



**Weltweites Wissensmanagement**  
-  
**Strukturierte Informationen im Internet**

Hausarbeit

an der

Privaten Fachhochschule für  
Wirtschaft und Technik  
Vechta/Diepholz/Oldenburg

im Rahmen des Studiengangs zum

Dipl.-Wirtschaftsinformatiker (FH)

in der Veranstaltung Knowledge Management (7. Semester)

vorgelegt von

Nils Löffler

aus

Hannover

Matrikel – Nummer: 2004308

Dozent: Dipl.-Betriebswirt Peter Recksiek

Abgabe: 28. Februar 2005

## **Inhaltsverzeichnis**

Abkürzungsverzeichnis.....	III
Abbildungsverzeichnis.....	IV
<b>1 Einleitung.....</b>	<b>1</b>
<b>2 Grundlagen.....</b>	<b>2</b>
2.1    Definition Wissensmanagement.....	2
2.2    Basiswissen Internettechnologie.....	3
<b>3 Voraussetzungen für die Strukturierung von Informationen.....</b>	<b>6</b>
3.1    Informationen mit Bedeutung.....	6
3.2    Vernetzung von Wissen.....	8
<b>4 Praktische Ansätze.....</b>	<b>9</b>
4.1    Semantic Web.....	9
4.2    XML Topic Maps.....	11
<b>5 Bewertung der praktischen Ansätze.....</b>	<b>13</b>
<b>6 Fazit und Ausblick.....</b>	<b>14</b>
Literaturverzeichnis.....	15

## Abkürzungsverzeichnis

ARPANET	Advanced Research Project Agency Network
Bit	Binary Digit
CERN	Conseil Européen pour la Recherche Nucléaire
FTP	File Transfer Protocol
HTML	Hypertext Markup Language
HTTP	Hypertext Transfer Protocol
IAB	Internet Architecture Board
ISO	International Organization for Standardization
NSFNET	National Science Foundation Network
OWL	Web Ontology Language
RDF	Resource Description Framework
RFC	Request for Comments
TCP/IP	Transmission Control Protocol/Internet Protocol
W3C	World Wide Web Consortium
WWW	World Wide Web
XHTML	Extended HTML
XML	Extensible Markup Language
XTM	XML Topic Maps

## **Abbildungsverzeichnis**

Abbildung 1	Begriffshierarchie Wissen	Seite 2
Abbildung 2	Differenzierungsmerkmale	Seite 3
Abbildung 3	Empfehlungen für das Internet von Morgen	Seite 4
Abbildung 4	Semantische Konflikte	Seite 6
Abbildung 5	Ausschnitt (Aspekt) einer Ontologie	Seite 7
Abbildung 6	RDF triple	Seite 10
Abbildung 7	Komponenten des Semantic Web	Seite 11

## 1 Einleitung

Schon Tim Berners-Lee<sup>1</sup> stellte im Jahr 1989 fest, dass die Verteilung von Informationen unter den Mitarbeitern des CERN (Conseil Européen pour la Recherche Nucléaire) nur unzureichend funktionierte.<sup>2</sup> Kurz gesagt entstand aus der Lösung dieses Problems, welches eigentlich nur einen begrenzten Bereich betraf, das World Wide Web (WWW) bzw. das, was heute gemeinhin als Internet bezeichnet wird.

Das World Wide Web bzw. das Internet hat in den letzten Jahren enorm an Bedeutung gewonnen. Ein Zeichen des Wandels zur heutigen Informationsgesellschaft. Angesichts des großen Umfangs dieses Netzes<sup>3</sup> stellt sich nun die Frage, wie die Informationen auf den tausenden von Webseiten strukturiert beschrieben und effizient genutzt werden können. Eine vollständig einheitliche Verwaltung der Informationen wird vermutlich aufgrund der dezentralen Struktur des Internets nie möglich sein. Gibt es aber vielleicht Möglichkeiten, das Internet zumindest auf einen entsprechenden Weg zu bringen ? Kurz gesagt: eine Art weltweites Wissensmanagement für das Internet, das sowohl von Menschen als auch von Computern verstanden werden kann.

Der Ansatz zur Schaffung einer Art künstlichen Intelligenz für das Internet – die sogenannte „Web Intelligence“<sup>4</sup> - soll an dieser Stelle der Vollständigkeit halber erwähnt aber in dieser Arbeit nicht weiter behandelt werden, da dies die nächste Evolutionsstufe darstellt und den Rahmen dieser Arbeit sprengen würde.

Diese Arbeit beschäftigt sich daher nur mit der oben genannten Fragestellung und ist wie folgt aufgebaut: Zuerst werden die grundsätzlichen Begrifflichkeiten und Zusammenhänge im Bereich Wissensmanagement und Webtechnologie erklärt. Danach werden die Voraussetzungen für die Strukturierung von Informationen bzw. Wissen erläutert. Darauf aufbauend werden zwei praktische Ansätze gezeigt, die in diese Richtung weisen. Abschließend erfolgen eine Bewertung und ein Vergleich dieser Ansätze. Die Arbeit endet dann schließlich mit einem Fazit und einem Ausblick in die Zukunft.

---

<sup>1</sup> der „Erfinder“ des World Wide Web; vgl. Abschnitt 2.2

<sup>2</sup> Vgl. Berners-Lee, T. (Information Management: A Proposal, 1989)

<sup>3</sup> Google durchsucht z.B. über 8 Mio. Seiten, vgl. <http://www.google.de>, Stand: Februar 2005

<sup>4</sup> weiterführende Informationen siehe: <http://wi-consortium.org> , Zhong, N. u.a. (Web Intelligence, 2003)

## 2 Grundlagen

### 2.1 Definition Wissensmanagement

Zu Beginn muss die Begriffsansammlung, die rund um den Begriff Wissen existiert, in eine logische Ordnung gebracht werden, damit im Folgenden die Abgrenzungen untereinander deutlich werden und man eine Vorstellung von dem Gegenstand der Betrachtungen dieser Arbeit hat.

