

Content Repurposing

Ein Überblick

Nima Mazloumi

Januar 2005

Seminararbeit
Content Delivery Networks

Lehrstuhl für Praktische Informatik IV
Prof. Dr. W. Effelsberg
Fakultät für Mathematik und Informatik
Universität Mannheim

Inhaltsverzeichnis

Inhaltsverzeichnis.....	2
Abkürzungsverzeichnis	3
1. Motivation	4
2. Problemstellung.....	4
Inhalte.....	4
Formate.....	5
Ausgabemedien	5
3. Begriffsbestimmung	6
4. Kategorien	6
Transcoding	6
Transducing	7
Transformation	7
Information Retrieval	8
Measurement	8
Intervention	9
Modellbasierte Ansätze	9
Repurposing im Autorenbereich	10
Avatare	10
5. Zusammenfassung	11
Literaturverzeichnis.....	12

Abkürzungsverzeichnis

ADL	Advanced Distributed Learning
BMP	Bitmap
DOC	Microsoft Word Document
GIF	Graphics Interchange Format
HTML	Hypertext Markup Language
HTTP	Hypertext Transfer Protocol
IANA	Internet Assigned Numbers Authority
IEEE	Institute of Electrical and Electronics Engineers
ISO	International Organization for Standardization
JPEG	Joint Photographic Experts Group
LOM	Learning Objects Metadata
MPEG	Moving Picture Experts Group
PC	Personal Computer
PDA	Personal Data Assistent
PDF	Portable Document Format
PIM	Platform Independent Model
PNG	Portable Network Graphics
PPT	Microsoft Powerpoint Document
PSM	Platform Specific Model
SCORM	Sharable Content Object Reference Model
UDDI	Universal Description, Discovery and Integration
WAP	Wireless Application Protocol
XLS	Microsoft Excel Document
XML	Extensible Markup Language
XHTML	Extensible HyperText Markup Language
RTF	Rich Text Format

1. Motivation

Die Untersuchung „How much Information 2003?“ an der Universität von Kalifornien hat für das Jahr 2002 ergeben, dass das Internet allein im Teilbereich Webinhalte (d.h. ohne Email und Instant Messaging) ein Datenvolumen von 92.017 Terabytes erreicht hat. Hinsichtlich der nicht öffentlich zugänglichen bzw. dynamisch generierten Inhalte (Deep Web) kann nur spekuliert werden [1].

Bei der Bereitstellung digitaler Inhalte spielen mehrere Faktoren eine Rolle. Der jeweilige Inhalt muss für eine bestimmte Zielsetzung selektiv ausgewählt und in ein geeignetes Format umgewandelt werden, welches vom jeweiligen Endgerät unterstützt wird.

Soll der Inhalt einem anderen Zweck dienen, in ein anderes Format und für ein anderes Medium aufbereitet werden, dann ist dies Gegenstand des Content Repurposing.

Nach einer Untersuchung des Problembereichs soll der Begriff des Content Repurposing untersucht und abschließend anhand von Fallbeispielen eine Kategorisierung nach Teilbereiche vorgenommen werden.

2. Problemstellung

Zur Verdeutlichung des zu Grunde liegenden Problems soll der Turmbau zu Babel als Gleichnis herangezogen werden [2].

„Es hatte aber alle Welt einerlei Zunge und Sprache. Und sie sprachen: Wohlauf, laßt uns eine Stadt und einen Turm bauen, dessen Spitze bis an den Himmel reiche, damit wir uns einen Namen machen; Daher heißt ihr Name Babel, weil der HERR daselbst verwirrt hat aller Länder Sprache und sie von dort zerstreut hat in alle Länder.“

Übertragen auf das Internet deutet dies, dass ein und dieselbe digitale Botschaft bzw. Information über eine Vielzahl von Trägerformaten (Kodierungen) auf unterschiedlichste Endgeräte bereitgestellt werden kann. Dies kann bei der Umwandlung der Information zu Datenverlusten und Fehlinterpretationen führen. Auf diese drei Faktoren soll im Folgenden näher eingegangen werden.

