern Wizard wurde eingeführt, um ungenaue Patentnummern suchen zu können. Eine Suche nach chemischen Synonymen ist über ChemIDplus (National Library of Medicine) möglich. Inzwischen kann auch die Suchstrategie editiert werden.

Weitere Tipps für die Verfeinerung der Recherche sind z. B. die Feldfunktion CCD=EP*2009 (Nummern Wizard) für die Nummernsuche, die Nutzung des % Trunkierungssymbols, für die Suche nach Prioritätsjahr und Land mit PR=DE2000*, für die Update Codes UP= und UE= und für die Ermittlung der Zitierungen in einer Trefferliste die Codes CTFn für die Vorwärtszitierungen und CTBn für die Rückwärtszitierungen, wobei n die Anzahl der Suchschritte angibt.

Patent Archiv

Die Nutzung eines eigenen Patent Archivs hat bei Minesoft große Bedeutung. So können Kunden Infos und Rechercheergebnisse im eigenen Archiv ablegen, verwalten, bewerten, weitergeben, in verschiedenen Formaten editieren und Daten per E-Mail aus dem Archiv versenden. Das Hosting kann bei Minesoft oder beim Kunden erfolgen. Alle Lösungen können an die Bedürfnisse des Kunden angepasst werden. Dies ermöglicht eine einfache und effiziente Verteilung von Patentinformationen in einem Unternehmen.

Für Profis gibt es sogar eine eigene weitere Schnittstelle. Aus dem Archiv sind auch die üblichen Überwachungsvorgänge zu initiieren.

Minesoft wurde anlässlich eines Empfangs bei der englischen Königin für das Jahr 2009 der Preis für international tätige erfolgreiche englische Unternehmen überreicht.

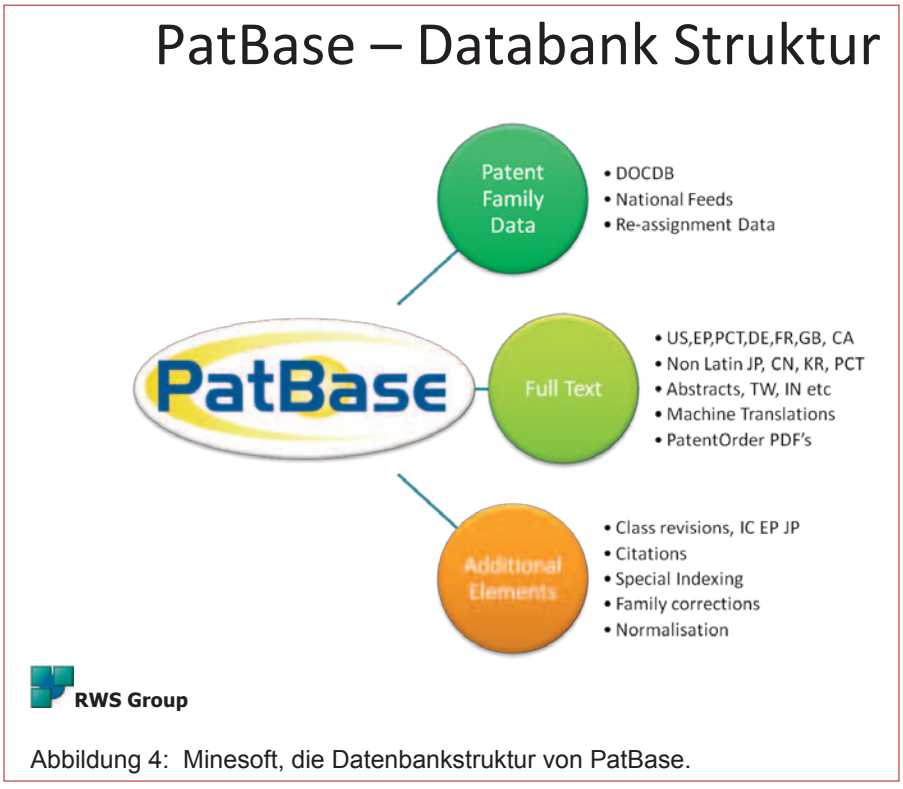


Abbildung 4: Minesoft, die Datenbankstruktur von PatBase.

Abbildung 4: Minesoft, die Datenbank

Thomson Reuters

Thomson Reuters hat heute rd. 50.000 Beschäftigte und ist in 93 Ländern tätig. Das Unternehmen legt hohe Maßstäbe bei der Unterstützung von innovativen Unternehmen und beim IP-Management in einem integrierten Prozess an, es hält die größten Sammlungen von IP-Informationen vor, stellt Recherchewerkzeuge zur Verfügung und bietet Dienstleistungen von Experten an.

Datenbanken

Folgende Datenbanken werden angeboten:
Thomson Innovation®, Thomson IP Manager®, der Derwent World Patents Index®, Aureka®, Delphion®, iEPROS®, Thomson Data AnalyzerTM, IP Payments, File Histories und Normen.
Ferner stehen im Bereich der Marken folgende Datenbanken zur Verfügung:
SERION®, SAEGIS®, ebenso der Thomson IP Manager®, Marken Bedeutungshinweise, Markenüberwachung und Markenverwaltung.
Im Bereich der Rechtswissenschaften sind es die Datenbanken Westlaw Patents® und der IP Monitor sowie Thomsons „Analytische Veröffentlichungen" (McCarthy's).
Im Lebenszyklus eines Schutzrechts gibt es viele Ansatzpunkte für die Unterstützung von Kunden (s. Abb. 5), z. B. um Innovationen zu erleichtern, das Schutzrechtsportfolio zu verwalten, eine korrekte Kommunikation zu führen, richtige strategische Entscheidungen zu treffen, beste Lösungsansätze zu finden und Risiken zu minimieren.
Der Schutzrechtsmanager bietet zunächst eine zentrale Stelle für das Management des Portfolios, dazu sichere Investitionen im Rechtsschutzbereich durch die Nutzung von IP Rules Calculations™ und Links zu den Verwaltungsdiensten, wie z. B. die Bezahlung von IP-Gebühren u.ä.m.
Der „Workflow Prozess Architekt™" wurde verbessert. So kann ein bestimmter Vorgang definiert, der Ablauf visualisiert und ein Ablaufschema vorbereitet werden.
Thomson Reuter hat eine verbesserte Plattform geschaffen für alle Belange der Rechtsschutzinformationen (s. Abb. 6) und dies von der Recherche über die Verwaltung, Überwachung und Bearbeitung bis zur Dokumentenlieferung im Volltext.

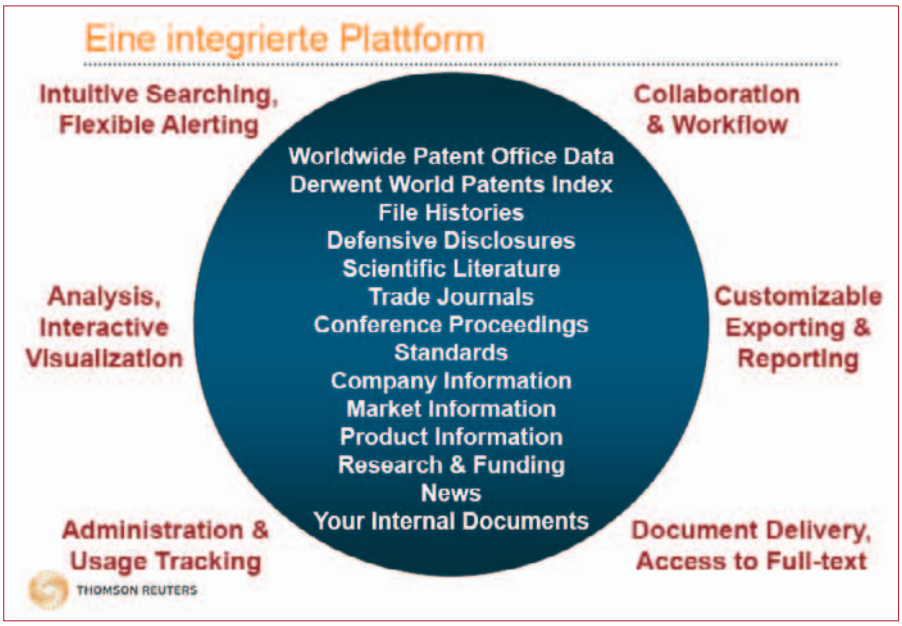


Abbildung 6: Die integrierte Plattform von Thompson Reuters für alle Bereiche des *gewerblichen Rechtschutzes.*

Asiatische Patentdaten werden nun, je nach Bedarf des Rechercheurs für bestimmte Felder und Bereiche in einer Volltextübersetzung in neun Sprachen angeboten und dies „on the fly". Auch die japanischen FI-Codes und F-Terms sind suchbar und exportierbar und auch im DWPI zu finden.
Als Analysewerkzeuge stehen bereit, die Zitatekartierung, das Text Clustering und der Aufbau von Themenlandschaften für die Analyse von Unternehmensportfolios, die Entscheidung über Lizenzfragen und die Markt-, Technologie- und Trendanalysen.