Auf der untersten Ebene stehen die *Zeichen* bzw. ein Zeichenvorrat, der aus numerischen oder alphanumerischen Zeichen besteht. Ein Zeichenvorrat sind z.B. die Binärzeichen, die nur aus zwei Elementen – nämlich 0 und 1 – bestehen; solch ein Element wird hier Bit (Binary Digit) genannt. Acht Bits sind ein Byte und stellen die kleinste für Computer verständliche Einheit zur Zeichendarstellung dar.<sup>5</sup> Setzt man diese Zeichen nun nach bestimmten Regeln – der sogenannten Syntax – zu Wörtern und Sätzen zusammen, erhält man *Daten*.<sup>6</sup> Die Daten werden zu *Information*, wenn man sie in einen bestimmten Kontext stellt, d.h. man assoziiert einen Begriff oder eine reale oder theoretische Vorstellung damit. Hierdurch wird den Daten eine Bedeutung (Semantik) zugeordnet. Vernetzt man nun diese Informationen untereinander, um eine bestimmte Wirkung zu erzielen (Pragmatik) kann man von *Wissen* sprechen.<sup>7</sup> Die folgende Abbildung verdeutlicht diese logische Ordnung.

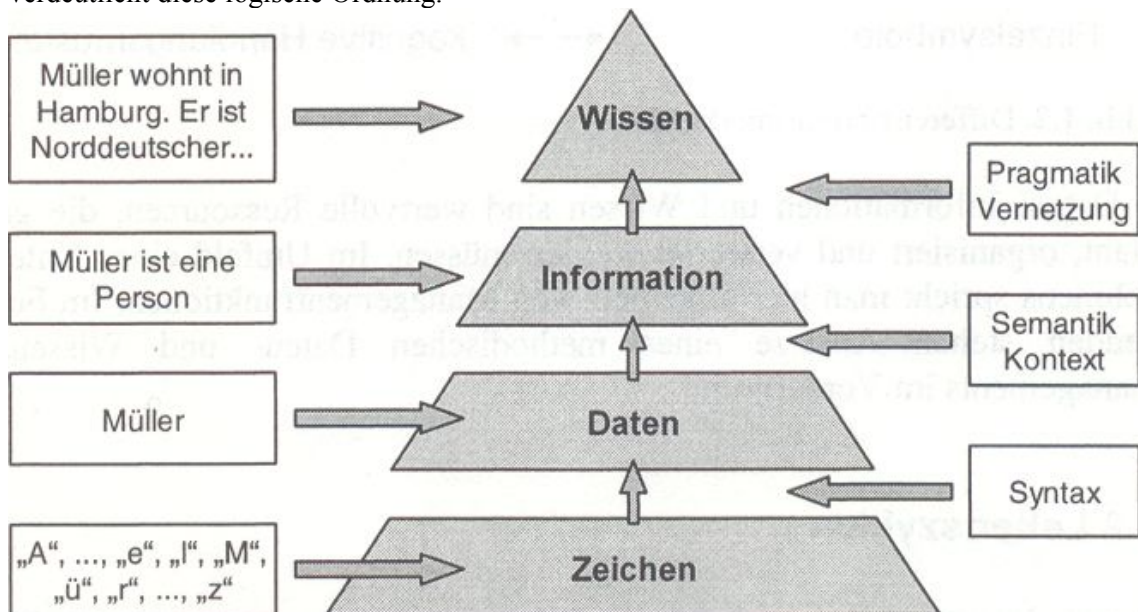


Abbildung 1: Begriffshierarchie Wissen

Quelle: Bodendorf, F. (Daten- und Wissensmanagement, 2003), S. 1

<sup>5</sup> Vgl. Schwarze, J. (Einführung in die Wirtschaftsinformatik, 2000), S. 40

<sup>6</sup> Vgl. Schwarze, J. (Einführung in die Wirtschaftsinformatik, 2000), S. 41; Bodendorf, F. (Daten- und Wissensmanagement, 2003), S. 1

<sup>7</sup> Vgl. Bodendorf, F. (Daten- und Wissensmanagement, 2003), S. 1 f.

Eine scharfe Abgrenzung zwischen Daten, Informationen und Wissen ist jedoch nicht immer möglich<sup>8</sup>. Die folgende Abbildung zeigt daher fünf Eigenschaften, nach denen Daten, Informationen und Wissen auf einer stufenlosen Skala eingeordnet werden können. Alle drei Arten stellen jedoch wertvolle Ressourcen dar, die geplant, organisiert und verwertet werden müssen; man kann daher auch allgemein von Wissensmanagement sprechen.<sup>9</sup>

<b>Daten</b>	<b>Information</b>	<b>Wissen</b>
strukturiert	←→	unstrukturiert
isoliert	←→	vernetzt
kontextunabhängig	←→	kontextabhängig
geringe Verhaltenssteuerung	←→	starke Verhaltenssteuerung
Einzelsymbole	←→	kognitive Handlungsmuster

Abbildung 2: Differenzierungsmerkmale

Quelle: In Anlehnung an Bodendorf, F. (Daten- und Wissensmanagement, 2003), S. 2

## 2.2 Basiswissen Internettechnologie

Da sich diese Arbeit mit dem Management von Wissen im Internet beschäftigt, müssen nun auch einige Grundlagen der dem Internet zugrunde liegenden Technologie erläutert werden.