Inhalte

Um Inhalte in geeigneter Form umwandeln zu können, muss folglich die Bedeutung der Information bekannt sein. Die Information aber ist bereits zum Zeitpunkt Ihrer Erstellung selektiv. Einem Menschen kann sich nachträglich mittels seiner physischen und geistigen Sinne die Bedeutung des Inhalts erschließen – einer Maschine jedoch nicht. Was sie allerdings kann ist mittels vorgegebener Algorithmen und beschreibender Daten diese Informationen aufbereiten. Diese Daten über Daten (Bits about Bits) [3] bzw. Meta-Daten sind nachträglich in geeigneter Qualität schwer zu beschaffen und bedürfen häufig der manuellen Aufbereitung,

sobald eine Umwandlung der Daten notwendig ist, die mehr Information braucht, als aus den originären Daten zu gewinnen ist.

Formate

Zur Verwendung digitaler Inhalte bedarf es eines Trägerformats. Prinzipiell sind sämtliche bei der Internet Assigned Numbers Authority (IANA) gelisteten Formate in den acht Hauptklassen Audio (62 Formate), Anwendungen (399 Formate), Grafiken (36 Formate), Message, Model, Multipart, Video (35) und Text (38) als Informationsträger im Internet denkbar [4 - 6]. Es wurde lediglich eine Zählung für die Multimediaformate durchgeführt. Sowohl innerhalb der Hauptkategorien als auch zwischen den Kategorien ist eine Umwandlung denkbar. Die Anzahl der Umwandlungsmöglichkeiten zwischen diesen möglichen Formaten beträgt daher bei N Formaten N^2-N .

Die Menge der Formate lässt sich jedoch auf eine gängige Anzahl reduzieren. Eine Untersuchung der Lernplattform dotLRN [7] mit täglich 435 Megabyte Datentransfer ergab für die Universität Mannheim folgende gängigen Formate:

61.61%	HTML (statisch, dynamisch)
18.93%	Adobe Portable Document Format
8.7%	Grafikformate JPEG, GIF, PNG, BMP
5.14%	Archive (ZIP)
2.16%	Office (DOC, XLS, PPT)
0.78%	Video (AVI)
0.51%	Anwendungen (EXE, JAR)
2.17%	Sonstige Formate

Zudem wurden 97,69% aller Inhalte von Windows-, nur 1,77% von Unix- und 0,7% der Inhalte von Macintosh-Systemen nachgefragt. Aufgrund der Besonderheit von dotLRN ist diese Untersuchung nicht notwendigerweise repräsentativ, allerdings lassen sich dennoch wesentliche Vereinfachungen ableiten:

- Für den größten Teil der Benutzer sind nur die Formate von Bedeutung, die die größte Verbreitung haben.
- Es ist davon auszugehen, dass das Internet zu einem wesentlichen Teil redundante Informationen in unterschiedlicher Kodierung und Granularität enthält.
- Das Internet enthält Inhalte, die lediglich von der Referenzierungsfunktion von Hypertext Gebrauch machen, sodass spezielle Anwendungen notwendig sind, damit der Informationssuchende diese Daten auswerten kann (z.B. Microsoft Office, Adobe Acrobat).

Es lässt sich somit eine Konsolidierung im Bereich der Formate beobachten.

Ausgabemedien

Bis vor einigen Jahren war das Internet im Wesentlichen für PCs mit integrierten Browsern zugänglich. Mit dem Durchbruch funkbasierter Netzwerke wird das Internet hunderten von neuen Geräten zugänglich [8a, nach 8b]. Insbesondere handelt es sich um mobile Geräten wie Handys, Personal Data Assistents (PDAs), Tablet- und Pocket-PCs, digitale Posters und Terminals [18, 19].

Die Geräte stellen unterschiedliche Anforderungen und Voraussetzungen [14 - 19] an die Verarbeitung der Inhalte. Sie unterscheiden sich beispielsweise hinsichtlich der Displaygröße, Farbtiefe, Interaktionsschnittstelle, Multimediafähigkeit, Bandbreite sowie verfügbarer Anwendungen. Diese Merkmale lassen sich in so genannte Geräteprofile zusammenfassen und können zur Anpassung der Inhalte dienen [14]. Nach Ansicht des Autors wird mit der Verbesserung und Verbreitung dieser Systeme eine ähnliche Konsolidierung im Bereich der eingesetzten Betriebssysteme und installierten Anwendungen erfolgen. Aufgrund der Eigenheiten dieser Geräte lässt sich jedoch das Hypertext-Paradigma nicht ohne weiteres auf diese Geräte übertragen [16 - 19].