Derwent World Patents Index

Beim DWPI wurden ständig Erweiterungen vorgenommen mit einem entsprechenden Qualitätsmanagement. Rund eine Million Datensätze sind im letzten Jahr hinzugekommen. Innerhalb von 20 Tagen werden die Daten der Ämter im DWPI veröffentlicht und verarbeitet. Im ersten Quartal 2009 wurden die Erfindernamen für kanadische und koreanische Patente nach dem Derwent Familienstandard ergänzt. Es folgte die Erweiterung und Ergänzung ab Juli 2009 mit Patent- und Gebrauchsmusterdaten aus Taiwan. Die Klassifikation des japanischen Patentamtes wurde übernommen. Die koreanischen Dokumente werden drei Wochen nach der Veröffentlichung in DWPI First angeboten.
Im 2. Quartal 2009 wurden die Schriftartencodes der Patentanmeldungen und der Patentschriften der Schweiz ergänzt.
Die Patent- und Gebrauchsmusterdaten von Spanien sowie die Daten von brasi-

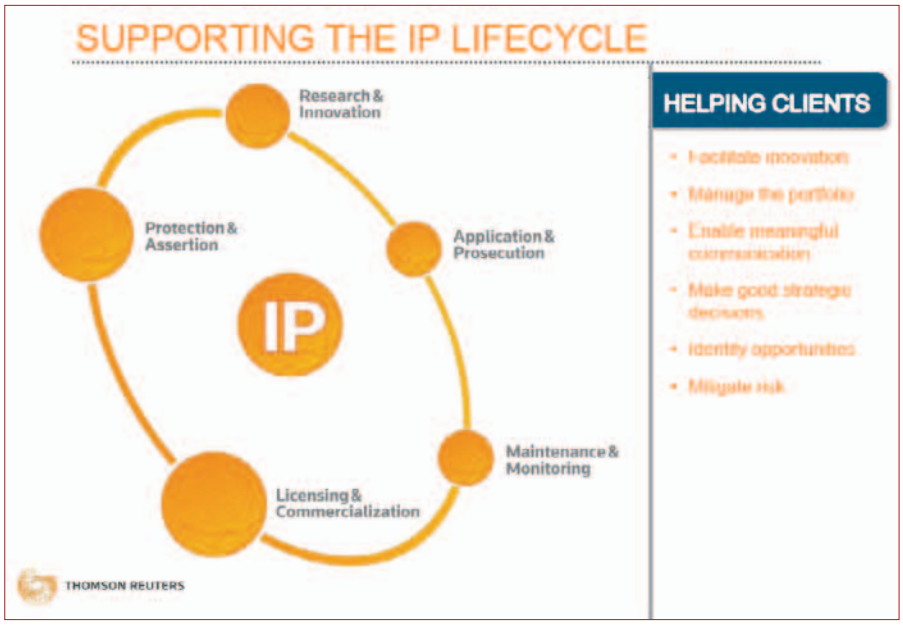


Abbildung 5: Thomson Reuter Hilfen beim Lebenszyklus von Schutzrechten.

lianischen Patentanmeldungen und Gebrauchsmustern wurden erfasst.
Vorgesehen ist eine Erweiterung der Patentdaten um die Länder Malaysia, Vietnam, Hongkong und die Länder des mittleren Ostens.

iEPROS

TR hat mit seinem Dienstleistungsangebot iEPROS ein Werkzeug geschaffen, mit dem die Integration von Datensätzen in eine Anwendung bei einem Unternehmen vorgenommen werden kann.
Es schafft Mehrwert bei der Versorgung mit Patentinformationen auch für die Nutzung in eigenen Datenbanken und stellt Werkzeuge und Software bzw. Schnittstellen für Recherchen bereit mit dem ganzen Umfeld der TR Patentverwaltungswerkzeuge. Ein ausgefeiltes Workflow Management ist eingebunden.

Questel Orbit

Questel-Orbit hat eine neue Integration aller Online-Dienste für Patent- u. Design-Informationen auf einer Ebene vorgenommen. Die Integration von Daten und Diensten wurde verbessert und führt zu umfassenderen, schnelleren Rechercheergebnissen mit großer Flexibilität für Anzeige, Export und Sitzungsverwaltung. Damit nimmt die Individualisierbarkeit des Recherchesystems zu und wird so noch besser auf die Bedürfnisse der Kunden zugeschnitten.

Die FamPat Datenbank enthält nun auch die Volltexte der US-Dokumente, der EP-Dokumente und der PCT-Dokumente sowie einen raschen Updatezyklus innerhalb von ein bis zwei Wochen nach der Publikation für Abstract, Anmelder und Titel in englischer Maschinenübersetzung für die Dokumente aus CN, IN, JP, KR und TW.
Neue Volltext-Datenbanken in Maschinenübersetzung, z. T. in Originalsprache, wurden aufgelegt oder ergänzt für Brasilien ab 1974, für Kanada (Ansprüche) ab 1978, für China ab 1985, für Dänemark ab 1895, für Finnland ab 1944, für Japan ab 2004, für die nicht englischsprachigen PCT-Dokumente ab 1978, sowie für Russland/GUS ab 1924, für Spanien ab 1980 und für Schweden ab 1885.

Rechercheebene

Die schon erwähnte Integration aller Online-Dienste für Patent- und Designinformationen bietet alles unter einem Dach wie die Patentrecherche, Analyse von Dokumenten, Archivierung, Schutzrechtsüberwachung, Rechtsstandsüberwachung und Designrecherche.
Alle Dienste sind auf einer benutzerfreundlichen Rechercheoberfläche angezeigt und abrufbar. Die Recherche und die Auswertung der Ergebnislisten sind individuell möglich mit einer entsprechenden Gestaltung der Suchstrategie, der Anzeige, des Exports, der Archivierung, der Suchhistorie und des Trefferlisten-Managements.

QPat kombiniert Volltextanzeige und Familien, sortiert die Trefferliste nach Relevanz und zeigt die Suchhistorie an.
Neue Anzeigemodi sind hinzugekommen, so dass beliebige Details der Treffer angezeigt werden können. Der Text des Dokumentes und das entsprechende PDF-Dokument können auf einem Bildschirm dargestellt werden. Der Export einer Datei ist im Format .TXT, .PDF, .RTF, .XLS, .CSV, .XML oder auch per E-Mail möglich. Es besteht die Option auf Templates, mit denen alle Exportformate und Felder für die individuelle Bearbeitung und Speicherung ausgewählt werden können.

Rechercheergebnisse, Trefferlisten und Bewertungen können in einem eigenen Patentarchiv (PatentExaminer) zusammengefasst und gespeichert werden mit der Ergänzung um eigene Felder oder auch mit den entsprechenden Tags, ob noch unbearbeitet, ob wichtig oder ob für die Weitergabe freigegeben usw.

Große Mengen von Dokumenten können in eine Überwachungsliste aufgenommen werden. Dies gilt für die Überwachung von Anmeldern, Erfindern oder Patentklassen, wie auch für die weitere Entwicklung des Verfahrensstandes.

Designrecherche

Die neue Nutzeroberfläche eröffnet die Möglichkeit nach Titel, Klassensymbolen, Anmeldern, Anmeldenummern, Veröffentlichungsnummern oder auch nach Prioritätsländern zu suchen. Die Ergebnislisten stellen die zugehörigen Abbildungen in allen Varianten bereit.
Dadurch ist eine schnelle und verlässliche Übersicht über Designanmeldungen möglich.
Die Betaversion des Portals steht seit Januar 2010 zur Verfügung.
Weitere Details bei rgissler@questel.com.

Unycom

Unycom hilft ihren Kunden mit Software und Beratungsleistungen, Schutzrechte zu verwalten und den Wert der Innovationen zu steigern.

Dienstleistungsangebote

Unycom bietet folgende Dienste an:

- Unycom IPMS - ein System zur Verwaltung von großen Schutzrechtsportfolios mit Bewertung und Steuerung,
- Unycom iPat - das integrierte, interaktive Informationssystem für Anwaltskanzleien und
- Unycom PatentPilot - das Online-Patentinformationssystem zur Überwachung mit Bearbeitung und Verwaltung von Fremdpatenten.

Zu den Kunden zählen eine große Anzahl von Konzernen und Großbetrieben weltweit.
Unycom arbeitet mit der Fa. Treparel (Niederlande) zusammen, die Spezialisten sind auf dem Gebiet des Textmining, der graphischen Darstellung und der Analyse von Informationen.
Ein Rechercheur hat heute fast immer die Nadel in einem Heuhaufen zu suchen, wenn es darum geht, aus einer Trefferliste von Hunderten oder gar Tausenden von Dokumenten die relevanten Dokumente zu finden.
Unycom versucht den so genannten „mikroskopischen Blick“ auf die Daten. Dies gilt auch, um Technologietrends auszumachen oder die Aktivitäten von Mitwettbewerbern zu überprüfen.
Nach der Recherche mit inzwischen verfeinerten Booleschen Operatoren werden weitere Fragen gestellt, die mit Hilfe von „Phrase Searches“ erweitert werden können.
Ganz davon abgesehen, dass Informationsspezialisten immer noch traditionelle Recherchen bevorzugen, da sie große Erfahrung haben in der Nutzung von ihnen bekannten Datenbanken, haben sie trotz einer erfolgreichen Recherche immer einen hohen Zeitaufwand bei der Durchsicht von Treffern in Listenform, die in der Regel wenig über Zusammenhänge aussagen. Die Folgen sind hohe Recherchekosten und die Befürchtung, dass eine immer stärkere Verlagerung der Informationsbeschaffung in den F&E Bereich (siehe Google usw.) erfolgt.
Alternativen sind die Volltextsuche mit entsprechenden Hilfsmitteln, wie z. B. eine statistische Auswertung oder Ranking-Methoden, sowie die Machine Learning Methode, eine semantische Suche mit automatischer oder halbautomatischer Klassifizierung.
Der Einsatz einer „Vektor Maschine“, mit rein mathematischem Ansatz zur Feststellung der Ähnlichkeiten von Datensätzen, benötigt nur wenig Input (20 bis 30), um Ähnlichkeiten festzustellen. Wenn die Trainingsdokumente hinterlegt sind, wird der Originaltext nach einer Tokenisierung und der Eliminierung von Stoppwörtern, Endungen und Erweiterungen (stemming) vektorisiert. Nach der Eingabe von Trefferlisten werden diejenigen Treffer, die eine hohe Relevanz beinhalten, gekennzeichnet und verkürzen die Recherchezeit.
Alternativ dazu kann ein Clustering, ein Unycom Ranking oder ein Textmining erfolgen.