Das Internet (in Langform *Interconnected Networks*) ist das heute wohl bekannteste Kommunikationsnetz. Es ist ein System unterschiedlicher, miteinander verbundener Netze, die als einheitliche Protokollfamilie TCP/IP (Transmission Control Protocol/Internet Protocol) zum Datenaustausch verwenden. Es entstand Anfang der 60er Jahre in den USA aus militärischen Zwecken in Form des ARPANET<sup>10</sup>; welches ein robustes Kommunikationsnetz für den Fall eines nuklearen Krieges darstellen sollte. Dieses Netz wurde Anfang der 90er Jahre durch das wissenschaftlich geprägte NSFNET (National Science Foundation Network) ersetzt, welches die Basis für das heutige Internet war. Da das Internet aus vielen autonomen Einzelnetzen besteht, gibt es keinen Betreiber oder verantwortliche Organisation dafür. Auf nationaler und internationaler Ebene gibt es jedoch Organisationen, die die Kooperation im Hinblick auf Protokolle und Standards fördern. Hierzu zählen z.B. das Internet Architecture Board (IAB)<sup>11</sup>, welches unter anderem die Requests for Comments (RFC) veröffentlicht – eine Reihe von

<sup>8</sup> Aufgrund dieser Tatsache und der unterschiedlichen Verwendung der Begriffe in den verwendeten Quellen, werden die Begriffe Daten, Informationen und Wissen in dieser Arbeit weitgehend synonym verwendet, wobei die einzelnen Definitionen dieser Begriffe nicht vergessen werden sollten.

<sup>9</sup> Vgl. Bodendorf, F. (Daten- und Wissensmanagement, 2003), S. 2

<sup>10</sup> ARPA=Advanced Research Project Agency des Verteidigungsministeriums der USA

<sup>11</sup> <http://www.iab.org>

technischen oder organisatorischen Vorschlägen zum Internet<sup>12</sup> - oder das World Wide Web Consortium (W3C)<sup>13,14</sup>. Das W3C ist ein Zusammenschluss aus über 350 Mitgliedsorganisationen aus Wirtschaft, Forschung, Regierung und privatem Umfeld, der im Oktober 1994 gegründet wurde. Seine Ziel ist durch die Entwicklung von interoperablen Technologien<sup>15</sup> das Internet weiterzuentwickeln und sein volles Potential auszuschöpfen.<sup>16</sup> Die folgende Abbildung veranschaulicht die geplante Weiterentwicklung des Internets hinsichtlich dieser Standards.

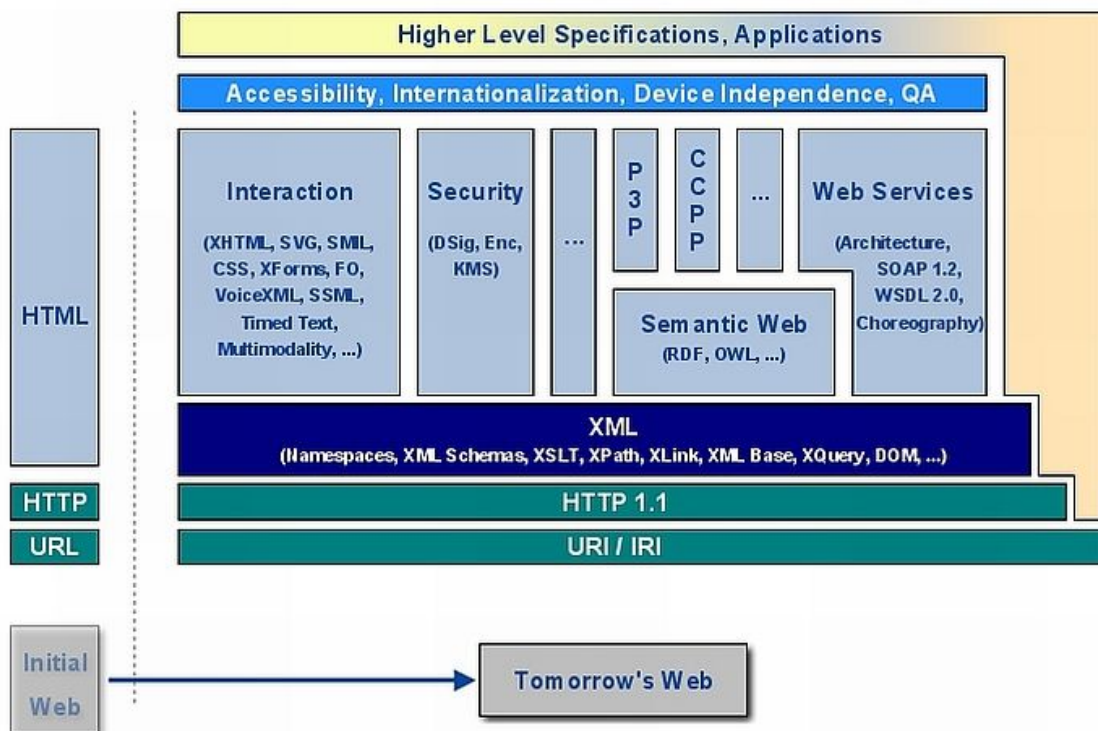


Abbildung 3: Empfehlungen für das Internet von Morgen

Quelle: About the World Wide Web Consortium (W3C), <http://www.w3c.org/Consortium/>, Stand: 15.02.2005

Der wichtigste Dienst im Internet stellt neben E-Mail und Dateiübertragung via File Transfer Protocol (FTP) das World Wide Web (WWW) dar. Die Datenübertragung zwischen WWW-Servern und -clients basiert hier auf dem Hypertext Transfer Protocol (HTTP).<sup>17</sup> Der Begriff Hypertext wurde 1965 von Ted Nelson eingeführt; er beschreibt ihn als non-sequential writing. Seine Grundidee war die Schaffung eines Informationssystems, das die assoziative Funktionsweise des menschlichen Gehirns hervorhebt und die Begrenzung herkömmlicher Texte in Papierform aufhebt.<sup>18</sup> Dieses Konzept inspirierte Tim Berners-Lee im Jahr 1991 schließlich zur

<sup>12</sup> Vgl. o.V. (The Requests for Comments (RFCs))

<sup>13</sup> <http://www.w3c.org>

<sup>14</sup> Vgl. Schwarze, J. (Einführung in die Wirtschaftsinformatik, 2000), S. 121 f.

<sup>15</sup> dies sind im Einzelnen Spezifikation, Standards, Richtlinien und Software-Werkzeuge

<sup>16</sup> Vgl. Jacobs, I. (About the World Wide Web Consortium (W3C), 2005)

<sup>17</sup> Vgl. Schwarze, J. (Einführung in die Wirtschaftsinformatik, 2000), S. 123 f.