3. Begriffsbestimmung

Der Begriff des Content Repurposing lässt sich wie folgt übersetzen: die Umwandlung eines Inhalts gemäß eines neuen Zwecks.

Nach Doerksen ist Repurposing ein Teilgebiet der Wiederverwendung von Inhalten (Content Reuse) [11]. Entlang einer diskreten Transformationsachse wird die Wiederverwendung in die Teilbereiche Sharing (also ohne nennenswerte Veränderung), Multipurposing (für ein bestimmtes Ausgabemedium [14 - 19]) und Repurposing (entsprechend den Informationsbedürfnissen des Nutzers [27, 30 - 37]) unterteilt.

Bei [12] ist eine Unterscheidung zwischen Multipurposing und Repurposing nicht zu finden. Hier steht die Automatisierung der Umwandlung und die Verwendung einer einzigen Datenquelle (Single Sourcing) im Vordergrund. Einigkeit besteht jedoch darin, dass eine Umwandlung aufgrund neuer Anforderungen (Zweck) notwendig wird.

4. Kategorien

Dreh- und Angelpunkt beim Content Repurposing ist folglich der zu Grunde gelegte Zweck. Eine Untersuchung der verschiedenen Ansätze hat jedoch ergeben, dass der Zweck der Umwandlung sehr breit gefasst wird, sodass Content Repurposing sich auf viele Bereiche beziehen kann. Im Folgenden soll in Anlehnung an [16] der Versuch einer erweiterten Kategorisierung vorgenommen werden.

Transcoding

Hierbei handelt es sich um die Überführung der Daten von einem Quell- in ein Zielformat. Wie zuvor erläutert ist die Zahl der Umwandlungsmöglichkeiten sehr hoch. Diese lassen sich entsprechend der Medientypen in die Kategorien Audio, Grafik, Nachricht, Modell, Multipart, Video [5] und Anwendungen [6] unterteilen. Die tatsächliche Umwandlung erfolgt durch spezialisierte Konverter.

Da Webseiten Inhalte mit unterschiedlichen Formaten enthalten, schlagen [13] einen service-orientierten Ansatz beim Transcoding vor. Dabei wird das Transcoding in die einzelnen Konversionsschritte heruntergebrochen. Diese sollen durch entsprechende

Web Services umgesetzt werden. Die Web Services werden im Anschluss katalogisiert (über einen sog. UDDI-Dienst) und damit anderen Anwendungen bereit gestellt.

Einen Teilbereich des Transcoding im Bereich von Grafiken und Videodaten bildet die Skalierung (Scaling), die sich mit der Anpassung von Inhalten an das jeweilige Geräteprofil beschäftigt.

[14] führt beispielsweise die automatische Anpassung von Grafiken anhand des Geräteprofils des Nutzers durch. Die Besonderheit liegt darin, dass das Transcoding nicht mehr autonom durch ein Proxy erfolgt, sondern der Server dem Proxy durch Anweisung über Art und Umfang des Transcoding informiert. Hierzu wird ein Kommunikationsprotokoll basierend auf HTTP-Request-Headerzeilen vorgeschlagen.

In [15] wird der Webinhalt als Miniaturansicht auf dem PDA dargestellt. Die Interaktion mit der Seite erfolgt durch spezielle Werkzeuge, die dem Benutzer zur Verfügung stehen. Diese umfassen u.a. das Ausschneiden von Inhalten, Zoomen in die Seite, Manövrieren über die Seite und Verfolgen von Links.

Transducing

Wird der Inhalt strukturell aufbereitet, spricht man von Transducing [16]. Hierbei wird die Struktur einer Webseite verändert, um beispielsweise den Einschränkungen aufgrund der Displaygröße des Endgeräts Rechnung zu tragen.

[17] stellen beispielsweise ein multimodales Navigationssystem vor. Der Fahrer wählt zunächst die Nummer des Dienstbetreibers und übermittelt mittels Spracheingabe Informationen zum Stand- und Zielort. Das System leitet dann diese Anfrage an ein öffentliches Routensystem im Internet weiter und übermittelt das Ergebnis nach dessen Aufbereitung an den PDA des Fahrers. Dieser Ansatz erscheint nur dann sinnvoll, wenn kein Navigationssystem zum Einsatz kommen kann.

