

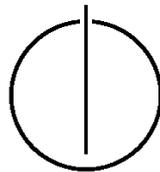
FAKULTÄT FÜR INFORMATIK

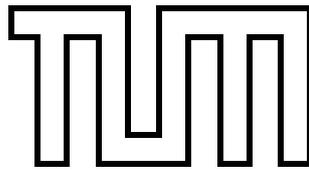
DER TECHNISCHEN UNIVERSITÄT MÜNCHEN

Bachelorarbeit in Wirtschaftsinformatik

**Ein Vergleich von Web-basierten Ansätzen
zum Strukturieren von Informationen - eine
Literaturstudie**

Matthias Hunecker





FAKULTÄT FÜR INFORMATIK

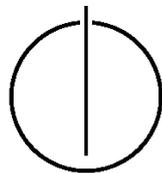
DER TECHNISCHEN UNIVERSITÄT MÜNCHEN

Bachelorarbeit in Wirtschaftsinformatik

Ein Vergleich von Web-basierten Ansätzen zum
Strukturieren von Informationen - eine Literaturstudie

A Comparison of Web-based Approaches for Structuring
Information - A Literature Analysis

Bearbeiter: Matthias Hunecker
Aufgabensteller: Prof. Dr. Florian Matthes
Betreuer: Dipl.Inf. Christian Neubert
Datum: 18. November 2011



Ich versichere, dass ich diese Bachelorarbeit selbständig verfasst und nur die angegebenen Quellen und Hilfsmittel verwendet habe.

München, 18. November 2011

Matthias Hunecker

Kurzbeschreibung

Trends wie Web 2.0 und in der Folge Enterprise 2.0 haben dazu geführt, dass Social Software wie Wikis oder Blogs neben dem privaten Einsatz auch zunehmende Bedeutung im Bereich des Informations- und Wissensmanagements in Unternehmen gewinnt. Wikis zeichnen sich vor allem durch ihre Einfachheit und Offenheit aus. Die Kehrseite der Medaille ist die fehlende Möglichkeit zur erweiterten Strukturierung von Information. Gerade im organisatorischen Kontext ist dies ein großes Manko, da es nicht möglich ist, Daten nach einem einheitlichen Schema zu beschreiben, zu filtern und zu exportieren. Eine intensive Nutzung unter diesen suboptimalen Voraussetzungen führt zur Bildung unübersichtlicher, heterogener Inhalte und letztendlich zu zusätzlichen Problemen bei der Navigation und Suche.

Methoden wie das Tagging, Konzepte wie das Templating, Visionen wie das Web of Data sowie Web-Standards wie RDF(S) oder OWL erlauben die Erstellung reichhaltiger Strukturen, die eine detaillierte Beschreibung und Verknüpfung von arbiträren Objekten der realen Welt, bis hin auf Bedeutungsebene, zulassen. Viele Tools greifen in Hinblick auf die angedeutete Problematik inzwischen auf diese Möglichkeiten zurück. Das Kategorisieren von Inhalten mit einfachen Schlagwörtern wird erweitert auf die Ontologieentwicklung, um leichtgewichtige, domänenspezifische Ontologien zu erstellen, zu optimieren und wiederzuverwenden.

Ein Ansatz, der auf die emergente Entwicklung von strukturierter Information setzt, ist Hybrid Wikis, eine am Software Engineering für betriebliche Anwendungen (sebis) Lehrstuhl der Technischen Universität München (TUM) seit 2009 entwickelte Erweiterung für das Wiki der Enterprise Collaboration-Plattform Tricia. Unter dessen Einbezug wird ein Vergleich mit weiteren Ansätzen zur Web-basierten Strukturierung von Informationen, die über eine Literaturstudie unter methodischer Vorgehensweise gesammelt wurden, durchgeführt. Diese werden analysiert und anhand ihrer festgestellten Konzepte und Methoden eingeordnet. Die gewonnenen Erkenntnisse werden im Anschluss dazu verwendet, um Gestaltungsempfehlungen für zukünftige Erweiterungen von Hybrid Wikis zu geben.

Schlagwörter: *Social Software, Wiki, Strukturierung von Inhalten, Web Kollaboration, Literaturstudie, Literaturanalyse, Vergleich, Hybrid Wikis, Gestaltungsempfehlungen*

Abstract

Trends like Web 2.0 and in consequence Enterprise 2.0 lead to the fact that Social Software like wikis or blogs recently gained importance beyond private usage in terms of information and knowledge management in organizations. Wikis are well-known for their openness and simpleness. In contrast, they lack possibilities for the structured representation and modelling of information. Especially in the enterprise area this is a big downside. It becomes impossible to describe, filter and export data following a certain schema. Intense use under these circumstances leads to the creation of noisy, heterogeneous contents and to additional problems regarding navigation and search.

Methods like tagging, concepts like templating, vision like the Web of Data and Web-standards like RDF(S) and OWL allow the construction of rich structures that make detailed descriptions and connections of arbitrary objects of the real world up to the level of meaning possible. Lots of tools utilize those to solve the outlined problems. The categorization of contents with the use of simple keywords is extended to the so called ontology engineering to create, optimize and reuse light-weight, domain-specific ontologies.

An approach that focuses on the emergent creation of structured information is Hybrid Wikis. Developed at the sebis chair of TUM since 2009, it represents an extension for the wiki of the enterprise collaboration platform Tricia, which was also created there. Hybrid Wikis gets compared to similar approaches that allow Web-based structuring of informations which were collected following the methods of a structured literature review. They get analyzed and sorted in according to their recognized concepts and methods. Additionally, the collected facts are used to provide design recommendations for future extensions of Hybrid Wikis.

Keywords: *Social Software, Wiki, Structured Information, Structuring of Content, Web Collaboration, Literature Review, Literature Analysis, Comparison, Hybrid Wikis, Design Recommendations*

Inhaltsverzeichnis

Kurzbeschreibung	vii
Abstract	ix
Abkürzungsverzeichnis	xv
Abbildungsverzeichnis	xvii
Tabellenverzeichnis	xix
1. Einleitung	1
1.1. Motivation	1
1.2. Forschungsfragen	2
1.3. Vorgehensweise	2
2. Grundlagen - die Entwicklung des Webs	5
2.1. Web 2.0	5
2.1.1. Ursprung	5
2.1.2. Definition	5
2.1.3. Tags und Folksonomien	7
2.1.4. Wikis	8
2.1.4.1. Ursprung	9
2.1.4.2. Die Wiki-Prinzipien	9
2.1.4.3. