

Xinnovations e.V. (Hrsg.)

Studie Semantic Media Web

erstellt unter Federführung von

Dr. Joachim Quantz

und Mitarbeit von

Wiebke Lang, Felix Daub, Stefan Grill, Rainer Thiem

GEFÖRDERT VOM



Bundesministerium
für Bildung
und Forschung

INNOVATIONSFOREN
UNTERNEHMEN 
Die BMBF-Innovationsinitiative
Neue Länder **REGION**

IMPRESSUM

HERAUSGEBER:

Xinnovations e. V.

Kleiststr. 23-26, D - 10787 Berlin

Tel.: +49.30.21001-470

Fax.: +49.30.21001-555

Email: info@xinnovations.org

www.xinnovations.org

Amtsgericht Charlottenburg HRB VR 21550 B

VORSTANDSVORSITZENDER:

Rainer Thiem

Email: rainer.thiem@xinnovations.org

PRESSE / BESTELLUNGEN:

Stefan Grill, PR-Manager Xinnovations e. V.

Email: stefan.grill@xinnovations.org

GESTALTUNG :

Wiebke Lang und Dr. Joachim Quantz, Berlin

PROGRAMMIERUNG:

3pc GmbH Neue Kommunikation, Berlin

BILDNACHWEIS FÜR ALLE ABBILDUNGEN:

Xinnovations e. V.

Die Studie wurde finanziert im Rahmen des Programms »Innovationsforen« als Teil der Initiative »Unternehmen Region« des Bundesministeriums für Bildung und Forschung (BMBF), eine Fördermaßnahme des BMBF in Zusammenarbeit mit dem Land Berlin

BERLIN 2013

ISBN 978-3-00-041115-1

Executive Summary

Die vorliegende Studie wurde im Rahmen des vom BMBF geförderten Innovationsforums „Semantic Media Web“ erstellt und untersucht das Potenzial des Semantic Media Web (SMW) in Bezug auf neue Geschäftsmodelle für die Medienbranche und Wissensarbeiter. Sie stützt sich zum einen auf Recherchen der Autoren, zum anderen auf Vorträge und Diskussionen im Rahmen der zweitägigen Forenveranstaltung im Oktober 2012 und des Statusseminars im Dezember 2012 mit insgesamt 300 Teilnehmern. Darüber hinaus wurden 20 etwa einstündige Interviews mit ausgewählten Experten aus unterschiedlichen Branchen geführt.

Anhand der Anwendungsfelder Verlage, Kulturgüter digital, Bibliotheken, Bewegtbild, Open Data/Government und Corporate Communication untersucht die Studie, welche Herausforderungen sich der Medienbranche aktuell stellen und welche Lösungsmöglichkeiten Technologien im Bereich Semantic Media Web hierzu anbieten können. Sie beleuchtet Berufsfelder von Verlegern, Redakteuren und Journalisten, Wissenschaftlern, Kuratoren, Archivaren, Bibliothekaren, Filmemachern und ähnlichen publizierenden oder vermittelnden Tätigkeiten. Dabei wurden insbesondere auch die Auswirkungen des Einsatzes von Technologien und Lösungen des SMW auf Anforderungsprofile für Wissensarbeiter, Prozessabläufe im Arbeitsalltag und Wertschöpfungsketten in der Medienbranche betrachtet.

Für den Zweck der Studie wurde eine sehr weite Interpretation des Begriffs Semantic Media Web gewählt, die alle Lösungen umfasst, in denen Technologien zur Veredelung von digital erstellten Inhalten zum Einsatz kommen, um diese besser suchen, bündeln und auswerten zu können. Unabhängig davon, ob diese Technologien RDF oder OWL verwenden oder als semantische Technologien bezeichnet werden.

Die Analyse der einzelnen Anwendungsfelder zeigt, dass das Semantic Media Web ein sehr hohes Potenzial besitzt. Es eröffnet der Medienbran-

che und Wissensarbeitern neue Möglichkeiten der Wertschöpfung und die nachhaltige Etablierung innovativer Mehrwertdienste. Dies beinhaltet insbesondere auch die Optimierung von Prozessen und die (Teil-)Automatisierung einzelner Prozessschritte. Darüber hinaus unterstützt das Semantic Media Web aktuelle Konvergenztrends und die Entwicklung hin zu integrierten Multimediaformaten. Schließlich ermöglicht es eine Personalisierung, die theoretisch beliebig feingliedrig sein kann. Inhalte können somit für die jeweiligen Zielgruppen optimiert werden, wobei eine Reduktion der Zielgruppengröße bis auf Einzelindividuen möglich ist. Das Semantic Media Web bietet Lösungsansätze für eine der umfassendsten Herausforderungen unserer digitalen Zukunft – also dabei, die zunehmende Informationsflut durch individualisiertes Aggregieren und Filtern handhabbar zu machen.

„Wissensarbeiter könnten das Tor zum Semantic Web sein.“ Kommunikationsbranche

Auch wenn es bereits eine ganze Reihe vielversprechender Ansätze für das Semantic Web gibt, sind noch diverse Herausforderungen zu meistern, die so unterschiedliche Aspekte wie technologische Reife, Kosten-Nutzen-Verhältnis, rechtliche Rahmenbedingungen, offene Formate oder neue Wertschöpfungsketten umfassen. Ein grundsätzlicher Trend geht dabei in Richtung komplexer und umfassender, formatübergreifender Dienstleistungen, weg vom Vertrieb vergleichsweise einfacher Produkte. Dabei werden traditionelle, lineare Wertschöpfungsketten aufgebrochen und durch komplexe Ökosysteme von Dienstleistern ersetzt.

In diesem Zusammenhang wurde insbesondere deutlich, dass die Kommunikation zwischen den einzelnen Akteuren, die teilweise sehr unterschiedliche Hintergründe aufweisen, nicht immer reibungsfrei verläuft. Netzwerkaktivitäten, die die unterschiedlichen Akteure des Semantic Media Web zusammenbringen und den interdisziplinären Austausch zwischen Technologie, Inhalt und Gestaltung kontinuierlich verbessern, kommt daher eine besondere Bedeutung zu. So wie die Medienbranche von einem besseren Wissen über technologische Möglichkeiten profitieren kann, so notwendig ist für Technologieanbieter ein besseres Verständnis von qualitativ hochwertigem Content, Kundenbedürfnissen und involvierten Geschäftsprozessen.

Inhaltsverzeichnis

1. Einleitung	8
2. Kontext und Hintergrund der Studie	10
2.1 Medienbranche	11
2.2 Wissensarbeit	14
2.3 Neue Geschäftsmodelle	18
2.4 Semantische Technologien	21
3. Semantic Media Web	24
3.1 Semantische Suche	24
3.2 Ontologien, Thesauri, Annotationen	26
3.3 Kategorisierung und Clustering	28
3.4 Informationsextraktion	29
3.5 Veredelung von Inhalten	30
4. Ausgewählte Anwendungsfelder für SMW	32
4.1 Verlage	32
4.2 Kulturgüter digital	38
4.3 Bibliotheken	46
4.4 Bewegtbild	50
4.5 Open Data/Open Government	55
4.6 Corporate Communication	61
5. Analyse und Bewertung	72
5.1 Herausforderungen	72
5.2 Potenzial	77
6. Handlungsempfehlungen	79
6.1 Ansätze für Technologieanbieter	79
6.2 Ansätze für Wissensarbeiter	80
6.3 Ansätze für Anbieter von Inhalten	81
6.4. Ansätze für die Politik	82
6.5. Forschungsförderung	83
7. Anhang	84
7.1 Glossar	84
7.2 Impulsgeber und Gesprächspartner der Interviews	90

1. Einleitung

Die vorliegende Studie wurde im Rahmen des vom BMBF geförderten Innovationsforums „Semantic Media Web“ erstellt und untersucht das Potenzial des Semantic Media Web (SMW) in Bezug auf neue Geschäftsmodelle für die Medienbranche und Wissensarbeiter.

In Abschnitt 2 werden zunächst Kontext und Hintergrund der Studie genauer erläutert. Neben einer allgemeinen Beschreibung des Gegenstandsbereichs, der Zielstellung und der verwendeten Methodologie umfasst dies insbesondere eine kurze Charakterisierung der Medienbranche (Abschnitt 2.1) und der Wissensarbeit (Abschnitt 2.2). In Abschnitt 2.3 wird dann diskutiert, was neue Geschäftsmodelle konkret bedeuten, bevor Abschnitt 2.4 einen groben Überblick über semantische Technologien gibt.

Abschnitt 3 skizziert dann die allgemeinen technologischen Lösungsansätze, die dem Semantic Media Web zugrunde liegen. Dabei werden die Grundideen der semantischen Suche (Abschnitt 3.1), von Ontologien, Thesauri und Annotationen (Abschnitt 3.2), der Informationsextraktion (Abschnitt 3.3) sowie der Kategorisierung und des Clustering (Abschnitt 3.4) beschrieben und anhand allgemeiner Use Cases veranschaulicht. Abschnitt 3.5 fasst diese Lösungen unter dem allgemeineren Begriff der Inhaltsveredelung zusammen, der in dieser Studie als generische Beschreibung der heterogenen Funktionalitäten des Semantic Media Web verwendet wird.

Abschnitt 4 bildet den inhaltlichen Schwerpunkt der Studie und analysiert Herausforderungen und Potenzial des Semantic Media Web anhand sechs ausgewählter Anwendungsfelder: Verlage (Abschnitt 4.1), Kulturgü-

ter digital (Abschnitt 4.2), Bibliotheken (Abschnitt 4.3), Bewegtbild (Abschnitt 4.4), Open Data/Government (Abschnitt 4.5) und Corporate Communication (Abschnitt 4.6). Diese Anwendungsfelder decken ein breites Spektrum der Medienbranche ab und bildeten auch die Grundlage für die Strukturierung des Innovationsforums im Oktober 2012 in einzelne Workshops.

Abschnitt 5 fasst die Analysen der Anwendungsfelder zusammen und bewertet Herausforderungen und Potenzial des Semantic Media Web. Aufbauend auf diesen Ergebnissen werden in Abschnitt 6 Handlungsempfehlungen für Technologieanbieter, Medienunternehmen, Wissensarbeiter und Politik präsentiert. Ein Glossar enthält kurze Erläuterungen und weiterführende Links für die wichtigsten Begriffe, die in dieser Studie verwendet werden. Literaturhinweise finden sich jeweils am Ende eines Abschnitts.

„Das Berufsbild unserer Redakteure hat sich komplett geändert, Technologie und Multimedia erhalten eine immer größere Bedeutung.“ Verlagswesen

„Wissensarbeiter sind zentral für die Interpretation von Daten - dabei spielt Semantik immer eine Rolle.“ Agentur

„So wie sich das Social Web wie ein Layer über Kommunikation und Verhalten unserer Gesellschaft gelegt hat, wird sich auch das Semantische Web darüber legen.“

Kreativwirtschaft

2. Kontext und Hintergrund der Studie

Diese Studie wurde im Zeitraum Juli bis Dezember 2012 für das vom BMBF geförderte Innovationsforum „Semantic Media Web“ erstellt. Sie stützt sich zum einen auf Recherchen der Autoren, zum anderen auf Vorträge und Diskussionen im Rahmen der zweitägigen Forenveranstaltung im Oktober 2012 und des Statusseminars im Dezember 2012 mit insgesamt 300 Teilnehmern. Darüber hinaus wurden 20 etwa einstündige Interviews mit Protagonisten der beteiligten Branchen und Anwendungsfelder geführt. Eine Liste mit Impulsgebern und Gesprächspartnern der Experteninterviews findet sich im Anhang (Abschnitt 7.2). Da die Kreativbranche als Innovationstreiber in der deutschen Wirtschaft eine zunehmend wichtige Rolle spielt und darüber hinaus an der Schnittstelle zwischen Medienprodukten und Technik tätig ist, wurden auch Designer und Konzeptioner in die Befragung einbezogen.¹

Im Rahmen der Studie sollten insbesondere folgende Fragestellungen berücksichtigt werden:

- Für welche Herausforderungen, die sich der Medienbranche aktuell stellen, bieten Technologien im Bereich Semantic Media Web Lösungsmöglichkeiten?
- Welche Auswirkungen hätte der Einsatz dieser Technologien und Lösungen auf Anforderungsprofile für Wissensarbeiter, Prozessabläufe im Arbeitsalltag und Wertschöpfungsketten in der Medienbranche?
- Welche Aufwände und Vorteile ergäben sich aus dem Einsatz (Kosten-Nutzen-Analyse)?

Dabei sollten unterschiedliche für Wissensarbeiter relevante Medienformate bzw. Unterbranchen oder Anwendungsfelder berücksichtigt werden, wobei im Zuge des Innovationsforums im Oktober 2012 folgen-

de Strukturierung in sechs ausgewählte Anwendungsfelder erfolgte, die auch die Basis für die vorliegende Studie bilden:

- Verlage
- Kulturgüter digital
- Bibliotheken
- Bewegtbild
- Open Data/Government
- Corporate Communication

Diese Anwendungsfelder decken ein sehr weites und heterogenes Spektrum ab und gehen deutlich über das übliche Verständnis der Medienbranche und Wissensarbeit hinaus. So wird im Bereich Corporate Communication beispielsweise untersucht, inwieweit aktuell bei (Groß-) Unternehmen ein Umbruch im Hinblick auf Kommunikation, PR und Strategieentwicklung stattfindet, der dazu führt, dass diese Aufgaben zunehmend von externen Agenturen übernommen wird. Dabei wird aber nicht das „klassische“ unternehmensinterne Wissensmanagement untersucht, sondern explizit ein potentiell neues Geschäftsfeld für die Medienbranche. Ein Beispiel für diesen Trend ist z.B. die Betreuung von Unternehmensseiten in sozialen Netzwerken wie Facebook – sowohl als Dialogmöglichkeit einer Marke mit ihren potentiellen Zielgruppen als auch das interne Management des unternehmenseigenen Wissens in wissensbasierten Branchen.

In den folgenden Abschnitte werden die unterschiedlichen Bereiche eingeführt, deren Spannungsfeld in dieser Studie untersucht wird, nämlich die Medienbranche (Abschnitt 2.1), Wissensarbeit (Abschnitt 2.2), neue Geschäftsmodelle (Abschnitt 2.3) und semantische Technologien (Abschnitt 2.4).

2.1 Medienbranche

Zur Medienbranche gehören alle Arbeitsfelder, in denen Inhalte über verschiedene Medienformate erstellt und vermittelt werden. Sie umfasst den Presse- und Buchmarkt, TV-/ Filmwirtschaft, Rundfunk und Musikwirtschaft, die Kunst- und Kulturindustrie mit Museen, Galerien und ähnlichen Kulturinstitutionen, Bibliotheken, Design- und Werbewirtschaft, Software- und Games-Industrie. Die Protagonisten dieser Bereiche sind Wissensarbeiter – also diejenigen, die die Inhalte generieren, über verschiedene Medienkanäle vermitteln und vertreiben. Dazu gehören

Journalisten, Wissenschaftler und andere Autoren, Korrespondenten- und Nachrichtenbüros, Zeitungs- und Zeitschriftenverlage, Buchverlage vom Sachbuch-, Kunst- und Fachverlag bis zu wissenschaftlichen und belletristischen Verlagen, Literaturagenten und Übersetzer, Bibliotheken, der Buchhandel mit Zeitschriftenhandel, Kuratoren, Galeristen und Archivare, Filmproduzenten und -verleiher sowie Kinobetreiber.²

Entscheidende Wachstumsträger der Branche in Deutschland sind aktuell die Games-/ IT- sowie die Designwirtschaft, während besonders das Verlagswesen an Umsatz eingebüßt hat.³ Basierend auf der Tatsache, dass Ende 2012 Zeitungen wie die „Financial Times Deutschland“ oder die „Frankfurter Rundschau“ eingestellt wurden bzw. Insolvenz anmeldeten, kündigte die Presse ein umfassendes „Zeitungssterben“ für die nächsten Jahre an. Zugleich hat hochwertig aufbereitete Information, also das wesentliche Verlagsprodukt, in der Wissensgesellschaft einen wichtigen Stellenwert erlangt; das öffentliche Interesse an tagesaktueller, tiefgreifender, glaubwürdiger und interessant aufbereiteter Information ist höher denn je. In diesem Spannungsfeld setzt die vorliegende Studie ihren Forschungsschwerpunkt auf den Presse- und Buchmarkt.

„In den letzten Jahren hat sich wenig im Verlagswesen bewegt, weil die Medienhäuser keine finanziellen und personellen Ressourcen für die Neuentwicklung und Erweiterung zeitgemäßer Medienformate zur Verfügung stellen. Ändert sich daran nichts, wird sich ein flächendeckendes, sinnvolles Annotieren von Informationen wohl kaum durchsetzen.“ Designbranche

„Museen werden nicht um die Digitalisierung herum kommen – aber um in die Zukunft zu denken, fehlen ihnen gegenwärtig die Ressourcen.“ Kulturarbeiter

Eine Herausforderung der Medien- und Kulturbranche sind die niedrigen Gehälter bei hohem Arbeitseinsatz. Diese Tatsache ist zum einen auf die große Konkurrenz, also das anhaltende, enorme Interesse an den Arbeitsfeldern der Branche, und zum anderen auf die sinkenden Absatzzahlen des Presse- und Buchmarktes zurückzuführen. Außerdem spielt auch

die Mentalität vieler Wissensarbeiter eine Rolle, die ein großes Bewusstsein für die Qualität des produzierenden Contents, seltener aber eines für ihren wirtschaftlichen Erfolg haben. Zugleich wird von den umsetzenden Wissensarbeitern höchste Qualität verlangt, die sich von der Konkurrenz und der schnellen Information des Internets abhebt.

Beispiel: Printjournalismus

Als besonders extrem lassen sich die ökonomischen Verhältnisse im Printjournalismus bezeichnen: Viele Medienhäuser haben ihre festen Redaktionen in den letzten Jahren reduziert, zusammengelegt oder weitgehend aufgegeben, um die Contentproduktion an freie Mitarbeiter in prekären Arbeitsverhältnissen auszulagern. Fest-freien (nicht selten scheinselfständig arbeitenden) Redakteuren zum Beispiel zahlen Zeitungs- und Zeitschriftenverlage häufig einen Tagessatz unter 200 Euro. Meist wird journalistische Arbeit nach Output (pro Artikel, Zeilen- oder Seitensatz) vergütet. Notwendige Vor- und Zusatzarbeiten wie Konzeption und Recherche, Bildredaktion, Layout etc. und sogar Spesen werden oft kaum oder gar nicht honoriert. Um ansatzweise ökonomisch zu arbeiten, lässt sich am einfachsten an der Recherche sparen: Informationen werden teilweise nicht vor Ort, sondern nur über Pressemeldungen, am Telefon oder im Internet eingeholt, auf Zweitmeinungen wird aus Zeitgründen verzichtet und mit Pressebildern gearbeitet. Gleichzeitig herrscht unter Wissensarbeitern die Auffassung

vor, dass die „Real Life“-Recherche, also die Arbeit am realen Objekt, der Besuch von Veranstaltungen, Life-Interviews am effizientesten und wichtigsten ist. Die Honorare für Webjournalismus sind oft noch niedriger – was wie Desinteresse vieler Autoren gegenüber dem Web wirkt, ist zum Teil also rein wirtschaftlichen Überlegungen geschuldet.

Laut einer Ver.di-Studie verdient lediglich ein Fünftel der freien Journalisten jährlich mehr als 25.000 Euro brutto, bei rund 25 Prozent liegt der Verdienst zwischen 15.000 und 25.000 Euro, bei 15 Prozent unter 7.000 Euro im Jahr. Nach einer DJV-Umfrage nahmen freie Journalisten 2010 im Durchschnitt knapp 17.000 Euro brutto pro Jahr ein, nach einer Studie der Künstlersozialkasse etwa 15.000 Euro. Zahlen wie diese lassen einen geringen Respekt der verantwortlichen Verleger für die eigene Arbeit, die eigenen Produkte – ihre Hauptressource zur Entwicklung von Qualitätsjournalismus – vermuten.

2.2 Wissensarbeit

Wissensarbeit bedeutet, Wissen zu generieren, zu verbreiten und anzuwenden. Dabei wird unterschieden zwischen Wissensarbeit, bei der individuelles, komplexes Erfahrungswissen („tacit knowledge“) kreativ zum Einsatz kommt, und den Routinetätigkeiten der Informationsverarbeitung („transactional labour“) wie etwa administrative Aufgaben.

Das Fraunhofer Institut für Arbeitswirtschaft und Organisation (IAO), Stuttgart, definiert Wissensarbeit anhand von drei Merkmalen: inhaltliche Neuartigkeit (inklusive Notwendigkeit zur Aktualisierung und Erweiterung von Wissen sowie Veränderung der Rahmenbedingungen), Komplexität der Aufgaben (Vielfalt, Schwierigkeitsgrad, Verantwortung) und Autonomie (eigenständige Gestaltung der Arbeitsinhalte, -prozesse und -zeit). Typische Wissensarbeiter sind Wissenschaftler, Journalisten und Verleger, Berater und Kreative.

Die Herausforderung: schwer greifbare Prozesse

Die eigentlichen Prozesse der Wissensarbeit sind oft nicht oder nur indirekt sichtbar. Konzeption und Recherche sind Vorarbeiten, die auf einem hohen Maß an Erfahrung und Vorbildung der Macher basiert. Die Auswahl an Quellen, also an Medien, Institutionen und Experten erfordert die tiefe Kenntnis der Branchen und Themen. Auch ein methodisches, handwerkliches Wissen ist notwendig, etwa dabei, wie sich die „richtigen“ Informationen herauskristallisieren lassen, welche und wie etwa Interviewfragen formuliert werden und wie die Ergebnisse für die entsprechenden Zielgruppen aufbereitet werden.

Einige dieser Vorarbeiten werden gar nicht im Endprodukt sichtbar: Zum Beispiel die Auseinandersetzung mit Quellen, die dem Überprüfen und Verifizieren von Aussagen dient; Gespräche mit Experten, die keine verwertbare Informationen liefern und im finalen Produkt ignoriert wer-

„Meiner Erfahrung nach haben dank der intuitiven Suchmöglichkeiten bei Google viele junge Leute schon jetzt die Fähigkeit verloren, eigenständig zu recherchieren und dabei auch mal kreativ um die Ecke zu denken.“ Verlagswesen

den; oder das Überwinden von Informationshürden wie Pressestellen, die neutrale Sachaussagen zugunsten eigener Interessen beschönigen.

Der Arbeitsprozess bei einem ganz wesentlichen Teil der Arbeit lässt sich besonders schwer greifbar machen: das Bewerten von Information, das Aufstellen neuer Thesen und Verknüpfen von Inhalten, das Herstellen relevanter Zusammenhänge geschieht vor allem in den Köpfen der Wissensarbeiter. Diese Arbeit manifestiert sich „nur“ im verdichteten Ergebnis der Verlagsprogramme, Filme und Publikationen, der Texte und Bilder. Für Außenstehende ist die Leistung dieser Arbeit oft schwer nachzuvollziehen.