Transformation

Bei der Transformation werden die Interaktionsmöglichkeiten mit dem Inhalt angepasst [16]. Da unterschiedliche Geräte entsprechend andere Bedienmetapher erwarten, kann das Hypertext-Paradigma nicht unmittelbar übernommen werden. Die Transformation beschäftigt sich mit der Umwandlung des Inhalts in neue Navigation- und Interaktionsformen.

[18, 19] schlägt daher für digitale Poster die Umsetzung des Model-View-Controller-Patterns [20] innerhalb der Browserkomponente vor. Durch einen objekt-orientierten Ansatz kann das Navigations- und Interaktionsverhalten durch Überschreibung der jeweiligen Methoden auf den jeweiligen Gerätetyp angepasst werden. So lassen sich unterschiedliche Bedienmetaphern umsetzen. Mit der Trennung der Darstellungsschicht können von Designern neue Benutzeroberflächen (sog. Skins) erstellt werden, wodurch verschiedene Konfigurationen für die unterschiedlichen Geräte möglich werden (sog. Customizing).

Für mobile Geräte mit kleinen Displays schlagen [16] die Übersetzung der Webseiten in navigierbare Verzeichnisse vor. Diese bestehen entweder aus Knoten

oder Blätter. Neben bestehenden Referenzen werden zusätzlich virtuelle Links erzeugt. Die werden durch entsprechende Detektoren erstellt, die auf bestimmte Inhalte spezialisiert sind, wie beispielsweise für die Extraktion von Telefonnummern und Anschriften. Die Interaktion mit den einzelnen Knoten und Blättern erlaubt ein adaptives Menü mit Funktionen wie „Lesen“, „Öffnen“, „Versenden“, „Drucken“. Die Aktionen sind nicht fest in das System implementiert. Sie werden vielmehr mittels XML mit den jeweiligen Linktypen verknüpft.

Information Retrieval

Mit dem Anwachsen des Datenvolumens verliert nach Ansicht von [21] der Inhalt an wirtschaftlicher Bedeutung wohingegen semantisch bedeutsame Daten über Daten wertschöpfend sind. Die Gewinnung dieser Daten aus bestehenden Inhalten ist ein breites Forschungsgebiet. Das Information Retrieval bietet sozusagen die Grundlage für das Content Repurposing.

Die Speicherung der gewonnenen Daten kann einerseits anhand selbstdefinierter Formate, andererseits durch den Einsatz verbreiteter Standards wie beispielsweise ISO MPEG-7 und MPEG-21 im Bereich von Video-Daten [28, 29], Dublin Core für die Katalogisierung digitaler Ressourcen oder IEEE LOM [22, 32] bzw. ADL SCORM [23] für Lernressourcen erfolgen.

Hinsichtlich der Datengewinnung wird die Unterscheidung zwischen „menschlicher Intervention“ (Intervention) und „maschinelle Messung“ (Measurement) vorgeschlagen [21].

Measurement

Bei der maschinen gesteuerten Gewinnung von Daten kommen unterschiedliche Verfahren zum Einsatz. Diese werden anhand von einigen Beispielen erläutert.

Mit der Verbreitung von Handykameras und der Übertragung aufgenommener Bilder über das Internet steigt die Notwendigkeit der bedeutsamen Verwaltung der Aufnahmen. [24] verfolgt einen kontextbasierten Ansatz zur Gewinnung von Metadaten zum Zeitpunkt der Erstellung der Aufnahmen. Informationen wie der Standort und das Benutzerprofil des Anwenders werden untersucht und ausgewertet.

[25] untersucht, inwieweit layout-basierte Formate wie HTML in semantische und strukturierte Daten (XML) umgewandelt werden können. Dabei wird das System mittels bestehender Quell-Ziel-Daten trainiert. Das System wird im Anschluss zur Gewinnung der Daten aus dem Quellsystem herangezogen. Dieser Ansatz eignet sich allerdings nur für Webseiten mit identischem Aufbau.