<sup>18</sup> Vgl. Bichler, M. (Aufbau unternehmensweiter WWW-Informationssysteme, 1997), S. 19



Gründung des WWW-Projekts.<sup>19</sup> Die Sprache, in der solche interaktiven Dokumente im WWW erstellt werden können, nennt sich Hypertext Markup Language (HTML), eine Textauszeichnungssprache, die es ermöglicht, diese Dokumente mit Elementen wie Überschriften und Absätzen zu strukturieren.<sup>20</sup> Die obige Grafik (Abb. 3) deutet bereits an, dass in Zukunft verstärkt weitere Auszeichnungssprachen wie XML (Extensible Markup Language), XHTML (Extended HTML) und RDF (Resource Description Framework) in WWW-Dokumenten<sup>21</sup> eingesetzt werden sollen. Letzteres steht im direkten Zusammenhang mit dem Semantic Web, welches in Kapitel 4 näher erläutert wird.

Im nächsten Kapitel werden nun die grundsätzlichen Voraussetzungen für die Strukturierung von Informationen in Ergänzung zu den in diesem Kapitel vorgestellten Auszeichnungssprachen behandelt.

---

<sup>19</sup> Vgl. Berners-Lee, T. (WorldWideWeb: Summary, 1991)

<sup>20</sup> Vgl. Bichler, M. (Aufbau unternehmensweiter WWW-Informationssysteme, 1997), S. 19

<sup>21</sup> auch als Webseiten bezeichnet

### 3 Voraussetzungen für die Strukturierung von Informationen

#### 3.1 Informationen mit Bedeutung

Im zweiten Kapitel wurde ja bereits erwähnt, dass Daten zu Informationen werden, wenn man ihnen eine Bedeutung zuordnet. Hiermit beschäftigt sich die Semantik<sup>22</sup>. Sie liefert somit einen wichtigen Beitrag zur Strukturierung der Informationen bzw. des Wissens im Internet.

Für den Wert des Wissens ist es entscheidend, dass die Anbieter und Nachfrager<sup>23</sup> von Informationen eine gemeinsame Abstimmung auf drei Ebenen erzielen, um semantische Konflikte, wie in der folgenden Abbildung dargestellt, zu vermeiden: bei den verwendeten Terminologien, den schematischen Schemata und den Ontologien.<sup>24</sup>

Konsens-ebene	Konsens-risiko	Beispiel	Konsens-werkzeug
Symbole	Inkompatibles Vokabular	"APFEL" ≠ "APPLE"	Terminologien
Zuordnung	Inkonsistente Zuordnung zu Konzepten	"APPLE" → "FRUCHT" "APPLE" → "UNTERNEHMEN"	Semantische Schemata
Konzepte	Inkonsistente Bedeutung (semantische Analyse)	"UNTERNEHMEN" = physisches Objekt, ... "UNTERNEHMEN" = rechtliches Gebilde, ...	Ontologien

Abbildung 4: Semantische Konflikte

Quelle: Bodendorf, F. (Daten- und Wissensmanagement, 2003), S. 107

Mithilfe von Terminologien erfolgt eine Einigung auf eine gemeinsame Menge von Symbolen<sup>25</sup>, die jedoch meist auf einen bestimmten Bereich (z.B. Englisch oder Deutsch) begrenzt und somit inkompatibel zu anderen Bereichen sein können, da die verwendeten Symbole nicht übereinstimmen. Darüberhinaus kann ein Computer diesen Symbolen auf dieser Ebene keine Bedeutung zuordnen, wodurch es zu Fehlinterpretationen durch den Benutzer kommen kann. Erst mit semantischen Schemata – quasi maschinenlesbaren „Begleittexten“ zu diesen Terminologien – kann eine Zuordnung der Symbole zu Konzepten<sup>26</sup> erfolgen. Ausserdem können terminologisch inkompatible Systeme durch ein gemeinsames semantischen Schema, welches die unterschiedlichen Symbole einem eindeutig definierten gemeinsamen Konzept zuordnet, sinnvoll miteinander kommunizieren. So können die Symbole von einem in das andere System übersetzt werden; dies könnte z.B. die Übersetzung vom Englischen ins Deutsche sein (Zwei

<sup>22</sup> Semantik=Bedeutungslehre oder Lehre von der inhaltlichen Bedeutung einer Sprache; vgl. Schwarze, J. (Einführung in die Wirtschaftsinformatik, 2000), S. 41

<sup>23</sup> Anbieter und Nachfrager können sowohl Maschinen als auch Menschen sein

<sup>24</sup> Vgl. Bodendorf, F. (Daten- und Wissensmanagement, 2003), S. 107

<sup>25</sup> Symbole können als Daten verstanden werden

<sup>26</sup> semantische Kategorie des Beschreibungsgegenstandes

unterschiedliche Symbole bzw. Wörter mit der gleichen Bedeutung). Semantische Schemata sind somit für Interoperabilität und Integration verteilter Systeme wie z.B. auch dem Internet von großer Bedeutung; für die Definition semantischer Schemata eignet sich z.B. die Auszeichnungssprache XML. Allerdings können auch auf dieser Ebene noch semantische Konflikte auftreten, wenn die Konzepte von den Beteiligten unterschiedlich interpretiert werden; d.h. einem Symbol werden verschiedene Bedeutungen zugeordnet. Dem Computer ist eine Interpretation der definierten Konzepte generell nicht möglich. Nur der Mensch kann in dieser Situation noch eine Unterscheidung aufgrund des Kontexts vornehmen. Zur Lösung dieser Problematik müssen Konzeptdefinitionen in Form von sogenannten Ontologien<sup>27</sup> festgelegt werden. Diese bilden die Bedeutung der Beschreibungsgegenstände maschinenlesbar ab, so dass sich dem Rechner die Bedeutung der zugeordneten Symbole erschließt. Eine Ontologie verknüpft hierzu die Konzepte über Relationen miteinander, welche die Beziehungen der Konzepte untereinander in beliebiger Komplexität darstellen können. Hierdurch wird die Bedeutung des Konzepts eingegrenzt. Zu den Konzepten und Relationen können noch Axiome ergänzt werden; dies sind Nebenbedingungen, die die Zulässigkeit von bestimmten Relationen zwischen Konzepten bestimmen (es darf z.B. nur eine Relation zwischen zwei Konzepten bestehen, wenn auch eine Relation zu einem bestimmten anderen Konzept besteht). Schließlich kann der Computer so die zugrunde liegenden Konzepte interpretieren, wodurch die Übertragbarkeit und Wiederverwendbarkeit von mit Ontologien formulierten Wissen sichergestellt wird.<sup>28</sup>