„Beim Taggen von Informationen müssen Geschichten erzählt werden – ähnlich wie beim journalistischen Recherchieren und Vermitteln“ Kommunikationsbranche

Hinzu kommt, dass eine Vielzahl an Information im Internet schnell und dank intelligenter Suchalgorithmen der Suchmaschinen einfach zugänglich ist, so dass Recherche auf relativ hohem Niveau im Prinzip auch für Laien möglich wird. Doch die Beschaffung von Information ist nur der erste Schritt einer Recherche, auf die ihre Selektion und Bewertung folgen sollte. Diese wird in der Masse der täglich neuen, digital erstellten Inhalte schwieriger. Zugleich lässt sich ein Mangel an Medienkompetenz vieler, gerade jüngerer Internet-Nutzer feststellen: Die Qualität von Beiträgen und Quellen, welcher Medienabsender für welche Haltung oder Motivation steht, wird oft nicht hinterfragt. Es ist üblich geworden, dass sich Laien, die Bilder oder Texte in Print- oder digitalen Medien nach Belieben zusammenstellen, zum „Kurator“ ernennen. Eine fundierte verlegerische oder journalistische Hintergrundarbeit, etwa das Einordnen und Bewerten in historischem oder soziokulturellen Kontext, scheint dabei an Bedeutung zu verlieren. Diese Tendenz zur Kritiklosigkeit verstärkt die Tatsache, dass die finanzielle und gesellschaftliche Anerkennung für den Arbeitsaufwand und den Wert professioneller, unabhängiger Wissensarbeit sinkt. Ein resignierter, ironischer Kommentar unter Journalisten, die zum wiederholten Male ein indiskutables Honorarangebot erhalten haben, lautet: „Ein Text, der drei Minuten zum Lesen benötigt, kann natürlich auch nicht mehr als fünf Minuten zum Schreiben erfordern.“

Zugleich ist in der aktuellen Debatte um den gegenwärtigen Wandel des Journalismus vor allem von der überfälligen Notwendigkeit die Rede, „Qualitätsjournalismus“ ins Web zu übertragen und für digitale Anwendungen nutzbar zu machen. Wie diese Qualität zu den im vorigen Kapitel genannten Bedingungen zustande kommen soll, wird selten diskutiert.

In dem Leitartikel der ZEIT vom 22. November 2012 etwa, also kurz nach dem Bekanntwerden von Insolvenzen und Einsparungen einiger deutscher Traditionsblätter, werden 15 Geschäftsführer, Herausgeber und Chefredakteure großer deutscher Zeitungen und Zeitschriften nach Fehlern der Vergangenheit und den Herausforderungen für die Zukunft gefragt. Einhellig wird die bisher fehlende Offenheit für die digitalen Medien für die aktuelle Zeitungskrise verantwortlich gemacht. Allein bei Frank Schirrmacher, Herausgeber der FAZ, beginnt die Fehlersuche in seiner Branche weit vor diesem Punkt: „Wie kann man von Fehlern reden, wenn eine ganze Branche sich einigt, dass geistige Arbeit keinen materiellen Wert hat?“⁴

„Journalisten sollten Informationen beschaffen und vermitteln können, die nicht jedermann erhält. Voraussetzung dafür ist eine Recherche, die über die vorhandenen Ressourcen im Netz und Presstexte hinausgeht. Dazu gehört auch das Hinterfragen und Überprüfen von Informationen, das Aufdecken neuartiger und das Vermitteln komplexer Zusammenhänge, sowie eine eigene Meinung. Semantische Webtechnologien, mit denen relevante Inhalte vorselektiert werden, könnten ihnen dabei helfen, sich wieder aufs Wesentliche zu konzentrieren.“ Kommunikationsbranche

Zukunftschancen

Will sich die publizierende Branche weiterhin am Markt behaupten, wird sie sich zunächst auf ihre Kernkompetenz – also professionelle, neuartige, eigenständig analysierende, tiefgreifende differenzierte, überprüfte Inhalte zu liefern – zurückbesinnen müssen und diese für digitale Medien aufbereiten. Entgegen dem Anschein, dass Recherche durch das Internet einfacher wird, ist es im Gegenteil immer schwieriger, Quellen, Absendermotivation und Wahrheitsgehalt einer Information im Web aus-

findig zu machen. Durch intensiveres, fundierteres Filtern, Überprüfen und Bewerten von Inhalten werden ausgebildete Wissensarbeiter künftig einen Mehrwert gegenüber den schnellebigen, oft oberflächlichen Webinformationen herstellen können. Gute Journalisten praktizieren das seit jeher – die Feuilleton-Journalistin der ZEIT Katja Nicodemus etwa sagte in einem Interviewratgeber über ihre eigene Gesprächsvorbereitung: „Wenn ich einen Schauspieler oder einen Regisseur treffe, schaue ich mir im Idealfall eine Woche lang alle seine Filme an, lese so viel wie möglich von dem, was er in den letzten Jahren in Interviews alles gesagt oder in Büchern geschrieben hat. Wenn man das alles im Kopf hat, hat man die Freiheit, im Gespräch reagieren zu können: ‚Da dachten sie vor 30 Jahren aber anders.‘ ...“⁵

In diesem Zusammenhang werden technische Tools und Services, die digitale Information intelligent vorselektieren und veredeln, an Bedeutung gewinnen. Zum einen werden sich Wissensarbeiter dadurch wieder verstärkt um die exklusive „Real Life“-Recherche kümmern können, zum anderen wird das Interesse der Endverbraucher an hochwertigen, auf individuelle Informationsbedürfnisse abgestimmten Medien wachsen.

Auch beim Erstellen der Metadaten ist das Fach- und Branchenwissen notwendig, das eine hochwertige, kreative Recherche in Journalismus und Wissenschaft voraussetzt. Wissensarbeiter könnten Medienhäusern und Unternehmen künftig ein fachspezifisches, differenziertes Annotieren als Dienstleistung anbieten.

„Wer sich der Digitalisierung von Kultursammlungen verweigert, nimmt sich viele Zukunftschancen. Zahlreichen Protagonisten fehlt dafür noch das Bewusstsein – dabei vereinfachen digitale Kataloge viele interne Abläufe wie der kuratorischen Arbeit und der Öffentlichkeitsarbeit enorm.“ Kulturbranche

„Wir können unsere Inhalte nicht einfach über neue Kanäle vertreiben, vielmehr müssen die Inhalte für diese Kanäle neu produziert oder aufbereitet werden.“ Verlagswesen

2.3 Neue Geschäftsmodelle

Auf den ersten Blick mag der Aufbau eines digitalen Angebots für Verlage trivial erscheinen: Die bereits für Printprodukte erstellten Inhalte müssen lediglich über das Web zugänglich gemacht werden, so eine gängige Haltung. Bei diesem Ansatz werden Internet und Web nur als zusätzlicher Vertriebskanal genutzt, ohne dass ihre spezifischen Herausforderungen und Potenziale adressiert bzw. genutzt werden.

Grob ließe sich dies vergleichen mit einer sehr eingeschränkten Digitalisierung des Versandhandels, bei der lediglich der ausgedruckte Katalog durch ein PDF ersetzt wird und das Bestellformular am Ende des Katalogs durch ein Online-Formular. Online-Händler wie Amazon nutzen digitale Möglichkeiten aber in sehr viel größerem Umfang, beispielsweise in Form von Suchfunktionen, Empfehlungen, Verlinkungen oder Nutzerbewertungen. Anbieter von Online-Auktionen wie eBay nutzen ebenfalls die vielfältigen und komplexen Möglichkeiten, die sich durch den technologischen Fortschritt ergeben haben, konsequent aus.

Aus Erfahrungen im eBusiness und eCommerce lernen

Auch wenn sich Erfahrungen aus den Bereichen eBusiness oder eCommerce nicht eins zu eins auf die Medienbranche übertragen lassen, kann ein Rückblick auf die Entwicklungen in diesen Bereichen wertvolle Anhaltspunkte für die zu erwartenden Veränderungen, Herausforderungen und Möglichkeiten liefern. Eine Erkenntnis, die sich dabei unmittelbar aufdrängt, ist, dass sich neue Anbieter häufig erfolgreicher am Markt etablieren können, als Unternehmen aus der „Offline-Welt“, die versuchen,

„Wikis haben in der Vergangenheit nicht dazu beigetragen, Abläufe in Unternehmen übersichtlicher zu machen. Semantisch angereicherte Informationsplattformen mit ihren Suchfunktionen hingegen eignen sich gut, um Geschäftsprozesse zu vereinfachen.“

Technologieanbieter

„Da es sich hier um disruptive Technologien handelt, mit entsprechend komplexen Änderungen der Wertschöpfung, würde ich weniger auf etablierte Akteure setzen als auf neue Player.“

Technologieanbieter

ihr Geschäftsmodell in die Online-Welt zu übertragen. Letztere haben vielfach das Problem, dass sie Gefahr laufen, mit ihrem Online-Angebot ihr Kerngeschäft zu „kannibalisieren“. Daher werden Investitionen oft nur zögerlich angegangen oder es wird versucht, die existierenden Geschäftsprozesse und -modelle möglichst unverändert in die Online-Welt zu übertragen. Neue Anbieter haben dagegen den Vorteil, dass sie keine Rücksicht auf solche „Altlasten“ nehmen müssen und ihr Angebot konsequent auf die digitale Nutzung hin ausrichten und optimieren können.

Interaktivität und Echtzeit

In der Medienbranche bieten sich bei dieser digitalen Nutzung insbesondere Möglichkeiten im Hinblick auf Echtzeitaktualisierungen, Interaktivität, Verlinkungen, nicht-sequentielle Strukturen, Multimedia und Personalisierung. Da in der Online-Welt zeitintensive Zwischenschritte wie Drucken und Ausliefern entfallen, verkürzt sich die Zeit zwischen Produktion und Nutzung von Inhalten drastisch. Aktuelle Ereignisse werden in vielen Online-Zeitungen mit Live-Blogs verfolgt. Auch Aktualisierungen oder Korrekturen publizierter Inhalte sind möglich, ohne dass größere Kosten anfallen. Die grundsätzliche Interaktivität von Online-Medien ermöglicht zum einen die einfache Integration von Nutzerkommentaren – quasi als Erweiterung der traditionellen Zeitungsrubrik „Leserbriefe“ – sowie von User Generated Content allgemein.

Zum anderen bieten interaktive Infografiken die Möglichkeit, komplexe Inhalte und Zusammenhänge kompakt und intuitiv zu vermitteln. Über Verlinkungen ganz allgemein können darüber hinaus Verknüpfungen mit weiterführenden oder verwandten Inhalten hergestellt werden. Dies geschieht zum Beispiel durch die Verlinkung eines Satzes oder Satzteils in einem Online-Artikel auf einen entsprechenden anderen Artikel oder durch die Einblendung von Teasern für verwandte Artikel im Seitenumfeld.

Konvergenz und Multimedia

Auch die Integration von Videos oder Audiodateien in Texte ist in der Online-Welt einfach möglich, beispielsweise bei Musikkreuzungen. Generell ist hier eine zunehmende Multimedialisierung der Angebote zu beobachten, nicht zuletzt weil sich verfügbare Bandbreiten und Leistungsfähigkeit auch mobiler Endgeräte in den letzten Jahren erheblich weiterentwickelt haben.

Interaktivität und Multimedia ermöglichen prinzipiell auch die Produktion und Nutzung von Inhalten in nicht-sequentiellen Strukturen. Versuche hierzu gab es in der Offline-Welt beispielsweise in Form „programmierter“ Lehrbücher, bei denen in Lernkontrollen je nach gewählter Antwort auf die Seite verwiesen wurde, mit der fortzufahren ist. Bei falschen Antworten führen diese Verweise auf weitergehende Erläuterungen und Übungen, bei richtigen Antworten auf neue Lerninhalte. Ein Beispiel für nicht-sequentielle Strukturen in der Online-Welt ist Googles Knowledge Graph.

„Neben unserem Redaktionssystem benutzen unsere Mitarbeiter auch noch andere Tools, etwa für Recherchen oder Analysen. Solche ‚Helferlein‘ werden in Zukunft sicherlich weiter an Bedeutung gewinnen und immer besser werden.“

Verlagsbranche

Obleich der Bereitstellung entsprechender Inhalte für die Zukunft grundsätzlich ein hohes Potenzial zugesprochen werden kann, sind die Details für ihre Herstellung und Nutzung noch nicht abschließend abzusehen, zumal sie einen radikalen Bruch mit traditionellen Publikationsgewohnheiten darstellen.

Schließlich können Online-Angebote sehr einfach personalisiert bzw. unterschiedlichen Zielgruppen in variabler Form angeboten werden. Auch hier liegen große Zukunftschancen, wobei erste Erfahrungen zeigen, dass das Interesse bzw. die Nachfrage nach entsprechenden Personalisierungsmöglichkeiten auf Nutzerseite begrenzt ist. Neben diesen technologischen Aspekten, aus denen sich entsprechende Möglichkeiten für neue Wertschöpfungen ergeben, besteht die zentrale Frage im Kontext neuer Geschäftsmodelle darin, wie in der Online-Welt überhaupt Geld verdient wird bzw. werden kann. Während die meisten Online-Angebote von Tageszeitungen nach wie vor kostenfrei sind, experimentieren einige Anbieter mit sogenannten Bezahlschranken. Im Extremfall sinken dadurch aber die Besuchszahlen drastisch, was wiederum zu Mindereinnahmen über Online-Werbung führt. Darüber hinaus setzen viele Verlage auf kostenpflichtige Mobile Apps für Smartphones und Tablets, die zumindest den Vorteil einfacher Abrechnungs- und Bezahlfunktionen haben.

Komplexe Dienstleistungen statt Produktvertrieb

Ein grundsätzlicher Trend in Bezug auf zukünftige Geschäftsmodelle geht in Richtung komplexer und umfassender Dienstleistungen, weg vom Vertrieb vergleichsweise einfacher Produkte. Die folgenden Beispiele, die nicht aus der Medienbranche stammen, veranschaulichen diesen substantiellen Wandel:

- Ein Fabrikant von Bohrmaschinen bietet seinen Kunden ganzheitliche Dienstleistungen rund um „Löcher“ statt konkreter Produkte an;
- Ein Automobilhersteller wandelt sich zum Dienstleister für Mobilität;
- Eine KFZ-Versicherung bietet im Schadensfall ein durchdachtes Komplettpaket inklusive Ersatzfahrzeug, Reparatur, etc. an.

Diesen Beispielen ist gemeinsam, dass Produkthanbieter sich Gedanken darüber machen, in welchen Kontexten Kunden ihre Produkte nutzen und wie sich diese Nutzungskontexte holistisch unterstützen lassen.

Mit der Entwicklung neuer Geschäftsmodelle gehen dabei häufig auch Veränderungen in den Wertschöpfungsketten einher. Zum einen werden traditionelle Wertschöpfungsketten aufgebrochen – beispielsweise werden für Websites von Zeitungen keinen Drucker oder Zeitungsausträger benötigt – zum anderen können komplexe Ökosysteme von Dienstleistern entstehen.

2.4 Semantische Technologien

Der Begriff Semantic Web wurde von Sir Tim Berners-Lee geprägt, der häufig als Erfinder des World Wide Web bezeichnet wird.⁶ Die Grundidee des Semantic Web besteht darin, Daten in einer Form bereit zu stellen und zu verknüpfen, dass sie maschinell verarbeitet und „verstanden“ werden können. Damit knüpft das Semantic Web mehr oder weniger direkt an Forschungsarbeiten an, die sich bis zum Anfang der 60er Jahre zurück verfolgen lassen und mit Begriffen wie Semantische Netze, Wissensrepräsentation und Künstliche Intelligenz assoziiert sind. Diese wiederum standen zumeist in der philosophischen Tradition formaler Logik.

Auch die vom W3C (World Wide Web Consortium) im Bereich Semantic Web unterstützten Standards wie RDF (Resource Description Framework) und besonders OWL (Web Ontology Language) sind in dieser Tradition zu sehen. Semantic Web im engeren Sinne ist daher untrennbar mit symbolischen Repräsentationssprachen wie RDF oder OWL verknüpft.

Idealisiert gehen diese Ansätze von eindeutigen Begriffen aus, die in einer hierarchischen Struktur angeordnet sind und über eindeutige Relationen miteinander verbunden sind. Das Projekt des Semantic Web besteht dann darin, umfassende Ontologien zu erstellen und statt notorisch mehrdeutiger natürlichsprachlicher Wörter und Sätze, formal eindeutige Begriffe und Aussagen zu verwenden (siehe auch Abschnitt 3.1).

Während viele semantischen Ansätze einen hohen manuellen Modellierungsaufwand beinhalten, gibt es auch eine Reihe von alternativen Forschungsansätzen, die sich eher auf numerische, insbesondere statistische Verfahren zur Textanalyse stützen, beispielsweise für Themenerkennung, Kategorisierung, Clustering oder Informationsextraktion. Da sich diese Anwendungen mehr oder weniger direkt auf die in den Texten enthaltenen Bedeutungen beziehen, ist in diesem Zusammenhang meist von semantischen Technologien die Rede.

„Redakteure sind in der Regel gut vernetzt und bekommen aktuelle Informationen schnell zugespielt. Die semantische Suche wird eher dazu benötigt, um herauszubekommen, was in unserem Contentangebot zu einem bestimmten Thema bereits veröffentlicht wurde.“ Verlagsbranche

Schließlich gibt es noch das weite Feld sogenannter „intelligenter“ Lösungen – im Englischen bietet sich hier der etwas weniger kontroverse Begriff „smart solutions“ an. Hierzu gehören insbesondere Verfahren zur Auswertung (sehr) großer Datenmengen (Big Data).

Aus technologischer und wissenschaftlicher Sicht sind die oben genannten Unterschiede in Bezug auf Verfahren und Formalismen (z.B. zwischen RDF und OWL oder zwischen logischen und statistischen Ansätzen) von großer Wichtigkeit und Gegenstand detaillierter Forschung und kontroverser Diskussion. Aus Anwendersicht spielen diese Unterschiede aber meist nur eine untergeordnete Rolle. Hier ist wichtiger, welchen unmittelbaren Nutzen eine Lösung für den Anwender bietet - wie die Lösung technisch genau realisiert wird, interessiert den Anwender dagegen in der Regel kaum.

Für den Zweck dieser Studie wählen wir daher eine sehr weite Interpretation des Begriffs Semantic Media Web, die alle Lösungen umfasst, in denen Technologien zur Veredelung von Inhalten zum Einsatz kommen

- unabhängig davon, ob diese Technologien RDF oder OWL verwenden oder explizit als semantische Technologien bezeichnet werden. Im nächsten Abschnitt werden die wichtigsten Lösungsansätze des Semantic Media Web kurz vorgestellt.

QUELLEN:

- 1) Die Kultur- und Kreativwirtschaft in der gesamtwirtschaftlichen Wertschöpfungskette - Wirkungsketten, Innovationskraft, Potenziale. Bundesministerium für Wirtschaft und Technologie (BMWi) /Fraunhofer ISI, 2012
- 2) www.kultur-kreativ-wirtschaft.de/KuK/Navigation/kultur-kreativwirtschaft,did=349672.html
- 3) Monitoring zu ausgewählten wirtschaftlichen Eckdaten der Kultur- und Kreativwirtschaft 2010. BMWi (Hrsg.) 2012
- 4) „Wie guter Journalismus überleben kann - Haltung statt Gleichmut. Deutschlands wichtigste Medienmanager sprechen über ihre Erwartungen, Einsichten – und ihre Fehler.“ 22.11.2012, DIE ZEIT, www.zeit.de/2012/48/Medien-Zeitung-Internet-Journalismus-Medienmanager/komplettansicht
- 5) Peer Teuwsen, Das gute Gespräch - Wie man erfolgreich fragt. Ein Lehrbuch. Echtzeit Verlag, Basel 2009, S. 81
- 6) Tim Berners-Lee, James Hendler, Ora Lassila, The Semantic Web, Scientific American, Mai 2001, www.scientificamerican.com/articlecfm?id=the-semantic-web

3. Semantic Media Web

In diesem Abschnitt werden die wichtigsten technologischen Lösungen vorgestellt, die dem Semantic Media Web zugrunde liegen. Dabei geht es zunächst darum, die Lösungen ganz allgemein zu erläutern und anhand beispielhafter Use Cases grob zu veranschaulichen. Eine Konkretisierung mit branchenspezifischen Anwendungsszenarien erfolgt dann in der Analyse der ausgewählten Anwendungsfelder in Abschnitt 4.

Abschnitt 3.1 erläutert zunächst das Konzept der semantischen Suche. In Abschnitt 3.2 werden dann Ontologien, Thesauri, Begriffshierarchien, kontrollierte Vokabulare und das Thema Annotationen behandelt. Abschnitt 3.3 beschäftigt sich mit Inhaltsextraktion, Abschnitt 3.4 mit Kategorisierung und Clustering. In Abschnitt 3.5 wird das allgemeine Konzept der Inhaltsveredelung vorgestellt, das in dieser Studie als generische Beschreibung der heterogenen Funktionalitäten des Semantic Media Web verwendet wird.

3.1 Semantische Suche

Im Bereich Semantic Web wird als ideales Anwendungsszenario für semantische Technologien häufig auf eine „semantische Suche“ verwiesen. Die Grundidee hierbei ist, dass eine semantische Suchmaschine bei der Erstellung von Ergebnislisten auch semantisches Wissen berücksichtigt und somit bessere Ergebnisse liefert als traditionelle Suchmaschinen.

Inwieweit dieses Versprechen auch in der Praxis eingelöst werden kann, ist aber nach wie vor offen. Zwar gibt es erste Ansätze wie z.B. Ask Jeeves (ask.com) oder WolframAlpha (wolframalpha.com), aber bisher keine etablierte semantische Suchmaschine mit ähnlicher Abdeckung oder Nutzung wie Google.

Bei der Suche können diverse Szenarien unterschieden werden. Wird etwa nach einer ganz bestimmten Information gesucht (z.B. der Name der Währung Myanmars), reicht es aus, wenn in den ersten Suchergebnissen eine Seite angezeigt wird, die diese Information beinhaltet. Idealerweise ist die Information sogar im Suchergebnis selber sichtbar, so dass die ursprüngliche Ergebnisseite gar nicht mehr aufgerufen werden muss.

Daneben gibt es aber eine komplexere Suche, die eher einer Recherche ähnelt (z.B. Welche Veröffentlichungen gibt es zu genetisch modifizierten Lebensmitteln?). Hier wäre es wünschenswert, unter den ersten Suchergebnissen möglichst viele unterschiedliche und relevante Webseiten zu finden. Außerdem sind hier sehr spezialisierte Recherchertools für bestimmte Berufsgruppen und Gegenstandsbereiche denkbar und sinnvoll.

Semantische Technologien können dabei in sehr unterschiedlichen Verarbeitungsschritten eingesetzt werden, insbesondere bei der Erstellung oder Indexierung von Inhalten (z.B. in Form von Anreicherung der Inhalte mit Metadaten), bei der Interpretation der Suchanfrage (z.B. durch das automatische Hinzufügen relevanter Suchterme in der internen Suchabfrage), oder bei der Ausgabe der Suchergebnisse (z.B. durch Kategorisierung).

Es scheint, als würde der Bedarf an besseren Suchmaschinen bzw. die Unzulänglichkeit existierender Suchmaschinen von Nutzerseite deutlich verhaltener eingeschätzt als von Anhängern semantischer Suchmaschinen. Die meisten Experten sind mit dem aktuellen Stand beispielsweise der Google-Suche durchaus zufrieden und verweisen auf die positive Entwicklung der letzten Jahre.

Von daher eignet sich die semantische Suche nicht unbedingt als vermeintliche „Killer App“ des Semantic Media Web. Vielmehr ist zu erwarten, dass semantische Verfahren kontinuierlich in die Verbesserung der Suchverfahren einfließen werden bzw. für die Realisierung spezialisierter Mehrwertdienste für Professionals genutzt werden.

