

Masterarbeit

Entwicklung eines Editors zur semantischen Auszeichnung von Texten

von

Canan Hastik

Darmstadt, den 10.08.2009

Referent:
Prof. Dr. Bernhard Thull

Koreferent:
Prof. Dr. Gerhard Knorz

Betreuer:
Dr. Achim Steinacker
Intelligent views gmbh
Darmstadt



Erklärung

Hiermit erkläre ich, dass ich die vorliegende Arbeit selbständig erstellt und keine anderen als die angegebenen Hilfsmittel und Quellen benutzt habe. Soweit ich auf fremde Materialien, Texte oder Gedankengänge zurückgegriffen habe, enthalten meine Ausführungen vollständige und eindeutige Verweise auf die Urheber und Quellen.

Alle weiteren Inhalte der vorgelegten Arbeit stammen im urheberrechtlichen Sinn von mir, soweit keine Verweise und Zitate erfolgen. Mir ist bekannt, dass ein Täuschungsversuch vorliegt, wenn die vorstehende Erklärung sich als unrichtig erweist.

Ort, Datum

Unterschrift

Erklärung

Bitte ankreuzen:

- Mit der Ausleihe der Diplomarbeit sind wir einverstanden.
- Mit der Ausleihe sind wir nicht einverstanden, die Diplomarbeit ist gesperrt.

Studentin / Student _____
Datum Unterschrift

Betreuender Professor _____
Datum Unterschrift

Betreuende Firma _____
Datum Unterschrift



Danksagung

An dieser Stelle soll den Personen gedankt werden, ohne die diese Arbeit nicht möglich gewesen wäre und die entscheidend zum Niveau beigetragen haben.

Mein Dank richtet sich zunächst an die intelligent views gmbh in Darmstadt und im speziellen an meinen Betreuer Dr. Achim Steinacker, für die Zusammenarbeit und die Möglichkeit der praxisnahen Tätigkeit.

Ich bedanke mich bei Prof. Dr. Bernhard Thull für seine Betreuung und sein entgegengebrachtes Vertrauen. Zu den richtigen Zeitpunkten hat er die richtigen Hinweise gegeben.

Ein besonderer Dank geht an Prof. Dr. Gerhard Knorz für die Übernahme des Koreferats.

Unendlicher Dank gebührt meinem Mann und meiner Tochter, dafür dass sie durchgehalten haben.

Ein großer Dank geht an meinen Lektor Peter Hastik und an meine Eltern für ihren Zuspruch.



M.C. Escher [1898-1972] – Belvedere
“Different views reflecting a different system”
[Tekinerdođan, 2000]



Inhaltsverzeichnis

Abbildungsverzeichnis	11
Tabellenverzeichnis	14
Abkürzungsverzeichnis	15
Kurzfassung	16
I	Einleitung
1	Ausgangslage und Methodik 17
1.1	Globale Probleme 17
1.2	Aufgabenstellung 19
1.3	Social Web und Semantic Web 21
1.4	Aufbau der Arbeit 23
II	Grundlagen, Analyse und Stand der Technik
2.	Analyse der Wissensarbeit 25
2.1	Begriffsbestimmung 25
2.2	Individuelle Wissensarbeit 26
2.3	Wissen und Lernen 29
2.3.1	Wahrnehmung, Lernen und Wissenserwerb 30
2.3.2	Interaktive Kommunikation & Kollaboration 33
2.3.3	Aspekte der Motivation 38
2.4	Anforderungen an eine Wissensinfrastruktur 42
3	Social Semantic Web 44
3.1	User Experience, Trends und Technik 45
3.1.1	Tagging und die Wissensarbeit 46
3.1.2	Folksonomy und die Wissensarbeit 50
3.1.3	(Hyper)-Text 52
3.1.4	Potentiale semantischer Technologien 55
3.1.5	„Think global, act local“ 60
3.1.6	Erfolgsfaktoren und Probleme 62

4	Stand der Technik	64
4.1	Empirische Untersuchungen	64
4.1.1	Usabilitytest MediaWiki und SMW+	65
4.1.2	Native Nutzerbeobachtung	72
4.1.3	Evaluation des ELWMS	72
4.1.4	Forschungsergebnisse	73
III	s2w Konzeption	
5	Web 2.0 trifft auf Semantic Web	74
5.1	Benutzerspezifische Anforderungen	74
5.1.1	Konzeptanforderungen	74
5.1.2	Anwendungskontext	76
5.1.3	Typisierung der Benutzergruppe	78
5.1.4	Konzeptziele	79
5.2	s2w-EDENtag Design Studie	81
5.2.1	Funktionsaufruf des EDENtag Dialogs	83
5.2.2	Interaktionsmechanismen	85
5.2.2.1	Aufbau und Struktur des Dialogs	86
5.2.2.2	Dateneingabe und Rückmeldung	88
5.2.3	Navigation und Orientierung	92
5.2.3.1	Facetten-Hierarchie	92
5.2.3.2	EDENtag Graph	94
5.2.4	s2w-EDENtag Design Resümee	95
5.3	s2w-Eden Service Architektur	96
5.3.1	s2w-EDENtag Ontologie	96
5.3.2	s2w-EDENtag Store	99
5.3.3	Nutzungspotential	100
6	Proof of Concept	101
6.1	WYSIWYG-Javascript-Editor	101
6.1.1	EDENtag Plug-In-Template	102
6.1.2	Funktionsimplementierung	104
6.1.2.1	jQuery Plug-In: Autocomplete	105
6.1.2.2	MAX_LEV_DIS Algorithmus	106

	6.1.2.3 Facetten-Hierarchie	107
	6.1.2.4 Graphenbasierte Darstellung	108
	6.1.3 Ergebnisse	108
6.2	Prototyp	108
	6.2.1 K-Infinity	108
	6.2.2 Demonstrator kp3	109
	6.2.3 Machbarkeitsprüfung PHP unter Tomcat	113
	6.2.4 EDENtag Editor Plug-In	114
	6.2.5 Anbindung des Net-Navigator	115
6.3	Evaluation	117
	6.3.1 Testkriterien und Methoden	117
	6.3.2 Testumgebung	118
	6.3.3 Test-Szenario	118
	6.3.4 Ergebnisse der Benutzerbeobachtung	120
	6.3.5 Ergebnisse des standardisierten Interviews	124
IV	Diskussion und Ausblick	
	7 Zusammenfassung	127
	7.1 Diskussion der Ergebnisse	127
	7.1.1 Konzeption	127
	7.1.2 Prototypische Implementierung	130
	7.2 Ausblick	131
V	Literaturverzeichnis	
	8 Quellenangaben	132
	8.1 Monographien und Periodika	132
	8.2 Online Quellen	143
VI	Anhang	
	Anhang A Fragebogen: Ermittlung des persönlichen Stils	148
	Anhang B Fragenkatalog Benutzertest	151
	Anhang C Test-Szenario: Dialognutzung	158
	Anhang D Demonstrator	161

Abbildungsverzeichnis

1| Ausgangslage und Methodik

Abb. 1.1	Netz 2.0 [Angermeier, 2005]	S.22
----------	-----------------------------	------

2| Analyse der Wissensarbeit

Abb. 2.1	Von der Wahrnehmung zum Wissen	S.27
Abb. 2.2	Hauptaufgaben des Wissensarbeiters	S.27
Abb. 2.3	Umgebungskontext des Wissensarbeiters	S.28
Abb. 2.4	Wahrnehmung und interaktiver Kommunikation	S.32
Abb. 2.5	Semiotisches Dreieck [Staab, 2001]	S.33
Abb. 2.6	The Human Information Processor [Card, 1983]	S.36
Abb. 2.7	Bedürfnispyramide [Maslow, 1954]	S.40

3| Social Semantic Web

Abb. 3.1	Information Management [Lee, 1998]	S.44
Abb. 3.2	Grafische Darstellung von Webseiten [Wiki, 2009]	S.45
Abb. 3.3	Semantisches Netzwerk [Collins, 1969]	S.49
Abb. 3.4	Flickr Tag Cloud – Screenshot	S.49
Abb. 3.5	Reciprocal Relationship: People & Content [Morville, 2002b]	S.50
Abb. 3.6	Navigating the Talmud [Hirshon, 2002]	S.54
Abb. 3.7	Semantic Annotation [Davies, 2007]	S.56
Abb. 3.8	Spectrum of Formalism [Bergmann, 2008]	S.57
Abb. 3.9	Classifying Ontologies [Ontology Summit, 2007]	S.58
Abb. 3.10	RDF-Beschreibung	S.59