Aus dem Unycom PatentPilot und aus PatBase von Minesoft stehen Exporte in den Formaten XLS, CVS und XML zur Verfügung und dazu die Literaturdatenbanken wie z. B. Medline. In Planung sind Exporte im XML-Format aus TotalPatent. Das KMX-Format kann an alle Datenformate angepasst werden. Weitere Details bei anja.schmitt-bender@unycom.com.

Dialog

Dialog hat das neue Informationstool Innography eingeführt. Es handelt sich dabei um ein Analysewerkzeug, das dem Kunden erlaubt, nach Patent- und Wirtschaftsinformationen zu recherchieren und die Ergebnisse zusammen zu analysieren und graphisch darzustellen. Was die Wirtschaftsdaten betrifft, so wird eine unterschiedliche Perspektive und ein anderer Weg beschritten, um die Ergebnisse zu bearbeiten.
Dialog hat einen sehr fortschrittlichen Ansatz entwickelt, um Lösungen für IP-Probleme zu erarbeiten.
Es beginnt bei der Patentrecherche (Abb. 7) und führt über eine Schlüsselwortsuche, die Suche nach Anmelder- und Rechtsstandsdaten zu den Patentdokumenten und weiter über die Patentanalyse mit Textminingwerkzeugen und Dokumentgruppierungen zur so genannten „IP Intelligenz" mit einer Korrelation von Patent-, Wirtschafts-, Rechts- und Marktdaten. Man erhält eine vorausschauende Analyse mit wirtschaftsrelevanten Daten.

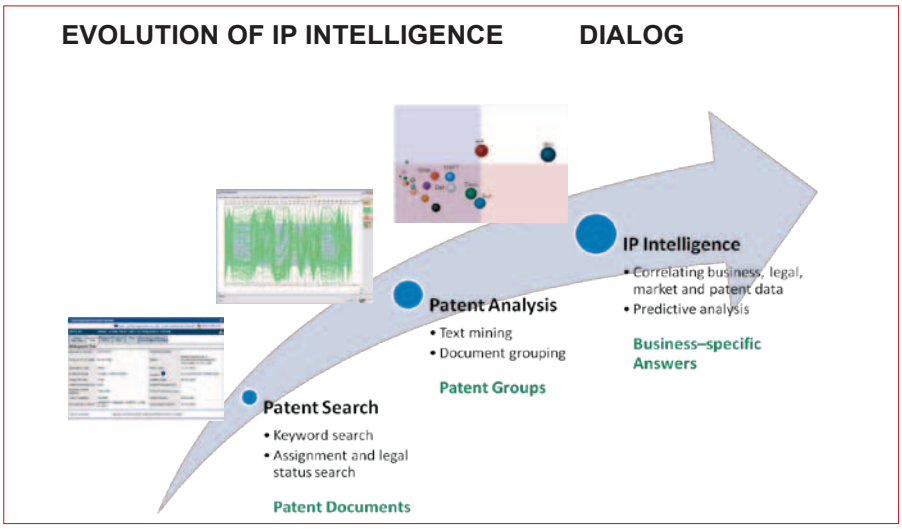


Abbildung 7: DIALOG Der Fortschritt der IP Intelligenz.

Das Innography Modul erstellt nach der Korrelation und Visualisierung ein kundenspezifisches Ergebnis.

Dazu stehen viele Module zur Verfügung, so für Verletzungsfragen, Lizenzfragen, Produktfragen, Technologiefragen, Wettbewerbsfragen und Marktfragen mit der Devise, ob Investitionen auch sinnvoll sind. Weitere Details bei: r.wonner@dialog.com.

Die nächste Arbeitssitzung der Arbeitsgruppe „Elektronische Medien in der Patentinformation" wird vom 3. bis 5. November 2010 im TIZ des DPMA in Berlin stattfinden.

Tagungsbericht, Patentinformation, Informationsdienst

Der Autor

BD Dipl. Ing. Dieter Geiß

ist Schriftführer der AGM und Berichterstatter zu Patentfragen der DGI.

Mittelstraße 33
44575 Castrop-Rauxel
d.geiss@t-online.de

IMPRESSUM

Gegründet von H.-K. Soeken †
unter dem Titel Nachrichten für Dokumentation (NfD)
Herausgegeben von der Deutschen Gesellschaft für Informationswissenschaft und Informationspraxis e.V. (DGI)
Präsident: Prof. Dr. Stefan Gradmann
Hanauer Landstraße 151-153
60314 Frankfurt am Main
Telefon: (0 69) 43 03 13
Telefax: (0 69) 4 90 90 96
mail@dgi-info.de
www.dgi-info.de
Mitteilungsblatt des Normenausschusses Bibliotheks- und Dokumentationswesen im DIN Deutsches Institut für Normung e.V., der Fachgruppe Dokumentation im Deutschen Museumsbund und der Arbeitsgemeinschaft der Spezialbibliotheken (ASpB)

Redaktionsbeirat
Dr. Sabine Graumann, München (Informationswirtschaft)
Prof. Dr. Hans-Christoph Hobohm, Potsdam (Management von Informationseinrichtungen)
Prof. Dr. Rainer Kuhlen, Konstanz (Informationswissenschaft)
Prof. Dr. Dirk Lewandowski, Hamburg (Suchmaschinen, Internet)
Prof. Dr. Wolfgang Ratzek, Stuttgart (Informationspraxis)
Prof. Dr. Ralph Schmidt, Hamburg (Newcomer Report, Medien)
Prof. Dr. Wolf G. Stock, Düsseldorf (Wissensrepräsentation und Information Retrieval)

Redaktion
Deutsche Gesellschaft für Informationswissenschaft und Informationspraxis e.V.
Marlies Ockenfeld (verantwortlich)
Viktoriaplatz 8, 64293 Darmstadt
Telefon: (0 61 51) 99 71 17
Telefax: (0 69) 4 90 90 96
ockenfeld@dgi-info.de
Daniel Ockenfeld (Redaktionsassistenz)

Verlag
Dinges & Frick GmbH
Greifstraße 4
65199 Wiesbaden
Postfach 1564
65005 Wiesbaden
Telefon: (06 11) 9 31 09 41
Telefax: (06 11) 9 31 09 43
Bankverbindung:
Wiesbadener Volksbank
BLZ 510 900 00, Kto-Nr. 714 22 26
Postbank Frankfurt
BLZ 500 100 60, Kto.-Nr. 267 204-606

Objektleitung
Erwin König,
e.koenig@dinges-frick.de

Anzeigenservice
Ursula Hensel Anzeigenservice
Hermann-Schuster-Straße 39
65510 Hünstetten-Wallbach
Telefon: (0 61 26) 57 08 82
Telefax: (0 61 26) 58 16 47
ursula.hensel@unitybox.de
Rocco Mischok
Verlag Dinges & Frick GmbH
Greifstraße 4
65199Wiesbaden
Telefon: (06 11) 3 96 99-60
Telefax: (06 11) 3 96 99-30
r.mischok@dinges-frick.de

Gestaltung
Meinhard Zielke, Wiesbaden

Druck
Dinges & Frick GmbH
Greifstraße 4
65199 Wiesbaden
Postfach 2009
65010 Wiesbaden
Telefon: (06 11) 3 96 99-0
Telefax: (06 11) 3 96 99-30
Leonardo: (06 11) 93 20 79
Twist: (06 11) 9 10 23 78
df@dinges-frick.de

Hinweis
Die Aufsätze stellen ausschließlich die Meinung der Autoren dar. Der Inhalt wurde sorgfältig und nach bestem Wissen erarbeitet. Dennoch kann von Verlag und Redaktion eine Gewährleistung auf Richtigkeit und Vollständigkeit nicht übernommen werden. Die Beiträge und die grafischen Darstellungen unterliegen dem Urheberrecht. Nachdruck und Vervielfältigung jeglicher Art bedürfen der Genehmigung des Verlages und der Autoren.

Erscheinungsweise/ Bezugspreise
Sieben Hefte jährlich (Doppelausgabe September/Oktober)
Jahresabonnement EUR 198,–
Schüler/Studenten EUR 148,–
Einzelheft EUR 34,–
inkl. Versandkosten/Porto im Inland,
Versand in Europa + EUR 28,–
Versand Welt / Luftpost auf Anfrage.
Das Abonnement gilt für mindestens ein Jahr und kann danach bis sechs Wochen zum Ende des Bezugszeitraums gekündigt werden.

Redaktionsschluss für
Heft 5 26. Juni 2010
Heft 6-7 23. August 2010

25. Oberhofer Kolloquium gibt frische Impulse

Marlies Ockenfeld, Darmstadt

47 Jahre ist es her, dass die Berater des US-amerikanischen Präsidenten Kennedy, in ihrem Bericht „Science, Government and Information" die Empfehlung aussprachen, in jedem Forscherteam gleichberechtigt „*research scientists*" und „*information scientists*" zusammen zu spannen, eine Forderung, die heute wieder neu gestellt werden muss. Sowohl die Deutsche Gesellschaft für Dokumentation (DGD) als auch das Zentralinstitut für Information und Dokumentation (ZIID) ließen diesen sogenannten Weinberg Report seinerzeit ins Deutsche übersetzen. Die Dokumentationsbewegung in ganz Deutschland erhielt politischen Rückenwind. Ein Jahr zuvor, am 16. und 17. November 1962, fand auf Initiative des Leiters des Instituts für Dokumentation und Patentwesen der Hochschule für Elektrotechnik, Ilmenau, Prof. Dr. Felix Weber, in Ilmenau ein Kolloquium „Dokumentation/Information" statt. Diese Veranstaltung bildete den gemeinsame Ursprung von zwei Veranstaltungsreihen, die es bis in die Gegenwart gibt: Das Internationale Kolloquium zur Praxis der Informationsvermittlung, kurz „Oberhofer Kolloquium" sowie das „Kolloquium der TU Ilmenau über Patentinformation und gewerblichen Rechtsschutz", kurz „PATINFO".