„Die Internetsuche beispielsweise über Google hat sich im letzten Jahrzehnt deutlich verbessert - mit dem aktuellen Stand sind wir ganz zufrieden.“ Agentur

3.2 Ontologien, Thesauri, Annotationen

Ein weiterer Forschungsschwerpunkt des Semantic Web liegt im Bereich Ontologien. Im weitesten Sinne erfüllen Ontologien ähnliche Zwecke wie auch Thesauri, Begriffshierarchien, kontrollierte Vokabulare, Metadaten oder Tags, die alle die Annotationen von Texten oder Textteilen ermöglichen und somit den Einbezug von Weltwissen. Die Unterschiede zwischen den jeweiligen Verfahren liegen im jeweiligen Grad der Formalisierung.

Der grundsätzliche Nutzen einer Ontologie lässt sich wie folgt veranschaulichen. Der Eigenname „Armstrong“ kann zum Beispiel für den Musiker Louis Armstrong, den Astronauten Neil Armstrong oder den Radfahrer Lance Armstrong stehen. Durch die Annotation des Wortes Armstrong mit einem Verweis auf eine eindeutige Ressource kann diese Mehrdeutigkeit aufgelöst werden.

Die Auflösung von Mehrdeutigkeiten, auch Desambiguierung genannt, ist insbesondere für die maschinelle Sprachverarbeitung eine zentrale Herausforderung, wie sich am Beispiel der maschinellen Übersetzung zeigen lässt. Für die meisten Wörter einer Ausgangssprache gibt es in der Zielsprache mehrere Übersetzungsmöglichkeiten. Sogenannte Interlingua-Ansätze überführen daher Sätze der Ausgangssprache zunächst in eindeutige formale Repräsentationen, aus denen dann Sätze der Zielsprache generiert werden. Dabei werden den potentiell mehrdeutigen Wörtern einer Sprache „sprachunabhängige“ eindeutige Konzepte einer Ontologie zugeordnet.

Nehmen wir als Beispiel das Wort „Überführung“. Dies kann etwa eine Brücke bezeichnen, eine Umwandlung, den Transport eines Leichnams oder den Schuldnachweis eines Täters. Für Menschen ist aus dem Kontext meistens sofort klar, welche Bedeutung jeweils gemeint ist. Aber für die maschinelle Verarbeitung ist eine explizite Desambiguierung notwendig.

Sehr schnell zeigt sich dabei, dass die Granularität der Desambiguierung und somit der Ontologiemodellierung von der jeweiligen Aufgabenstellung abhängt – „eindeutig“ und „mehrdeutig“ sind also nicht unbedingt absolute Begriffe, sondern selber kontextabhängig. Die Verwendung eindeutiger Bezeichnungen für Ressourcen ist eine wesentliche Grundlage des Semantic Web bzw. dafür, dass Daten maschinell verarbeitet werden können.

Die Verwendung von Formalismen erlaubt darüber hinaus auch die maschinelle Ableitung weiterer Informationen, sogenannte Inferenzen oder Schlussfolgerungen. Dies lässt sich am einfachsten anhand hierarchischer Beziehungen veranschaulichen. Wird in einer Ontologie etwa modelliert, dass ein „VW Golf“ ein „Auto“ ist und ein „Auto“ ein „Fahrzeug“, kann diese Information für Schlussfolgerungen verwendet werden. Wird ein Objekt etwa in einer Videoszene dann mit der Metainformation „VW Golf“ annotiert, so kann diese Szene auch als Suchergebnis zurückgeliefert werden, wenn ein Nutzer nach den Begriffen „Auto“ oder „Fahrzeug“ sucht.

Somit ergeben sich im Bereich der Ontologien eine ganze Reihe grundsätzlicher Herausforderungen. Zunächst einmal erfordert die Erstellung und Pflege einer Ontologie einen beträchtlichen manuellen Aufwand. Hinzu kommt in der Regel ein Abstimmungsaufwand zwischen den Personen und Organisationen, die an der Ontologiemodellierung beteiligt sind. Dieser kann sich beträchtlich erhöhen, wenn die beteiligten Akteure sehr unterschiedliche Sichtweisen auf den zu modellierenden Gegenstandsbereich haben oder die Ontologie für unterschiedliche Aufgaben einsetzen wollen.

Dies verdeutlicht z.B. das Klassifikationssystem `eCl@ass` (`eclass.de`) für Warengruppen und Warenmerkmale, das von Käufern und Verkäufern entwickelt wurde, um die elektronische Abwicklung von Beschaffungsvorgängen zu ermöglichen. Dabei wurden schnell sehr unterschiedliche Modellierungsanforderungen aus Käufer- bzw. Verkäufersicht deutlich.⁷

Gerade in sensiblen Bereichen geht es bei Modellierungsfragen implizit auch um die „Deutungshoheit“ über Begriffe, beispielsweise bei der Kategorisierung von Parteien oder Gruppierungen als extrem, terroristisch, etc. Vor diesem Hintergrund scheint es grundsätzlich wünschenswert, die Entwicklung von Ontologien als gemeinschaftliche und gesellschaftlich relevante Aufgabe aufzufassen, da es sich hier um eine grundlegende inhaltliche Infrastruktur des Semantic Media Web handelt.

„DBpedia ist teilweise zu wissenschaftlich und daher nicht immer alltagstauglich. Bottom-Up-Ansätze wie `schema.org` sind da häufig zielführender, weil pragmatischer.“

Dienstleister

Im Idealfall gibt es für das Semantic Media Web genau eine Ontologie, die von unabhängigen Standardisierungsgremien entwickelt und von allen beteiligten Akteuren genutzt wird. Diese Ontologie kann dabei durchaus verteilt entwickelt und gepflegt werden, in Formalismen wie RDF oder OWL gibt es dafür entsprechende Unterstützung und auch Linked Data geht von einer dezentralen Informationsspeicherung aus.

Tagging ohne festgelegtes, übergreifendes Gliederungssystem bietet im Prinzip ähnliche Möglichkeiten wie Ontologien, allerdings fehlen hier meist die hierarchischen Beziehungen zwischen den Tags. Teilweise gibt es auch kein vorgegebenes Tag-Vokabular, d.h. zur Annotierung können beliebige Tags verwendet werden. Damit entfällt zwar der Aufwand der Ontologie-Erstellung, dafür ist aber auch keine Konsistenz und Eindeutigkeit mehr gewährleistet.

Ontologien oder ähnliche Verfahren bilden eine gute Grundlage für Personalisierung: Nutzer können Profile anlegen, in denen sie festlegen, welche Konzepte oder Tags sie interessieren, und erhalten dann genau die Inhalte, die Informationen hierzu enthalten.

„Das semantische Web muss darauf reagieren können, dass die Nutzer zu Beginn ihrer Recherche oft selbst nicht genau wissen, was sie suchen.“ Technologieanbieter

3.3 Informationsextraktion

Verfahren zur Informationsextraktion bilden einen wesentlichen Bestandteil des Semantic Media Web. Sie werden auch bereits in der Praxis eingesetzt, sowohl als vollautomatische Verfahren, als auch in semi-automatischen Varianten, in denen die maschinellen Ergebnisse manuell bestätigt oder korrigiert werden.

Ein einfaches Beispiel für die Informationsextraktion ist die Entity Recognition (auch Entity Extraction oder Named-Entity Recognition genannt). Dabei werden in einem Text vorkommende Eigennamen identifiziert bzw. Wörter und Phrasen, die sich auf Personen, Orte, Zeitpunkte oder -intervalle, Zahlen, etc. beziehen. Idealerweise werden die jeweiligen Objekte gleich mit einer eindeutigen Bezeichnung, zum Beispiel einer URI, versehen. Die Informationsextraktion nutzt dabei die im vorigen Abschnitt beschriebenen Wissensquellen wie Ontologien oder Thesauri.

Das Spektrum reicht hier von einfachen Verfahren, die z.B. Schreibvarianten ausländischer Namen erkennen, bis hin zu komplexer Sprachverarbeitung mit Auflösung von Personalpronomen oder deskriptiven Phrasen wie „der siebenfache Tour-de-France-Sieger“. Das visionäre Ziel der Informationsextraktion ist die Überführung von Texten in eindeutige formale Repräsentationen, die den Textinhalt sprachunabhängig abbilden.

Auch im Kontext von Bildern und Videos können Verfahren der Informationsextraktion eingesetzt werden, etwa zum Erkennen und Tracken von Personen, Gegenständen oder Orten in einer Szene. Ein Spezialfall ist dabei die Gesichtserkennung.

„Interessant wird das semantische Web dort, wo Annotationen die Bedürfnisse der Nutzer interpretieren.“ Technologieanbieter

3.4 Kategorisierung und Clustering

Abschließend sollen noch kurz Verfahren zur Kategorisierung und zum Clustering erläutert werden. Bei der Kategorisierung geht es darum, Inhalte (semi-)automatisch vorgegebenen Kategorien, z.B. den in einer Ontologie modellierten Begriffen, zuzuordnen. Beim Clustering werden Inhalte in Gruppen mit ähnlichen Eigenschaften eingeteilt. Die Methoden, die für die Realisierung dieser Verfahren eingesetzt werden, sind grundsätzlich ähnlich, ein wesentlicher Unterschied besteht darin, dass für das Clustering kein Kategoriensystem vorgegeben werden muss.

Sowohl Kategorisierung als auch Clustering bieten somit Strukturierungen an und sind daher insbesondere für die Suche und die Navigation nützlich. Die Entwicklung der Suchmaschinen zeigt aber auch, dass der Bedarf an diesen Strukturierungen sehr differenziert betrachtet werden muss. So hat etwa Yahoo! (yahoo.de) in der Anfangszeit eine sehr ausgefeilte Themenhierarchie redaktionell gepflegt und verwendet – ausgewählte Websites wurden dann manuell in diese Hierarchie eingehängt. Google wiederum hat eine Zeitlang eine Clustering-Funktion angeboten, mit der Suchergebnisse gruppiert wurden.

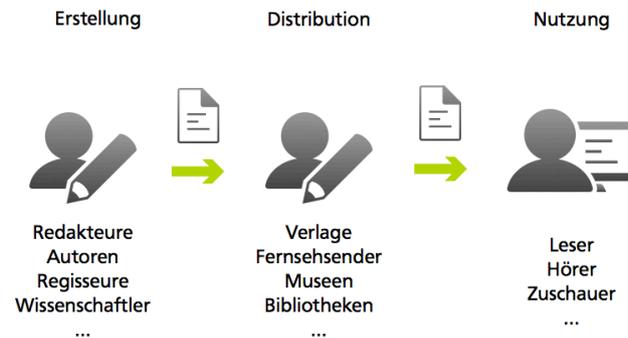
Dass diese Funktionalitäten in der aktuellen Internetsuche kaum noch angeboten werden, deutet auf ein unbefriedigendes Kosten-Nutzen-Verhältnis für dieses Anwendungsszenario hin. Für professionelle Rechercheszenarien, insbesondere in gut abgegrenzten Gegenstandsbereichen,

liefern Verfahren zur Kategorisierung und zum Clustering aber erheblichen Mehrwert und eine gute Grundlage für die semantische Suche.

3.5 Veredelung von Inhalten

Die in den letzten vier Abschnitten skizzierten Lösungsansätze unterscheiden sich sowohl in Hinblick auf die jeweils eingesetzten Technologien als auch in Bezug auf die Funktionalität, die sie zur Verfügung stellen. Allen gemeinsam ist aber, dass sie zur Veredelung von Inhalten beitragen, d.h. für Inhaltsobjekte zusätzliche Informationen bereitstellen, die wiederum für die Entwicklung neuer Mehrwertdienste genutzt werden können. In den folgenden Abschnitten werden daher im Kontext des Semantic Media Web prinzipiell alle Verfahren berücksichtigt, die zur Veredelung von Inhalten beitragen.

Klassische Wertschöpfung in der Medienbranche (stark vereinfachtes Schema)



Es gibt eine ganze Reihe von Parametern, in denen sich die Einsatzszenarien von Verfahren zur Inhaltsveredelung voneinander unterscheiden können. Wie sich in den nächsten Abschnitten zeigen wird, haben die jeweiligen Ausprägungen dieser Parameter erheblichen Einfluss auf Herausforderungen und Potenzial des Semantic Media Web. So kann eine Inhaltsveredelung beispielsweise manuell erfolgen, semi-automatisch oder vollautomatisch, mit entsprechenden Konsequenzen für Qualität, Konsistenz und Aufwand. Die Veredelung kann direkt bei der Inhaltserzeugung

erfolgen, beispielsweise wenn ein Online-Redakteur seinen Artikel selber verschlagwortet und kategorisiert, oder aber nachträglich durch externe Dienstleister oder Crowdsourcing hinzugefügt werden. Ein weiterer wichtiger Parameter betrifft das Ziel bzw. die Zielgruppe der Veredelung. Richtet sich eine Veredelung beispielsweise an bestimmte Berufsgruppen wie Mediziner oder Juristen, mit dem Ziel, deren Arbeitsprozesse zu unterstützen und zu optimieren, kann auch mit einer entsprechenden Zahlungsbereitschaft kalkuliert werden.

Zukünftige Wertschöpfung in der Medienbranche (stark vereinfachtes Schema)



Wenn Inhaltserzeugung und Veredelung von unterschiedlichen Institutionen durchgeführt werden, stellt sich natürlich unmittelbar die Frage nach Nutzungs- und Verwertungsrechten. Die Veredelung ermöglicht in der Regel Mehrwertdienste ist aber andererseits auf die Basisinhalte angewiesen. Vor diesem Hintergrund ist zu erwarten, dass sich im Semantic Media Web entsprechende „Ökosysteme“ ausbilden werden, die Inhaltserzeuger, Veredler und Nutzer umfassen.

QUELLEN:

- 7) Joachim Quantz, Thorsten Wichmann, E-Business-Standards in Deutschland - Bestandsaufnahme, Probleme, Perspektiven, Berlin, April 2003, www.bmwi.de/BMWi/Redaktion/PDF/E/e-business-standards-endbericht,property=pdf,bereich=bmwi2012,sprache=de,rwb=true.pdf

4. Ausgewählte Anwendungsfelder

In diesem Abschnitt werden sechs ausgewählte Anwendungsfelder im Hinblick auf ihr jeweiliges SMW-Potenzial genauer untersucht. Dabei geht es weniger um eine detaillierte und ausführliche Analyse der Anwendungsfelder, die den Rahmen dieser Studie sprengen würde. Vielmehr soll ein möglichst breites Spektrum abgedeckt werden, so dass Gemeinsamkeiten und Unterschiede zwischen den einzelnen Anwendungsfeldern herausgearbeitet werden können. Somit wird eine allgemeine Charakterisierung der Herausforderungen und des Zukunftspotenzials des Semantic Media Web möglich..

Die sechs ausgewählten Anwendungsfelder sind Verlage (Abschnitt 4.1), Kulturgüter digital (Abschnitt 4.2), Bibliotheken (Abschnitt 4.3), Bewegtbild (Abschnitt 4.4), Open Data/Open Government (Abschnitt 4.5) und Corporate Communication (Abschnitt 4.6).

4.1 Verlage

Verlage drängen sich als naheliegendes Anwendungsfeld für Semantic Media Web und neue Geschäftsmodelle für Wissensarbeiter geradezu auf. Zum einen sind die Veränderungen und Umbrüche in dieser Branche sowie das Spannungsfeld zwischen klassischen Print-Produkten und digitalen Online-Angeboten unübersehbar (siehe Abschnitt 2.1). Zum anderen sind Verlage und ihre Produkte den meisten Menschen grundsätzlich vertraut, so dass auch ein grobes, zumeist allerdings etwas naives Bild der involvierten Arbeitsabläufe verbreitet ist.

Heterogenes Anwendungsfeld

Bei genauerer Betrachtung stellt sich dann schnell heraus, dass das Anwendungsfeld Verlage sehr heterogen ist und entsprechend differenziert betrachtet werden muss. Zum einen gibt es verschiedene Formate wie Zeitungen, Zeitschriften und Bücher, zum anderen sehr unterschiedliche, teilweise wissenschaftliche (Fach-)Inhalte und entsprechend heterogene Leserschaft und Zielgruppe.

Viele Zeitungen und Zeitschriften haben recht früh ein Online-Angebot als Ergänzung zu ihren gedruckten Ausgaben aufgebaut. Beispiele hierfür sind Spiegel Online (seit 1994), taz.de (seit 1995), Zeit Online (seit 1996) oder FAZ.Net (seit 2001). In vielen Fällen gibt es für das Online-Angebot eine eigene Redaktion, darüberhinaus werden aber auch Artikel aus den Print-Ausgaben übernommen. In den letzten Jahren wurden darüber hinaus spezielle Formate für mobile Endgeräte entwickelt, zunehmend auch eigene Apps für Smartphones und Tablets. Dass hier die klassischen Grenzen zwischen Print und audio-visuellen Medien zunehmend verschwinden, zeigt sich beispielsweise an der Klage von Zeitungsverlagen gegen die „Tagesschau“-App der ARD.

„Natürlich ist der digitale Wandel unabwendbar und wir müssen neue Geschäftsmodelle schneller hochfahren als alte absterben. Aber es besteht die Angst, etablierte Geschäftsmodelle schneller kaputt zu machen als notwendig.“ Verlagsbranche

„Redakteure sind mit ihren IT-Systemen nie zufrieden: Entweder sind sie Ihnen nicht leistungsfähig genug oder sie sind zu komplex und schwer zu bedienen.“ Verlagsbranche

Kannibalisierung des Kerngeschäfts

Traditionsverlage stehen dabei im Gegensatz zu Websites wie Politico (politico.com) oder Huffington Post (huffingtonpost.com) vor dem grundsätzlichen Dilemma, dass ihre Online-Angebote potenziell ihr Print-Geschäft kannibalisieren. Inwieweit Einnahmen durch Online-Werbung, Bezahlschranken oder kostenpflichtige Apps für Mobilgeräte

mittel- und langfristig Einnahmeverluste im klassischen Print-Geschäft kompensieren können, ist dabei nach wie vor unklar.

In diesem Zusammenhang muss allerdings auch darauf hingewiesen werden, dass die Leserzahlen vieler Tageszeitungen schon seit den 90er Jahren rückläufig sind und daher wohl nicht allein auf digitale Konkurrenz zurückzuführen sind. Und bereits lange vor dem Crowdfunding-Hype ist die „tageszeitung“ ihrer notorischen Finanznot sehr erfolgreich begegnet, indem sie von ihren Lesern über ein Genossenschaftsmodell Geld eingesammelt hat. Auch mit der Aktion „taz-zahl-ich“ geht die taz einen eigenen Weg – statt eine generelle Bezahlschranke einzuführen können Leser über diverse Bezahlarten freiwillig für einzelne Artikel zahlen.

„Im Bereich Mainstream-Nachrichten sind inzwischen über 70% aller Meldungen mehr oder weniger deutlich Doubletten von Agenturmeldungen.“ Technologieanbieter

„Einfach nur Agenturmeldungen durchzuschleifen ist langfristig keine gute Idee, da dies die Sichtbarkeit in Suchmaschinenergebnissen verringert. Für eine gute Positionierung in Ergebnislisten braucht man Unique Content.“ Verlagsbranche

Qualitätsjournalismus und Edelblogger

Von Verlagsseite wird in diesem Zusammenhang gerne vor einem Verschwinden des Qualitätsjournalismus gewarnt. Journalisten werden dabei als Garanten für eine qualitativ hochwertige Berichterstattung gesehen, in der Fakten vor der Veröffentlichung überprüft und Argumente abgewogen werden. Es ist allerdings fraglich, inwieweit die Realität in aktuellen Online-Publikationen diesem hehren Anspruch tatsächlich noch gerecht wird. Ständiger Zeit- und Aktualitätsdruck sowie die schlechte Bezahlung von Online-Redakteuren führen dazu, dass viele Online-Artikel nur noch aus Agenturmeldungen bestehen, die mehr oder weniger ungefiltert an den Leser weitergeleitet werden.

Eine etwas positivere Sichtweise sieht die Zukunft des Journalismus in Blogs und spricht in diesem Kontext etwa von „Edelbloggern“. In diesem Szenario verschwinden die Verlage – die Blogger publizieren direkt

für ihre Leser (Follower). Auch hier stellt sich aber die Frage, ob eine angemessene Bezahlung dieser Form der Wissensarbeit über Werbeeinnahmen, flatter oder ähnliche Bezahlmodelle realistisch ist. Grundsätzlich kann aber davon ausgegangen werden, dass Leser eher bereit sind, für originelle Beiträge zu bezahlen, als für Beiträge, die in an anderer Stelle in ähnlicher Form auch kostenfrei zu erhalten sind.⁸

„Wir sehen Self Publishing nicht wirklich als Bedrohung, weil dabei die Marke fehlt und die Markteintrittsbarrieren recht hoch sind. Als Verlag haben wir den Vorteil, dass wir unsere Kunden gut kennen und auch verstehen.“ Verlagsbranche

Wissenschaftsverlage

Etwas besser stellt sich die Situation für Wissenschaftsverlage dar. Diese haben zum einen wohldefinierte und häufig auch zahlungskräftige Zielgruppen, zum anderen bieten sie Inhalte an, die in der Regel nicht anderweitig kostenfrei verfügbar sind. Aus Sicht des Semantic Media Web haben wissenschaftliche Publikationen darüber hinaus den Vorteil, dass sie sich zumeist auf einen wohldefinierten und häufig schon systematisierten Gegenstandsbereich beziehen. Dies führt zu einer vergleichbar günstigen Kosten-Nutzen-Relation was die Erstellung und Verwendung von Ontologien bzw. kontrollierten Vokabularen angeht.

Aufgrund des schwierigen Marktumfeldes sowie sinkender Auflagen und Einnahmen verfügen die meisten Verlage gegenwärtig nicht über die notwendigen Mittel, um umfangreiche Investitionen zu tätigen. Ohne diese Investitionen können aber die innovativen Angebote nicht entwickelt werden, die dringend nötig wären, um sich vom Wettbewerb abzuheben oder neue Einnahmen zu erschließen. Verstärkt wird dieser Teufelskreis durch die teilweise geringe Innovationsbereitschaft und Technologieaffinität einiger Verlage.

„Meine Vision ist eine wirklich neutrale Datenhaltung, die eine industrielle Inhaltsproduktion ermöglicht, also quasi ein Content Stream, der die Entwicklung von Nachrichten abbildet, den alle Redakteur abgreifen können und der die Grundlage für die Erstellung konkrete Print- oder Online-Artikel bildet.“ Verlagsbranche

Beispiel: NewsMap

NewsMap (ontonym.de/newsmap.html) ist ein Dienst des Berliner Unternehmens Ontonym, der lokale Nachrichten aus großen Berliner Tageszeitungen in einer Kartendarstellung visualisiert. Somit ist NewsMap eine semantische Augmented Reality Anwendung, die semantische Technologien, Textanalyse und Augmented Reality verknüpft, um Lesern einen einfacheren Zugang zu lokalen Nachrichten zu ermöglichen. Dazu werden aus den Nachrichten Ortsbezeichnungen extrahiert. Aus diesen ermittelt ein kombiniertes Clustering- und Desambiguierungsver-

fahren, das geografisches Hintergrundwissen nutzt, diejenigen Orte, an denen die Nachricht anzuzeigen ist. Abschließend lokalisiert eine konventionelle Geocodierung diese Orte zur Anzeige auf Karten oder in Augmented Reality Apps.