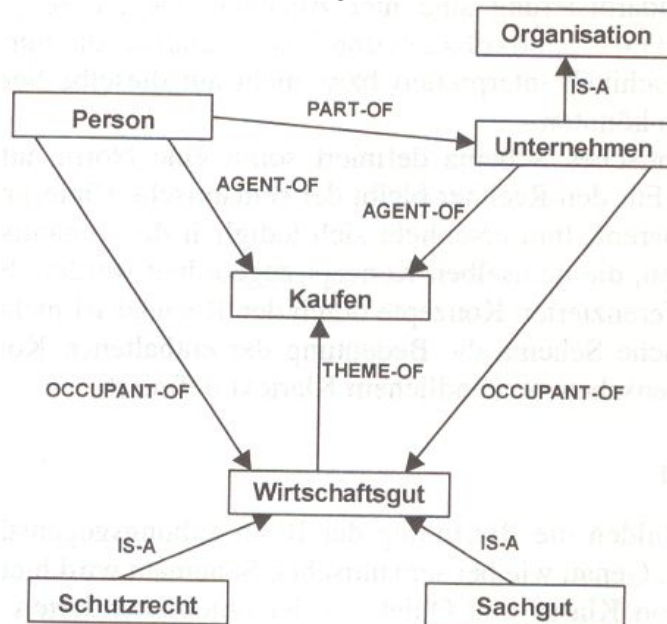


Abbildung 5: Ausschnitt (Aspekt) einer Ontologie  
Quelle: Bodendorf, F. (Daten- und Wissensmanagement, 2003), S. 112

<sup>27</sup> Ontologie=Zweig der Philosophie, der das grundlegende Wesen existierender Dinge zu ergründen versucht; vgl. Bodendorf, F. (Daten- und Wissensmanagement, 2003), S. 112

<sup>28</sup> Vgl. Bodendorf, F. (Daten- und Wissensmanagement, 2003), S. 108 ff.

### 3.2 Vernetzung von Wissen

Anfangs wurde ja bereits erwähnt, dass durch die Vernetzung von Informationen Wissen geschaffen wird. Somit kann die Vernetzung von Informationen und Wissen auch zur inhaltlichen Strukturierung des Internets beitragen.

Die Vernetzung von Wissen beruht auf der Funktionsweise des menschlichen Gehirns, welches Wissen in Strukturen ablegt und verwaltet, die weder linear noch hierarchisch gegliedert sind. Es verbindet Wissen und das Wissen über die Beziehungen zwischen Wissen in vielfältigen assoziativen Verknüpfungen miteinander. Ein Ansatz zum Wissensmanagement – wie hier in Bezug auf das Internet –, der diesen Aspekt nicht berücksichtigt, kann demnach die Arbeitsweise des menschlichen Gehirns nicht nachbilden. D.h. erfolgt die Verwaltung von Wissen (z.B. Dokumenten) in linearen oder hierarchischen Strukturen, führt dies zu Problemen bei der Einordnung dieses Wissens; insbesondere können mehrdimensionale Verbindungen nicht ohne Redundanzen<sup>29</sup> in diesen Strukturen abgebildet werden. Dieses Problem kann durch die Verwaltung von Wissen als „Knoten“ eines Netzwerks vermieden werden, da potenziell alle Knoten miteinander verknüpft werden können. Auf Basis dieses Netzwerks ist eine assoziative Navigation, die der Arbeitsweise des menschlichen Gehirns nahe kommt, zwischen den Wissens-elementen möglich.<sup>30</sup>

Diese Vernetzung von Wissen kann z.B. durch das bereits erwähnte Konzept des Hypertextes aber auch durch Auszeichnungssprachen wie XML oder RDF gefördert werden.

Da die Grundlagen und Voraussetzungen für die Strukturierung von Informationen bzw. Wissen im Internet erläutert wurden, werden in den folgenden Kapiteln zwei praktische Ansätze in Bezug auf diese Problematik vorgestellt und bewertet.

---

<sup>29</sup> das Wissen muss mehrfach in unterschiedliche Kategorien eingeordnet werden

<sup>30</sup> Vgl. Bodendorf, F. (Daten- und Wissensmanagement, 2003), S. 114

## 4 Praktische Ansätze

### 4.1 Semantic Web

Der am weitesten verbreitete Ansatz ist das Konzept „Semantic Web“. Daher wird es an dieser Stelle zuerst vorgestellt.

„The Semantic Web is an extension of the current Web in which information is given well-defined meaning, enabling computers and people to work in better cooperation“.<sup>31</sup> Diese Beschreibung von Tim Berners-Lee macht deutlich was mit dem Semantic Web beabsichtigt wird: eine Erweiterung des Internets um ein Konzept, dass den Informationen eine Bedeutung geben soll, die sowohl von Menschen als auch von Maschinen verstanden werden kann.