Das „Oberhofer Kolloquium" mit dem Thema „Recherche im Google-Zeitalter – vollständig und präzise?! Die Notwendigkeit von Informationskompetenz" fand vom 22. bis 24. April zum 25. Mal statt. Etwa hundert ehemalige, aktive und künftige praktisch und wissenschaftlich Tätige aus dem weit gespannten Bereich der Informationsaufbereitung und -vermittlung waren der Einladung nach Barleben bei Magdeburg gefolgt.

Zur Eröffnung der Jubiläumsveranstaltung konnte DGI-Präsident Stefan Gradmann neben einigen langjährigen Besuchern und Referenten des Oberhofer Kolloquiums auch eine Gruppe von Studenten und Studentinnen des Schwerpunkts Informationsmanagement von der Hochschule Anhalt in Köthen mit ihrem Hochschullehrer begrüßen, ferner Professor Hans-Jürgen Manecke und Fritz Schael, ohne deren Engagement die traditionsreiche DDR-Tagung nach 1989 wahrscheinlich nicht fortgeführt worden wäre, sowie den VDI Bezirksverband Magdeburg, vertreten durch seinen Vorsitzenden Herrn Busch, die Firma regiocom, vertreten durch ihren Geschäftsführer Herrn Gutmann, und das Leibniz-Institut für Neurobiologie, vertreten durch Herrn Löw, dem ersten Mitglied der DGD aus den Neuen Bundesländern nach der Wende, denen er für ihre großzügige ideelle und finanzielle Unterstützung der Tagung dankte.

Zum Auftakt gab Siegfried Rosemann, Vorsitzender des AKI Magdeburg, einen kurzen Abriss über die Geschichte des Oberhofer Kolloquiums, der ausführlicher im Tagungsband nachzulesen ist.

Dem eingeladenen Eröffnungsvortrag von Professor Walther Umstätter, in dem er unter anderem – im Tagungsband einzeln nachvollziehbare – Berechnungen zur Anzahl der Dokumente im Word Wide Web anstellte, folgte eine engagierte Diskussion, bei der die Meinung vorherrschte, dass eine Berechnung aufgrund der Gesamttrefferangaben von Suchmaschinen grundsätzlich unmöglich sei, weil diese Angaben nahezu willkürlich seien und von den Suchmaschinen nicht korrekt berechnet würden.

Dank der routinierten Organisation durch das Ortskomitee, insbesondere Herrn Rosemann und Herrn Löw vom AKI Magdeburg, war wieder ein reibungsloser Ablauf der Tagung gewährleistet. Den sieben von neun Mitgliedern des DGI-Vorstands, von denen einige während der gesamten Dauer der Fachtagung anwesend waren, offenbarte sich eine diskussionsfreudige, zukunftsorientierte und problembewusste Teilnehmerschar, die den Eindruck verstärkte, dass es zusätzlich zur jährlichen Frankfurter Tagung anlässlich der Buchmesse und dem alle drei Jahre stattfindenden großen Leipziger Kongress für Information und Bibliothek unbedingt weiterhin spezifische Treffen geben muss,

Eine aufmerksame Zuhörerschaft

Networking am Abend ...

... und in der Kaffeepause

Gute Laune beim Abschlusspodium ...

in denen der persönliche und fachliche Austausch rund um eine durchgehende Plenumsveranstaltung im Vordergrund steht. Die Zusammenkunft in einem abgeschiedenen Tagungshotel, wenngleich für einige Großstädter als Ghetto empfunden und etwas gewöhnungsbedürftig, beförderte durch die beiden Abendveranstaltungen den Austausch zwischen den Tagungsgästen.

Eingebettet in das Programm war eine Diskussion zur angestrebten Satzungsreform und zur künftigen Positionierung der DGI. Sie zeigte, dass es zu den Vorschlägen des Vorstands durchaus dezidierte Meinungen gibt, das Angebot einer persönliche Diskussion aber offensichtlich auch in Zeiten der Netzwerkgesellschaft und im Kreis von Informationsprofis eher angenommen wird als die Möglichkeit, in Wikis oder Mailinglisten computervermittelt zu kommunizieren, wo die Gefahr von Missverständnissen und Fehlinterpretationen selbst bei sorgfältiger Formulierung ungleich größer ist. Bedauert wurde das Herausdriften des Hochschulverbandes Informationswissenschaft aus der DGI.

Außerdem konnten alle während der gesamten Tagung an aufgestellten Pinwänden Gedanken, Ideen, Beobachtungen und Vorschläge zum Verlauf der Veranstaltung machen. Diese wurden am Schluss im Rahmen einer Podiums- und Plenumsdiskussion unter der Moderation von Matthias Ballod erörtert, um eine Bilanz der diesjährigen Tagung zu ziehen und sich über die künftige Ausrichtung des Oberhofer Kolloquiums auszutauschen. Selbstverständlich gab es dabei zum Teil widersprüchliche Auffassungen, doch ist als Grundtenor festzuhalten, dass Erfahrungsaustausch und Weiterbildung im Mittelpunkt stehen sollen, dabei Wissenschaft und Praxis im ausgewogenen Verhältnis zur Sprache kommen sollen, viel Zeit zur Diskussion bleiben soll, eher eine Beschränkung und Konzentration anstelle einer Ausweitung gewünscht wird und neue Entwicklungen im inhaltlichen Programm berücksichtigt werden sollen.

Eines der diskutierten Themen war die künftige Ausrichtung dieser Zeitschrift im Spannungsfeld von Informationswissenschaft – für manchen gelegentlich an der Grenze der Lesbarkeit – und Informationspraxis. Mehrheitlich wurde die Auffassung vertreten, grundsätzlich keine Trennung in eine wissenschaftliche und eine praxisorientierte Zeitschrift vorzunehmen, sondern die Brückenfunktion der IWP beizubehalten. Über die Ausgestaltung im einzelnen, begutachtete Schwerpunktausgaben, Preisgestaltung, Open-Access, elektronische und gedruckte Ausgabe, Verlags- oder Eigenpublikation u.a. soll ein Votum der Redaktionsbeiräte eingeholt werden. Ob es einen Markt für zwei Zeitschriften gibt, darf bezweifelt werden. Auch die Zeitschrift cogito ging nach zwölf Jahren in der nfd/IWP auf, nachdem sie nie schwarze Zahlen geschrieben hatte. Derzeit führt eine Masterstudentin der HdM Stuttgart einige Experteninterviews durch. Gerne nehmen Vorstand und Redaktion auch unabhängig davon Anregungen aus der Leser- und Autorenschaft entgegen.

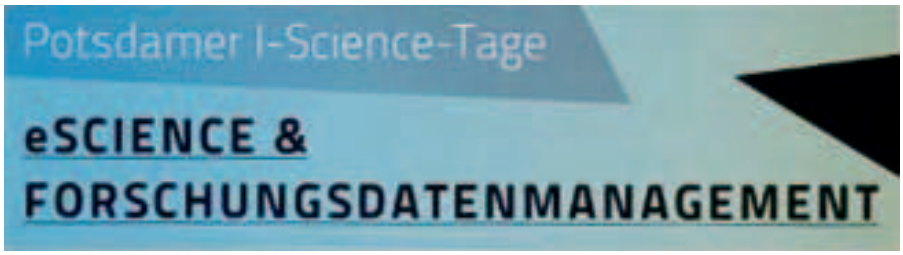
Alle achtzehn Beiträge der Tagung sind im Tagungsband versammelt, der bei der DGI-Geschäftsstelle bestellt werden kann; darunter auch der viel Applaus und anerkennende Worte erntende Vortrag von Bärbel Mende, die anschaulich und detailreich ihren schließlich gewonnenen Kampf um die Frage, ob sie als Information Brokerin freiberuflich oder gewerblich tätig ist, beschrieb.

Es sollen daher hier nur einige Aussagen von Miriam Lorenz wiedergegeben werden, die über erste Ergebnisse einer Befragung der Beschäftigten der Fraunhofer-Gesellschaft berichtete, die bei Redaktionsschluss des Tagungsbandes noch nicht vorlagen. Aus dem Rücklauf von 2196 Antworten, was etwa 13 Prozent entspricht, mit etwa 600 freien Kommentaren werden sich einigermaßen zuverlässige Aussagen treffen lassen. Nicht ganz unerwartet, meinen viele der wissenschaftlich Tätigen, eine Institutsbibliothek werde nicht mehr gebraucht, sofern alle Informationen, einschließlich der benötigten Zeitschriften im Volltext online verfügbar wären. Dass viele der von den Befragten vermissten Fachinformationen längst im durchgängig schlecht beurteilten Intranet verfügbar sind, ist eine auch in anderen Unternehmen immer wieder feststellbare Lücke zwischen gut gemeintem Angebot und tatsächlicher Nutzung. 1600 Antwortende hätten gerne eine Metasuchmaschine, die Hochwertigkeit der Informationen ist den meisten wichtiger als ein einfacher Zugriff. Der vollständige Bericht mit den Ergebnissen der Befragung soll im Herbst 2010 frei zugänglich veröffentlicht werden.

... unter der Moderation von Matthias Ballod

Die Datenflut im Forschungsprozess erfordert neue Informationsstrategien

Lars Müller, Potsdam, und Ute Rusnak, Karlsruhe



Über 180 Informationsexpertinnen und Informationsexperten sind der Einladung der FH Potsdam und von FIZ Karlsruhe zu den Potsdamer Information-Science-Tagen (kurz: I-Science-Tage) gefolgt, um sich vom 23. bis 24. März 2010 über die aktuellen Entwicklungen im Forschungsdatenmanagement zu informieren. Die Teilnehmenden aus Wissenschaft, Bibliotheken, Infrastruktureinrichtungen sowie wissenschaftlicher Softwareentwicklung nutzten am ersten Tag das gemeinsame Forum mit zahlreichen Expertenvorträgen zu „eScience und Forschungsdatenmanagement" ausgiebig dazu, grundlegende Fragestellungen des Datenmanagements im Forschungsprozess zu diskutieren.[1]
Am zweiten Tag gab es zwei Workshops für speziell Interessierte: den eSciDoc-Workshop für Anwendungsentwickler zur Realisierung konkreter Lösungen auf Basis der eSciDoc[2]-Software und einen Erfahrungsaustausch im Rahmen des Projekts „Wibaklidama"[3] mit über 20 Einrichtungen aus dem Raum Brandenburg-Berlin, die im Rahmen der Klimaplattform[4] miteinander vernetzt sind.