NewsMap ist verfügbar für alle Endgeräte auf denen die Augmented Reality Apps Wikitude oder Junaio App laufen, in der Testphase jedoch nur für Berlin, deutsch-sprachige Nachrichten und die großen Berliner Tageszeitungen. Erweiterungen für andere Metropolregionen Deutschlands sind in Planung.

Beispiel: Jurion

Jurion (jurion.de) ist eine 2011 gelaunchte Plattform des Verlages Wolters Kluwer Deutschland, die juristische Fachinformationen zur Verfügung stellt und insbesondere die Recherchearbeit von Juristen unterstützen soll. Dabei liegt ein Schwerpunkt auf der nahtlosen Integration in den Arbeitsalltag und die typischen Prozessstrukturen von Journalisten. In Jurion

kommen bereits Technologien des Semantic Web zum Einsatz, in Zukunft soll insbesondere die semantische Suche noch weiter ausgebaut werden. Jurion veranschaulicht den Wandel von Wolters Kluwer von einem traditionellen Wissenschaftsverlag hin zu einem Informationsdienstleister.

Fazit

Buch-, Zeitungs- und Zeitschriftenverlage befinden sich gegenwärtig in einem tiefgreifenden Strukturwandel und bilden so grundsätzlich ein ideales Anwendungsfeld für das Semantic Media Web. Hier existieren auch bereits erste Ansätze für den Praxiseinsatz von Semantic-Web-Technologien. Gleichzeitig tun sich viele Verlage noch sehr schwer mit der Entwicklung und Umsetzung neuer Geschäftsmodelle. Wissenschaftsverlage mit „Professionals“ als Zielgruppe haben es hier deutlich einfacher als beispielsweise breiter aufgestellte Zeitungsverlage. Auch gegenüber Open Access verhalten sich viele Verlage verständlicherweise eher skeptisch. Wünschenswert wäre hier ein verstärkter Austausch mit Technologieanbietern, Kommunikationsagenturen und anderen Dienstleistern sowie eine höhere Bereitschaft, mit neuen Ansätzen zu experimentieren.

Das größte Potenzial für das Semantic Media Web im Verlagswesen liegt daher im Bereich bereits etablierter Online-Angebote mit großer Reichweite sowie neuer Dienstleistungsangebote von Wissenschaftsverlagen. In diesen beiden Anwendungsfeldern stellt sich das Kosten-Nutzen-Verhältnis für die Integration innovativer Technologien und Lösungen grundsätzlich positiv dar.

4.2 Kulturgüter digital

Kulturinstitutionen wie Museen und Galerien verfügen über enorme Datensammlungen in unterschiedlichen Medienformaten, die eine Art kulturelles Gedächtnis darstellen: historische und wissenschaftliche Exponate wie zwei- und dreidimensionale Kunstwerke, ergänzende Literatur, Foto- Film- und Tondokumente, Modelle, Prototypen und Vorentwürfe aus Nachlässen und Forschungsarbeiten. Bislang ist dieses Wissen, solange es nicht nach Themen gegliedert in Ausstellungen oder Katalogen präsentiert wird, meist hinter Museums- und Archivmauern versteckt und erfordern oft eine aufwendige Recherche mit langen Wegen. Doch längst existiert eine Erwartungshaltung in der Wissenschaft und der Medienwelt, Kulturschätze auf einer digitalen Instituts- und medienübergreifenden Plattform der Öffentlichkeit zugänglich zu machen.

Digitale Sammlungskataloge

Digitale Sammlungskataloge wären auch für die Institutionen selbst von unschätzbarem Nutzen: Erstens ist es die ureigenste Aufgabe der nicht-kommerziellen Einrichtungen, ihre Arbeit für Öffentlichkeit und Wissenschaft zugänglich zu machen - was im digitalen Zeitalter selbstverständlich auch über das Web geschehen muss. Zweitens können digitale Sammlungen Arbeitsprozesse und Projektplanung massiv vereinfachen. Beim Konzipieren und Gestalten von Ausstellungen oder Publikationen, bei Umzug und Reorganisation von Archiven oder Museumshäusern reduziert ein digitales Archiv Zeit- und Arbeitsaufwand enorm. Drittens dient ein Onlinekatalog auch Marketingzwecken und PR, denn Digitalisate können lassen sich damit problemlos anderen Medien zur Veröffentlichung zur Verfügung stellen.

„Die Digitalisierung von Sammlungskatalogen kann die komplette Museumsproduktion lahmlegen.“ Kulturbranche

Beispiel: Virtuelle Museums- und Galerieführer

Neue Endgeräte wie Tablet-PCs zum Beispiel versprechen eine emotionale, bequeme Kulturvermittlung. So kann der Rezipient nach Belieben - sozusagen vom Sofa aus - Ausstellungsrundgänge vornehmen, dabei Inhalte nach Belieben mit Zusatzinformationen vertiefen und mit spielerischen Elementen anreichern. Museums-Apps, wie sie etwa mit der Medienplattform AppShelves vom Peter Delius Verlag entwickelt werden, bieten zum Beispiel individualisierbare Führungen für Museumsausstellungen oder umfassende Kulturevents. Sie binden Bewegtbild und Sound ein und liefern Zusatzservices wie e-Cards, Adressverzeichnisse, Restaurant-Tipps, Taxi-Ruf, Terminkalender, Stadtplanverortung und aktuelle Social Media-Meldungen zu den Kulturevents. Die Services lassen sich beliebig weiterdenken, etwa mit Bilderkennung und Datenvisualisierungen. Verknüpfungen mit dem Social Web

ermöglichen durch Userbewertungen Feedback über Ausstellungen und Kunstwerke, die sich anschließend in die Planung der Museen integrieren lassen. Dieses nonlineare, intuitive nach individuellen Rezeptionsbedürfnissen ausgerichtete Erleben von Ausstellungen bietet völlig neue Chancen der Kulturvermittlung. Zudem lassen sich neue Zielgruppen wie Kinder, junge, bildungsfernere Personengruppen oder kranke und immobile Menschen erschließen.

Semantische Webtechnologien können hierbei nicht nur helfen, die heterogene Datenmasse zu bewältigen, einen emotionalen Zugang zu Kultur zu bieten und intelligente Suchmaschinen zu bieten, sondern auch, diese Daten zu analysieren. Vor allem die Auswertung des Rezeptionsverhalten der Anwender lässt sich für die Ausstellungs- und Publikationsplanung sowie für das Marketing der Institutionen verwerten.

Zuletzt bieten digitale Sammlungen neue Einnahmequellen für Kulturinstitutionen, etwa durch den Verkauf digitaler Abbildungen. Eine intelligent und umfassend annotierte Datenbank kann als Basis für die Entwicklung zunehmend wichtiger digitaler Produkten wie mobile Apps dienen. Indem die vorhandenen Kulturprodukte in kleinste Informationseinheiten unterteilt und für innovative Anwendung neu zusammengestellt werden, lassen sie sich für Mehrfachverwertungen, die Ballung bzw. Neukombination unterschiedlicher monetarisierbarer Services nutzen.

Herausforderungen beim digitalen Katalogisieren

Die Realität in vielen Kulturinstitutionen sieht allerdings häufig anders aus. Aufgrund des Umfangs der Materialsammlungen bei gleichzeitig personellen Engpässen vieler Kulturinstitutionen sind viele Archive bislang nur teilweise digital dokumentiert. Der Initialaufwand für die Digitalisierung von Kultursammlungen ist sehr hoch, erfordert Disziplin und Fachkompetenz von den Institutionen.

Katalogisiert werden die Daten meist nach Künstlernamen, Titeln, Entstehungsjahr, Lebensdaten der Künstler und Wissenschaftler, Stilepochen, Entstehungsort, Materialien und Techniken, Medienformaten und -größen, Motiven und Themen sowie Lagerungsort. Diese Gliederungen sind oft nicht konsequent oder fehlerhaft durchgeführt und nur lückenhaft mit Inhalten gefüllt. Regeln für Schreibweisen etc. müssen also vereinheitlicht werden, und trotzdem sollten zum Beispiel alle druckgrafischen Werke in der Trefferliste erscheinen, gleich ob der Nutzer das Stichwort „Druckgrafik“ oder „Druckgraphik“ eingibt. Bei der Eingabe des Stichwortes „Albrecht Dürer“ etwa muss der Nutzer Dürer den Älteren als auch den Jüngeren spezifizieren können. Auch Werke, die von dem genannten Künstler inspiriert sind, sollten auftauchen und als solche gekennzeichnet werden. Im Voraus geklärt werden müssen auch die Filter. Der Filter „Künstler“ zum Beispiel kann unterschiedlichste Absender-Arten der heterogenen Objekte umfassen: Künstler, Gestalter, Handwerker,

„Dass viele digitale Sammlungen nicht funktionieren, ist weder ein technisches noch ein fachliches Problem, sondern ein kommunikatives.“ Kreativwirtschaft

„Die Karteikartensysteme vieler Kulturarchive sind für Außenstehende eine Black Box.“ Kulturareiter

Wissenschaftler, Autoren, Verlage und Hersteller. Gleichzeitig gilt es, trotz möglichst präziser Suchoptionen auch Zufallsfunde im Archiv zuzulassen. Bislang werden diese Eingabearbeiten meist von Praktikanten vorgenommen, die Ergebnisse allerdings oft nicht überprüft und nachgehalten. Eine relativ umfassende, gut funktionierende Suchmaschine bietet heute etwa das British Museum in London (britishmuseum.org/research/search_the_collection_database.aspx). Die Institution brauchte zur Entwicklung der Datenbank allerdings 10 Jahre.

Beispiele digitaler Sammlungskataloge

Je breiter gefächert die Themenfelder eines Archivs, desto komplexer sind die Anforderungen an die Suchmaschine und ihre Suchfilter. Als Vorzeige-Projekt gilt die Europeana (europeana.eu), ein von der EU finanzierter digitaler Sammlungskatalog von etwa 1500 europäischen Museen, Archive, Bibliotheken und audiovisuellen Sammlungen. Hier lassen sich Digitalisate europäischer Kultursammlungen nach Medientypen (Bild, Text, Film und Sound), nach Sprache und Datum, Land, Anbieter und Ausstellungen suchen. Nach einem ähnlichen Prinzip arbeitet die Ende 2012 gelaunchte die Deutsche Digitale Bibliothek, ein Gemeinschaftsprojekt von 30.000 Kultur- und Wissenschaftseinrichtungen Deutschlands, mit dem sich Bücher, Denkmäler, Kunstwerke, Tondokumente, Filme und Noten suchen lassen. Bislang funktioniert allerdings noch keine Datenbank reibungslos.

In Deutschland sind es insgesamt erstaunlich wenige Institutionen, die ein umfassendes digitales Sammlungsmanagement betreiben. Eine recht fortgeschrittene Datenbank beherbergt die Staatliche Kunstsammlung Dresden (skd-online-collection.skd.museum/) mit 14 Institutionen. Hier sind etwa die völlig heterogenen Werke der Gemäldegalerie Alter und Neuer Meister neben Münz- und Porzellansammlung sowie Handpuppen und Mathematisch-Physikalischen Wissenschaftsexponaten gelistet.

„Der Nutzer steht im Denken vieler Museen, aber auch in dem von Programmieren noch zu selten im Vordergrund.“ Kreativwirtschaft

Usability

Zudem fehlt vielen Produkten und Werkzeugen der Kulturinstitutionen der Servicecharakter. Dabei wächst, bedingt durch den alltäglichen Umgang mit anwenderfreundlichen Suchmaschinen wie Google, auch für digitale Kulturvermittlungsprodukte die Notwendigkeit ansprechender Nutzeroberflächen mit simplen Eingabefeldern, intelligenten Suchfiltern und Lösungsvorschlägen. Endnutzer digitaler Anwendungen sind immer weniger gewohnt und gewillt, die komplexen Eingabefelder einer „Erweiterten Suche“ zu bedienen.

Ein entsprechendes Beispiel liefert der Getty Search Way (search.getty.edu/gateway/landing), eine Datenbank des Paul J. Getty Museums und des Getty Research Instituts. Die eingesetzten semantischen Technologien werden hier nicht mehr sichtbar, sie sorgen im Hintergrund für passende Suchergebnisse. Auffallend ist, dass Usability-Studien mit Testanwendern für digitale Archive - im Gegensatz zu anderen Ländern - in Deutschland kaum durchgeführt werden.

Auch internen Arbeitsprogrammen, etwa den Museumsdatenbanken, mangelt es gegenwärtig an Anwenderfreundlichkeit. Deutsche Museen arbeiten mit Datenbanken wie Adlib, Museum+ oder APS von Stegmann Systems. Oftmals sind ihre Interfaces sehr nüchtern gestaltet, umständlich und kompliziert zu bedienen. Zudem existieren in Deutschland, anders als im Bibliothekswesen, kaum einheitliche Standards für die Formate der Museumsdatenbanken. Dies erschwert den Austausch von Metadaten zwischen verschiedenen Institutionen.

Bei der Entwicklung digitaler Kulturvermittlung müssten Kulturinstitutionen also nicht nur enger mit Programmierern, sondern auch mit Web- und Kommunikationsdesignern zusammenarbeiten, die die Inhalte konzeptionell und visuell für die jeweiligen Zielgruppen aufbereiten.

Haltungsfragen: Ängste bewältigen, Perspektiven eröffnen

Darüber hinaus kann auch die persönliche Haltung vieler Archivare und Museumsangestellter eine reibungslose Kooperation bei der Digitalisierung von Sammlungen erschweren. Häufig hat diese Berufsgruppe nur ein geringes Interesse daran, den Zugang zu ihren Gliederungssystemen zu öffnen und sich mit den Möglichkeiten digitaler Medien auseinanderzusetzen. Gefürchtet wird erstens, dass sich die Archivarsarbeit durch digitale Archive selbst überflüssig macht, zweitens macht eine Übertragung der eigenen Ordnung ins Digitale auch die über Jahre hinweg entstanden Ungenauigkeiten und Fehler sichtbar.

„Parallel zum Interesse an digitaler Information lässt sich ein Interesse an Erfahrungen im realen Leben, an der Aura des Kunstwerks beobachten. Digitale und herkömmliche Museen werden sich künftig gut ergänzen.“ Kulturarbeiter

Viele Museen sind digitalen Sammlungen gegenüber auch deshalb zurückhaltend, weil sie vermuten, dass eine ihrer Hauptfunktionen, das Ausstellen, überflüssig wird und die Besucher fernbleiben. In aktuellen Befragungen aber berichten Museumsangestellte, die intensiv in digitale Kommunikation investieren, dass sich diese positiv auf Besucherzahlen auswirkt und neue Zielgruppen erschließt. Digitale Ausstellungsinhalte werden zur Vor- und Nachbereitung des Ausstellungsbesuches verwendet und lassen sich für Marketingmaßnahmen aufbereiten.

„Natürlich ist es nicht einfach, Museumssammlungen digital zugänglich zu machen. Denn indem Kulturarbeiter ihre Arbeit transparent machen, lassen sie die Öffentlichkeit auf ihren Schreibtisch schauen – und der ist eben oft nicht besonders aufgeräumt.“ Kulturbranche

Neue Arbeitsfelder für Kulturberufe

Um also umfassende Datenbanken für Museen zu erstellen, und die Digitalisate mit Zusatzinformation anzureichern, müssten in den Kulturinstitutionen eigene Stellen eingerichtet werden. Erforderliche Fertigkeiten sind wissenschaftliche Fachkenntnisse, Kenntnisse des Archivwesens und der Informatik sowie Gründlichkeit, Geduld und Ordnungssinn. In vielen Studiengängen würden dadurch auch neue Ausbildungsinhalte an der Schnittstelle zwischen IT und Geisteswissenschaften notwendig.

Klar ist auch, dass die extrem monotone Arbeit des Annotierens nur in Teilzeit- und Assistenzstellen übernommen werden kann. Ob Wissensarbeiter, die tendenziell eher am Entwickeln von Konzepten interessiert sind, freiwillig eine derartige Arbeit in größerem Ausmaß übernehmen würden, wird von einigen Kulturarbeitern bezweifelt. Die Übertragung analoger Karteikartensysteme lässt sich im besten Fall zum Teil auch durch Texterkennungsprogramme vereinfachen. Sinnvoll ist ein zusätz-

„Metadaten für Sammlungsdatenbanken einzugeben ist eine monotone Tätigkeit, die niemandem mehr als 4 Stunden täglich und nur wenige Tage in der Woche zuzumuten ist.“ Kulturarbeiter

liches Lektorat, das die Annotationen auf orthografische und inhaltliche Richtigkeit überprüft. Darüber hinaus lassen sich die Qualität der Veredelung durch Crowd Sourcing, also die Ergänzungen recherchierender Nutzer, optimieren. Gute Erfahrung damit macht etwa das British Museum, das für seine digitale Datenbank ein Feedbacksystem mit automatisierten Emails für die Behebung von Fehlern eingerichtet hat.

Daneben übernehmen Projektmanager eine verantwortungsvolle Aufgabe, sobald ein Projekt zur Digitalisierung einer Sammlung ins Leben gerufen wird. Diese Tätigkeit verlangt ein Verständnis der kuratorischen Arbeit, Grundwissen der digitalen Gestaltung als auch Informatikwissen. Zudem muss ein Projektmanager mit den Mentalitätsunterschieden der Kooperationspartner und ihren interdisziplinären Grabenkämpfen umgehen können.

Weitere Arbeitsprofile bieten sich bei der Konzeption und Erstellung der Datenbanken und ihrer visuellen Gestaltung: Im Bereich der Kunstgeschichte existieren bereits Studiengänge wie „Digitales Sammlungsmanagement“, aber auch für Bibliothekare und Webdesigner eröffnen sich dabei neue Tätigkeitsfelder.

„Studenten und Kunsthistoriker, die heute das Annotieren in Museumsdatenbanken übernehmen, lernen meiner Erfahrung nach schnell und entwickeln eigene Ideen für neue Muster und Ontologien.“ Kulturarbeiter

Fazit

An der Erstellung digitaler Sammlungskataloge geht kein Weg vorbei. Digitale Archive bieten Kulturinstitutionen einen großen Mehrwert für die internen Arbeitsprozesse, für Marketing und PR sowie für neue Erlösmodelle. Semantische Webtechnologien lassen sich hier sinnvoll einsetzen; klar ist aber, dass die Annotation der Metadaten ein hohes Maß an Fachwissen erfordert und deshalb eine manuelle Arbeit bleiben wird. Zudem müssen technologische Standards eingeführt werden, die die unterschiedlichen Plattformen kompatibel machen. Darüber hinaus nimmt die Bedeutung attraktiver, emotionaler Interfaces und intuitiver Bedienbarkeit digitaler Produkte zu. Das „digitale Museum“ mit virtuellen Museumsführungen und spielerischen Apps macht die Kulturvermittlung emotionaler, was eine engere Zusammenarbeit zwischen Technik, Inhalt und Gestaltung erfordert.

Bevor sich die Digitalisierung von Kultursammlungen flächendeckend durchsetzen kann, ist allerdings ein Mind Shift in der zum Teil technikfeindlichen Kulturbranche erforderlich. Dieser wird sich aber durch den anstehenden Generationenwechsel nach und nach von selbst vollziehen. Die brancheninterne Angst, dass digitale Archive die geisteswissenschaftliche Arbeit völlig überflüssig machen, wird sich nicht bestätigen - im Gegenteil können sich Kuratoren und andere Wissenschaftler stärker ihrer eigentlichen Arbeit, dem Selektieren, Bewerten, Interpretieren und Aufbereiten von Inhalten, zuwenden. Neue Arbeitsfelder wird die Konzeption neuer digitaler Kulturvermittlungsprodukte wie mobile Apps beinhalten. Voraussetzung dafür ist, dass Staat und Kulturförderer ein Problembewusstsein für die Notwendigkeit entwickeln, Sammlungskataloge so früh wie möglich ins Digitale zu übertragen, um langfristig zu profitieren.

Dazu werden zunächst umfangreiche Investitionen erforderlich, etwa für eigene Stellen zwecks semantischer Datenveredelung. Insgesamt eröffnet die Entwicklung digitaler Sammlungsarchive viele neue Tätigkeitsfelder und Ausbildungsinhalte für Bibliothekare, Redakteure und Kuratoren an der Schnittstelle zwischen Geisteswissenschaft und IT.

4.3 Bibliotheken

Bibliotheken kommt traditionell eine Vermittlerfunktion zwischen Inhaltsproduzenten und Nutzern von Inhalten zu. Mit der Digitalisierung verschiebt sich aber ihre konkrete Rolle weg von „Aufbewahrungs-Einrichtungen“ und auch ihre Bedeutung als Zugangsorte mit Quasimonopol für den Zugriff auf Bücher und Zeitschriften schwindet. Vor diesem Hintergrund ist aktuell ein rasanter Wandel zu beobachten sowohl in Bezug auf die Rolle der Bibliotheken und der Bibliothekare selber, als auch auf die damit zusammenhängende Ausbildung von Bibliothekaren. Es wird in Zukunft immer weniger darum gehen, Nutzern den Zugang zu physischen Trägermedien des Wissens zu ermöglichen, sondern vielmehr darum, ein möglichst effizientes und komfortables Auffinden relevanter Inhalte zu gewährleisten.

Systematische Ordnungssysteme

Grundsätzlich sind Bibliothekare sehr gut ausgebildet und trainiert in der Erstellung und Nutzung systematischer Ordnungssysteme. So werden die in einer Bibliothek vorhandenen Bücher und Zeitschriften nach strengen Vorgaben und Regeln katalogisiert und mit Metadaten ausgezeichnet. Dies umfasst zum einen Angaben zu inhaltlichen Themenbereichen, in sehr viel größerem Umfang aber nicht-inhaltliche Daten wie Erscheinungsjahr, Seitenzahl, Verlag, Erscheinungsort, Auflage, etc. Der Bibliothekswissenschaftler Prof. Stefan Gradmann spricht in diesem Zusammenhang daher vom traditionellen Fokus der Bibliotheken auf Container-Eigenschaften - also auf strukturelle Angaben - statt auf Inhalte und davon, dass sich mit dem Verschwinden des Trägermediums Buch „die Gutenberg'sche Klammer schließt“.

Anhand von Bibliotheken lässt sich auch sehr anschaulich verdeutlichen, wie komplex die Festlegung und Anwendung von Ordnungssystemen sein kann. So gibt es vermutlich mehr Varianten von Ordnungssystemen als Bibliotheken. Die Änderung von Ordnungssystemen im Laufe der Zeit mit entsprechenden negativen Konsequenzen für die Auffindbarkeit von Büchern, die nach veralteten Ordnungssystemen eingeordnet wurden, ist ein häufig wiederkehrender Topos in der Literatur.

Beispiel: Europeana

Europeana (europeana.eu) ist eine virtuelle Bibliothek, über die die digitalen Bestände in Europas Museen, Bibliotheken, Archiven und audio-visuellen Sammlungen online erkundet werden können. Europeana umfasst Bilder (Gemälde, Zeichnungen, Karten, Fotografien und Bilder von Museumsobjekten), Texte (Bücher, Zeitungen, Briefe, Tagebücher und Archivalien), Tonaufnahmen (Musik und Gesprochenes von Phonographenzylindern, Tonbändern, Schallplatten

und aus Rundfunkstendungen) sowie Videos (Filme, Wochenschauen und Fernsehsendungen) aus etwa 1.500 Institutionen. Auf nationaler Ebene entwickelt Deutschland mit der Deutschen Digitalen Bibliothek (DDB, www.deutsche-digitale-bibliothek.de) ein in Europeana eingebundenes, nationales Portal für digitalisierte Kunst- und Kulturgüter. Dieses ist gegenwärtig als Beta-Version nutzbar, weist jedoch bislang noch zahlreiche Lücken in den Suchmöglichkeiten auf.