4| Stand der Technik

Abb. 4.1	The User Experience Honeycomb [Morville, 2002]	S.64
Abb. 4.2	Histogramm über Wiki-Probleme [Désilets, 2004]	S.66
Abb. 4.3	Demonstrator für Testumgebung	S.68

Abb. 4.4	Navigation mit Button	S.69
Abb. 4.5	Navigation mit Link	S.69
Abb. 4.6	Semantic Toolbar	S.70
Abb. 4.7	Ontologie Browser	S.70
Abb. 4.8	Annotation Toolbar	S.71
Abb. 4.9	Semantic Toolbar mit Auto Completion	S.71

5| Web 2.0 trifft auf Semantic Web

Abb. 5.1	Typische Anwendungsfälle	S.77
Abb. 5.2	Design-Skizze	S.80
Abb. 5.3	GOMS Interaktionsmodell	S.81
Abb. 5.4	Elementares Kontextelement: FCKeditor	S.82
Abb. 5.5	Textauswahl	S.83
Abb. 5.6	Funktionsaufruf über Icon in der Toolbar	S.84
Abb. 5.7	Funktionsaufruf über rechte Maustaste	S.84
Abb. 5.8	Best Practice for Form Design [Wroblewski, 2008]	S.86
Abb. 5.9	Above and Left-Aligned Labels [Penzo, 2006]	S.87
Abb. 5.10	EDENtag Dialog	S.87
Abb. 5.11	Eingabebeispiel	S.88
Abb. 5.12	Auto Completion mit persönlicher Tag-History	S.89
Abb. 5.13	Integration von Synonymen	S.89
Abb. 5.14	Did you mean...	S.90
Abb. 5.15	Levenshtein-Beispiel	S.90
Abb. 5.16	Icon-Beispiele zur Textauszeichnung	S.91
Abb. 5.17	Ausgeblendeter OK-Button bei fehlender Eingabe	S.92
Abb. 5.18	Tag-Facetten	S.93
Abb. 5.19	Facetten-Hierarchie	S.93
Abb. 5.20	Mouse-Over	S.94
Abb. 5.21	EDENtag Graph	S.94
Abb. 5.22	Ergebnis einer semantischen Auszeichnung	S.95
Abb. 5.23	Basic Model of Tagging-System [Smith, 2008]	S.96
Abb. 5.24	RDF-Aussagen	S.97
Abb. 5.25	s2w EDENtag Modell	S.97
Abb. 5.26	XML-Serialisierung des EDENtag Modells	S.98
Abb. 5.27	EDENtag Ontologie	S.99
Tab. 5.28	EDEN_tagging Datenbank-Schema	S.99
Tab. 5.29	EDEN_labeling Datenbank-Schema	S.99

6| Proof of Concept

Abb. 6.1	Javascript-Wysiwyg-Editoren [Kreußel, 2007]	S.101
Abb. 6.2	A Sample Plug-In Definition File	S.103
Abb. 6.3	EDENplug	S.103
Abb. 6.4	Plug-In Konfiguration I	S.103
Abb. 6.4	Plug-In Konfiguration II	S.104
Abb. 6.5	Installationsanleitung EDENtag Plug-In	S.104
Abb. 6.6	jQuery Plug-in: Autocomplete	S.105
Abb. 6.7	SQL-Query aus get_tags.php	S.105
Abb. 6.8	Implementiertes jQuery Plug-In: Autocomplete	S.106
Abb. 6.9	Berechnung der Levenshtein-Distanz	S.106
Abb. 6.10	Abfrage „Related documents“	S.107
Abb. 6.11	Knowledge Builder	S.109
Abb. 6.12	Beispiel-Import „keywords INI Science Library“	S.110
Abb. 6.13	Semantisches Wissensnetz	S.110
Abb. 6.14	kp3-Wissensportal	S.111
Abb. 6.15	Integrierter FCK-Editor in K-Infinity	S.112
Abb. 6.16	Integrierter FCK-Editor mit Plug-In-Funktion	S.112
Abb. 6.17	php5-Servlet	S.114
Abb. 6.18	Konsolen-Batch	S.114
Abb. 6.19	mysql_connect	S.115
Abb. 6.20	Net-Navigator Request	S.115
Abb. 6.21	Graphische Darstellung der relevanten Facetten	S.116
Abb. 6.22	Inhaltserfassung: Scan-Pfad & Blickpunkte	S.120
Abb. 6.23	Fokus der visuellen Aufmerksamkeit (6 counts)	S.121
Abb. 6.24	Nutzung der Auto-Completion (Proband2)	S.122
Abb. 6.25	Related facetts	S.123
Abb. 6.26	Scan-Pfad EDENtag-Icons	S.123