Informationsinfrastruktur erfordert nicht nur Technik, sondern auch eine Vernetzung der Akteure

Die Feststellung, dass eine neue Generation von Informationsinfrastruktur über Disziplingrenzen hinweg benötigt wird, um die aus dem Forschungsprozess resultierenden Datenmengen bewältigen und deren Nachvollziehbarkeit und Nachnutzung gewährleisten zu können, zog sich als roter Faden durch die gesamte Tagung. Die Informationstechnik ist dabei ein wichtiges Werkzeug für die Wissenschaftlerinnen und Wissenschaftler, die mit ihrem Bedarf im Fokus der Entwicklung neuer Systeme stehen. Komplementär dazu bedarf es jedoch auch organisationeller Infrastrukturen, um eine gemeinsame Plattform für die Akteure im Forschungsdatenmanagement zu schaffen und deren Vernetzung, Kommunikation und Diskurse zu unterstützen. Sie bilden die Basis für die Integration von Datenmanagement und Datenpublikation in die Wissenschaftskultur. Die Potsdamer I-Science-Tage haben gezeigt, dass die Gestaltung einer solchen Informationsinfrastruktur die gemeinsame Erfahrung aller beteiligten Akteure aus den jeweiligen wissenschaftlichen Disziplinen, Bibliotheken, Infrastruktureinrichtungen und Informationswissenschaften braucht.

The Fourth Paradigm: Data-Intensive Scientific Discovery

Tony Hey, Vicepresident Microsoft Research, nutzte für seinen Keynote-Vortrag „The Fourth Paradigm: Data-Intensive Scientific Discovery" die informationstechnischen Möglichkeiten zur virtuellen Kommunikation: Hey referierte aus seinem Arbeitszimmer an der amerikanischen Westküste und diskutierte anschließend mit dem in Potsdam anwesenden Fachpublikum. Hey erläuterte in seiner Keynote den Paradigmenwechsel in der Wissenschaft, der sich während der letzten tausend Jahre vollzogen hat: Nach Empirie, Theorie und rechnergestützter Simulationsmodelle sind wir heute bei dem „Vierten Paradigma" angekommen: Der datenzentrierten Wissenschaft, deren Erkenntnisse auf der Auswertung ständig wachsender Datenbestände aus meist heterogenen Quellen beruhen. Hey stellte fest, dass die Zukunft wissenschaftlichen Arbeitens in der Unterstützung des gesamten Forschungsprozesses liegt. Die Entwicklung neuer Ansätze für das wissenschaftliche Datenmanagement im gesamten, von der Datengewinnung über die Datenanalyse, Datenvisualisierung und Datenauswertung bis hin zur Publikation der Forschungsergebnisse, ist gefordert. In einer zukünftigen E-Infrastruktur sollten alle Daten miteinander verknüpft sein, von den originären Daten über Metadaten bis hin zu semantischen Daten. Ein lokaler „Client" sollte den Zugriff auf Daten in der „Cloud" ermöglichen, um Speicherung und Analyse riesiger Datenmengen skalierbar zu machen.[5]

Paradigmenwechsel in den Informationswissenschaften – The Data Librarian

Prof. Stephan Büttner und Prof. Hans-Christoph Hobohm vom Fachbereich Informationswissenschaften der FH Potsdam skizzierten in ihrem Einführungsvortrag die grundlegenden Veränderungen, die sich aus der datenorientierten Wissenschaft für die beteiligten Akteure im Forschungsdatenmanagement ergeben. Eine ganzheitliche Betrachtung des Forschungsdatenmanagements liefert die vom britischen Digital Curation Centre (DCC) veröffentlichte Darstellung des „Curation Lifecycle Model".[6] Anhand des Lebenslaufes digitaler (Forschungs-) Daten werden Konzepte im Forschungs-

1 Videos von den Vorträge und die Präsentationen stehen zum Download bereit unter http://informationswissenschaften.fh-potsdam.de/programm.html
2 eSciDoc: https://www.escidoc.org/ [1.4.2010].
3 Projekt Wibaklidama – wissensbasiertes Klimadatenmanagement: http://wibaklidama.fh-potsdam.de/
4 Klimaplattform – Forschungsplattform zum Klimawandel: www.klimaplattform.de/
5 Tony Hey/Stewart Tansley/Kristin Tolle (Hg.): The Fourth Paradigm, Data-Intensive Scientific Discovery. Redmond 2009. URL: http://research.microsoft.com/en-us/collaboration/fourthparadigm/ [1.4.2010].
6 DCC Curation Lifecycle Model. URL: http://www.dcc.ac.uk/resources/curation-lifecycle-model [1.4.2010].

datenmanagement dargestellt, deren Anwendung die Definition neuer Akteure und Rollen erfordert (nach Donnelly[7]): Data Manager (Steuerung), Data Creator (Bearbeitung), Data Librarian (Unterstützung) und Data Scientist (Analyse). Es gilt nun, diesen Wandel in den bibliothekarischen und informationswissenschaftlichen Berufsfeldern zu berücksichtigen. An der informationswissenschaftlichen Einordnung des Themas in das DIKW-Kontinuum, das den Übergang von Daten zu Information, zu Wissen (knowledge) und zu Weisheit beschreibt, wird deutlich, wie sehr am Ende evidenzbasiertes Wissen auf Daten beruht.

Im weiteren Verlauf der Tagung stellten zwölf Expertenvorträge in den Themenblöcken „State-of-the-Art", „Praxis & Anwendung" sowie „Metadaten & Verlinkung" grundlegende Aspekte zum Thema Forschungsdatenmanagement heraus.

Über 180 Teilnehmerinnen und Teilnehmer aus Hochschulen und Forschungseinrichtungen nahmen an der Tagung teil.

State-of-the Art: Virtuelle Forschungsumgebungen und Qualitätssicherung für die Nachvollziehbarkeit und Nachnutzung von Forschungsdaten

Unter dem Stichwort „Enablement" stellte Dr. Maximilian Stempfhuber von der Hochschulbibliothek der RWTH Aachen die Bedeutung einer adäquaten Informationsinfrastruktur für Forschende an Hochschulen und Forschungseinrichtungen heraus, damit diese ihr Forschungs(daten)umfeld optimal nutzen und gestalten können. Die im September 2009 gegründete Arbeitsgruppe „Virtuelle Forschungsumgebungen"[8] der Deutschen Initiative für Netzwerkinformation e.V. (DINI) nimmt sich dieser Thematik umfassend an.Welche Bedeutung Qualitätssicherung von Forschungsdaten für deren Nachvollziehbarkeit und Nachnutzung hat, führte Dr. Hans Pfeiffenberger vom Alfred Wegner Institut in Bremerhaven aus. Da Forschungsdaten selbst die Basis für wissenschaftliche Erkenntnis sind, ist die Frage nach ihrer Vertrauenswürdigkeit existenziell, auch oder gerade im Sinne „guter wissenschaftlicher Praxis"[9]. Neben den technischen und organisatorischen Grundlagen einer sicheren Speicherung von Daten mit Authentizitätsnachweis wird zunehmend die Publikation von Datenbeständen im Peer-Review-Verfahren praktiziert und damit neben den formalen auch eine intellektuelle Qualitätsbewertungen von Datenbeständen durchgeführt.
Einen Bogen von den Anfängen der Datenspeicherung in den 1970er Jahren zur bestehenden Infrastruktur für sozialwissenschaftliche Daten schlug der Pionier des Forschungsdatenmanagement, Dr. h.c. Ekkehard Mochmann. Über die Data Documentation Initiative (DDI)[10] wurde in den Sozialwissenschaften ein de-facto-Standard für Metadaten etabliert, der den Zugriff auf gespeicherte Forschungsdaten über zeitliche, räumliche und disziplinäre Grenzen hinweg ermöglicht.

Problemstellungen und Lösungsansätze in der Medizin erläuterte Dr. Thomas Schrade von der Charité Berlin/FH Brandenburg am Beispiel des Projekts OpEN.SC[11], in dem ein Informationssystem entwickelt wurde, das den heterogenen Daten medizinischer Arbeit wie auch den hohen Datenschutzanforderungen gerecht wird.

Praxis & Anwendung: Die datenorientierte Wissenschaft erfordert neue Formen der Informationsinfrastruktur

Der Vortrag „Designing for User Experience in eScience" von Prof. Dr. Frank Heidmann, FH Potsdam, zeigte anhand zahlreicher Darstellungen wie mit Datenvisualisierungen intuitive Nutzeroberflächen gestaltet werden können. Durch geeignete visuelle Aufbereitung können Forschungsdatenbestände zeitsparender und auch für Nicht-Wissenschaftler verständlich gemacht werden.
Dass die einfache technische Speicherung von Daten nicht gleichzusetzen ist mit deren Archivierung, stellte Dr. Karin Schwarz, ebenfalls FH Potsdam, in ihrem Vortrag heraus. Datenmanager sollten auf bewährte Strategien der Archivwissenschaften zurückgreifen, um sicherzustellen, dass nur relevante Daten gespeichert werden, diese aber zeitlich unbegrenzt. Dies bedarf einer gezielten Strategie und neben der engen Zusammenarbeit mit den Fachwissenschaftlern auch einer hohen Datenkompetenz der Archivarinnen und Archivare. Wie schwierig und wichtig es ist, diese Auswahl zu treffen, illustrierte folgendes Beispiel aus dem Vortrag von Dr. Pfeiffenberger: Das erste Aufkommen des Ozonlochs über den Polen wurde zunächst für einen Messfehler und die Daten für wertlos gehalten. Erst Jahre später wurde die wissenschaftliche Bedeutung dieser Messdaten in ihrem ganzen Ausmaß erkannt.
Roland Bertelmann, Bibliotheksleiter am Potsdamer Wissenschaftspark Albert Einstein, überraschte die Anwesenden mit der Feststellung „Wir wissen, wie es geht" und stellte aktuelle Lösungsansätze zum Aufbau von Forschungsdateninfrastrukturen vor. Bis die Publikation und durchgängige Referenzierung von Forschungsdaten als normaler Vorgang im Wissenschaftsprozess etabliert sein wird, ist allerdings noch ein langer Weg. Solange Forschende für die Veröffentlichung ihrer Forschungsergebnisse an Reputation gewinnen, aber nicht für die Veröffentlichung der zugehörigen Forschungsdaten, bleibt dies schwierig. Die Schaffung entsprechender Anreizsysteme

7 Donnelly, M.: RDMF2 Core Skills Diagram. Research Data Management Forum, 17.12.2008. URL: http://data-forum.blogspot.com/2008/12/rdmf2-core-skills-diagram.html [1.4.2010].