Das Semantic Web ist eine Initiative des W3C, die 1998 von Tim Berners-Lee ins Leben gerufen wurde<sup>32</sup>. Es soll das Management persönlicher Informationen, die Integration von Unternehmens-Anwendungen und den globalen Austausch von kommerziellen, wissenschaftlichen und kulturellen Informationen effizient untereinander verbinden. Hierzu werden unter Führung des W3C entsprechende Technologien und Standards entwickelt. Als Schlüsseltechnologien haben sich die Auszeichnungssprache RDF und die Web Ontology Language (OWL) herausgestellt, da der Datenaustausch über diese Sprachen alle oben genannten Bereiche so vereinen kann, dass alle die gleichen Informationen teilen auch wenn sie nicht die gleiche Software teilen. Software für das Semantic Web, die RDF und OWL verwendet bzw. generiert, sollte demnach folgende Funktionen beinhalten:<sup>33</sup>

- Autoren von Web-Dokumenten können diesen Metadaten (z.B. Thema, Verfasser, Sprache und vieles mehr) hinzufügen, um die Suche nach diesen Dokumenten zu vereinfachen.
- Große Web-Sites<sup>34</sup> können entsprechend der Kategorien, den der Inhalt zugeordnet ist, verwaltet werden.
- Die Daten können von verschiedenen Quellen anwendungsübergreifend wiederverwendet werden, da RDF und OWL nicht-proprietäre Standards sind.

Im Folgenden wird noch einmal näher auf RDF und OWL eingegangen.

---

<sup>31</sup> Berners-Lee, T. / Miller, E. (The Semantic Web lifts off, 2002)

<sup>32</sup> Vgl. Berners-Lee, T. (Semantic Web Roadmap, 1998)

<sup>33</sup> Vgl. Miller, E. (Semantic Web Activity Statement, 2004)

<sup>34</sup> Web-Site = Sammlung von miteinander verbundenen Web-Seiten / -dokumenten

RDF ist eine Empfehlung des W3C<sup>35</sup> und dient als Rahmenkonzept zur Repräsentation von Informationen im WWW. Es basiert auf der Auszeichnungssprache XML und kann verschiedene Vokabularien verwenden; das wichtigste Vokabular ist das RDF Schema<sup>36</sup>. Das Ziel von RDF ist es, Web-Ressourcen von jedermann mit Meta-Daten über Inhalt und sonstige wichtige Eigenschaften der Ressource anzureichern. Dies soll unter Verwendung eines einfachen Datenmodells (RDF triple) aus Subjekten, die über ein Prädikat (= Eigenschaft) ein Objekt beschreiben, geschehen<sup>37,38</sup>.

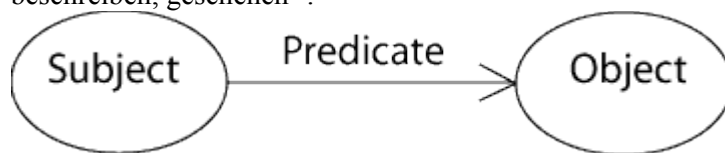


Abbildung 6: RDF triple

Quelle: Klyne, G. u.a. (Resource Description Framework (RDF): Concepts and Abstract Syntax, 2004)

Als Beispiel kann man Folgendes annehmen:

Man hat ein Subjekt bzw. eine Resource in Form einer Web-Site (<http://www.beispiel.de>); deren Eigenschaft (Property) ist z.B. der Autor oder dessen E-Mail-Adresse. Der Name des Autors oder die konkrete Adresse sind Objekte (Statement). Diese Beziehungen können nun im HTML-Quelltext der jeweiligen Webseite über sogenannte Tags<sup>39</sup> hinzugefügt (annotiert) werden.<sup>40</sup> In Ergänzung hierzu kann das Metadaten-Set Dublin Core gesehen, welche ebenfalls standardisierte Elemente zur Beschreibung der oben genannten Informationen einer Webseite ermöglicht.<sup>41</sup>

OWL hat noch mehr Möglichkeiten die Semantik in Webseiten auszudrücken als RDF. D.h. diese Sprache macht den Seiteninhalt für den Computer noch wesentlich verständlicher. Der Einsatz von OWL ist dann sinnvoller, wenn die Informationen aus Web-Dokumenten in Computer-Anwendungen verarbeitet werden sollen anstatt sie nur dem Menschen zu präsentieren. Die Sprache dient letztendlich dazu, Ontologien wie in Abschnitt 3.1 beschrieben, darzustellen.<sup>42</sup>

<sup>35</sup> Aktuelle Version: <http://www.w3.org/TR/rdf-concepts/>

<sup>36</sup> Aktuelle Version: <http://www.w3c.org/TR/rdf-schema/>

<sup>37</sup> wird auch Statement, Property und Resource bezeichnet; vgl. van Harmelen, F. / Fensel, D. (Practical Knowledge Representation for the Web, 1999), S. 3

<sup>38</sup> Vgl. Klyne, G. u.a. (Resource Description Framework (RDF): Concepts and Abstract Syntax, 2004)

<sup>39</sup> die genaue Syntax kann in den jeweiligen Empfehlungen des W3C nachgelesen werden

<sup>40</sup> Vgl. van Harmelen, F. / Fensel, D. (Practical Knowledge Representation for the Web, 1999), S. 3 f.

<sup>41</sup> Vgl. o.V. (Dublin Core Metadata Element Set, Version 1.1: Reference Description, 2004)

<sup>42</sup> Vgl. McGuinness, D. L. / van Harmelen, F. (OWL Web Ontology Language Overview, 2004)

Die folgende Grafik veranschaulicht noch einmal den ungefähren Aufbau des Semantic Web:

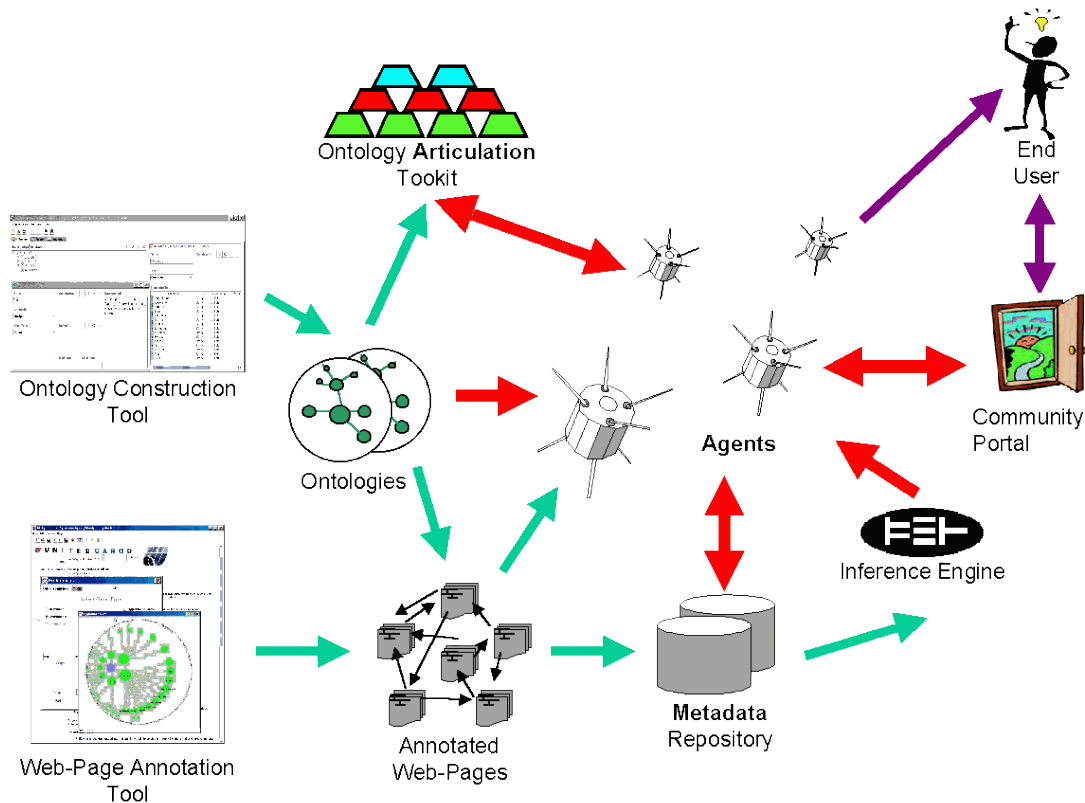


Abbildung 7: Komponenten des Semantic Web  
Quelle: o.V (About SemanticWeb.org)

## 4.2 XML Topic Maps

XML Topic Maps gehören ähnlich wie das Semantic Web auch zu den semantischen Netzen. Mit ihrer Hilfe können Wissensstrukturen und Assoziationen in Informationsbeständen standardisiert beschrieben werden.<sup>43</sup> Der Standard „XML Topic Maps (XTM) 1.0“ wurde unter anderem von Steve Pepper entwickelt. Er ermöglicht die Darstellung von Informationsstrukturen in Form von so genannten Topics und Beziehungen (associations) zwischen diesen Topics. Topics sind abstrakte Objekte, die durch Namen, Informationsressourcen oder Beziehungen charakterisiert werden können. Diese Charakteristika gelten innerhalb eines begrenzten Kontextes (scope) im Rahmen des Topics. Die letzte Schlüsselkomponente sind die Instanzen (occurrences), in denen relevante Informationen zu dem Topic abgelegt werden können; diese müssen entweder über eine Adresse (z.B. eine Webadresse) referenzierbar sein

<sup>43</sup> Vgl. o.V. (Was sind Topic Maps?)

oder in Form von einfachen Zeichen darstellbar sein. Schließlich werden ein oder mehrere Dokumente, die diese Darstellungsweise verwenden, als Topic Map bezeichnet.<sup>44</sup>

Der Standard XTM wurde im Jahr 2000 durch die TopicMaps.Org Authoring Group<sup>45</sup> veröffentlicht. Die Ursprünge gehen zurück auf eine Idee der Davenport Group aus dem Jahre 1993. Das Konzept wurde erstmals in Form des ISO-Standards 13250<sup>46</sup> niedergeschrieben.<sup>47</sup>

Mit XTM werden unter anderem die folgenden Ziele verfolgt:<sup>48</sup>

1. Es soll überall im Internet verwendbar sein
2. Es soll eine Vielzahl von Anwendungen unterstützen
3. Es ist kompatibel mit XML
4. Es soll einfach sein XTM-Dokumente zu erstellen

Dieses Kapitel kann nur einen oberflächlichen Überblick über die beiden praktischen Ansätze geben, da ein detaillierter Einblick – insbesondere in Bezug auf die verwendete Syntax der Auszeichnungssprachen – den Rahmen dieser Arbeit sprengen würde. Zur Vertiefung werden die entsprechenden Quellen des W3C bzw. von TopicMaps.Org empfohlen.

---

<sup>44</sup> Vgl. Pepper, S. / Moore, G. (XML Topic Maps (XTM) 1.0, 2001)

<sup>45</sup> <http://www.topicmaps.org>

<sup>46</sup> <http://www.topic-maps.at/downloads/iso13250-2nd-ed-v2.pdf>

<sup>47</sup> Vgl. Pepper, S. / Moore, G. (XML Topic Maps (XTM) 1.0, 2001)

<sup>48</sup> Vgl. Ebd.



## 5 Bewertung der praktischen Ansätze

Im Rahmen einer Bewertung kann man feststellen, dass die Voraussetzungen für die Strukturierung von Informationen von beiden Ansätzen verfolgt werden. So ermöglichen beide Konzepte die Zuordnung von Zusatzinformationen bzw. Metadaten, um den eigentlichen Informationen eine bestimmte Bedeutung zuzuweisen. Desweiteren wird durch die Möglichkeiten beider Modelle, Beziehungen zwischen Informationsobjekten darzustellen, auch dem Aspekt der Vernetzung Rechnung getragen. Dies haben beide Ansätze gemeinsam.