Beispiel: Dublin Core Metadata Initiative

Dublin Core (dublincore.org) legt ein Grundvokabular für die Beschreibung von Dokumenten und Objekten mit Metadaten fest. Ziel war dabei unter anderem, durch die Verwendung von Metadaten die Ergebnisse von Suchmaschinen zu verbessern. Die Entwicklung von Dublin Core begann Mitte der 90er Jahre und kann somit als Vorläufer weitergehender

Entwicklungen wie RDF oder Linked Data gesehen werden. Es ist daher nicht verwunderlich, dass Bibliotheken bei der Entwicklung von Standards wie z.B. Dublin Core maßgeblich beteiligt sind und dass Formalismen des Semantic Web wie RDF oder OWL inzwischen an vielen Lehrstühlen der Bibliothekswissenschaft zur Basisausbildung gehören.

Sonderstellung der Bibliotheken

In Bezug auf das Semantic Media Web nimmt das Anwendungsfeld Bibliotheken in zweifacher Hinsicht eine Sonderstellung ein. Zum einen werden Bibliotheken als neutral angesehen, was sie zusammen mit ihrer natürlichen Affinität zur Systematik geradezu idealerweise für eine führende Rolle in der Qualitätssicherung prädestiniert. Zum anderen ist ihre Grundfinanzierung mehr oder weniger gesichert, in jedem Fall sind sie nicht unmittelbar darauf angewiesen, ihre Aufwände über die erfolgreiche Positionierung von Produkten oder Dienstleistungen und entsprechende Marktanteile zu refinanzieren. Von daher sind Bibliotheken auch hervorragend als Pilotanwender semantischer Technologien in Forschungs- und Entwicklungsprojekten geeignet.

Das Geschäftsmodell verändert sich dabei - wie auch in anderen Anwendungsfeldern - von vergleichsweise einfachen Produkten und Dienstleistungen hin zu komplexen Angeboten, d.h. konkret weg von der Bereitstellung von Inhalten hin zur Bereitstellung umfassender Mehrwertdienste wie Suchen, Zusammenstellen und Übersetzen von Informationen - also einer kreativen, hochwertigen Recherche. Dies kann auch bedeuten, dass Bibliotheken Medien, die nicht vor Ort verfügbar sind, kostenpflichtig zusammentragen und gegebenenfalls auch durch Aufbereitung oder Übersetzung veredeln. Als Erlösmodelle für solche Dienste kommen dabei dauerhafte Nutzungsgebühren für bestimmte Services in Frage aber auch Varianten, die auf Pay per Use basieren – analog zu bereits heute existierenden, kostenpflichtigen Diensten wie Spezialausleihe, Kopierservice, Recherche oder Übersetzungsdienste.

Der Fokus der Tätigkeit von Bibliothekaren würde sich damit noch stärker in Richtung einer fachlichen Erschließung von Informationen und Wissensgebieten verschieben sowie hin zur intelligenten Strukturierung und Vernetzung dieser Wissensbereiche, inklusive Dokumentation sowie interdisziplinärer und interkultureller Unterstützung. Aus dem traditionellen Bibliothekar würde somit mehr und mehr ein Wissensmann.

„Die Zusammenarbeit und Harmonisierung auf europäischer Ebene funktioniert recht gut. Brüche gibt es eher zwischen unterschiedlichen Akteuren - Bibliotheken, Archiven oder Museen - nicht so sehr zwischen den Ländern. Deutschland hat in einigen Bereichen Vorreiterfunktion und liegt bei der Digitalisierung von Archiven allgemein im guten Mittelfeld.“ Archivwesen

„Als öffentliche Einrichtung haben wir als Archivare einen Serviceauftrag, dem wir unter Anderem auch durch Open Access nachkommen wollen.“ Archivwesen

FAZIT

Bibliotheken können bei der nachhaltigen Etablierung des Semantic Media Web eine tragende Rolle spielen. Zum einen verfügen sie über eine Grundfinanzierung und sind daher nicht von einer unmittelbaren Refinanzierung über Produkte und Marktanteile abhängig. Zum anderen haben sie eine sehr lange Tradition in der Arbeit mit systematischen Ordnungssystemen. Hinzu kommt, dass sie nachhaltig als Orte des Wissenszugangs etabliert sind und traditionell eine Mittlerrolle zwischen Verlagen und Lesern einnehmen.

Zudem werden schnell erhältliche und zugleich glaubwürdige, transparent recherchierte Inhalte im Informationszeitalter wichtiger. Und da Recherche im digitalen Zeitalter nicht, wie oft behauptet, einfacher, sondern dank häufig fehlender Quellenangaben und unsachgemäßen Informationen im Web komplizierter wird, ist davon auszugehen, dass derartige Dienstleistungen an Bedeutung gewinnen.

4.4 Bewegtbild

Aufgrund der technologischen Entwicklung der letzten Jahre gewinnt der Bereich Bewegtbild zunehmend an Bedeutung. Die kontinuierlichen Fortschritte bei Übertragungsbandbreiten, Rechenleistung (CPU, Grafikkarten), Speicherkapazität und Datenkomprimierung sowie Flatrate-Angebote von Internet-Providern haben dazu geführt, dass Videos inzwischen auch in hohen Auflösungen über das Internet gestreamt werden können und auch zunehmend auf mobilen Endgeräten nutzbar sind.

Angebote in diesem Bereich umfassen Websites von Fernsehsendern, Videoportale wie YouTube (youtube.com) oder Vimeo (vimeo.com), NetTV-Angebote wie freshmilk.tv (freshmilk.de) sowie Plattformen für Video-on-Demand-Streams wie Netflix (netflix.com). Darüber hinaus werden Videos im Zuge der Konvergenz hin zu Multimediaformaten auch mehr und mehr in Artikel von Online-Zeitungen und Blogs eingebunden. Dies umfasst zunehmend die Einbeziehung von User Generated Content.

Wie auch im Anwendungsszenario Verlage gibt es dabei etablierte Anwender wie beispielsweise die Fernsehsender, die ihr Kernangebot um ein Online-Angebot erweitern, und neue Anbieter, die sich ganz auf Online konzentrieren und sich nicht mit „Altlasten“ und der Kannibalisierung ihres Kernangebots konfrontiert sehen.

„Man könnte Annotationen nicht nur auf Clip-Ebene, sondern auch auf Szenen-Ebene vornehmen. Im Durchschnitt wird die Betrachtung eines Clips auf YouTube aber nach 30 Sekunden abgebrochen. Interaktive Elemente und kontextuelle Informationen durch semantische Annotation erscheinen ohne passende Technologie noch sehr aufwendig, würden Verweildauer und Involvement aber stark verbessern.“ Web-TV-Betreiber

Rechtliche Fragen

Auch im Bereich Bewegtbild kommt rechtlichen Fragen in der aktuellen Diskussion eine große Bedeutung zu, wie der seit Jahren andauernde Konflikt zwischen GEMA und YouTube als Negativbeispiel veranschaulicht. In gewisser Weise durchläuft die Bewegtbild-Branche gegenwärtig eine ähnliche Entwicklung wie die Musikbranche einige Jahre zuvor. Durch die Digitalisierung der Inhalte ist prinzipiell die technologische

Beispiel: Deutsche Kinemathek

Die Deutsche Kinemathek (deutsche-kinemathek.de) setzt für Archivierung, Katalogisierung und Kontextualisierung das webbasierte Sammlungsmanagementsystem CollectiveAccess (collectiveaccess.org) ein, das von zahlreichen wissenschaftlichen und kulturellen Institutionen zur Dokumentation und Präsentation digitaler Objekte verwendet wird. Zentrale Anforderungen der Deutschen Kinemathek waren dabei insbesondere

die Unterstützung kollaborativer Prozesse, zentraler Standards, zeitbasierter Verschlagwortung und flexibler Datenmodellierung. Konkret hat die Deutsche Kinemathek u.a. die Projekte Lost Films (lost-films.eu) und Wir waren so frei ... (wir-waren-so-frei.de) mit OpenAccess umgesetzt. Die Annotierungen erfolgen manuell mit entsprechend hohem Aufwand.

Beispiel: Klickfilm

Klickfilm (klickfilm.com) ist ein Spin-Off der Freien Universität Berlin, das neue Verfahren für die Anreicherung von Bewegtbild mit Informationen entwickelt und über die Einbindung innovativer Technologien wie Ontologien, Linked Data, Crowdsourcing oder maschinelles Lernen Monetarisierungsalternativen anbietet. Durch die teilautomatisierte Markierung und Beschreibung von Szenen und Objekten in Videos und die vollautomatisierte Verknüpfung dieser Objekte mit Informationen, kann auch eine nutzergesteuerte Werbung erfolgen, die einen nutzerfreundlichen

neuen Werbekanal öffnet. Streuverluste für Werbetreibende und unbeliebte Push-Werbung könnte durch zielgruppenspezifische und durch Nutzer selbst abrufbare Informationen (Pull) verhindert werden. So kann beispielsweise auf Produkte aus Online-Shops verlinkt werden, die den im Bild aktuell dargestellten Kleidungsstücken, Möbeln, Gebrauchsgegenständen direkt oder indirekt entsprechen oder diese assoziieren. Für die automatische Bestimmung passender Produkte sollen semantische Informationen auf Basis von Ontologien verwendet werden.

Beispiel: Freshmilk

Freshmilk NetTV (freshmilknet.tv) kreiert und produziert Branded Entertainment Konzepte und betreibt WebTV-Sender mit einem spezialisierten Fokus auf Musik, Mode und urbanen Lifestyle. Die Inhalte, die aktuell mehr als 4.000 Videos umfassen, werden von einem internationalen Redaktionsteam produziert und täglich aktualisiert. Das Angebot ist klar auf die urba-

ne Lifestyle-Elite als Zielgruppe ausgerichtet und zielt auf eine optimale Marken-Inszenierung ab. Über das Urban NetTV Network werden täglich mehrere Tausend Zuschauer erreicht, wobei auch mobile Endgeräte und internetfähige Fernseher unterstützt werden. Die durchschnittliche Sehdauer liegt bei 11 Minuten.

Grundlage geschaffen, diese ohne Qualitätsverlust beliebig häufig zu kopieren. Und wie die Musikbranche beklagt die Filmbranche Umsatzeinbußen aufgrund der Nutzung von Tauschbörsen oder Filehoster. Gegenwärtig experimentiert sie deshalb mit Verfahren zum Kopierschutz und Digital Rights Management (DRM).

Dabei ist in beiden Branchen, wie übrigens auch in der Spielebranche, das Verschwinden eines physikalischen Datenträgers (CD, DVD, Blu-ray) zu beobachten. Musikstücke oder Filme werden zunehmend als Dateien gekauft und über das Internet auf die eigenen Endgeräte geladen oder über Streaming-Dienste genutzt. Darüber hinaus gibt es zunehmend Flatrate-Angebote, bei der nicht mehr für die Nutzung einzelner Inhalte bezahlt wird, sondern in Form beispielsweise monatlicher Raten für die Nutzung eines Dienstes, der den Zugriff auf ein breites Inhaltsangebot bereitstellt. Es gibt aber auch entscheidende Unterschiede zur Musikbranche. Zum einen werden Musiktitel in der Regel mehrfach gehört, während Filme im Normalfall meist nur einmal gesehen werden. Zum anderen wurde Musik traditionell als Album oder CD verkauft, während durch die Online-Nutzung der Kauf einzelner Titel möglich wurde. Für einen solchen Wechsel der Kaufgewohnheiten gibt es in der Filmbranche keine Entsprechung.

Automatische Bildverarbeitung

Im Anwendungsfeld Bewegtbild können Technologien zur automatischen Bildverarbeitung und -erkennung zur Veredelung von Inhalten eingesetzt werden. Dies umfasst ein Spektrum von der Szenenerkennung über die Erkennung und das Verfolgen von Objekten in einer Szene bis hin zur Gesichtserkennung. Entsprechende Annotationen von Hand durchzuführen ist prinzipiell auch möglich, aber mit hohem Aufwand verbunden.

Im Unterschied zu Texten, die ebenfalls eine sequentielle Struktur aufweisen, lassen sich Filme sehr viel schwieriger „überfliegen“. Daher ist auch der Nutzen von Mehrwertdiensten auf Basis von Metadaten grundsätzlich höher einzustufen, beispielsweise in Form einer semantischen Suche für Filmsequenzen mit bestimmten Eigenschaften (Schauspieler, Objekte, Dialoginhalte, Orte, etc.).

Obgleich solche Mehrwertdienste für professionelle Nutzer sicherlich sehr nützlich sind, ist unklar, wie relevant sie für die eher dominierenden Nutzungsszenarien im Kontext von Videoplattformen wie YouTube oder Vimeo sind. Wahrscheinlich werden solche Plattformen Schritt für Schritt Verfahren integrieren, die Clips und Szenen automatisch mit Metadaten annotieren und somit die Suche nach Clips und Szenen für Endnutzer komfortabler machen. Da hierbei keine manuellen Aufwände anfallen, bleiben die Kosten überschaubar – die Refinanzierung erfolgt dann nicht direkt durch die Nutzer, sondern indirekt, z.B. durch die Verbesserung der Wettbewerbsfähigkeit, Alleinstellungsmerkmale gegenüber konkurrierender Angebote oder Steigerung der Besucherzahlen.

„Ein Problem bei Open Data ist die fehlende Nachverfolgbarkeit der Datennutzung - wo landen die Daten denn im Endeffekt?“

Bibliothekarswesen

„Das föderalistische System in Deutschland macht die Datenbereitstellung nicht unbedingt einfacher, da diese häufig nicht zentral geregelt werden kann, sondern auf Länder- oder sogar Bezirksebene erfolgen muss.“ Dienstleister

Zusatzinformationen über Layer-Technologien

Ein weiteres Anwendungsszenario für das Semantic Media Web im Bereich Bewegtbild erschließt sich durch die Bereitstellung zusätzlicher Informationen zu Inhalten, Darstellern, Orten bis hin zu einzelnen Objekten. Dies kann auf einem zusätzlichen Gerät in Anwendungen wie Second Screen oder z.B. über Layer-Technologien, also eigene Ebenen, direkt im Bild geschehen. Wie das Beispiel Klickfilm zeigt, können über diesen Ansatz sehr spezifische, kontextuelle Zusatzinformationen eingebunden werden, die Nutzer adäquat zum Videotext bequem abrufen können.

Fazit

Das Anwendungsfeld Bewegtbild bietet aus mehreren Gründen ein sehr hohes Potential für das Semantic Media Web. Zum einen entstehen hier gegenwärtig innovative Angebote, zum anderen ist eine Technologieunterstützung und die Verwendung automatischer Verfahren naheliegender und unverzichtbarer als bei der Arbeit mit Texten. Auch die zunehmende Medienkonvergenz und der Trend hin zu Multimedia und vor allem der stark zunehmende Einsatz von bewegtbild in der Kommunikation kann in diesem Zusammenhang adäquat vom Semantic Media Web unterstützt werden.

4.5 Open Data /Open Government

In diesem Abschnitt wird das Anwendungsfeld Open Data beleuchtet. Die Grundidee von Open Data besteht darin, Datenbestände ohne Einschränkung zur freien Nutzung, Verbreitung und Weiterverwendung zugänglich zu machen. Da die Basisdaten, die Open-Data-Anwendungen zugrunde liegen, häufig aus öffentlichen Verwaltungen stammen und auch viele Anwendungen den Bereich öffentlicher Dienstleistungen betreffen, ist das Anwendungsfeld Open Data eng mit dem weiter gefassten Thema Open Government verknüpft. Darüber hinaus gibt es Gemeinsamkeiten mit anderen „Open-Bewegungen“ wie Open Source, Open Content oder Open Access.

Datenbereitstellung

Die Bereitstellung der Daten, gerade von Seiten öffentlicher Verwaltungen, verläuft aber nicht immer ohne Widerstände. Zum einen sehen viele Verwaltungen die Daten als ihr „Eigentum“ an und fürchten einen Verlust an Einfluss und Kontrolle, wenn diese frei verfügbar gemacht werden. In offiziellen Diskussionen finden sich statt dieser unterschweligen Bedenken meist Verweise auf grundsätzliche Problematiken hinsichtlich Datenqualität und Datenschutz.

Gerade im Hinblick auf die Datenqualität ist Open Data aber eher Teil der Lösung und nicht des Problems. Durch die automatische Verarbeitung von Daten können Fehler und Inkonsistenzen erkannt und beseitigt werden. Ähnliche Erfahrungen haben viele Unternehmen bei der Einführung von E-Business-Lösungen gemacht: Wenn Stammdaten nicht mehr nur im eigenen System gespeichert werden, sondern elektronisch mit Zulieferern und Kunden ausgetauscht werden, führen Inkonsistenzen, die in ausgedruckten Katalogen häufig übersehen werden, zu automatischen Fehlermeldungen.

Das Thema Datenschutz ist dagegen deutlich vielschichtiger. Dies betrifft zunächst grundsätzlich alle personenbezogenen Daten und die damit verbundene Herausforderung der Datenanonymisierung. Teilweise wird aber auch argumentiert, dass eine freie Verfügbarkeit und Verknüpfbarkeit beliebiger Daten problematische Konsequenzen haben kann. Als Beispiel kann hier eine hypothetische GIS-Anwendung genannt werden,

Beispiel: „Apps& The City“

Unter dem Titel „Apps & The City“ (appsandthecity.net) haben der Arbeitskreis Open Data Nahverkehr Berlin zusammen mit der Senatsverwaltung für Wirtschaft, Technologie und Forschung, der VBB Verkehrsverbund Berlin-Brandenburg GmbH, berlinonline.de und der Open Knowledge Foundation am 29. November 2012 zu einem Entwicklertag eingeladen.

An diesem Tag wurden Datensätze und Problemstellungen von Entwickler/innen bearbeitet und Prototypen umgesetzt. Über Konzepte und Prototypen sollte das Potential von Open Data aufgezeigt und weitere

Entwicklungen angeschoben werden. Dabei ging es um Fragen wie: Was kann man mit den Daten des öffentlichen Nahverkehrs tun, wenn sie offen für alle zugänglich sind? Welche Anwendungen, Visualisierungen und Apps können aus ihnen entstehen? Wie können diese dabei helfen den Service der Verkehrsunternehmen und den Alltag von Bürger und Bürgerinnen zu verbessern und zu vereinfachen?

Die Webseite appsandthecity.net enthält Daten, Apps, Beispiele und Links auf Berichte in verschiedenen Medien.

Beispiel: 5 Stufen der Offenheit

Sir Tim Berners-Lee hat 2010 vorgeschlagen, die Erhältlichkeit und Zugänglichkeit von Daten nach „5 Stufen der Offenheit“ zu unterscheiden. Diese bewertet er mit einem bis fünf Sternen (Siehe z.B. w3.org/DesignIssues/LinkedData.html):

1 Stern: Daten sind mit einer offenen Lizenz in irgendeinem Format im Web zugänglich (zum Beispiel als eingescannte Tabelle).

2 Sterne: Daten sind strukturiert und maschinenlesbar zugänglich (z.B. als Excel-Dateien).

3 Sterne: Daten liegen in einem nicht-proprietären

Format vor (zum Beispiel CSV, also Comma Separated Values, statt Excel).

4 Sterne: Offene W3C-Standards wie RDF oder SPARQL werden verwendet.

5 Sterne: Daten sind mit weiteren Daten aus anderen Quellen verlinkt.

Die Grundidee dieses Bewertungsschemas liegt darin, dass Daten umso einfacher zu verarbeiten und umso nützlicher sind, je mehr Sterne „erfüllt“ sind. Doch bereits Daten mit einem Stern sind nützlicher als Daten, die gar nicht offen zugänglich sind.

die die Korrelation zwischen Straftaten und Anteil von Einwohnern mit Migrationshintergrund visualisiert.

Die Grundidee von Open Data ist aber gerade, dass Daten generell und nicht für einen bestimmten vordefinierten Anwendungskontext bereitgestellt werden. Wie die bereits vorliegenden Erfahrungen zeigen, entstehen auf Basis frei verfügbarer Daten sehr schnell innovative Anwendungen, die nicht nur abstrakten Nutzen haben, sondern neue Wertschöpfungen oder Prozessoptimierungen ermöglichen - und somit auch unmittelbare positive Effekte auf betriebs- und volkswirtschaftlicher Ebene zeitigen.

Der Nutzen von Open Data

In seinem Vortrag „The Benefit of Open Data - offene Daten aus Verwaltungssicht“ vom 4. Dezember 2013 führt Dr. Wolfgang Both von der Senatsverwaltung für Wirtschaft, Technologie und Forschung des Landes Berlin eine ganze Reihe von Beispielen dafür an, welchen (monetären) Nutzen Open Data haben kann. Die Beispiele umfassen Verbesserungen in der öffentlichen Gesundheit, Abbau von Misstrauen und Korruption, Effizienzverbesserung und Ressourcenschonung sowie die Eröffnung neuer Geschäftsfelder und Wertschöpfungen.¹¹

„Das Ausprobieren neuer Ideen geht sehr einfach, wenn Daten erst einmal vorliegen. Nach unserer Erfahrung taugen 80% unserer Anwendungsideen nichts. Im Moment geht dabei noch viel zu viel Zeit für die Datenbeschaffung drauf.“ Dienstleister

„Wir nutzen frei verfügbare Datenquellen wie DBpedia und Open Street Map und auch selbst-entwickelte Stichwortlisten zur Informationsextraktion.“ Dienstleister

Big Data

Im Hinblick auf das Semantic Media Web und die zukünftige Rolle von Wissensarbeitern im Medienbereich kommt dabei den Themen Datenjournalismus und Big Data besondere Bedeutung zu. Während beim Datenjournalismus die Aufarbeitung, Interpretation und auch Visualisierung von Daten im Vordergrund steht, geht es bei Big Data in erster Linie um die automatische Verarbeitung sehr großer Datenmengen. Beides sind relativ neue Erscheinungen, die eine ganz neue Qualität darstellen und erst durch aktuelle technologische Entwicklungen möglich geworden sind. Ihr Potenzial in Bezug auf das Semantic Media Web und neue Wertschöpfungsketten sind dabei sehr hoch einzuschätzen.

Datenjournalismus

Das Blog *fivethirtyeight* von Nate Silver in der New York Times stellt ein sehr gutes und prominentes Beispiel für Datenjournalismus dar. Auf Basis aller vorliegenden Umfrageergebnisse zur Präsidentschaftswahl in den USA hat Nate Silver täglich das zu erwartende Wahlergebnis prognostiziert und seine Prognosen mit Wahrscheinlichkeiten qualifiziert. Seine Prognosen, die sich nur auf mathematische Verknüpfungen des vorhandenen Datenmaterials stützten und im Vorfeld der Wahl sehr kontrovers diskutiert wurden, erwiesen sich am Wahlabend als ausgesprochen exakt.