Tabellenverzeichnis

Tab.1	Konzeptvergleich Web 1.0 und Web 2.0	S. 46
Tab.2	Vor- und Nachteile: Tagging & Folksonomies	S. 63
Tab.3	Zufriedenheits-Barometer	S. 71
Tab.4	Anforderungsdefinition	S. 75
Tab.5	Nutzungsszenario: Innovationsprojekt	S. 76
Tab.6	Nutzungsszenario: Literaturrecherche	S. 77
Tab.8	Ergebnisse aus dem Fragenkatalog	S. 126

Abkürzungen

HTML	Hypertext Markup Language
www	World Wide Web
RDF	Ressource Description Framework
OWL	Web Ontology Language
W3C	World Wide Web Consortium
EXIF	Exchangeable Image File Format
PDF	Portable Document Format
s2w	Socio-Semantic-Web
URI	Unified Ressource Identfier
SKOS	Simple Knowledge Organisation System
REST	Representational State Transfer
SOAP	Simple Object Access Protocol
WSDL	Web Service Description Language
JSON	JavaScript Object Notation
API	Programmierschnittstelle
SMW	Semantic MediaWiki
SMW+	Semantic MediaWiki mit Halo-Extension
SHE	Syntax Highlighting Editor
WYSIWYG	What You See Is What You Get
IDC	internationale statistische Klassifikation der Krankheiten und verwandter Gesundheitsprobleme
KOM	Fachgebiet Multimedia Kommunikation der TU Darmstadt
ELWMS	E-Learning-Wissensmanagement-System
DOT	DOT Language
SPARQL	SPARQL Protocol and RDF Query Language

Kurzfassung

In dieser Arbeit wird eine Konzeption vorgestellt, die das persönliche und soziale Umfeld bei der Wissensarbeit in seinen Facetten erfasst und in Relation zu den technischen Veränderungen des Socio-Semantic-Web detailliert betrachtet. Aus dieser theoretischen Auseinandersetzung heraus wurde eine webbasierte leichtgewichtige Design- und Service Architektur entwickelt, die zeitgemäß, frei und flexibel in einem breiten Anwendungskontext zum Einsatz kommen kann und eine Schlüsselkomponente in der Erschließung von lokalem und globalem Wissen einnimmt. Durch die konsequente Ausrichtung auf die Bedürfnisse der Nutzer bei der Wissensarbeit und die Einhaltung von Prinzipien des Interface Design sowie Formalismen der Wissensrepräsentation wurde ein Ansatz verfolgt, der dem Wissensarbeiter die Möglichkeit bietet, sich frei, partizipativ, effizient und nachhaltig an der Entwicklung der Vision Socio-Semantic-Web zu beteiligen.

Keywords: Wissensarbeit, Tagging, Semantic Web, Text-Editor

Einleitung

1| Ausgangslage und Methodik

Die rasche Entwicklung der Informationsgesellschaft resultiert aus der fortlaufenden Beseitigung technischer Barrieren. Aktuelle Hindernisse beziehen sich dabei auf die universelle Verfügbarkeit von und dem Zugriff auf Information. Einer der wichtigsten Aspekte hierbei ist, gezielt semantische Zusammenhänge zu nutzen, um Information in einer maschinenzugänglichen Form zu integrieren.