[26] schlägt einen mehrstufigen Prozess für die funktionale Kategorisierung von Grafiken auf Webseiten vor. 7 Kategorien werden unterschieden: Grafiken der Kategorie Handlung (Story), Vorschau (Preview) und Veranstalter (Host) gehören zu dem übergeordnete Kategorien „Fotografien“, Werbung (Commercial) Icons und Logos, Überschriften und Formatierungselemente zu den „Grafiken“. Neben 18 Frequenzmaßen kommen auch 2 Farbmaße zum Einsatz. Bislang ist die Unterscheidung in die zwei übergeordneten Kategorien und die Trennung zwischen den Kategorien Handlung, Vorschau und Veranstalter gelungen. Die gewonnenen

Informationen können bei der Übertragung der Grafiken bei geringer Bandbreite hilfreich sein, da bedeutsamere Grafiken Vorrang bekommen.

[27] nutzt Texterkennungsalgorithmen im Vor- und Nachspann von Videodaten zur deren Indizierung. Diese Informationen können Interessenten bei der Recherche unterstützen.

Intervention

Es existieren aber auch Ansätze, bei denen Inhaltsersteller bei der Erzeugung von Metadaten involviert sind. Hierzu sollen auch hybride Ansätze, also jene, die maschinell unterstützt werden, gezählt werden.

In [28] wird bei der Anreicherung von Videodaten mit Metadaten die Spracherkennung eingesetzt. Nachträgliche Anmerkungen des Autors werden gemäß eines vordefinierten Vokabulars erkannt und mittels MPEG-7 erfasst. Diese Daten werden der Recherche und der Katalogisierung der Videodaten zur Verfügung gestellt.

MPEG-7 und MPEG-21 kommen bei [29] zum Einsatz. Nachdem Autoren das Videomaterial mit Metadaten angereichert haben, kann das System die Inhalte an die jeweiligen Anfrage des Benutzers und unter Verwendung seines Geräte- und Benutzerprofils anpassen.

Modellbasierte Ansätze

Eine Vielzahl von Ansätzen beschäftigt sich mit der modellbasierten Erstellung von Inhalten. Insbesondere im Bereich des Publishing sind diese Ansätze verbreitet [30, 31] Gleiches gilt für das E-Learning. Hierbei liegt der Fokus auf die Anpassung der Lerninhalte in Abhängigkeit vom Kenntnisstand des Studenten [32 - 34].

Auf Unix-Systemen ist der Einsatz von DocBook bei der Dokumentation von Softwareprojekten verbreitet [30]. Dazu wird die Dokumentation als ein XML-Dokument erstellt. Das Dokument bildet die Basis für die Erstellung unterschiedlichster Derivate wie (X)HTML, RTF, PDF, PostScript, JavaHelp und vieles mehr. Einen ähnlichen Ansatz beschreibt [31]. Gemeinsam ist beiden der Einsatz von XSLT zur Transformation der Daten in beliebige andere Formate.

[32] schlagen beispielsweise einen modellbasierten Ansatz für die dynamische Bereitstellung von Kursen vor. Die Bereiche Inhalt, Student und Lehrpfad werden zunächst getrennt voneinander modelliert. Zur Beschreibung der Inhalte kommt ein erweiterter IEEE LOM Standard zum Einsatz. Die Autoren aus den jeweiligen Fachbereichen definieren im Anschluss Lernpfade durch das entsprechende Themengebiet. Diese beziehen den Wissenstand des Studenten mittels fachbereichsspezifischer Vokabulare mit ein. Zur Laufzeit wird aus dem Benutzerprofil des Studenten, seines aktuellen Wissens und seiner Präferenzen bzgl. gewünschter Inhalte und Art der Darstellung ein möglichst passender Kurs zusammengestellt. Ähnliche Ansätze beschreiben [33] und [34].

Ebenfalls einen modellbasierten Ansatz verfolgen [35] für die Bereitstellung multimedialer Inhalte. Während bei der Umwandlung der Daten zwischen zwei Plattformen so genannte Bridges eingesetzt werden, zielt dieser Ansatz auf die

Verwendung eines Multimedia Metamodels, welches die Inhalte auf einem hohem Abstraktionsniveau beschreibt. Dieses plattformunabhängige Modell (PIM) kann durch ein entsprechendes Mapping bidirektional in ein plattformspezifisches Modell (PSM) überführt werden. Dadurch lassen sich ab vier Plattformen Umwandlungsschritte einsparen. Während bei der direkten Umwandlung auf der PSM-Ebene bei N Plattformen N^2 -N Brigdes denkbar sind, ist bei der Überführung aller PSM in das PIM lediglich $N*2$ Umwandlungen nötig, da das erzeugte PIM in beliebige andere PSM umgewandelt werden kann. Dieser Ansatz aus dem Bereich der modellbasierten Softwareentwicklung ist nach Ansicht des Autors für den Bereich der Inhaltserstellung nur bedingt einsetzbar, da er im Detail viel zu komplex ist.