Silvers Rechenmodell beim Auswerten der Wahlumfragen in den einzelnen Ländern bezieht unterschiedliche Faktoren ein: etwa die Spendenaukommen für die jeweiligen Kandidaten, demografische Aspekte, die Aktualität der Erhebungen, und Parteistatistiken. Hier wird deutlich, wie erst die spezifische journalistische Fähigkeit des Recherchierens, Bewerrens und Verknüpfens glaubwürdiger Quellen den Daten eine sinnvolle Struktur gibt.^{9,10}

Die Bedeutung des Visuellen: Datenvisualisierung

Zeitungen wie die New York Times, The Guardian und DIE ZEIT, setzen zunehmend auf Datenjournalismus-Blogs, da sich Informationen in Bildern effizient, pointiert und kurzweilig vermitteln lassen.

Die Ergebnisse aus Datenanalysen werden, zum Beispiel mittels Software wie Circos, Gephi, Google Refine Basic oder DataWrapper, in eine verständliche visuelle Form gebracht. Dabei bezeichnet man niederkomplexe Balken-, Kuchen- und Liniendiagramme als „Informationsgrafik“,

die komplexeren Grafiken als „Datenvisualisierung“. Da das öffentliche Interesse an Datenjournalismus zunimmt und die technischen Möglichkeiten das Darstellen komplexester Zusammenhänge und unterschiedlichster Quellen ermöglicht, etabliert sich das Thema Datenvisualisierung gegenwärtig als wichtige Disziplin im Design bzw. in der Visuellen Kommunikation - ebenso wie im Bildjournalismus.

„Interaktive Datenvisualisierungen stoßen auch in der Allgemeinheit immer stärker auf Interesse. Vermutlich deshalb, weil sie die User befähigen, vielschichtige Zusammenhänge nach unterschiedlichen Fragestellungen, entsprechend ihrer individuellen Interessenslage, zu beleuchten.“ Kreativwirtschaft

„Im Moment befindet sich das Thema Datenvisualisierung im Tal der Ernüchterung. Es wird deutlich, dass Algorithmen Daten zwar ansehnlich darstellen, aber nicht unbedingt analysieren und interpretieren können.“ Kreativwirtschaft

„Bilder werden zunehmend als eigenständiges, glaubwürdiges Informationsmedium wahrgenommen. Gestalter sollten sich also ihrer Verantwortung beim Visualisieren von Daten bewusst sein: Bildinformationen – Farben, Formen, Größenverhältnisse - bewerten Fakten genauso wie Worte, allerdings subtiler und emotionaler.“ Kreativwirtschaft

Die Herausforderungen beim Visualisieren komplexer Datensätze lassen sich im Idealfall in engen, transdisziplinären Kooperationen zwischen Designern, Journalisten und Programmierern lösen. Besonders gut sichtbar werden die Schwierigkeiten in „Generativen Grafiken“, die auf Basis beliebiger Datensätze automatisiert, mit spezieller Software wie etwa Processing, entstehen. Sie faszinieren oft mit ihren ästhetischen Qualitäten, präzise Aussagen werden in den komplexen, unübersichtlichen Netzwerkstrukturen allerdings nur selten erkennbar.

Zum Suchen und Artikulieren von relevanten Fragestellungen, Thesen und Geschichten werden also bei der Erstellung von Datenvisualisierungen klassische journalistische Kompetenzen benötigt, während Gestalter

diese Botschaften in pointierte, schnell verständliche, narrative Bilder transformieren. Darüber hinaus übernehmen letztere die wichtige Aufgabe, die Aussagekraft der Bilder verantwortungsvoll einzusetzen. Denn die Botschaften von Daten, mittels Formen visuell in ein Verhältnis gesetzt, wirken in der menschlichen Wahrnehmung schneller, emotionaler und subtiler als reiner Text.

Immer mehr Aufmerksamkeit erhalten interaktive Grafiken, die den Rezipienten involvieren, indem sie ihm individualisierte Datenvisualisierungen auf verschiedene Fragen zu einem speziellen Thema liefern. So können auch große, komplexe Datensätze verknüpft und nach unterschiedlichen Aspekten aufbereitet werden. Hier sind intelligente Algorithmen und Ontologien gefragt, die dem Nutzer ein möglichst freies, intuitives Formulieren seiner Fragen ermöglichen.

Fazit

Open Data bietet ein sehr gutes Anwendungsfeld für das Semantic Media Web. Zum einen gibt es bereits eine hohe Affinität zur Verknüpfung von Daten und damit zu Linked Data. Zum anderen sind viele Daten mit öffentlichen Mitteln erhoben worden, so dass ihre freie Bereitstellung für die Nutzung durch die Allgemeinheit grundsätzlich möglich ist. Mit vergleichsweise wenig Aufwand lassen sich sehr schnell nützliche Anwendungen generieren. Dabei stellt der offene Ansatz sicher, dass die in einer bestimmten Anwendung vorgenommenen Veredelungen von anderen Dienstleistungen weiterverwendet und optimiert werden können. So lassen sich schrittweise unterschiedliche Schichten entwickeln, die schrittweise aufeinander aufbauen und immer komplexere Anwendungsfälle ermöglichen (vgl. hierzu beispielsweise die Schichten oder Layer in OpenStreetMap - openstreetmap.org).

4.6 Corporate Communication

Im Bereich Corporate Communication setzen Großunternehmen zunehmend auf die Expertise externer Dienstleister wie PR-, Strategie- und Designagenturen, um Strategieentwicklung sowie externe und interne Kommunikationsmaßnahmen zu planen und begleiten. Klassische Aufgaben wie Markenkommunikation und Reputationsmanagement werden durch eine dialogorientierte Kommunikation in Sozialen Netzwerken ergänzt. Auch den Themen Datenvisualisierung sowie Verarbeitung und Auswertung sehr großer Datenmengen (Big Data) kommt eine stetig wachsende Bedeutung zu.

Gleichzeitig wird Kommunikation im Raum wichtiger, die durch individualisierte, digitale Medien erweitert werden – etwa auf Messen, in Markenausstellungen und -akademien sowie in weiteren Wissens- und Bildungsräumen. Für die interne Unternehmenskommunikation werden Aspekte relevant, die in der Vergangenheit zumeist unter dem Schlagwort Wissensmanagement bearbeitet wurden, um das komplexe Marken- und Produktions-Know How in global agierenden Firmen zu sammeln, bündeln und weiterzugeben.

Die Markenkommunikation in Unternehmen, also Werbung, Marketing und PR, hat sich durch die Digitalisierung und das Web 2.0 grundlegend verändert. Im Kampf um Aufmerksamkeit muss Markenkommunikation heute vor allem relevant für die Zielgruppe sein, dem Konsumenten einen individuellen Nutzen bieten. Deshalb investieren Unternehmen zunehmend in Content Marketing, also die multimediale Weiterführung von Corporate Publishing: in die Entwicklung eigener, medienübergreifender Publikationsformate wie Print- und Onlinemagazine, Blogs oder redaktionell bespielte Facebook-Auftritte, aber auch neue Serviceprodukte und Alltagshelfer wie mobile Apps. Diese werben nur indirekt für die jeweilige Marke, indem sie dem Leser mit Beiträgen zu Gesellschafts-, Service- und Lifestylethemen einen Nutzwert liefern. So baut sich im besten Fall ein positives Markenimage auf, das mit gesellschaftlicher Relevanz und Funktionalität einhergeht.

Individualisierte Medienplattformen

In der zunehmenden Masse an Corporate Medien werden künftig Medienplattformen relevanter, die nach individuellen Vorlieben der User Inhalte filtern und aufbereiten. Hier bieten semantische Webtechnologien

die Möglichkeit, auf einer Markenplattform Informationen aus den Weiten des Internets zu sammeln und nach beliebigen Kriterien zu filtern, etwa aus Webmagazinen, Blogs und Social Media-Plattformen, die dann nach den individuell gewünschten Aspekten abgerufen werden könnten.

Ein Beispiel für solch ein Corporate Social Publishing Tool ist etwa Anythingabout (anythingabout.de), das unter anderem vom Chemiekonzern BASF erfolgreich genutzt wird. Interessant ist hierbei, dass die Aggregationsleistung vieler dieser Servicetools auf maschinell vorselektierter und dann durch menschliche Kuratierung nachbearbeiteter Informationen basiert, also semi-automatisch erfolgt. Von der Vorstellung komplett automatisierter Verfahren des Semantic Web sind die technischen Entwicklungen aktuell noch weit entfernt.

Beispiel: Webmagazin Mooove

Die digitale Markenplattform „Mooove“, initiiert von der Audi Urban Future Initiative, thematisiert die Rolle der Mobilität in der zunehmend urban geprägten Zukunft. Darin wird ein redaktionell betreutes Online Magazin (mooove.com) mit einer offenen, kontextuellen Suchmaschine (we.mooove.com) kombiniert. Die Suche bezieht sich auf die kuratierten Magazinartikel und die Inhalte aus dem Web, das heißt vorselektierte Websites und Social Networks. Die semantisch generierten Suchergebnisse liefern aber nur bedingt zufriedenstellende Treffer.

Die Plattform wurde etwa über ein Jahr hinweg entwickelt; die Redaktion, betreut von dem Corporate

Publishing-Spezialisten Stylepark für Design- und Architekturthemen, stellte Audi zwei Monate nach dem Magazinlaunch ein. Letztlich steckt in derartigen Projekten ein für nichtkommerzielle Marketingprodukte hoher Entwicklungsaufwand, für den sich Marken bewusst entscheiden müssen. Die notwendigen personellen Ressourcen für eine professionelle Redaktion und die technische Umsetzung, ebenso wie die noch begrenzten technischen Möglichkeiten einer semantischer Suche, werden in Marketingabteilungen häufig unterschätzt – als Endprodukt entsteht dann „nur“ eine virtuelle Website mit Text und Bildern, die nicht das volle Potential des Semantic Media Web ausschöpft.

Social Media als Treiber der Business Intelligence

Semantische Webtechnologien vereinfachen im Bereich der Corporate Communication nicht nur das Aggregieren, sondern auch das Auswerten von Inhalten, die Analyse der Zielgruppen und ihrer Verhaltensweisen sowie Dialoge mit ihnen. "Social Media Intelligence" bezeichnet die Kombination aus Web Tracking, Business Intelligence und Social Media und kann in verschiedenen Bereichen zur Realisierung von Unternehmenszielen eingesetzt werden.

Dazu gehören etwa Wettbewerbsbeobachtung, Trend- und Marktanalyse, Identifikation von Meinungsführern und „Early Adopters“, Customer-Relationship-Management sowie Produkt- und Innovationsmanagement. Hilfreich sind dabei Social Media Monitoring Services wie etwa Brandwatch, IBM COBRA, Visible Intelligence und Dienstleistungen nichtkommerzieller Anbieter wie Bitly, Twittanalyser oder Facebook Insights und Open Facebook Search.

Viele dieser Bereiche lassen sich mit Social Media sogar kombinieren und damit zugleich vereinfachen: Indem Marken das Feedback ihrer Social Media-„Fans“ zum Beispiel umgehend in die Weiterentwicklung von Produkten und Marke einfließen lassen, kann Marketing und Produktentwicklung effizient kombiniert werden. Diese Methode wird etwa von vielen Car Sharing-Plattformen genutzt, deren Erfolg auf dem hohen Nutzwert ihrer Dienstleistung und dem Komfort ihrer digitalen Interfaces basiert.

Unternehmensarchive als Pool für Corporate Heritage

Neue Arbeitsfelder für Bibliothekare, Journalisten und Kunsthistoriker eröffnet die digitale Erschließung von Unternehmensarchiven großer Traditionsunternehmen. Unternehmensarchive zeichnen mit zeitgeschichtlich interessanten Firmenpublikationen, historischen Bild- und Textdokumenten die Firmengeschichte auf, sind aber bisher nur selten digital aufgearbeitet. Hier bieten sie erhebliche Potenziale für PR und die Vermittlung der „Corporate Heritage“. Konsumenten verlangen von Marken heute eine authentische, transparente und wertige Kommunikation – die eigene Tradition zu inszenieren ist dafür eine zunehmend gefragte Möglichkeit. Um die Inhalte dann effizient und komfortabel nutzbar zu machen, wären semantisch angereicherte digitale Archive eine notwendige und sinnvolle Voraussetzung.

Digitale Serviceprodukte in der Markenkommunikation

Einen immer wichtigeren Bereich der Corporate Communication nimmt die Entwicklung nutzwertiger Dienste auf Basis von Hard- und Software ein, die personalisierte Daten analysiert. Dazu gehören zum Beispiel Nike+, das NikeFuel Armband oder der MiCoach von adidas. Hierbei verbinden Marken ihre Expertise, zum Beispiel Sportartikel, mit einem weiteren, für den Alltag nützlichen Bereich, etwa das Gesundheitswesen.

Indem diese Services die Körperdaten des Nutzers beim Sport messen und über semantische Webtechnologien analysieren, dienen sie als Gesundheits- oder Trainingscoach. Solche Services tragen zum „Quantified Self“ bei, also der Tatsache, dass Menschen durch SelfTracking und Feedbackschleifen ihr Verhalten und ihre Leistung kontrollieren. In diesem Bereich steckt viel Potential – das spielerische Vorgehen dieser Tools bei der Selbstbeobachtung und -optimierung kann die intrinsische Motivation des Users aktivieren und eignet sich zum Beispiel für den Einsatz in Medizin oder Bildung. Die Studie „POWER OF OPENNESS - Chancen der Neuen Offenheit für Business, Konsum und Gesellschaft“ vom Zukunftsinstitut bezeichnet dieses Phänomen als eine der wesentlichen zukünftigen Kulturtechniken.

„Ein interessantes Aufgabenfeld könnte darin bestehen, die Archive von großen Unternehmen und Medienhäusern zu digitalisieren. Doch noch sind die Unternehmen oft skeptisch, weil sie keinen direkten Nutzwert erkennen.“ Kulturwirtschaft

Interdisziplinäre Kooperation in Content Marketing und Corporate Publishing

Die meisten Corporate Publishing-Projekte werden von den internen Marketingabteilungen in Kooperation mit Design- oder Kommunikationsagenturen betreut. Deren Kernkompetenz besteht genau darin, Inhalte und Technologie in innovative, ästhetische und funktionale Konzepte zu überführen, angepasst an die Markenwerte des jeweiligen Auftraggebers. Damit ein Produkt von Konsumenten nicht nur akzeptiert, sondern als besonders attraktive „Biografierequisite“¹² passend zum individuellen

Lifestyle und besonders nachhaltig in ihr Leben integriert wird, spielt Design eine wesentliche Rolle. Das führt dazu, dass viele Kommunikationsmedien im Corporate Publishing hochwertig und modern gestaltet sind und sich technisch am Puls der Zeit bewegen.

Einen wichtigen Part wird die Einbindung von Social Media einnehmen. Nicht nur, dass sich per Twitter, Facebook oder Pinterest gezielt junge Zielgruppen ansprechen lassen – der grundsätzlich interaktive Charakter dieser Medien ermöglicht auch die Auswertung von Feedback und die Einbeziehung von User Generated Content.

„Marken wollen unsere Freunde sein. Deshalb müssen sie uns im Alltag mit nützlichen Tools und Services zur Seite stehen.“

Kreativwirtschaft

„Mit intelligenten Corporate Publishing-Produkten schenken Marken ihren Kunden Zeit – die Zeit, die sie selbst an Recherche und Auswertung relevanter Daten sparen.“

Kommunikationsbranche

Doch auch wenn Marken intensiver als Verlage in innovative Medienformate zwischen Redaktion, Technologie und Gestaltung investieren, fehlt in vielen Unternehmen noch immer die Infrastruktur für die interne Umsetzung. Zudem mangelt es oft an Verständnis für den Zeit-, Arbeits- und Kostenaufwand einer fundierten redaktionellen Arbeit und der technischen Umsetzung digitaler Produkte. Voraussetzung für die Einrichtung neuer Stellen wäre ein langfristigeres und interdisziplinäres Denken in den Unternehmen. Dabei gilt es den Wert eigenständiger, hochwertiger und relevanter redaktioneller Arbeit – über kurzfristige, lautstarke Werbung hinaus – für die Unternehmen zu erkennen.

Medienmacht für Unternehmen?

Aller Wahrscheinlichkeit nach werden Presseprodukte aufgrund ihrer monetären Möglichkeiten künftig zunehmend von Marken vermittelt. Generell problematisch ist dabei die Vermischung freier Wissensvermittlung in Forschung und Journalismus mit zielgerichteter kommerzieller Werbung – mit entsprechend negativen Konsequenzen für eine unabhängige und kritische Berichterstattung.

Auch die Entwicklung digitaler Service-Produkte auf Basis personalisierter Daten, aus der Markenkommunikation heraus entwickelt als freundschaftliche Alltagshelfer, sollte kritisch beobachtet werden. Denn kommerzielle Unternehmen erhalten mit der Ansammlung persönlicher Datenmengen samt Analyse eine große politische und gesellschaftliche Macht. Zudem stellt sich die grundlegende Frage, inwieweit sich der Mensch ständig kontrollieren und optimieren lassen möchte. Medienkompetenz wird in Zukunft noch stärker bedeuten, als Rezipient und Konsument diese Mechanismen einzuordnen und Quellen aufmerksam zu hinterfragen.

Die Bedeutung interner Unternehmenskommunikation im Wissensmanagement

Neben der externen Kommunikation gewinnt auch die Kommunikation nach innen an Bedeutung. Je überregionaler Unternehmen agieren, je komplexer ihre Produkte und Prozesse werden, desto wichtiger wird auch der interne Austausch des Firmenwissens.

Zum einen gilt es gerade für wissensbasierte Branchen, globale technologische und wissenschaftliche, mediale, politische und gesellschaftliche Entwicklungen zu erfassen, um sie für eigene Produkt- und Service-Entwicklung – vom Automobil über Telekommunikation bis zum Medizinprodukt – sowie die Markenkommunikation zu adaptieren. Zum anderen müssen die internen Forschungen und Entwicklungen über alle Firmestandorte und Unternehmensbereiche hinweg kommuniziert werden, um einen optimalen Wissensstand zu ermöglichen.

„Mitarbeiter erkennen den enormen Mehrwert veredelter Inhalte erst, wenn eine brauchbare semantische Suche längst existiert. Deshalb ist es wichtig, sie von Beginn an dazu anzuhalten, ihre Web-Inhalte ständig und gründlich zu vertaggen.“ Kommunikationsbranche

Dafür installieren Unternehmen aus wissensbasierten Branchen zunehmend eigene Enterprise Social Networks wie etwa Yammer (yammer.com). Digitale Tools wie Clipboard (clipboard.com) oder Keeeb (keeeb.it) bieten als Mischform aus Diensten wie Pinterest, Delicious und Evernote die Möglichkeit, Bilder, Texte, Videos und Bookmark-Links zu sammeln, nach Themen zu gliedern und mit ausgewählten Personen zu teilen. Ein

enormer Mehrwert ergibt sich, wenn Social Networks mit Wikis, Adressbüchern sowie Projektmanagement- und Kollaborationstools wie Basecamp (basecamp.com) gebündelt und in einer Suchmaschine ganzheitlich verknüpft werden.

So können Informationen aus ganz heterogenen Quellen veredelt werden: externe Medieninhalte, aber auch die internen Kontaktlisten, Projekt-Inhalte wie Korrespondenzen, File-Sharing, Meilensteine und To-Do-Listen. Je mehr Quellen und Plattformen hier andocken können, desto größer der Nutzen. Das setzt aber voraus, dass beim Erstellen digitaler Inhalte alle eingegeben Informationen konsequent und diszipliniert mit allen Metadaten angereichert werden, die für das Unternehmen relevant sind oder werden könnten. Diese Metadaten können dann bei der weiteren Nutzung von allen Mitarbeitern kollaborativ ergänzt und verfeinert werden.

Herausforderungen

Nach wie vor leiden Geschäftsprozesse vor allem in großen Unternehmen unter einem lückenhaften bzw. mit Medienbrüchen behafteten Informationsfluss. So gibt es oft noch zu wenige Schnittstellen zwischen den unterschiedlichen Unternehmensbereichen und den jeweils verwendeten IT-Systemen. Das führt zu Missverständnissen und redundanten Arbeitsschritten bei Mitarbeitern und Kunden. Interessante Tätigkeitsfelder eröffnen sich also in der Konzeption solcher Systemzusammenführungen und intelligenten Verknüpfungen von Unternehmensbereichen.

„Beim Erstellen von Ontologien und Annotationen für Unternehmen ist die Fähigkeit zu strukturiertem Arbeiten wichtiger als tiefe Fachkenntnisse. Archivare und Bibliothekare könnten hier ihr optimal ihr Wissen einbringen.“ Technologieanbieter

„Qualitätsjournalismus in der Unternehmenskommunikation ist eine Frage der Firmenkultur. Um wirklich gelungene Corporate Publishing-Formate zu entwickeln, die freier Redaktion mehr als Werbung ähneln, müssen sich erst die Strukturen und Hierarchien in vielen Unternehmen ändern.“ Kommunikationsbranche

Beispiel: SemTalk/Semtation

Ein Tool zur Vereinfachung unternehmensinterner Arbeitsabläufe ist die Prozessmodellierungs-Plattform SemTalk (semtalk.com), entwickelt von dem Berliner Dienstleister Semtation. Kunden sind unter anderem Versicherungen, Maschinenbauunternehmen, Banken und Energieversorger. Indem sich über die Anwendung Informationen zu bestimmten Arbeitsprozessen suchen lassen, können Arbeitsabläufe verdichtet und beschleunigt werden. Semtation entwickelt für jeden Kunden eine individuelle Darstellung des Unternehmensprozesses – eine zwischen Unternehmensberatung, Informationsdesign und Programmierung angesiedelte Dienstleistung. Semtation übernimmt in enger Kooperation mit den Firmen auch das Modellieren der Begriffssysteme, der Ontologien und Annotationen.

Dabei stellt ein grafisches Interface die jeweiligen Arbeitsprozesse als Bild dar, die in den eigens erstellten Ontologien im Hintergrund – mehrere 100 Konzepte pro Prozess - ihre inhaltliche Entsprechung finden. Wichtig ist die Entwicklung parallel existierender, rollenspezifischer Ontologien, zwischen denen die Anwender je nach Arbeitsbereich wechseln können. Die Annotationen bestehen vor allem aus firmen- und branchenspezifischen Fachbegriffen, die die Mitar-

beit der Unternehmen erfordert. Per Crowdsourcing werden die Annotationen verbessert, dennoch müssen diese Tags von einem fachkundigen Verantwortlichen gesteuert und überprüft werden. Das heißt, das Ver-taggen bleibt eine manuelle Arbeit.

Interessant ist, dass das Unternehmen seine Leistung als „Prozessmodellierung“ verkauft und nicht unter dem Begriff „Semantisches Web“. Die Dienstleister haben die Erfahrung gemacht, dass das Schlagwort in Unternehmen zum Teil negativ konnotiert ist – möglicherweise aufgrund der geringen Erfolge der zeitaufwändigen Forschungen zu diesem Thema in der Vergangenheit. Um die Jahrtausendwende wurde das Semantische Web in Forschung und der IT-Branche als Wunderwaffe für die Zukunft des Internets propagiert und mit großem Einsatz gefördert, allerdings bis heute mit mäßig funktionierenden Ergebnissen. Dass die Wortwahl für die Semtation-Dienstleistung das Konzept in den Vordergrund rückt, demonstriert die Notwendigkeit, den Nutzen einer Anwendung und nicht ihre technische Grundlage in den Vordergrund zu stellen. Semantische Webtechnologie sollte ein Produkt also wie ein gut ausgebildeter Butler beflissen im Hintergrund vereinfachen.