1.1 Globale Probleme

Das Internet wurde 1989 von Tim Berners-Lee entwickelt, um einen neuartigen Austausch wissenschaftlicher Dokumente zu ermöglichen. Es ging darum, den Bedarf an Navigations- und Suchwerkzeugen zu unterstützen und die Idee der Hypertextfunktionalität zu integrieren, sodass Dokumente Verweise auf beliebige andere Dokumente enthalten konnten. Mit der Hypertext Markup Language (HTML) steht dem Nutzer ein einfaches Werkzeug zur Hand, einen begrenzten Sprachkern zu definieren. Das Internet wächst kontinuierlich zu einer breiten Sammlung von Dokumenten heran, die mit Einschränkungen allgemein zugänglich sind. Mit diesem Ursprung kann das World Wide Web (www) als gigantische, globale und vernetzte Bibliothek verstanden werden, da es sich real nur durch das Fehlen einer Instanz, die alle Information zentral verwaltet, sammelt, erschließt und verfügbar macht, von einer Bibliothek unterscheidet.

Das Internet ist einfach, offen, verteilt, dezentral und heterogen; diese freie Struktur ist entscheidend für die Beliebtheit. Das Web hat jedoch auch seine Schwächen: Es ist zwar nicht kompliziert, mit HTML (aktuelle Spezifikation HTML 5, 29. Juli 2009) Web-Dokumente zu erstellen, jedoch eignet sich HTML nur begrenzt für den direkten Zugriff auf spezifische Inhalte. Zudem ist eine standardisierte Erschließung aufgrund der Vielfalt an Formaten und Inhalten eine große Herausforderung. Viele Versuche allgemein anerkannte Standards zu entwickeln wurden nicht wahrgenommen oder konnten nicht mit der rasanten Entwicklung des Netzes mithalten.

Für die semantische Auszeichnung von Dokumenten im Internet ist die 1999 veröffentlichte Spezifikation Resource Description Framework (RDF) und die darauf aufbauende Web Ontology Language (OWL) bis heute von

besonderer Bedeutung. Im Jahre 2004 wurden beide Spezifikationen als Empfehlung des World Wide Web Konsortiums (W3C) mit dem Zweck

„...to publish and share sets of terms called ontologies, supporting advanced Web search, software agents and knowledge management“
[Lee, 1998]

veröffentlicht. Diese Strukturcharakteristik hat Einfluss auf die aktuelle Wissensordnung, was dazu führt, dass die Relevanz eines Dokuments zunehmend über den Inhalt selbst, sowie über seine Konnektivität, basierend auf einheitlichen semantischen Beschreibungen, ermittelt wird.

Das Hauptproblem dabei ist, dass es zu wenig standardkonforme Metadaten gibt, wie beispielsweise das Exchangeable Image File Format (EXIF) oder Datenformate verwendet werden, die sich semantisch nur schwer erklären lassen, wie Office-, PDF- oder Video-Dateien, und schließlich bestehende Standards syntaktisch falsch genutzt und somit unbrauchbar werden. Letzteres richtig zu machen erfordert, neben dem technischen Fachwissen, mindestens gute Kenntnisse auf dem Gebiet der formalen und inhaltlichen Erschließung. Im Rahmen der Formalerschließung (Descriptive Cataloguing) geht es nur um die formalen Metadaten eines Dokuments, dem durch die Beschreibung normierte Zugangspunkte zugewiesen werden (beispielsweise Dublin Core), bei der Sacherschließung (Subject Cataloguing) hingegen wird ein Dokument auf Grund seines Inhaltes durch Schlagwörter, Deskriptoren, Klassifikationen und Ontologien beschrieben. Der Kontext spielt dabei eine ausschlaggebende Rolle.

Mit dieser Form der komplexen Dokumentation sind herkömmliche Internetnutzer in der Regel völlig überfordert. Dennoch haben sich in den letzten Jahren durch die Entwicklung von Web 2.0-Technologien Möglichkeiten entwickelt, die jedermann die semantische Beschreibung von Web-Dokumenten gestatten. Die grundlegende Entwicklung in diesem Bereich ist das Prinzip des „Taggings“, bei dem statt dem kontrollierten Vokabular einer spezifischen Klassifikation freie Schlagwörter (Tags) verwendet werden können. Durch Zusammenarbeit und Partizipation entsteht dann eine implizite Klassifikation, die als „Folksonomy“ bezeichnet wird und über eine reine Erschließung hinaus geht.