Repurposing im Autorenbereich

Ausgehend von einer bestehenden Datenbasis erlauben bestimmte Autorensysteme die Erstellung von Metadaten, die den originären Daten eine neue Bedeutung geben.

Beispielsweise stellt [36] einen neuen Ansatz für die Erstellung von Lesezeichen vor. Während gängige Browser lediglich die Erstellung von Lesezeichen für eine ganze Webseite unterstützen, erlaubt dieses System durch den Einsatz eines Proxyservers die Referenzierung von Abschnitten der besuchten Webseiten. Diese lassen sich zu so genannten Collections zusammenführen. Die Zusammenstellungen können sich auf eine unbegrenzte Anzahl besuchter Seiten beziehen, ohne dass eine Kopie der Inhalte notwendig ist. Dadurch kann der Benutzer Webinhalte nach eigenen Vorstellungen zusammenstellen. Anschließend kann er diese Collection anderen Interessenten über einen einfachen Link zur Verfügung stellen.

Einen ähnlichen Ansatz verfolgt [37]. Ausgangslage bildet die Erstellung von Manuskripten, wie sie bei der Erstellung von wissenschaftlichen Arbeiten üblich sind. Mittels eines Autortools kann der Nutzer Informationseinheiten (Items) definieren. Diese verfügen über einen Titel, etwaige Annotationen und einer Referenz auf ein so genanntes Basisdokument in den Formaten HTML, PDF, Excel, Word. Diese Referenzen werden auch Markierungen genannt. Eine Markierung stellt einen Verweis auf einen bestimmten Abschnitt in der Datei dar. Items können in logische Gruppen zusammengefasst und verschachtelt werden. Daraus entsteht schließlich ein virtuelles Manuskript, welches als Ausgangslage für die Erstellung der wissenschaftlichen Arbeit dienen kann.

Im Bereich von Videodaten existieren ähnliche Ansätze [38]. Hierbei kann der Autor mittels einer virtuellen Kamera Sequenzen aus einem hochwertigen Videofilm aufzeichnen und diese mit Anmerkungen versehen. Diese virtuellen Aufzeichnungen werden auch Dives genannt. Es werden lediglich Informationen über Position, Größe und Dauer der Aufzeichnung gespeichert. Die Dives können interessierten Personen über das Internet verfügbar gemacht werden.

Avatare

In Anwendungen kommen verstärkt so genannte Avatare zum Einsatz. Hierbei handelt es sich um Dienste, die den Benutzer in seiner Arbeit unterstützen. Ein Teilbereich der Avatare bilden so genannte Kundschafter (Reconnaissance Agents). Diese sind auf die Beobachtung des Internetnutzers spezialisiert. Während gängige Browseranwendungen lediglich die Historie einer Internetsitzung speichern

untersuchen diese Kundschafter den Nutzer während einer Sitzung [39]. Zwei Typen von Kundschaftern werden unterschieden. Jene, die das nähere Umfeld einer besuchten Seite nach interessanten Informationen untersuchen und jene, die zentrale Informationssysteme verwenden (z.B. Suchmaschinen). In beiden Fällen werden dem Benutzer zusätzliche Links und Inhalte für das weitere Vorgehen vorgeschlagen. Aufgrund der gesammelten Informationen erstellen diese Anwendungen des Weiteren Benutzerprofile, die für künftige Erkundungen eingesetzt werden.

5. Zusammenfassung

Im Mittelpunkt des Content Repurposing steht der jeweilige Zweck, der die Umwandlung eines Inhalts bedingt. Unzählige Zielsetzungen sind denkbar – die Erschließung neuer Endgeräte, Katalogisierung von Daten, Adaption an das Benutzerverhalten oder dessen Vorwissen und die Erschließung neuer Einsichten.