Hier gibt es Überschneidungen zwischen dem Anwendungsfeld Corporate Communication und dem Bereich Geschäftsprozessoptimierung. Tools wie die Prozessmodellierungsplattform SemTalk des Berliner Dienstleisters Semtation stellen hier semantische Lösungen auf der Basis von Ontologien zur Verfügung, mit denen Arbeitsabläufe teilautomatisiert, verdichtet und beschleunigt werden können.

QUELLEN:

- 9) Nate Silver, For Romney to Win, State Polls Must Be Statistically Biased , 3.11.2012
<http://fivethirtyeight.blogs.nytimes.com/2012/11/03/nov-2-for-romney-to-win-state-polls-must-be-statistically-biased/>
- 10) Patrick Beuth, Nate Silver zeigt die Macht der Algorithmen, 08.09.2012
<http://blog.zeit.de/open-data/2012/11/08/nate-silver-usa-wahl-statistik/>
- 11) Beitrag Zum Innovationsforum Semantic Media Web
www.semantic-media-web.de/downloads/files/1_Dr_Wolfgang_Both_The_Benefit_of_Open_Data.pdf
- 12) Wolfgang Ullrich, Habenwollen: Wie funktioniert die Konsumkultur? Fischer Verlag, Frankfurt a.M.2006

„Für die Modellierung einer einfachen Ontologie mit 100 Begriffen für unseren Unternehmenskunden haben wir etwa einen Monat benötigt. Es gibt aber auch Anwendungsfälle, bei denen in zwei Monaten 10.000 Begriffe modelliert werden können.“ Technologieanbieter

Fazit

Im Bereich Corporate Communication entstehen aktuell eine Reihe potentieller Anwendungsmöglichkeiten für das Semantic Media Web. Viele Unternehmen haben nach und nach die Bedeutung externer Markenkommunikation sowie der Kommunikation zwischen Mitarbeitern und Kooperationspartnern erkannt. So sind die Verantwortlichen auch zunehmend bereit, in innovative digitale Produkte zu investieren.

Dabei gilt es erstens, digitale Tools und Plattformen für die Kommunikations- und Prozessoptimierung innerhalb der Firmen zu entwickeln. Gerade in wissensbasierten Branchen wächst die Erkenntnis, dass Wissenstransfer zwischen verschiedenen Unternehmensbereichen und -standorten und eine entsprechende Integration der jeweiligen IT-Systeme einen Großteil der Geschäftsprozesse signifikant beschleunigen und verbessern kann.

Zweitens geht es um die Konzeption nützlicher Tools und medienwirksamer Experimente in der externen Kommunikation – Technologien, die es der Markenkommunikation ermöglichen, mit neuartigen, nützlichen Publikationsformaten und Servicetools die Aufmerksamkeit der Konsumenten zu wecken und halten, kommen gerade zur richtigen Zeit. Denn zunehmend sind digitale Dienste als Alltagshelfer gefragt, die Marken positiv als mitdenkenden, persönlich unterstützenden „Freund“ darstellen. Dabei wird die Rolle von Social Media in immer mehr Unternehmensbereichen wichtig – als personalisierte Corporate Publishing Formate, im internen Wissensmanagement oder im Bereich Business Intelligence.

Die Funktion des Semantic Media Web wird in der Corporate Communication vor allem darin bestehen, Daten über mehrere Kommunikationsplattformen hinweg zu verknüpfen und auszuwerten und mit

semantischen, systemübergreifenden Suchmaschinen auszustatten. Im Bereich der „Social Business Intelligence“ werden darüber hinaus möglichst umfassend semantisch aufbereitete Social Media Monitoring Tools gefordert sein, um Unternehmensdaten und Zielgruppeninformationen zu aggregieren und analysieren. Klar ist, dass eine gut funktionierende Umsetzung semantisch aufbereiteter Kommunikationsprodukte vor allem da realistisch ist, wo die Datenarten und -inhalte verhältnismäßig übersichtlich abgegrenzt sind – je komplexer die Themenbereiche, desto schwieriger die semantische Websuche. Ein gutes Beispiel ist die Rezeptplattform Google Recipes; vergleichbar wäre das Tagging von Medizin-, Kosmetik- oder Modeprodukten mit Information und Bewertung der Inhaltsstoffe, Chemikalien, Zutaten.

Eine wichtige Rolle spielt hier schon heute die transdisziplinäre Kooperation zwischen inhaltlicher und technologischer Konzeption und Gestaltung. Anders als viele Verlage wissen Marken um den Wert einer ansprechenden Erscheinung mit maximaler Usability. Wissensarbeiter können sich zum Einen in der Konzeption von Publikations- und Serviceformaten einbringen, aber auch beim Annotieren der branchen- und fachspezifischen Metadaten. Gegenwärtig sammeln und filtern Aggregationstools wie Social Media Monitoring Services nicht komplett automatisiert, sondern mit menschlicher Hilfe – auch beim Nachbearbeiten der maschinellen Vorkuratierung von Datensammlungen sind Einsatzgebiete für Journalisten und Experten denkbar.

Erfolgreiche Beispiele zeigen dabei, dass nicht das Schlagwort „Semantic Media Web“ im Zentrum stehen sollte, sondern der konzeptionelle Nutzen der Endprodukte für den Kunden. Generell gilt: Semantische Technologien bleiben also als Hilfsmittel im Hintergrund, während im Vordergrund die ästhetische Qualität den hohen Produkt- und Markennutzen vermittelt.

5. Analyse und Bewertung

In Abschnitt 3 haben wir eher generisch orientierte Use Cases für Technologien des Semantic Media Web präsentiert und diese in Abschnitt 4 anhand konkreter Einsatzszenarien in unterschiedlichen Anwendungsfeldern genauer analysiert. In diesem Abschnitt werten wir diese Ergebnisse detaillierter aus und erläutern die wichtigsten Herausforderungen, mit denen sich das Semantic Media Web konfrontiert sieht (Abschnitt 5.1) sowie das Potenzial das das SMW mit sich bringt (Abschnitt 5.2).

5.1 Herausforderungen

Wie sich in den einzelnen Anwendungsfeldern gezeigt hat, gibt es bereits eine ganze Reihe vielversprechender Ansätze für das Semantic Media Web. Bevor dieses weitverbreitete Realität werden kann, sind aber noch diverse Herausforderungen zu meistern, die so unterschiedliche Aspekte wie technologische Reife, Kosten-Nutzen-Verhältnis, rechtliche Rahmenbedingungen, offene Formate oder neue Wertschöpfungsketten umfassen.

Forschungsprototypen sind keine fertigen Produkte

Im Hinblick auf technologische Aspekte ist zunächst festzuhalten, dass viele bereits vorhandene Lösungsansätze des Semantic Media Web im Rahmen von Forschungsprojekten entwickelt und eingesetzt wurden und daher lediglich Prototypcharakter aufweisen. Zwar kann ein Forschungsprototyp die grundsätzliche Machbarkeit eines Ansatzes oder Verfahrens demonstrieren, er sollte aber nicht mit einem marktfähigen Produkt verwechselt werden. Die Weiterentwicklung eines Prototyps zur Produktreife umfasst in der Regel erhebliche Entwicklungsarbeiten, insbesondere

bezüglich Bereichen wie Robustheit, Skalierung, Performanz, Usability oder Dokumentation.

Generell ist die Diskussion um das Semantic Web immer noch sehr akademisch geprägt. Die Vision des Semantic Media Web kann aber nur Realität werden, wenn zusätzlich zu den eher wissenschaftlich orientierten Forschungs- und Entwicklungsarbeiten auch verstärkt auf Fragen der Produktisierung und der Markteinführung fokussiert wird.

Technologisch liegen dabei die wesentlichen Herausforderungen in der Robustheit und Skalierbarkeit der verwendeten Verfahren. Robustheit umfasst Aspekte wie Fehlertoleranz (z.B. fehlerhafte Eingabedaten) ebenso wie Ergebnisqualität (z.B. Fehlerquoten). Skalierung bedeutet beispielsweise, dass Verfahren auch bei sehr große Datenmengen oder Zugriffszahlen effizient funktionieren.

Ein wichtiger Teilaspekt der Skalierbarkeit betrifft den manuellen Aufwand, der mit dem Aufbau des Semantic Media Web verbunden ist. In vielen Show Cases von Forschungsprojekten werden kleine Anwendungen für überschaubare Gegenstandsbereiche realisiert. Eine entsprechende Ausweitung auf praxisrelevante Einsatzszenarien würde dabei zu erheblichen Kosten für die Modellierung und Wartung führen.

Mögliche Lösungen für diese Herausforderungen sind die Verwendung (semi-)automatischer Verfahren bzw. die Wiederverwendung manuell erstellter Daten. Grundsätzlich müssen die Kosten für Aufbau und Betrieb des Semantic Media Web dabei in einem vernünftigen Verhältnis zum jeweiligen Nutzen stehen - im Idealfall in Form von Einsparungen an anderer Stelle oder zusätzlichen Umsätzen und Gewinnen.

Offene statt proprietärer Lösungen

Die Grundidee des Semantic Media Web geht von einem offenen Szenario aus, in dem verteilte Informationen miteinander verknüpft werden, z.B. in der Form von Linked Open Data. Theoretisch sind aber auch mehr oder weniger geschlossene SMWs denkbar, die sich jeweils auf die Verknüpfung von Informationen eines einzelnen Anbieters konzentrieren. Der Vorteil eines geschlossenen Ansatzes besteht zunächst darin, dass kein Abstimmungsaufwand mit anderen Partnern notwendig ist und die Inhalte und deren Verknüpfung besser kontrolliert werden können. Beispiele hierfür sind die Verwendungen selbstentwickelter Ontologien z.B. bei BBC oder New York Times.

Demgegenüber stehen offene Ansätze, bei denen Informationen aus sehr unterschiedlichen Quellen miteinander verbunden werden. Hier sind die Abstimmungsaufwände zwar entsprechend größer und Kontrollmöglichkeiten geringer, die Entwicklungsaufwände werden aber auf unterschiedliche Player aufgeteilt und im Idealfall werden Parallelentwicklungen vermieden. Sehr grob entsprechen diese Varianten dem Unterschied zwischen dem Apple- und dem Android-Ansatz im Bereich mobiler Endgeräte und Betriebssysteme.

Viele Beispiele in Abschnitt 4 haben deutlich gemacht, dass mit dem Einsatz neuer Technologien in der Medienbranche erhebliche Umbrüche bei Geschäftsmodellen und Wertschöpfungsketten einhergehen. Hier liegt sicherlich die größte Herausforderung für die erfolgreiche Ausbildung des Semantic Media Web. Zum einen ändern sich die etablierten Geschäfts- und Arbeitsprozesse, zum anderen müssen Geschäftsbeziehungen mit neuen Dienstleistungspartnern aufgebaut werden.

Mentalitätenunterschiede

Sowohl in den Expertengesprächen als auch in den Diskussionen im Rahmen der Forenveranstaltungen wurde deutlich, dass die Kommunikation zwischen den einzelnen Playern, die teilweise sehr unterschiedliche Hintergründe aufweisen, nicht immer reibungsfrei verläuft. Spitzt man die jeweiligen Extrempositionen karikierend zu, so stehen auf der einen Seite Verleger im Vorruhestandsalter, die auf die vermeintliche Kostenloskultur, Gratismentalität und das Internet als rechtsfreien Raum schimpfen, auf der anderen Seite freiberufliche Programmierer mit frisch absolviertem Bachelorabschluss, die felsenfest davon überzeugt sind, mit Open Access und Crowd Sourcing ließen sich alle potentiellen Probleme der Medienbranche von heute auf morgen aus der Welt schaffen. Nicht gerade die optimalen Voraussetzungen für eine konstruktive Zusammenarbeit, insbesondere vor dem Hintergrund, dass die meisten Verlage mit stark sinkenden Einnahmen zu kämpfen haben, während gleichzeitig Start-Ups auf Basis knapp skizzierter Geschäftsideen nicht unerhebliche Mengen an Venture Capital erhalten.

Dass der rasante technologische Wandel und die mit ihm einhergehenden Umbrüche alles andere als trivial in die eigene Geschäftsentwicklung zu integrieren sind, zeigen auch jüngste Beispiele wie Kodak, Nokia oder Research in Motion – Unternehmen, die im Vergleich zu Verlagen durchaus als innovations- und technologieaffin eingeschätzt werden können.

Vor diesem Hintergrund kommt Netzwerkaktivitäten, die die unterschiedlichen Player des Semantic Media Web zusammenbringen und die Kommunikation untereinander kontinuierlich verbessern eine besondere Bedeutung zu. So wie die Medienbranche von einem besseren Wissen über technologische Möglichkeiten profitieren kann, so notwendig ist für Technologieanbieter ein besseres Verständnis von Kundenbedürfnissen und involvierten Geschäftsprozessen.

Ängste bei den Contentproduzenten

Noch immer lässt sich bei vielen Medienmachern, Journalisten, Verlegern, Archivaren, Geisteswissenschaftlern Misstrauen gegenüber dem Internet erkennen. Dies liegt unter anderem an der Angst, dass ihre Position mit der Aufgabe von „Herrschaftswissen“ durch die digitalen Medien an Bedeutung einbüßen und die eigene Arbeit überflüssig machen könnte. Besonders semantisch veredelten Informationen stehen viele Wissensarbeiter skeptisch gegenüber: Ein Teil ihrer Arbeit – und damit die Kontrolle über die Inhalte – wird ihnen damit auf undurchschaubare Weise abgenommen, was dem Berufsethos einer möglichst „wahren“, neutralen Berichterstattung widerspricht.

„Ich brauche keine semantisch annotierten Inhalte, weil ich nicht will, dass undurchsichtige Algorithmen Informationen für mich vorselektieren.“ Journalismus

Technik als Allheilmittel?

Umgekehrt verhält es sich in der IT-Branche: Viele Programmierer wiederum übersehen aufgrund ihres Desinteresses am Wesen der Metadaten, dass beim Annotieren komplexer Webinhalte fundierte Recherchetechniken und inhaltliche Fachkenntnisse unabdingbar sind. Darüber hinaus erwarten einige Informatiker freien Zugang zu sämtlichen wissenschaftlichen und journalistischen Inhalten. Unter dem Vorwand einer demokratischen Informationsbeschaffung für alle geht es dabei aber tatsächlich im Grunde vor allem darum, digitale Anwendungen, die mit fremden Content arbeiten, schnell und einfach umsetzen zu können. Eine Zahlpflicht für Inhalte würde die Entwicklung solcher Tools verteuern und verkomplizieren.

„Das semantische Web für die Medienwelt würde schon jetzt problemlos funktionieren - hätten die Verlage nur ihre Hausaufgaben gemacht.“ IT-Branche

Designer - mehr als „Hübschmacher“

Auffallend ist auch, dass die Notwendigkeit, sich mit der ästhetischen Qualität und der Anwenderfreundlichkeit neuer Medientools auseinanderzusetzen, sowohl bei Fachleuten der technischen als auch der inhaltlichen Seite auf wenig Aufmerksamkeit stößt. Dabei übernehmen Digital- und Designagenturen, entgegen ihrem überholten Ruf als reine „Hübschmacher“, eine wichtige Rolle, wenn es darum geht, einem neuen digitalen Produkt am Markt über eine zeitgemäße, emotionale, anwenderfreundliche und funktionale Gestaltung Akzeptanz zu verschaffen. Für die Entwicklung neuer Geschäftsmodelle wäre es zunächst also wichtig, im Dialog Verständnis für die Arbeit und die Perspektiven der jeweils anderen Disziplinen zu schaffen.

„Designer brauchen wir nicht – wer soll die denn bezahlen?“ IT-Branche

Sharing als Geschäftsmodell?

Eine weitere Frage ist die nach dem Teilen von Information. Während, überspitzt dargestellt, die IT-Branche offen zugängliche Inhalte in allen Wissenschafts- und Medienbereichen fordert, wehren sich Kultur- und Medienberufe vehement dagegen - bis hin zur irrationalen Abschottung auch öffentlicher Datenarchive, etwa in Museen. Dahinter stehen Ängste vieler Medienmacher, ihre Arbeit durch die freie Veröffentlichung überflüssig zu machen - vielleicht auch ein verständlicher, aber in dieser Form wenig erfolgreicher Versuch, so Respekt für den Wert der Wissensarbeit zu schaffen. Klar ist, dass eine ausführliche Debatte über Bezahlmodelle für Inhalte und professionelle Inhaltproduzenten in Politik, Medien und Wirtschaft notwendig ist.

Rechtliche Situation

Schließlich stellt sich noch die Herausforderung der (unklaren) rechtlichen Rahmenbedingungen. Dies lässt sich beispielsweise am Leistungsrecht veranschaulichen. Aus Sicht eines Ökosystems profitieren Suchmaschinenbetreiber wie Google davon, dass Verlage Inhalte anbieten, nach denen die Nutzer von Suchmaschinen suchen. Umgekehrt profitieren die Verlage davon, dass ihre Inhalte über Suchmaschinen gefunden werden können und somit potenziell eine größere Leserschaft erreichen. Eine Konfrontation kann dagegen weder im langfristigen Interesse der Verlage, noch der Suchmaschinenbetreiber, noch der Endnutzer sein – ganz unabhängig davon wie berechtigt oder unberechtigt das Leistungsrecht sein mag.

5.2 Potenzial

Die Analyse der einzelnen Anwendungsfelder in Abschnitt 4 hat das hohe Potenzial verdeutlicht, das im Semantic Media Web steckt. Das SMW eröffnet der Medienbranche und Wissensarbeitern neue Möglichkeiten der Wertschöpfung und die nachhaltige Etablierung innovativer Mehrwertdienste. Dies beinhaltet auch die Optimierung von Prozessen und die (Semi-)Automatisierung einzelner Prozessschritte.

Wie Erfahrungen aus dem elektronischen Geschäftsverkehr zeigen, führen entsprechende Prozessoptimierungen nicht nur zu Kostensenkungen und Beschleunigung, sie haben zumeist auch positive Auswirkungen auf Qualität und Konsistenz. Während Papier bekanntlich geduldig ist, führen fehlerhafte Eingaben oder Datenübergaben an Programmierschnittstellen in der Regel zu entsprechenden Fehlermeldungen. Darüber hinaus können durch Beseitigung von Medienbrüchen sowie die (semi-) automatische Erzeugung von Inhalten und Metadaten weitere Fehlerquellen eliminiert werden.

Das Semantic Media Web unterstützt auch aktuelle Konvergenztrends und die Entwicklung hin zu integrierten Multimediaformaten. Die Verknüpfung von Daten, Texten, Grafiken, Fotos, Bildern, Filmen und Audio wird in Zukunft weiter zunehmen – mit dem Semantic Media Web bietet sich eine einfache und umfassende Möglichkeit, diese Verknüpfung formal zu unterlegen und die Abbildung entsprechender Querbezüge zu unterstützen. Dies entspricht dem Alltagsverständnis, dass sich unterschied-

liche Medienformate im Endeffekt auf dieselben Objekte in der Welt oder Konzepte beziehen.

Schließlich ermöglicht das Semantic Media Web eine Personalisierung, die theoretisch beliebig feingliedrig sein kann. Inhalte können somit für die jeweiligen Zielgruppen optimiert werden, wobei eine Reduktion der Zielgruppengröße bis auf Einzelindividuen möglich ist. In dieser Hinsicht erleben wir aktuell in der Medienbranche eine Entwicklung, die in naher Zukunft durch den Fortschritt bei 3D-Druckern auch in vielen Bereichen der Produktherstellung zu erwarten ist – mit ähnlich komplexen Folgen für Geschäftsmodelle, Wertschöpfungsketten und Lizenzfragen.

„Vielen Museen wollen ihre Schätze wahren,
aber nicht teilen. Die Vermittlung der
Inhalte wird immer noch stiefmütterlich
behandelt.“ Kulturbranche

6. Handlungsempfehlungen

Abschließend listen wir in diesem Abschnitt Handlungsempfehlungen für Technologieanbieter, Wissensarbeiter, Unternehmen der Medienbranche und Politik auf. Diese Empfehlungen basieren im Wesentlichen auf der Analyse der Anwendungsszenarien in Abschnitt 5 und den daraus abgeleiteten Einschätzungen zu Potenzial und Herausforderungen des Semantic Media Web in Abschnitt 5.

6.1 Ansätze für Technologieanbieter

Technologieanbieter sollten zum einen ihre Lösungen zur Produktreife weiterentwickeln und dabei insbesondere (semi-)automatische Verfahren unterstützen. Darüber hinaus sollten Nutzen und Skalierbarkeit anhand von Kundenreferenzprojekten demonstriert werden. Dabei sollten Fragen der Integration in Kundenprozesse und -systeme im Mittelpunkt stehen, d.h. der Projektschwerpunkt sollte auf den fachlichen Anwendungsszenarien des Kunden liegen und die technischen Lösungen lediglich als Hilfsmittel in diesem Kontext dienen.

In vielen Forschungsprojekten liegt der Schwerpunkt stattdessen auf

„Wir verkaufen unsere Dienstleistung als Prozessmodellierung und nicht unter dem Begriff ‚Semantic Web‘, weil das Schlagwort inzwischen negativ konnotiert ist.“

Technologieanbieter

der Technologieentwicklung, die anhand vergleichsweise kleiner anwendungsorientierter Show Cases illustriert werden. Auch solche Projekte sind sinnvoll und notwendig, im Bereich Semantic Media Web sollten diese Projekte aber in Zukunft verstärkt durch komplementäre kundenorientierte Entwicklungen ergänzt werden.

6.2 Ansätze für Wissensarbeiter

Wissensarbeitern eröffnen sich mit dem Semantic Media Web eine Reihe neuer Möglichkeiten und Chancen. In den Expertengesprächen wurde immer wieder betont, wie wichtig gut qualifizierte Wissensarbeiter für die Einführung und den Einsatz neuer Technologien sind. Gefragt sind nicht nur ihr Hintergrundwissen, sondern auch das Wissen um Recherchemethoden und -prozesse. Dabei wurde insbesondere auch deutlich, dass für die erfolgreiche Arbeit mit Ontologien, Annotationen, kontrollierten Vokabularen etc. eine gesunde Mischung aus formaler Disziplin für möglichst präzise Ergebnisse und pragmatischer Herangehensweise unabdingbar ist. Wie Erfahrungen aus dem Bereich Prozessmodellierung zeigen, verfügen nicht alle fachlich qualifizierten Mitarbeiter eines Unternehmens über solche Fähigkeiten.

„Journalisten, Programmierer und Designer müssen beim Visualisieren von Datensätzen künftig sehr intensiv zusammenarbeiten. Eine wesentliche Voraussetzung dafür ist der Respekt aller Beteiligten für die Arbeit der jeweils anderen Disziplinen.“

Kommunikationsbranche

„Meiner Erfahrung nach ist es nicht schwierig, die Datensammlungen von Museen sinnvoll veredeln zu lassen. Es gibt viele unterbeschäftigte Geisteswissenschaftler mit dem notwendigen Hintergrundwissen, und durch Open Source Anwendungen können die Zusatzinformationen weiter veredelt werden.“

Kulturbranche

Darüber hinaus sind ihre Kompetenzen aber auch bei der Konzeption und Vermittlung neuer Produkte und Services gefragt, also bei der Entwicklung von multimedialen Publikationsformaten, Self Tracking Devices und anderen Apps und Tools.