Diese Art der kollaborativen Schlagwort-Systeme werden von den Web-Nutzern weitgehend als nützliche und leistungsfähige Werkzeuge zum Organisieren, Suchen und Veröffentlichen persönlicher Ressourcen angenommen [Quintarelli, 2006] und haben eine grundlegende Veränderung in der gesellschaftlichen Wahrnehmung und dem Selbstverständnis mit sich gebracht. Es ist davon auszugehen, dass

„die zukünftige Entwicklung semantischer Anwendungen immer mehr von vernetzten, kontextabhängigen Ontologien geprägt sein wird“,
[Studer, 2008a]

basierend auf Ansätzen der Informations- und Applikationsintegration, der semantischen Suche und der Weiterentwicklung semantischer Web-Services, sowie unter Nutzung von Synergien und Kollaborationspotentialen. Immer wichtiger wird es also auch für bibliographische Dienste, Fachportale und Kataloge semantisch erschließbar, für die Nutzer handhabbar und für Maschinen lesbar zu machen. Aus dieser Motivation heraus soll eine Konzeption entwickelt werden, die flexibel in verschiedenen Bereichen eingesetzt werden kann.

1.2 Aufgabenstellung

„Die Arbeitswelt verschiebt sich immer weiter in Richtung informationsgestützter Tätigkeit“ [Baumeister, 2007],

während im Jahr 2000 der Anteil der Wissensarbeiter bereits bei 62% lag, wird bis 2020 eine Zunahme auf über 75% erwartet. Das Aufgabenspektrum der Wissensarbeiter erweitert sich vom Informationsverbraucher hin zum Informationsproduzenten. Organisationsübergreifend spielt Dokumentation und Vermittlung von Information eine immer wichtigere Rolle.

Ein Wissenschaftler beispielweise, der in ein Forschungsprojekt involviert ist, welches sich über ein kollaboratives Werkzeug organisiert, nutzt für seine wissenschaftliche Arbeit u.a. den Online Katalog der fachspezifischen Institutsbibliothek. Man kann davon ausgehen, dass er darüber hinaus noch weitere verteilte Informationsquellen nutzt, die mehr oder weniger komplex aufgebaut sind. Der Web-Browser ist zur Universalschnittstelle für individuell abgestimmte Wissensdienste (Knowledge Services) geworden. In diesem Szenario erlebt jeder den Prozess der Informationsverarbeitung unterschiedlich und entwickelt eine persönliche Denkfähigkeit, eigene Kriterien der Bewertung und Wahrnehmungspräferenzen. So kann er Ressourcen managen, kontinuierlich revidieren, optimieren und daraus Kompetenzen entwickeln.

Es ist wichtig, diese Unterschiedlichkeit zuzulassen und eine Basis zu schaffen, auf der gute Kommunikation stattfinden kann, sodass Denkmodelle

und individuelle Konzepte wechselseitig ausgetauscht werden können. Denn implizites und explizites Feedback sind elementare Kommunikationsbausteine und ein Instrument zur Beurteilung von Inhalten, Darstellung neuer Perspektiven und Unterstützung des sozialen Lernens [Bierhoff, 2006]. Social Software und Semantische Technologien machen das agile Management von Wissen möglich, unterstützen die persönliche Wissensarbeit und die Kommunikation, die in organisatorische Arbeitsprozesse eingebettet ist. Web 2.0 Methoden erlauben Anwendungen mit niedrigen Einstiegschwellen, während semantische Technologien die notwendige Intelligenz liefern [Studer, 2008b].

Die Aufgabenstellung dieser Masterarbeit basiert auf Trends und Aktivitäten in den Forschungsbereichen des Web 2.0 im Übergang zum Web 3.0 und nutzt Erkenntnisse und Potentiale, um daraus eine benutzerfreundliche soziale Infrastruktur für ein partizipatives verteiltes Schlagwortsystem zu konzipieren. Der Nutzer soll die Möglichkeit haben, seine persönliche Wissenskollektion selbständig zu strukturieren, effizient zu nutzen und nach Bedarf mit anderen zu teilen. Mit dem eigenen Vokabular soll der Nutzer in der Lage sein, seine Wissenswelt zu verwalten und relevante Textteile mit einer inhaltlichen Bedeutung zu versehen, während er Zugriff auf valide Ontologien von Experten und auf implizite Klassifikationen anderer Nutzer hat.

Ziel der Arbeit ist es ein Konzept zu entwickeln, das ein individuelles Wissensmanagement unterstützt und durch Kollaborationsfunktionen Wissen verteilt, neues Wissen erzeugt, sowie bestehende Wissensbestände anreichert. Intelligente Filtertechniken und Retrieval-Operationen sollen potentiell bessere Ergebnisse aggregieren und den qualitativen Zugriff auf Ressourcen ermöglichen. Erreicht werden soll dies mit geeigneten Interface-Design-Techniken und einer optimierten Informationsarchitektur, die es dem Nutzer ermöglicht, intuitive Zugriffs- und Nutzungsmöglichkeiten zu entwickeln.

Zum Arbeitspaket der Masterarbeit gehören insbesondere folgende Meilensteine:

- Analyse von Grundlagen und Konzepten rund um die Schnittmenge zwischen persönlichem Wissensmanagement, Web 2.0 Trends und Web 3.0 Architekturen
- Ermittlung des Angebotscharakters (Affordance) von wissensbasierten kollaborativen Systemen durch empirische Methoden und native Alltagsbeobachtung

- Konzeptionelle Entwicklung einer webbasierten Benutzeroberfläche und sozialen Informationsarchitektur für die Erweiterung klassischer Wissensmodelle durch freies Tagging
- Prototypische Implementierung der Konzeption und Machbarkeitsprüfung auf Integration in ein ontologiebasiertes Informationssystem
- Szenariobasierte Validierung des Prototypen durch die Zielgruppe
- Diskussion des Konzepts mit Blick in die Evaluationsergebnisse und Ausblick respektive Auswirkungen auf das kollaborative Informations-Management und offene Forschungsfragen.

1.3 Social Web und Semantic Web

Die Idee des Web 2.0 entstand aus einem Brainstorming zwischen O'Reilly und MediaLive International und hält als eine der wirklichen Erfolgsgeschichten an. So lieferte Google im Juli 2009 zu diesem Begriff über 550 Millionen Treffer. Die Grundidee des Konzepts Web 2.0 ist die Infrastruktur eines für die Gemeinschaft offenen Informationsnetzes. Es ist ein organisches Gebilde, das durch die Interessen und den Einsatz der Webnutzer stetig wächst und für diese von großem Nutzen ist.

Während es für den einen keinen Unterschied zwischen dem Web 1.0 und dem Web 2.0 gibt [Lee, 2006] oder der Begriff nur ein Marketing-Schlagwort darstellt, ist es für andere

„...the new conventional wisdom“
[O'Reilly, 2005].

Tim O'Reilly prägte und dokumentierte den Begriff, den er mit Hilfe einiger Praktiken und Prinzipien beschrieb. Dabei geht es immer um die Beteiligung der Nutzer und um die kollektive Macht vieler kleiner Informationseinheiten (Chunks), die den Hauptbestandteil des Web 2.0 bilden. Die folgende Grafik im Web 2.0-Stil (Abb. 1) verdeutlicht die Komplexität des Themas und visualisiert die wichtigsten Buzzwörter rund um den Begriff in einer Schlagwortwolke (Tag Cloud).

In dieser Wolkenkarte werden die Schwerpunkte

- Standardisierung (standardization)
- Weiterverwendbarkeit (remixability)

machen, mit anderen zu kommunizieren, zu kollaborieren und schließlich soziale Netzwerke aufzubauen und zu pflegen. Bevor das Web 2.0 zur Normalität werden kann, beginnt die Ära

„Social Web meets Semantic Web“.
[Gruber, 2006]

In den Vordergrund der Betrachtung rücken intelligente Kollektionen, sowie die Komposition, Integration und Aggregation von Inhalten. Diese basieren auf den Funktionsprinzipien der sozialen Interaktion und vielschichtigen Netzstrukturen. Dabei geht es um eine Mischung aus strukturierten, maschinenlesbaren Daten und unstrukturierten Daten aus menschlichem Input. Das Semantische Web (Semantic Web) bietet einen gemeinsamen Rahmen, der es möglich macht, Daten gemeinschaftlich, plattformunabhängig und unternehmensübergreifend zu teilen und wiederzuverwenden. Relevant sind dabei vor allem das implizite Wissen der Benutzer, sowie die erfolgreiche Informationsintegration und -Kombination aus verschiedensten Quellen.