Einen Einfluss auf das jeweilige Verfahren haben das Zielformat, das jeweilige Endgerät aber auch der Benutzerprofil. Allen Verfahren ist gemeinsam, dass die Qualität der Umwandlung durch die Bereitstellung von Metadaten oder durch ihre nachträgliche Generierung verbessert werden kann.

Während bei den modellbasierten Ansätzen das Repurposing von oben nach unten erfolgt (Top-Down), kommen bei fremden Datenbeständen Bottom-Up-Ansätze zum Tragen.

Schließlich existieren neben automatisierten Ansätzen auch Autorensysteme, die Nutzern die Umwidmung bestehender Inhalte erlauben.

Literaturverzeichnis

- [1] P. Lyman, Hal R. Varian, et al., *How Much Information 2003?*, Universität Kalifornien, SIMS Berkley, http://www.sims.berkeley.edu/research/projects/how-much-info-2003/printable_report.pdf
- [2] Altes Testament, Genesis, Kapitel 11,1
- [3] N. Negroponte – *Being Digital*, Vintage Books, New York, 1996
- [4] Media Mime Types, Internet Assigned Numbers Authority (IANA), <http://www.iana.org/assignments/media-types/>
- [5] Network Working Group, Request for Comments: 1521 - *MIME (Multipurpose Internet Mail Extensions)- Teil 1*, <http://rfc.sunsite.dk/rfc/rfc1521.html>
- [6] Network Working Group, Request for Comments: 2046 - *Multipurpose Internet Mail Extensions (MIME) – Teil 2*, <http://rfc.sunsite.dk/rfc/rfc2046.html>
- [7] Universität Mannheim, Access Log Report, <http://dotlrn.uni-mannheim.de/log/traffic>
- [8a] G. Singh, A. Das, R. Suh – *Content Repurposing for Damage Control onboard Navy Ships* – 2003, IEEE Call for Papers (eingereicht, noch nicht veröffentlicht)
- [8b] R. J. Suh – *Wireless Content Repurposing Architecture for DC Command and Control*, Thesis, Naval Postgraduate School, September 2003
- [11] T. Doerksen - *Content Repurposing Theory*, University of Calgary <http://socrateslearning.com/media/CRT.ppt>
- [12] G. Singh - *Content Repurposing*, Multimedia, IEEE , Band: 11 , Ausgabe: 1 , Jan-Mar 2004, S. 20 - 21
- [13] M. S. Hossain, M. A. Rahman, A. El Saddik - *A Framework for Repurposing Multimedia Content*, Electrical and Computer Engineering, 2004. Canadian Conference on Electrical and Computer Engineering, Niagara, Ontario, Canada, Band 2, Mai 2004, S. 0971 - 0974
- [14] B. Knutsson, H. Lu, J. Mogul, B. Hopkins - *Architecture and Performance of Server-Directed Transcoding*, ACM Transactions on Internet Technology (TOIT), Volume 3 Issue 4, November 2003, S. 392 - 424
- [15] Jacob O. Wobbrock, Jodi Forlizzi, Scott E. Hudson, Brad A. Myers - *WebThumb: Interaction Techniques for Small-Screen Browsers*, Proceedings

vom 15th Annual ACM Symposium on User Interface Software and Technology, Oktober 2002, S. 205 – 208