8 DINI-Arbeitsgruppe Virtuelle Forschungsumgebungen (vForum). www.dini.de/ag/vforum/ [1.4.2010].

9 Vorschläge zur Sicherung guter wissenschaftlicher Praxis. Empfehlungen der Kommission „Selbstkontrolle in der Wissenschaft". Denkschrift. Weinheim: Wiley-VCH, 1998. www.dfg.de/download/pdf/dfg_im_profil/reden_stellungnahmen/download/empfehlung_wiss_praxis_0198.pdf [1.4.2010].

10 Data Documentation Initiative. www.ddialliance.org/ [1.4.2010].

11 Projekt Open European Nephrology Science Center OpEN.SC. http://opensc.charite.de/ [1.4.2010].

und Maßnahmen zur Sensibilisierung der Forschenden sind eine wichtige Voraussetzung.
Am Beispiel des Projekts BW-eLabs[12], zeigte Dr. Matthias Schulze, Universitätsbibliothek Stuttgart, wie Daten und Dokumente aus allen Schritten des Forschungsprozesses betrachtet und in die Gesamtprozesskette experimenteller Umgebungen im Bereich der Nanotechnologie eingebettet werden können. Damit sollen alle relevanten Forschungsdaten für Forschungs- und Ausbildungszwecke nachvollziehbar und nachnutzbar gemacht werden. Die eSciDoc-Infrastruktur wird hierbei zur Archivierung der heterogenen Forschungsdaten eingesetzt und mit dem OPUS Repository-System[13], das für das Publikationsmanagement eingesetzt wird, verknüpft.
Eine umfassende Lösung zum Forschungsdatenmanagement stellte Matthias Razum von FIZ Karlsruhe mit der Open-Source-Software eSciDoc vor. Das gemeinsam von der Max-Planck-Gesellschaft und von FIZ Karlsruhe entwickelte eSciDoc-System stellt umfangreiche Dienste zur Verfügung, die das Datenmanagement im gesamten Forschungsprozess unterstützen. Zugriffssteuerung auf Objektebene und Datenspeicherung unter Berücksichtigung von Aspekten zur Langzeitverfügbarkeit stellen wichtige Elemente des Systems dar. Darauf aufsetzend können passgenaue Anwendungen für den individuellen Bedarf der Wissenschaftler entwickelt werden, die sowohl Forschungsdaten als auch Publikationen verwalten. Die Programmierung von konkreten eSciDoc-Lösungen wurde am zweiten Tag in einem Workshop vertieft.

Metadaten & Verlinkung – Standardisierung, Vernetzung und Persistenz sind wichtige Voraussetzungen für die zukünftige Nutzung heutiger Forschungsdaten

Thomas Severiens vom Institute for Science Networking, Oldenburg zeigte anhand seiner Erfahrungen mit dem Portal PhysNet,[14] einer RDF-basierten Datenbank für die Physik, wie wichtig die Vergabe und Standardisierung von Metadaten für Forschungsdaten ist. Die Erstellung der Metadaten kann über Web2.0-Instrumente durch eine Fach-Community oder durch die Systembetreiber selbst erfolgen. Je komplexer die Daten sind, desto eher sollte das System aus eigener Kraft gepflegt werden. Sein Fazit: Aktuelle und qualitätsgesicherte Daten erfordern Zeit und Aufwand.
Die Weiterentwicklung des World Wide Web zum „Semantischen Netz“ ist ein wichtiger Schritt, um eine Orientierung in der Datenflut zu erhalten. Das Konzept von „Linked Data“[15] ist auch in Bezug auf Forschungsdaten geeignet, um ein umfassendes Wissensnetz zu weben, davon sind Prof. Dr. Günther Neher und Prof. Dr. Felix Sasaki, beide FH-Potsdam, überzeugt.
Dr. Jan Brase von der TIB Hannover stellte mit „DataCite“[16] ein internationales Konsortium vor, das durch die Vergabe von persistenten Identifiern für Forschungsdaten deren langfristige Zitierfähigkeit sichert.

Hands-On-Workshop: Einstieg in die Programmierung von eSciDoc-Solutions

Am zweiten Tag vermittelten Referenten von FIZ Karlsruhe in einem ganztägigen Workshop 20 Softwareentwicklerinnen und Softwareentwickler, wie der technische Rahmen für Forschungsdatenmanagement auf Basis der eSciDoc-Infrastruktur installiert wird und wie darauf eigene eSciDoc-Lösungen entwickelt bzw. eingebunden werden können.

Erfahrungsaustausch zum Datenmanagement in der Klimaplattform

Im parallel stattfindenden Workshop zum „Erfahrungsaustausch Datenmanagement“ wurden organisatorisch orientierte Fragen diskutiert. Eine höhere Sensibilität für die Bedeutung eines strategischen Datenmanagements in den Institutionen und Wissenschaftsverwaltungen wurde gefordert. Obwohl professionelles Datenmanagement eine ökonomische Notwendigkeit ist (Vermeidung von Mehrfacherhebungen, schneller Zugriff auf bestehende Daten) stehen für diese Aufgaben oft zu wenig Mittel bereit. Durch eine verstärkte Kooperation, zunächst im Raum Brandenburg-Berlin, wollen Datenmanager das Thema stärker in den Fokus rücken und die Entwicklung einer modernen Forschungsdateninfrastruktur weiter verbessern.

Fazit

Die zahlreiche Teilnahme und positive Resonanz auf die Potsdamer I-Science-Tage machte deutlich, dass „eScience und Forschungsdatenmanagement“ ein aktuelles und wichtiges Themenfeld in der Wissenschaftslandschaft darstellt. Die Verknüpfung von informativen Vorträgen am ersten Tag mit praktisch ausgerichteten Workshops am zweiten Tag zeigte deutlich, welchen Beitrag die Informationswissenschaften und Informationseinrichtungen leisten, um Strategien zur Bewältigung der Datenflut zu entwickeln und aktiv mitzugestalten.

Tagungsbericht,
Datendokumentation, Forschung

Die Autoren

Ute Rusnak

leitet die Abteilung Öffentliche Forschung und Lehre bei Marketing und Vertrieb von FIZ Karlsruhe. Die Diplom-Biologin mit Zusatzausbildung im IT-Bereich ist seit 1987 bei FIZ Karlsruhe tätig und konnte in verschiedenen Führungspositionen umfangreiche Erfahrung bei der Neu- und Weiterentwicklung von wissenschaftlichen Informationsdienstleistungen in einem interdisziplinären Umfeld erwerben.
FIZ Karlsruhe
Hermann-von-Helmholtz-Platz 1
76344 Eggenstein-Leopoldshafen
Telefon: +49 (0)7247 808 433
Fax: +49 (0)7247 808 259
Ute.Rusnak@fiz-karlsruhe.de
www.fiz-karlsruhe.de

Lars Müller

ist akademischer Mitarbeiter am Fachbereich Informationswissenschaften der FH Potsdam und arbeitet im Projekt „Wissensbasiertes Klimadatenmanagement (Wibaklidama)“. Zuvor war der studierte Europäische Ethnologe mit einem Master in Bibliotheks- und Informationswissenschaften als Mitarbeiter der Universitätsbibliothek der Humboldt-Universität zu Berlin in dem Aufgabenfeld elektronischer Dienstleistungen tätig.
Fachhochschule Potsdam
FB Informationswissenschaften
Friedrich-Ebert-Straße 4, 4467 Potsdam
lars.mueller@fh-potsdam.de
http://wibaklidama.fh-potsdam.de/

12 BW-eLabs, Wissensmanagement in virtuellen und remote Laboren. www.ub.uni-stuttgart.de/wirueberuns/projekte/bw-elabs/ [1.4.2010].
13 OPUS. URL: http://samos.bsz-bw.de/ [1.4.2010].
14 PhysNet. URL: http://de.physnet.net/PhysNet/ [1.4.2010].
15 http://linkeddata.org [1.4.2010].
16 www.datacite.org/ [1.4.2010].