Vergleicht man die Ansätze Semantic Web und XML Topic Maps bzw. den Standard XTM und die Auszeichnungssprache RDF miteinander stösst man aber auch auf Unterschiede, die im Folgenden aufgeführt werden sollen:<sup>49</sup>

- Topic Maps und RDF wurden von verschiedenen Organisationen entwickelt (ISO und W3C), die allgemein hin als Wettbewerber angesehen werden. Somit können auch die beiden Standards als Konkurrenten betrachtet werden; wobei durchaus Synergieeffekte zwischen beiden bestehen (z.B. können durch RDF Topic Maps generiert werden).
- Topic Maps werden zum Wissensmanagement aus der Sicht des Menschen eingesetzt; wohin gegen RDF zum Wissensmanagement aus der Sicht von Computern eingesetzt werden soll.
- Topic Maps ermöglichen eine detailliertere Modellierung als RDF. Während es in RDF nur Ressourcen gibt, die bestimmte Werte als Eigenschaften haben, können Topics sehr verschiedenartige Charakteristika zugeordnet werden; was eher der menschlichen Denkweise entspricht.

Im Endeffekt hängt die Entscheidung für den ein oder anderen Ansatz vom geplanten Einsatzgebiet ab. Für das Internet wird das Konzept des Semantic Web weiter an Bedeutung gewinnen, da seine Entwicklung durch einflussreiche Organisationen und Personen weiterentwickelt wird.

---

<sup>49</sup> Vgl. Pepper, S. (Ten Theses on Topic Maps and RDF, 2002)

## **6 Fazit und Ausblick**

Als Fazit dieser Arbeit kann man nun feststellen, dass es durchaus sinnvolle Ansätze für ein Wissensmanagement im Internet gibt. Diese Standards sind auch schon seit einiger Zeit etabliert und werden aktiv gefördert. Insbesondere das Konzept des Semantic Web wird schon als Erweiterung des bisherigen Internets eingeplant.

Die Anreicherung der Informationen mit semantischen Metadaten liegt jedoch im Grunde in der Verantwortung des jeweiligen Autors des Web-Dokuments. Dies muss flächendeckend auch konsequent durchgeführt werden, damit möglichst viele Information strukturiert erfasst werden können. Weiterhin sollten Anwendungen entwickelt und eingesetzt werden, die die semantische Auszeichnung der Information für den Autor vereinfachen bzw. diese Arbeit möglichst weitgehend automatisieren.

Aber auch schon das heutige Internet bietet durch den Einsatz von Hypertexten mit entsprechenden Links den Aspekt der Vernetzung von Informationen, der der Funktionsweise menschlichen Denkens nahe kommt. So werden hierdurch schon heute die Recherche nach Wissen und E-Learning-Maßnahmen sinnvoll unterstützt.

Schließlich bleibt zu beobachten, wie sich das Internet in Zukunft entwickelt und inwiefern es eine Art künstliche Intelligenz entwickelt, die dem Menschen in seinem täglichen Leben unterstützt, wie es in Visionen von einigen Experten beschrieben wird.

## **Literaturverzeichnis**

### **Berners-Lee, T. (1989):**

Information Management: A Proposal, <http://www.w3.org/History/1989/proposal.html> ,  
abgerufen am: 15.02.2005

### **Berners-Lee, T. (1991):**

WorldWideWeb: Summary, <http://groups.google.com/groups?selm=6487%40cernvax.-cern.ch> , abgerufen am: 15.02.2005

### **Berners-Lee, T. (1998):**

Semantic Web Roadmap, <http://www.w3c.org/DesignIssues/Semantic> , abgerufen am:  
16.02.2005

### **Berners-Lee, T. / Miller, E. (2002):**

The Semantic Web lifts off, [http://www.ercim.org/publication/Ercim\\_News/enw51/berners-lee.html](http://www.ercim.org/publication/Ercim_News/enw51/berners-lee.html) , abgerufen am: 16.02.2005

### **Bichler, M. (1997):**

Aufbau unternehmensweiter WWW-Informationssysteme, Braunschweig u.a. 1997

### **Bodendorf, F. (2003):**

Daten- und Wissensmanagement, Berlin u.a. 2003

### **Jacobs, I. (2005):**

About the World Wide Web Consortium (W3C), <http://www.w3.org/Consortium/about-w3c.html> , abgerufen am: 16.02.2005

### **Klyne, G. u.a. (2004):**

Resource Description Framework (RDF): Concepts and Abstract Syntax, <http://www.w3.org/TR/2004/REC-rdf-concepts-20040210/> , abgerufen am: 16.02.2005

**McGuinness, D. L. / van Harmelen, F. (2004):**

OWL Web Ontology Language Overview, <http://www.w3.org/TR/2004/REC-owl-features-20040210/> , abgerufen am: 17.02.2005

**Miller, E. (2004):**

Semantic Web Activity Statement, <http://www.w3c.org/2001/sw/Activity> , abgerufen am: 16.02.2005

**o.V.:**

The Requests for Comments (RFCs), <http://www.rfc-editor.org> , abgerufen am: 15.02.2005

**o.V.:**

About SemanticWeb.org, <http://www.semanticweb.org/about.html#bigpicture> , abgerufen am: 17.02.2005

**o.V.:**

Was sind Topic Maps?, <http://www.topic-maps.at/topicmap.html> , abgerufen am: 17.02.2005

**o.V. (2004):**

Dublin Core Metadata Element Set, Version 1.1: Reference Description, Dublin Core Metadata Initiative, <http://dublincore.org/documents/2004/12/20/dces/> , abgerufen am: 17.02.2005

**Pepper, S. / Moore, G. (2001):**

XML Topic Maps (XTM) 1.0, <http://www.topicmaps.org/xtm/1.0/xtm1-20010806.html> , abgerufen am: 17.02.2005

**Pepper, S. (2002):**

Ten Theses on Topic Maps and RDF, <http://www.ontopia.net/topicmaps/materials/rd-f.html> , abgerufen am: 18.02.2005

**Schwarze, J. (2000):**

Einführung in die Wirtschaftsinformatik, 5. Aufl., Herne u.a. 2000

**van Harmelen, F. / Fensel, D. (1999):**

Practical Knowledge Representation for the Web, <http://www.cs.vu.nl/~frankh/post-script/IJCAI99-III.pdf>, abgerufen am: 17.02.2005

**Zhong, N. u.a. (Hrsg.) (2003):**

Web Intelligence, Berlin u.a. 2003

Pepper, S. (, 2002)