- [16] B. N. Schilit, J. Trevor, D. M. Hilbert, T. K. Koh - *Web Interaction Using Very Small Internet Devices*, Computer , Band: 35 , Ausgabe: 10 , Oktober. 2002, S. 37 – 45
- [17] A. Morde, R. S. Kashi, M. K. Brown, D. Grove, J. L. Flanagan - *A Multimodal System for Accessing Driving Directions*, DAS 2002, LNCS 2423, 2002, S. 556 – 567
- [18] L. Nelson, E. F. Churchill, L. Denoue, J. Helfman, P. Murphy - *Gooey Interfaces: An Approach for Rapidly Repurposing Digital Content*, Human Factors and Computing Systems Konferenz, April 2004, S. 1293 - 1296
- [19] E. F. Churchill, L. Nelson, L. Denoue, J. Helfman, P. Murphy - *Sharing Multimedia Content with Interactive Public Displays: A Case Study*, Proceedings der Konferenz Designing Interactive Systems: Processes, Practices, Methods, and Techniques August 2004, S. 7 - 16
- [20] F. Buschmann, R. Meunier, H. Rohnert, P. Sommerlad, M. Stal - *Pattern-Oriented Software Architecture: A System of Patterns*. John Wiley & Sons Ltd., 1996.
- [21] K. Haase - *Context for Semantic Metadata*, Proceedings der 12th Annual ACM International Conference on Multimedia, Oktober 2004, S. 204 - 211
- [22] Intelligent Manufacturing Systems - *Learning Objects Metadata, Question & Test Interoperability Specification und Simple Sequencing*, <http://www.imsglobal.org>
- [23] Advanced Distributed Learning - *Sharable Content Object Reference Model (SCORM) 2004 2nd Edition*, <http://www.adlnet.org/index.cfm?fuseaction=SCORMDown>
- [24] R. Sarvas, E. Herrarte, A. Wilhelm, M. Davis, *Metadata Creation System for Mobile Images*, Proceedings der 2nd International Conference on Mobile Systems, Applications, and Services, Juni 2004
- [25] B. Chidlovskii, J. - *Supervised Learning for the Legacy Document Conversion*, ACM Symposium on Document Engineering, Oktober 2004, S. 220 - 228
- [26] J. Hu, A. Bagga - *Categorizing Images in Web Documents*, Multimedia, IEEE , Band: 11 , Issue: 1 , Januar - März 2004, S. 22 - 30
- [27] R. Lienhart - *Automatic Text Recognition for Video Indexing*, Proceedings der 4th ACM international Conference on Multimedia, Februar 1997, S. 11 - 20
- [28] D. Poncelon, S. Srinivasan, A. Amir, D. Petkovic, D. Diklic - *Key to Effective*

- Video Retrieval: Effective Cataloging and Browsing*, Proceedings der 6th ACM International Conference on Multimedia, September 1998, S. 99 - 107
- [29] B. L. Tseng, C.-Y. Lin, and J. R. Smith - *Using MPEG-7 and MPEG-21 for Personalizing Video*, IEEE Computer Society, 2004, S. 62 - 71
- [30] *The DocBook Document Type, Working Draft 4.4CR3*, 17 Dec 2004, <http://www.docbook.org/specs/index.html>
- [31] M. Dunn - *Single-Source Publishing with XML*, IT Professional , Band: 5 , Ausgabe: 1 , Januar 2003 S. 51 – 54
- [32] O. Conlan, V. Wade , C. Bruen, M. Gargan - *Multi-model, Metadata Driven Approach to Adaptive Hypermedia Services for Personalized eLearning*, Malaga, Spain, Mai 2002, S. 100 - 111
- [33] F.P. Rokou, Y. Rokos, E. Rokou - *Infolab: a Web Learning Pedagogical Based Content Repurposing Approach*, Advanced Learning Technologies, 2003. Proceedings der 3rd IEEE International Konferenz on , 9-11 Juli 2003, S. 150 - 154
- [34] P. de Bra, P. Brusilovsky, R. Conejo - *Adaptive Hypermedia and Adaptive Web-Based Systems: Second International Conference*, AH 2002 Malaga, Spain, Mai 29 - 31, 2002. Proceedings, August 2003
- [35] Z. Obrenovic, D. Starcevic, B. Selic – *A Model-Driven Approach to Content Repurposing*, Multimedia, IEEE , Band: 11 , Ausgabe: 1 , Januar - März 2004, S. 62 - 71
- [36] M. C. Schraefel, Y. Zhu, D. Modjeska, D. Wigdor, S. Zhao - *Hunter Gatherer: Interaction Support for the Creation and Management of Within-Web-Page Collections*, Proceedings der 11th International Conference on World Wide Web, Mai 2002, S. 172 - 181
- [37] S. Murthy, D. Maier, L. Delcambre - *Querying Bi-level Information*, Proceedings der 7th International Workshop on the Web and Databases, Juni 2004, S. 7 - 12
- [38] R. Pea, M. Mills, J. Rosen, K. Dauber, W. Effelsberg , E. Hoffert. - *The Diver Project: Interactive Digital Video Repurposing*, Multimedia, IEEE , Band: 11 , Ausgabe: 1 , Januar - März 2004, S. 54 - 61
- [39] H. Lieberman, C. Fry, L. Weitzman - *Exploring the Web with Reconnaissance Agents*, Communications of the ACM, Band 44, Ausgabe 8, August 2001, S. 69 - 75