Journal of Documentation

London, Vol.65, No.6, 2009

Yi, Kwan: Linking folksonomy to Library of Congress subject headings: an exploratory study, pp.872
Golub, Koraljka; *Lykke*, Marianne: Automated classification of web pages in hierarchical browsing, pp.901
Julien, Heidi; *Genuis*, Shelagh: Emotional labour in librarians' instructional work, pp.926
Tang, Muh-Chyun: A study of academic library users' decision-making process: a Lens model approach, pp.938
Winget, Megan: Describing art: an alternative approach to subject access and interpretation, pp.958
Shachaf, Pnina: The paradox of expertise: is the Wikipedia Reference Desk as good as your library?, pp.977
Sukovic, Suzana: References to e-texts in academic publications, pp.997
BOOK REVIEW
Jansová, Linda: Managing Information Services: A Transformational Approach (2nd ed.)
Urquhart, Christine: Text Editing, Print and the Digital World
Stricevic, Ivanka: Reader Development in Practice: Bringing Literature to Readers
EDITORIAL
Bawden, David: Naming of parts (and things)
ADWARDS FOR EXCELLENCE
2009 Awards for Excellence
COMPARATIVE REVIEW
Bade, David: Comparative review

Journal of Information Science

East Grinstead, Vol.35, H.1, December 2009

Wang, Wei-Tsong; *Belardo*, Salvatore: The role of knowledge management in achieving effective crisis management: a case study, pp.635
Nicholas, David; *Clark*, David; *Rowlands*, I.; *Jamali*, Hamid R.: Online use and information seeking behaviour: institutional and subject comparison of UK research, pp.660
Hu, Ya-Han; *Chen*, Yen-Liang; *Tang*, Kwei: Mining sequential patterns in the B2B environment, pp.677
Huvila, Isto: Ecological framework of information interactions and information infrastructures, pp.695
Na, Jin-Cheon; *Thet*, Tun Thura: Effectiveness of web search results for genre and sentiment classification, pp.709
Park, Jung-ran; *Childress*, Eric: Dublin Core metadata semantics: an analysis of the perspectives of information professionals, pp.727
Hara, Noriko; *Shachaf*, Pnina; *Stoerger*, Sharon: Online communities of the practice typology revisited, pp.740
Huberman, Bernado; *Romero*, Daniel M.; *Wu*, Fang: Crowdsourcing, attention and productivity, pp.758

Online

Weston, Vol.33, No.5, September/October 2009

FEATURES
Becker, Danielle A.: A Website Redesign Owner's Manual, pp.14
Hammond, Richard: Visualizing Business Rules: SVG in Action, pp.20
Rollo, Pamela: Disruption and Vision: Allies of Innovation, pp.26
Ojala, Marydee: Delores Meglio and the Information Generations, pp.32
Fernandez, Joe: A SWOT Analysis for Social Mediain Libraries, pp.35
Ojala, Marydee: The Australiasian Journals Issue, pp.38
COLUMNS
Notess, G.R.: On the Net. Bing and Increases in Data-Based Search, pp.40
Ojala, Marydee: The Dollar Sign. Searching for Corporate Responsibility, pp.44
Badke, William: InfoLit Land. Media, ICT, and Information Literacy, pp.47
Jascó, Péter: Péter's Picks & Pans. CISDOC, MedlinePlus, CISILO, pp.50
Wisniewski, Jeff: Control-Shift. Next-Gen OPACs: No Time Like Present, pp.54
Crawford, Walt: Crawford at Large. PermacOpyright: A Modest Proposal, pp.61
Wiley, Deborah Lynne: Searching 2.0 / Google: The Digital Gutenberg / E-Learning 2.0: Proven Practices and Emerging Technologies to Achieve Real Results / The Mobile Marketing Handbook: A Step-by-Step Guide to Creating Dynamic Mobile Marketing Campaigns, pp.61
Bates, Mary Ellen: Online Spotlight. Do I Look Like a Librarian?, pp.64
DEPPARTMENT
Ojala, Marydee: HomePage. I'll Know It When I See It, pp.5
Sabroski, Suzanne: Industry News, pp.8
Notess, Greg R.: Search Engine Update. New Search Features, Developments, and Content, pp.13

Weston, Vol.33, No.6, November/December 2009

Herther, Nancy K.: Digital Natives and Immigrants: What Brain Research Tell us, pp.14
Stuart, David: Social Media Metrics, pp.22
Levine, Emil: Policing Goes Online: Introducing POLIS, pp.26
Bhatnagar, Alka: Web Analytics for Business Intelligence: Beyond Hits and Sessions, pp.32
Dorms, Carol: Promoting Advertising: Advertising Database Express and World Advertising Research Center (WARC), pp.36
Arnold, Stephen E.: Real-Time Search: Where Retrieval and Discovery Collie, pp.40
COLUMNS
Notess, Greg R.: On the Net. Yahoo!'s Long Strang Journey, pp.42
Ojala, Marydee: The Dollar Sign. Searching for Business Trends and Trending Topics, pp.45
Badke, William: InfoLit. The Great Research Disaster, pp.48
Jacsó, Péter: Péter's Picks & Pans. A Final Joy Write: Panorama of Past Pans, pp.51
Wisniewski, Jeff: Control-Shift. Location, Location, Location, pp.54
Crawford, Walt: Crawford at Large. Agreeable Disagreement and the Excluded Middle, pp.58
Wiley, Deborah Lynne: Hardcopy. Free: The Future of a Radical Price / YouTube. An Insider's Guide to Climbing the Charts / e-Riches 2.0: Next Generation Marketing Strategies for Making Millions Online / When Search Meets Web Usability, pp.61
Bates, Mary Ellen: Online Spotlight. The Search Tool Arms Race, pp.64
DEPARTMENTS
Ojala, Marydee: HomePage. The Power of Three, pp.5
Sabroski, Suzanne: Industry News, pp.8
Notess, Greg R.: Search Engine Update. New Search Features, Developments, and Content, pp.13
Index to Advertisers, pp.63

Dok. Technologien, Strategien & Services für das digitale Dokument

Offenbach, H.3, September 2009

THEMA
Sack, H.: Such-Trends im World Wide Web, S.8
Sander, Philipp: Next Generation Dokumenten-Management: das Tor zum Unternehmenswissen, S.10
Scholtes, Johannes: Interview: „Ein Terabyte Daten ist schnell ein Terabyte Junk", S.16
Freisberg, Oswald: Nur einen Klick vom richtigen Treffer entfernt, S.18
Hepfner, Andreas: Enterprise Search-Wissensinseln in Unternehmen verbinden, S.22
BETRIEBLICHE PRAXIS
Röhrborn, Dirk. Microblogging goes Business: als integraler Bestandteil der Kommunikation, S.24
Haberl, Gerhard: Wie Unternehmen ihr Wissen kartografieren, S.28
Käfer, Gerhard; *Heiss*, Michael: Wissensnetze als Basis für Enterprise 2.0: ein Erfahrungsbericht, S.32
Moosreiner, Georg; *Neuberger*, Helmut: Interview: „Zarafa hat längst zu den großen Collaboration Systemen aufgeschlossen", S.36
Eitz, Gerhard; *Imhof*, Stephan: Dokumenten-Management in den Vergabeworkflows, S.38
Günther, Ralpf: „Informationsfabrik Krankenhaus": der Weg zum Dokumenten-Management, S.40
Ferling, André: Die Wirtschaftlichkeit von Enterprise Content Management, S.43
Amm, Martin: Ordnung vs. Chaos: Rollen und Rechte im Enterprise 2.0, S.46
Vogler, Jürgen: 400 Jahre Zeitgeschichte sicher ins 21. Jahrhundert bringen, S.49
Dietrich, Nicole: Lückenlose Sicherheit beim Umgang mit vertraulichen Dokumenten, S.52
Kehm, Edgar: Produktdaten-Management: Brücken statt Inseln, S.54
SPECIAL
Statements, Termine und Meldungen zur DMS EXPO: Capture- & Sprachtechnologiren / BPM & DMS / Outsourcing & Managed Services / Managed Services / Wann ist Software-as-a-Service marktreif? / SAP + DMS-Studie – ECM & E-Discovery / ECM & CMIS / Studie Cloud Computing – Studie E-Akte / Wissensmanagement mit SharePoint, S.56
TECHNOLOGIE
Kess, Matthias: Interview: "Für jeden Anwender von jedem Arbeitsplatz aus erreichbare Sicherheit", S.68
Tochtermann, Klaus: Neue Gesellschaftsmodelle im Semantic-Web und die Wertschöpfungskette der Daten, S.70
Kafiz, Willi: Integrierter Datenschutz macht Kollaborationsprozesse sicherer, S.73
Balfanz, Judith: Elektronische Zeitstempel – Absicherung für elektronische Geschäftsprozesse aller Art, S.76
Pröhl, Winfried: Datenlopgistik für Storage-Systeme, S.79
Wenzel, David: E-Mail-Recording mit einem „Flugschreiber" für die E-Mails im Unternehmen, S.82
Cohen, Victor: Computer, die denken können? Eine Einführung in Meaning Based Computing, S.84

Wissensmanagement

Reutlingen, Jg.11, H.6, August/September 2009

Lehnert, Oliver: Web 1.0 – Web 2.0 – Web 3.0: Aller guten Dinge sind drei? Editorial, S.3
PRAXIS WISSENSMANAGEMENT
Schmitz, Karl-Werner: Wissenstransfer mit allen fünf Sinnen, S.10
Kässer, Jürgen: Gefühl hilft Kopf: Die Rolle der Intuition in der Wissensarbeit, S.14
Held, Carsten: Den Herausforderungen im Bankensektor begegnen, S.16
TITELTHEMA
Albrecht, Jörg: Social Software im Unternehmen: Chancen nutzen, Risiken managen, S.18
Gattwinkel, Henning: Auf dem Weg zum intelligenten Internet: Wohin führt das Web 3.0?, S.22
Schütt, Peter: Twittern in Unternehmen, S.26
Tuchel, Susan: Schichtwechsel: Eine neue Ära der Suchmaschinen kündigt sich an, S.29
Henkel, Markus; *Gruber*, Stephan: Haben Google & Co. Bald ausgedient?, S.32
DOKUMENTATION + KOMMUNIKATION
Glitsch, Simone: Effiziente Geschäftsprozesse nach ISO 9000 – geht das?, S.34
Körber, Nicole: Rechnungsprozesse automatisieren & Compliance-konform klassifizieren, S.36
HUMAN RESOURCES
Dörhöfer, Steffen; *Eberling*, Wolfgang: Wissensaustausch sichtbar machen, S.42
Schmidt, Annette; *Haitzer*, Axel: War for Talents: Kann Ihre Firma High Potentials punkten?, S.46
THEMA
Kübler, Anette; *Strobel*, Thomas: Das Geschäft mit dem Wissen, S.48
Schmid, Simone: ITIL & Co.: Der Faktor Wissen – ein oft vernachlässigter Aspekt, S.51
WISSENSWERTES
Termine: Fachveranstaltungen für Wissensmanager, S.54
Buchvorstellung: Wissen als Mittelpunkt des Wettbewerbs, S.56
Vollmar, Gabriele: Wissensmanagement mit 2.0? Ein Stoßseufzer! Kolumne, S.57