Vor diesem Hintergrund ist ein wachsender Bedarf an Wissensarbeitern zu erwarten, die mit den in Kapitel 3 skizzierten technologischen Lösungen des Semantic Media Web grundsätzlich vertraut sind und deren Potenzial und Grenzen realistisch einschätzen können. Als Orientierung können hier die bereits zu beobachtenden Veränderungen in der Ausbildung von Bibliothekaren dienen. Das Einsatzspektrum muss sich dabei nicht allein auf die Arbeit mit entsprechenden Tools beschränken, sondern kann auch Planungsberatung, Einführungsunterstützung, sowie Training oder Schulung neuer Nutzer umfassen. Möglichkeiten für Wissensarbeiter werden sich somit voraussichtlich nicht nur in der Medienbranche selber, sondern auch bei Technologieanbietern, Agenturen und externen Dienstleistern aus den Bereichen Design und Kommunikation ergeben.

6.3 Ansätze für Anbieter von Inhalten

Anbieter von Inhalten sollten zum einen innovative Angebote entwickeln, die die Möglichkeiten neuer Technologien möglichst optimal ausnutzen und unterstützen und ihren Kunden signifikanten Mehrwert gegenüber traditionellen Angeboten bieten. Dies umfasst insbesondere auch die Entwicklung neuer, umfassender und kundenorientierter Dienstleistungen.

Dabei sollten Inhalteanbieter auch die enge Kooperation mit Dienstleistern und Technologieanbietern suchen, um von deren Kompetenz zu profitieren und ihre eigenen Entwicklungsarbeiten auf ihre jeweiligen Kernkompetenzen konzentrieren zu können.

Schließlich sollten auch die Entwicklungen im Bereich Open Data und Open Access sowie die damit verbundenen Wertschöpfungsketten und Geschäftsmodelle sehr genau beobachtet werden. Die grundsätzliche Skepsis von Inhalteanbietern gegenüber den „Open-Bewegungen“ ist zwar verständlich, da sie ihr Kapital darstellen, in dessen Produktion sie viel Zeit und Energie investieren. Doch es verhindert im schlimmsten Fall die Entwicklung innovativer Mehrwertdienste und somit die Erschließung neuer Märkte und Umsatzmöglichkeiten, von denen sie mittelfristig selbst profitieren können.

Einige Anbieter experimentieren auch bereits mit einer „Moving Wall“ statt mit einer strikten Bezahlschranke. Dabei werden bestimmte Inhalte kostenfrei angeboten, z.B. ab einer bestimmten Zeit nach ihrer Veröffentlichung, während andere kostenpflichtig sind. Generell scheint eine Flexibilisierung bzw. ein breiteres Angebot von Bezahl- und Abrechnungsmodellen sinnvoll zu sein, insbesondere um erste Erfahrungen bezüglich Akzeptanz und Potenzial neuer Dienstleistungen zu sammeln.

6.4. Ansätze für die Politik

Der Politik kommt in diesem Kontext zunächst einmal die sicherlich nicht triviale Aufgabe zu, möglichst optimale Rahmenbedingungen und ein hohes Maß an Rechtssicherheit zu gewährleisten. In Expertengesprächen und Diskussionen wurde dabei deutlich, dass hier zwar akuter Handlungsbedarf wahrgenommen wird, dass man aber gleichzeitig davon ausgeht, dass eine Besserung in naher Zukunft nicht wirklich zu erwarten ist. Neben der üblichen Problematik unterschiedlicher Interessen und damit verbundener Lobbyarbeit – wie beispielsweise beim Leistungsschutzrecht oder der GEMA-Google-Auseinandersetzung deutlich wird – sehen viele Experten die aktuelle Politikergeneration schlichtweg überfordert, adäquat auf die rasanten technologischen Entwicklungen zu reagieren.

Dieser weit verbreitete Pessimismus, der sich vielfach auf Erfahrungen der letzten Jahre stützt, wie zum Beispiel die Debatten zum Jugendschutzvertrag, ist für ein innovationsorientiertes Land wie Deutschland sicherlich ein Hindernis und sollte als Alarmsignal verstanden werden – hier besteht ganz offensichtlich erheblicher und akuter Verbesserungsbedarf.

„Wissensarbeiter könnten das Tor zum
Semantic Web sein.“ Kommunikationsbranche

Verhältnismäßig gut wird die Situation im Bereich Open Data eingeschätzt. Bei der Bereitstellung offen zugänglicher und maschinell verarbeitbarer Daten wurde Politik und Verwaltung eine durchaus positive Entwicklung mit Vorreiterrolle und Vorbildfunktion bescheinigt. Tatsächlich dürften die bereits vorliegenden und in naher Zukunft zu erwartenden

Ergebnisse den Nutzen von Linked Open Data eindrucksvoll demonstrieren. Auch wenn sich diese Erfahrungen und insbesondere die damit verbundenen Geschäftsmodelle nicht eins zu eins auf die gesamte Medienbranche übertragen lassen, können sie helfen, das Potenzial des Semantic Media Web anschaulich zu verdeutlichen und Anhaltspunkte für Aufwände und Nutzen liefern.

6.5. Forschungsförderung

Schließlich hat die Politik noch die Möglichkeit, die Entwicklung des Semantic Media Web über die Forschungsförderung positiv zu beeinflussen. Großprojekte wie Europeana oder THESEUS wurden in Diskussionen und Expertengesprächen grundsätzlich positiv beurteilt, wobei auch immer wieder vor überzogenen Erwartungen bzgl. der Marktreife entsprechender Forschungsergebnisse gewarnt wurde.

Dabei sollte es weniger um reine Technologieförderung gehen, sondern zunehmend um den Einsatz innovativer Technologien in realistischen Anwendungsszenarien, die Robustheit und Skalierbarkeit des Semantic Media Web sowie Kosten-Nutzen-Verhältnisse anhand von Pilotlösungen anschaulich demonstrieren. Als Anwendungsfelder bieten sich dabei insbesondere Multimedia incl. Bewegtbild und Big/Open Data an.

Auch für die Entwicklung einer inhaltlichen Infrastruktur, z.B. in Form frei nutzbarer Ontologien oder kontrollierter Vokabulare für ausgewählte Anwendungsfelder wäre eine Unterstützung durch Forschungsförderung sinnvoll und hilfreich. Schließlich sollte noch die bessere Vernetzung der relevanten Akteure und die Entwicklung eines Ökosystems unterstützt werden, um hier eine dringend notwendige Grundvoraussetzung für den kommerziellen Erfolg des Semantic Media Web zu schaffen.

Derartige Förderprojekte kommen auch der Medienbranche zugute, die als wesentlicher Bestandteil einer Demokratie in der Pflicht steht, sich weiterhin als starkes, relevantes Sprachrohr der Gesellschaft zu positionieren. Das Zusammenspiel von Contentproduktion und Technologie versetzt die klassische Wissensarbeit in die Lage, sich für die Herausforderungen der Zukunft zu wappnen.

7. Anhang

7.1 Glossar

ANNOTATION

Die Auszeichnung von Objekten oder Ressourcen (z.B. Texten, Textelementen, Bildern) mit zusätzlichen Informationen, die zumeist als Metadaten bezeichnet werden. Damit werden Inhalte, die gewöhnlich nur für Menschen verständlich sind, auch von Maschinen gelesen werden. So könnte zu den Annotationen für den Begriff „Jaguar“ sowohl „Tier“, und „Wildkatze“ als auch „Auto“ und Designklassiker“ gehören.

ASKJEEVES

Ask Jeeves, inzwischen Ask (ask.com, ask.de), ist eine Suchmaschine, die auf Anfragen nicht nur Ergebnisse, sondern auch Antworten liefern soll.

BEGRIFFSHIERARCHIE

In einer Begriffshierarchie werden Begriffe oder Konzepte in einer hierarchischen Struktur von Ober- und Unterbegriffen angeordnet, also z.B. der Begriff „Auto“ als Unterbegriff von „Fahrzeug“.

BUSINESS INTELLIGENCE

Verfahren und Lösungen zur elektronischen Auswertung von Unternehmensdaten (z.B. regionale Umsätze, Bearbeitungszeiten, Produktionsabläufe, Auslastungen), zumeist mit dem Ziel, Markt- und Wettbewerbsentwicklungen besser zu verstehen. Digitale Analysetools unterstützen das Sammeln und Analysieren dieser Daten.

BIG DATA

Sehr große Datenmengen, die spezielle Anforderungen an Speicherung, Verarbeitung, Analyse oder Visualisierung stellen, und mit herkömmlichen Verfahren nicht adäquat behandelt werden können. Im Kontext dieser Studie z.B. Auswertungen von News, Facebook-Postings oder Twitter-Tweets. 2011 wurden weltweit etwa zwei Zettabyte (zwei Billionen Gigabyte) Daten generiert.

CLUSTERING

Beim Clustering werden Objekte oder Ressourcen (z.B. Texte, Bilder oder Suchergebnisse) in Gruppen mit ähnlichen Eigenschaften eingeteilt. Im Gegensatz zur Kategorisierung wird hierfür kein vorgegebens Kategoriensystem benötigt.

DATENVISUALISIERUNG

Datenvisualisierungen stellen Zusammenhänge und Verhältnisse ausgewählter Daten mit visuellen, grafischen Mitteln dar. Einfache Infografiken können aus Balken-, Torten oder Liniendiagrammen bestehen, komplexere Datenvisualisierungen bilden ganz unterschiedliche Bilder ab: etwa Wolken oder Netzwerke, individuelle narrative oder künstlerische grafische Aufbereitungen. Die Grafiken werden meist mit entsprechender Software oder von ausgebildeten Gestaltern erstellt.

Interaktive Datenvisualisierungen werten Daten nach individuellen Kriterien und Fragestellungen aus, die sich über ein Webinterface eingeben lassen. Dank wahrnehmungspsychologischer Prinzipien werden komprimierte Bildinformationen vom menschlichen Gehirn schneller, effizienter und emotionaler erfasst als Textinformation.

DATENJOURNALISMUS

Als eine Form des Online-Journalismus gewinnt Datenjournalismus zunehmend an Bedeutung. Dabei geht es darum, verschiedene Daten (vor allem offene Daten) zu sammeln, in einen Kontext zu bringen und aufzubereiten, zu analysieren und in klassischen journalistischen Darstellungsformen oder als Datenvisualisierung zu publizieren. Viele Datenjournalisten stammen aus der Statistik, andere aus dem Journalismus. Mit der zunehmenden digitalen Datenflut reagieren auch immer mehr Ausbildungsgänge auf die neuen Anforderungen an Datenjournalismus. Große Zeitungen wie The New York Times, der Guardian oder Die Zeit führen eigene Infografik-Blogs, die insbesondere zu aktuellen Ereignissen wie Wahlen stark frequentiert werden.

DBPEDIA

DBpedia (dbpedia.org) wird gemeinsam von der Universität Leipzig, der Freien Universität Berlin und OpenLink Software entwickelt und stellt Web-Anwendungen aus Wikipedia extrahierte Informationen in RDF zur Verfügung.

DIGITALISAT

Das Endprodukt eines Digitalisierungsvorgangs, gewöhnlich eine Datei. Im Museumskontext zum Beispiel gehören digitale Abbildungen von Kunstwerken, Literatur, Forschungsobjekten, Fotodokumenten und Zeichnungen dazu.

EUROPEANA

Europeana (europeana.eu) ist eine virtuelle Bibliothek über die die digitalen Bestände in Europas Museen, Bibliotheken, Archiven und audiovisuellen Sammlungen online erkundet werden können.

GEMA-GOOGLE-STREIT

Beim Streit zwischen Google (als Betreiber von YouTube) und der GEMA geht es um das Vergütungsmodell für den Abruf geschützter Musikvideos. Da Google und GEMA hier bislang keine Einigung erzielen konnten, sind viele Videos auf YouTube für den Aufruf aus Deutschland gesperrt („Dieses Video ist in Ihrem Land nicht verfügbar“).

GIS (GEOINFORMATIONSSYSTEM)

Ein Geoinformationssystem stellt Funktionalität zur Arbeit mit räumlichen bzw. geografischen Informationen zur Verfügung, wie z.B. die Visualisierung von Informationen in Karten oder 3D-Stadtmodellen.

KATEGORISIERUNG

Bei der Kategorisierung werden Objekte oder Ressourcen (z.B. Texte oder Bilder) einer oder mehreren vorgegebenen Kategorien zugeordnet - z.B. den in einer Ontologie oder einer Begriffshierarchie modellierten Begriffen. Eine Kategorisierung kann entweder manuell oder (semi-)automatisch erfolgen.

KNOWLEDGE GRAPH

Mit dem Knowledge Graph entwickelt Google eine strukturierte Wissensbasis unter anderem mit dem Ziel seine Suchergebnisse um semantische Informationen zu erweitern. Nach eigenen Angaben verwendet Google dabei Quellen wie Freebase (www.freebase.com), Wikipedia und das CIA World Factbook.

KONTROLLIERTES VOKABULAR

Ein kontrolliertes Vokabular ist eine Sammlung von Bezeichnungen, zum Beispiel Glossare und Fachwortverzeichnisse, die eindeutig Begriffen (Deskriptoren) zugeordnet sind. Die Bezeichnung stammt aus den Dokumentationswissenschaften. Im semantischen Web werden per RDF über IRIs Begriffe indexiert werden. Verwaltet werden kontrollierte Schlagwörter in Thesauri oder Normdateien.

KURATIERUNG

Ähnlich wie in der Arbeit eines Kurators im Museum, werden zukünftig Positionen benötigt, die die Fülle und die heterogenen Formate an Informationen im Netz selektieren, filtern, analysieren und interpretieren und über verschiedenen Plattformen hinweg vermitteln. Auf diese menschlichen Filter wird man auch in Zukunft nicht verzichten können.

LEISTUNGSSCHUTZRECHT-DEBATTE

Aktuell gibt es eine Diskussion zwischen Verlagen und Suchmaschinenbetreibern wie Google darüber, inwieweit eine Anzeige von Textauszügen in Suchergebnissen nur bei Zahlung entsprechender Lizenzgebühren erlaubt werden sollte.

LINKED OPEN DATA (LOD)

Im WWW frei verfügbare Daten, die über URIs eindeutig identifiziert sind, über HTTP abgerufen werden können und mittels URIs mit anderen Datenbeständen verknüpft sind.

METADATEN

Daten, mit denen Objekte oder Ressourcen (z.B. Texte, Textteile, Bilder, etc.) weiter beschrieben werden.

ONTOLOGIE

Formale Modellierung von Begriffen und zwischen ihnen bestehenden Beziehungen für einen bestimmten Gegenstandsbereich.

OPEN ACCESS

Freier Zugang zu wissenschaftlicher Literatur und anderen Materialien, zumeist über das Internet.

OPEN DATA

Freie Verfügbarkeit und Nutzbarkeit von Daten, meist öffentlicher Herkunft.

OPENSTREETMAP

Freies Projekt zur Sammlung und Verfügbarmachung von Geodaten, z.B. in Form von Spezialkarten und Navigationsanwendungen.

OWL (ONTOLOGY WEB LANGUAGE)

OWL (Ontology Web Language) ist ein Standard des W3C (World Wide Web Consortium) zur formalen Modellierung von Ontologien. OWL setzt auf der RDF-Syntax auf ist aber ausdrucksmächtiger als RDF.

PERSONALISIERUNG

Anpassung eines Angebots oder einer Funktionalität an den aktuellen Nutzer z.B. auf Basis expliziter oder impliziter Nutzerprofile.

QUANTIFIED SELF

Dank Datensammlern wie den allgegenwärtigen Smartphones vermessen sich Menschen zunehmend selbst. Das „Quantified Self“, das quantifizierte Ich, kontrolliert und korrigiert sein Verhalten mittels Self Tracking selbst. Siehe auch SELF TRACKING.

RDF (RESSOURCE DESCRIPTION FRAMEWORK)

RDF (Ressource Description Framework) ist ein Standard des W3C (World Wide Web Consortium) zur Beschreibung von Metadaten und grundlegender Baustein des Semantic Web.

SCHEMA.ORG

schema.org (schema.org) ist ein Metadatenstandard für Webseiten, der insbesondere von Bing und Google unterstützt wird und die automatische Verarbeitung von Seiteninhalten erleichtern soll.

SECOND SCREEN

Nutzung eines zweiten Bildschirms (z.B. Smartphone oder Tablet) parallel zum laufenden Fernsehprogramm. Auf dem Second Screen können dann Zusatzinformationen eingeblendet werden oder Anbindungen an Social Media erfolgen.

SEMANTIC WEB

Vision von Sir Tim Berners-Lee und Aktivität des W3C (World Wide Web Consortium) zur Weiterentwicklung des Webs hin zu einer Informationsammlung, die auch maschinell verarbeitet werden kann und somit neue Mehrwertdienste möglich macht.

SELF TRACKING

SelfTracking bedeutet, das eigene Verhalten mittels (meist digitaler) Datenanalysen zu überprüfen. Persönliche Daten werden dank technologischer Möglichkeiten oft in Echtzeit analysiert und lassen sich über Feedbackschleifen überprüfen. Diese Selbstbeobachtung soll die intrinsische Motivation des Nutzers anspornen, um sein Verhalten zu ändern oder Leistungen zu optimieren. SelfTracking wird oft im Bereich Gesundheit und Sport eingesetzt, Potenzial bietet das Prinzip zum Beispiel auch in der Bildung oder im Finanzwesen. SelfTracking Services, als Datensammlungen gebündelt in unseren Smart Phones, führen zum „Vermessenen Selbst“. Siehe auch QUANTIFIED SELF.

TAGGING

Begriffe im semantischen Web zu „taggen“ bedeutet, sie mit inhaltlich ergänzenden Schlagworten zu verschlagworten. Beim Tagging sind keine übergeordneten Ontologien notwendig. Siehe auch ANNOTATION.

THESAURUS

In der Dokumentationswissenschaft besteht ein Thesaurus aus kontrollierten Vokabularen, also vernetzten, systematisch geordneten Begriffssammlungen. Diese beschreiben Schlagworte zu einem Themengebiet und sind über inhaltliche Relationen miteinander verbunden. Verwaltet werden in erster Linie Synonyme, aber auch Ober- und Unterbegriffe zu den einzelnen Inhalten.

THESEUS

THESEUS ist ein inzwischen abgeschlossenes Forschungsprogramm des BMWi mit dem Ziel, den Zugang zu Informationen zu vereinfachen, Daten zu neuem Wissen zu vernetzen und die Grundlage für die Entwicklung neuer Dienstleistungen im Internet zu schaffen.

URI (UNIFORM RESSOURCE IDENTIFIER)

URIs sind einheitliche Bezeichner für Ressourcen, d.h. Zeichenfolgen, mit denen abstrakte oder physische Ressourcen eindeutig identifiziert werden können.

USER-GENERATED CONTENT

User Generated Content (UGC) sind Inhalte in (meist digitalen) Medienplattformen, die nicht von deren Anbietern, sondern von ihren Nutzern erstellt werden. Das können Kommentare auf Blogs sein, Webforen, Wikis, Youtube und Flickr. Der Begriff hat sich in Zusammenhang mit dem Schlagwort „Web 2.0“ etabliert. Bekanntestes funktionierendes UGC-Webangebot ist die Wissensplattform Wikipedia.

WEB 2.0

Das Schlagwort Web 2.0 wurde vom Internetpionier Tim O'Reilly geprägt. 2004 hatte er die Konferenz „Web 2.0“ initiiert, mit der er eine neue, dialogorientierte Ära des Internet einläutete, die von der Partizipation der Nutzer lebt. Siehe auch USER GENERATED CONTENT.

WOLFRAMALPHA

WolframAlpha (wolframalpha.com) ist eine semantische Suchmaschine.

7.2 Impulsgeber und Gesprächspartner der Interviews

Reinhard Altenhöner, Deutsche Nationalbibliothek
Armin Berger, 3pc GmbH Neue Kommunikation
Dr. Wolfgang Both, SenWTF Berlin
Felix Daub, KLICKFILM
Verena Dauerer, Stylepark AG / Mooove.com
Peter Delius, Peter Delius Verlag GmbH & Co KG
Christian Dirschl, Wolters Kluwer Deutschland GmbH
Holger Düwiger, Neofonie GmbH
Katja Eck, Wolters Kluwer Deutschland GmbH

Dr. Kai Eckert, Universität Mannheim
Prof. Dr. Ralph Ewerth, Ernst-Abbe-Fachhochschule Jena
Dr. Matthias Flügge, Fraunhofer FOKUS
Martin Fröhlich, PaperC GmbH
Dr. Seven Fund, Walter de Gruyter GmbH
Prof. Dr. Stefan Gradmann, Humboldt-Universität zu Berlin
Dr. Rita Gudermann, QUAGGA Media UG (haftungsbeschränkt)
Max Gutscher, CAMAO AG
René Herzer, Neuland + Herzer GmbH
Prof. Dr. Heiko Hartmann, Akademie Verlag
Dr. Thomas Hoppe, Ontonym GmbH
Jürgen Keiper, Deutsche Kinemathek/iRightsLab
Gudrun Knaus, Beraterin für Digitales Sammlungsmanagement
Dirk Krause, Pixelpark AG
Jens Landvogt, Neofonie GmbH
Wiebke Lang, freie Redakteurin und Designerin
Christoph Link, Cornelsen Schulverlage GmbH
Markus Luczak-Rösch, Freie Universität Berlin
Christoph Lüscher, iA Information Architects
Dr. Malgorzata Mochól,]init[AG
Lorenz Matzat, OpenDataCity
Dr. Knud Möller, Berater
Matthias Ort, ORTEC GmbH
Peter Neumann, Urban Media/Online Tagesspiegel-Gruppe
Dr. Joachim Quantz, ART+COM/Xinnovations e.V.
Prof. Dr. Felix Sasaki, FH Potsdam/DFKI Berlin/W3C
Dr. Volker Schwarz, Berliner Wissenschafts-Verlag
Regine Stein, Bildarchiv Foto Marburg
Rainer Thiem, Xinnovations e.V.
Alexandru Todor, Freie Universität Berlin
Prof. Dr.-Ing. Robert Tolksdorf, Freie Universität Berlin
Dr. Jutta Weber, Staatsbibliothek zu Berlin - Preußischer Kulturbesitz
Dr.-Ing. Frauke Weichhardt, Semtation GmbH
Sacha Weinberg, Condat AG
Kai Wermer, Freshmilk Creative Media
Gerd Zechmeister, Semantic Web Company GmbH
Anonym, Automobilbranche
Anonym, Freie Journalistin
Anonym, Freier Kurator

Die vorliegende Studie beleuchtet das Potenzial des Semantischen Webs für die Medienbranche, also Wissensarbeiter wie Verleger, Journalisten und Redakteure, Regisseure, Kuratoren und Archivare, Bibliothekare, und Wissenschaftler. Sie zeigt beispielhaft bestehende Projekte und Produkte aus Wirtschaft, Medien und Forschung, aber auch die Fragen bei der Entwicklung funktionierender Anwendungen und Herausforderungen der Zukunft auf. Dabei stehen sowohl die Perspektiven der Contentproduzenten, als auch die der Technologieanbieter im Fokus.

Die Analysen stützen sich auf Vorträge und Diskussionen im Rahmen einer zweitägigen Forenveranstaltung und eines Statusseminars mit insgesamt 300 Teilnehmern, auf Experteninterviews und journalistische Recherchen. Die Studie wurde 2012 für das von Xinnovations e.V. initiierte und vom BMBF geförderte Innovationsforum „Semantic Media Web“ erstellt.