“Properly designed, the Semantic Web can assist the evolution of human knowledge as a whole”
[Lee, 2001].

1.4 Aufbau der Arbeit

Der Erfolg webbasierter interaktiver Anwendungen ist abhängig von der Benutzerfreundlichkeit (Ease of Use) und dem Nutzererlebnis (User Experience) der Anwender. Es gibt zwar kein sicheres Rezept, jedoch wirksame, bewährte Vorgehensweisen, die miteinander kombiniert fundamental für die Anwendungsentwicklung sind. Der Kern dieser Arbeit stellt eine Konzeption dar, die aus oben genanntem Ursprung heraus zentrale benutzerspezifische Aspekte aufgreift und diese für die Entwicklung einer Applikation innerhalb eines komplexen Szenarios nutzt.

In dieser Arbeit wird der konzeptionelle Entwurf, die Gestaltung und prototypische Umsetzung wie Evaluierung einer semantischen Benutzerschnittstelle beschrieben und diskutiert. Die Thematik gliedert sich in vier Kapitel, die im Folgenden kurz dargestellt werden.

Kapitel 1 bietet eine Einführung in die vorherrschende Situation bezogen auf globale Probleme bei der Entwicklung des Internets zu einem semantischen Netz. Mit dieser Einleitung wird die Komplexität des Themas dargestellt, Motive für die Aufgabenstellung abgeleitet und ein Lösungsprozess skizziert, der in den nächsten Kapiteln ausgearbeitet wird.

In Kapitel 2 werden in einer umfassenden Analyse Grundlagen und Aspekte herausgearbeitet, die für die Wissensarbeit im Umfeld des Social Semantic Web von Relevanz sind und Rückschlüsse auf die Anforderungen und Bedürfnisse der spezifischen Benutzergruppe von semantischen Web-Systemen erlauben. Dabei richtet sich der Fokus auf kognitive Prozesse des Wissensarbeiters in seinem Umgebungskontext, die analog zu den technischen Auswirkungen webbasierter, sozialer, semantischer Wissensinfrastrukturen und Schlüsseltechnologien betrachtet werden. Wichtige Erfolgsfaktoren und Probleme werden durch ein breites Spektrum an empirischen Untersuchungen ermittelt und belegt.

Das erarbeitete Basiswissen formt zusammen mit den Forschungsergebnissen die Methodik, mit welcher in Kapitel 3 das Design- und Service-Konzept erarbeitet wird. Der Schwerpunkt bei dieser Studie richtet sich gezielt auf die benutzerspezifischen Anforderungen und das Nutzererlebnis im Rahmen eines definierten Anwendungskontexts. Unter Einhaltung von Designprinzipien und Grundsätzen der Mensch-Maschine-Interaktion wird die aufgabenorientierte prozedurale Interaktion des Wissensarbeiters in einem angemessenen Dialogdesign zum Ausdruck gebracht. Die intelligente Systemunterstützung auf Basis der semantischen Verständigung (Semantic Agreement) zwischen heterogenen Quellen lässt den Wissensarbeiter die Komplexität der Anwendung nicht spüren und gewährleistet eine effiziente Interaktion, während die Potentiale der Architektur beliebig erweiterbar bleiben.

In einem weiteren Schritt wird die Konzeption prototypisch implementiert und eine technische Machbarkeit in einer realitätsnahen Umgebung überprüft. In diesem frühen Entwicklungsstadium findet zudem eine Evaluation statt, die aufzeigt, wie sich die Akzeptanz der Benutzer gegenüber dem Werkzeug und seiner Funktionalität darstellt.

Im letzten Kapitel wird die Arbeit mit einer Diskussion abgeschlossen. Kritisch werden Ergebnisse der Evaluation aufgezeigt und ein Ausblick gegeben.

Im Anhang sind darüber hinaus alle ergänzenden Test-Dokumente aufgeführt, sowie der Prototyp auf einer CD-Rom beigelegt.