Reutlingen, Jg.11, H.7, Oktober 2009

Lehnert, Oliver: Technische Dokumentation – dem Anwender zuliebe. Editorial, S.3
PRAXIS WISSENSMANAGEMENT
Metzmann, Freddy: Die Bedeutung des Nicht-Wissens, S.10
Bettray, Silke: Web 2.0 – Neuland für die Investitionsgüterindustrie, S.13
Richter, Uwe: Der steigenden Arbeitsbelastung von Wissensmanagern begegnen, S.16

2010

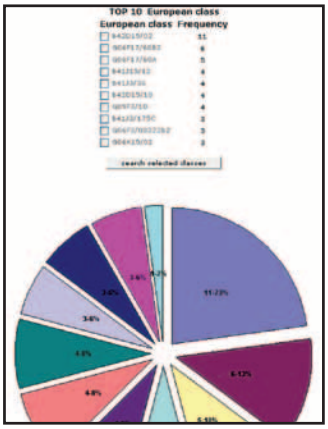
10. bis 11. Juni *Ilmenau*	PATINFO 2010 - 32. Kolloquium der Technischen Universität Ilmenau über Patentinformation „Methoden und Werkzeuge gegen Barrieren und Rückstände im gewerblichen Rechtsschutz“	Astrid Schieck, Technische Universität Ilmenau, PATON, Landespatentzentrum Thüringen, Postfach 10 05 65, 98684 Ilmenau, Telefon 03677 69-4572, Fax 03677 69-4538, patinfo@tu-ilmenau.de, www.paton.tu-ilmenau.de
14. Juni *München*	10. Single-Source-Forum Effiziente Prozess-Dokumentation: Mit dem richtigen Anfang zum Erfolg	Comet Communication GmbH, Goethestraße 17, 80336 München, Telefon 089 59989260, Fax 089 59989269, info@comet-comm.de, www.single-source-forum.de
15. bis 16. Juni *München*	Kongress „Next Generation Communication – Herausforderungen für die digitale Gesellschaft“	Münchner Kreis, Tal 16, 80331 München, Telefon 089 223238, Fax 089 225407 office@muenchner-kreis.de, www.muenchner-kreis.de
26. bis 30. Juli *Tilburg, Niederlande*	International Ticer School Digital Libraries à la Carte	Tilburg University, Library and IT Services, Ticer, Warandelaan 2, P.O. Box 90153, 5000 LE Tilburg, The Netherlands, Tel. +31 13 466 8310, Fax +31 13 466 8383, ticer@uvt.nl, www.tilburguniversity.nl/ticer/2010/
10. bis 15. August *Göteborg, Schweden*	Weltkongress Bibliothek und Information: 76. IFLA Generalkonferenz und Generalversammlung „Open access to knowledge - promoting sustainable progress”	Hella Klauser, Sekretariat des IFLA-Nationalkomitee, c/o Deutscher Bibliotheksverband e.V., Straße des 17. Juni 114, 10623 Berlin, Telefon 030 64498 9916, Fax 030 64498 9929, klauser@bibliotheksverband.de, www.bibliotheksverband.de
3. bis 8. September *Berlin*	IFA 2010	Messe Berlin GmbH, Messedamm 22, 14055 Berlin, Telefon 030 3069-6924, Fax 030 3038-2401, ifa@messe-berlin.de
6. bis 10. September *Glasgow, Großbritannien*	ECDL2010 12th European Conference on Research and Advanced Technology for Digital Libraries	Department of Computing Science, University of Glasgow, 18 Lilybank Gardens, Glasgow G12 8QQ, Great Britain, info@ecdl2010.org, www.ecdl2010.org
22. bis 24. September *Essen*	Jahrestagung der Association Internationale des Bibliothèques, Archives et Centres de Documentation Musicaux, Gruppe Bundesrepublik Deutschland	Verena Funtenberger, Stadtbibliothek Essen, Musikbibliothek der Zentralbibliothek, Hollestr. 3 (Gildehof), 45127 Essen, Telefon 0201 88-42280, verena.funtenberger@stadtbibliothek.essen.de, www.aibm.info
27. September bis 2. Oktober *Leipzig*	40. Jahrestagung der Gesellschaft für Informatik e.V. INFORMATIK 2010 - Service Science - Neue Perspektiven für die Informatik	Universität Leipzig, Fakultät für Mathematik und Informatik , Alexandra Gerstner, Telefon 0341 97-32223, Fax 0341 97-32329, info@informatik2010.de, www.informatik2010.de
29. September bis 1. Oktober *Middelburg, Niederlande*	Society of Indexers Conference 2010 The challenging future of indexing	Society of Indexers, Woodbourn Business Centre, 10 Jessell Street, Sheffield S9 3HY, Telefon +44 (0)114 244 9561, Fax +44 (0)114 244 9563, 2010conference@indexers.org.uk, www.indexers.org.uk
29. September bis 2. Oktober *Dresden*	80. Deutscher Archivtag 2010	VdA - Verband deutscher Archivarinnen und Archivare e.V., Geschäftsstelle, Wörthstraße 3, 36037 Fulda, Telefon 0661 29109-72, Fax 0661 29109-74, info@vda.archiv.net
6. bis 10. Oktober *Frankfurt am Main*	Frankfurter Buchmesse Ehrengast Argentinien	Ausstellungs- und Messe GmbH, Buchmesse Frankfurt, Reineckstraße 3, 60313 Frankfurt am Main, Telefon 069 2102-0, Fax 069 2102-227/-277, info@book-fair.com, www.buchmesse.de
7. bis 9. Oktober *Frankfurt am Main*	62. Jahrestagung der DGI DGI-Konferenz 2010 Semantic Web & Linked Data	DGI-Geschäftsstelle, Hanauer Landstraße 151-153, 60314 Frankfurt am Main, Telefon 069 430313, Fax 069 4909096, mail@dgi-info.de, www.dgi-info.de
11. bis 13. Oktober *Berlin*	Herbsttagung der Fachgruppe Dokumentation im Deutschen Museumsbund	Prof. Monika Hagedorn-Sauppe, Institut für Museumsforschung, In der Halde 1, 14195 Berlin, Telefon 030 8301-460, Fax 030 8301-504, m.hagedorn@smb.spk-berlin.de, www.museumsbund.de
4. bis 6. November *Berlin*	Arbeitssitzung der Arbeitsgruppe „Elektronische Medien in der Patentinformation“	DGI-Geschäftsstelle, Hanauer Landstraße 151-153, 60314 Frankfurt am Main, Telefon 069 430313, Fax 069 4909096, mail@dgi-info.de, www.dgi-info.de
8. bis 10. November *Jülich*	WissKom2010 – „eLibrary — den Wandel gestalten“	Edith *Salz*, Forschungszentrum Jülich, 52425 Jülich, Telefon 02461 61-2907, Fax 02461 61-6103, zb-konferenz@fz-juelich.de, www.wisskom2010.de
10. bis 11. November *Stuttgart*	Deutscher Multimedia Kongress 2010 Digitale Wirtschaft	MFG Medien- und Filmgesellschaft Baden-Württemberg mbH, Breitscheidstraße 4, 70174 Stuttgart, Telefon 0711 90715-300, Fax 0711 90715-350, info@dmmk.de, www.mfg-innovation.de
16. bis 17. November *Stuttgart*	6. Stuttgarter Wissensmanagement-Tage	Oliver Lehnert, wissensmanagement, Schertlinstraße 21, 86159 Augsburg, Telefon 0821 45360-55, Fax 0821 450360-10, service@wima-tage.de, www.wima.tage.de
30. November bis 2. Dezember *London, England*	Online Information 2010	VNU Exhibitions Europe, 32-34 Broadwick Street, London, W1A 2HG, UK, lorna.candy@vnuexhibitions.co.uk, www.online-information.co.uk/
1. bis 3. Dezember *Berlin*	Online EDUCA Berlin 2010 16. Internationale Konferenz für technologisch gestützte Aus- und Weiterbildung	ICWE GmbH, Leibnizstraße 32, 10625 Berlin, Telefon: 030 327 6140, Fax 030 324 9833, info@online-educa.com, www.online-educa.com

Designed for patent searchers everywhere as a one-stop shop

PatBase®

- **Over 28 million patent families since early1900's**
- **Abstracts and full-text**
- **IPC, USPC, ECLA, JP classifications**
- **Updated weekly on publication days with a rapid Alert Service**
- **Integrated PDF copies, legal status, registers, translations, file histories**

Advanced Keyword Highlighting
Patent Family Explorer
Priority Analysis Maps
Citation Finder
....and more



Competitively priced unlimited annual subscription - including unlimited use of specially designed patent viewing and analysis tools:

minesoft
www.minesoft.com

THE QUEEN'S AWARDS FOR ENTERPRISE: INTERNATIONAL TRADE 2009

PatBase has been developed in partnership with RWS GROUP www.rws.com

Try PatBase for Free

Ask for your free trial, free web demonstration or more details

www.patbase.com
info@patbase.com

contact Minesoft on:
+44(0)20 8404 0651 (UK),
(401)737 6499 (USA),
+81 3 5403 9191 (Japan),
+49 (0) 211 15924489 (Germany).

Offices also in China, France, Germany, India, Israel, Italy, Korea, Switzerland, Taiwan.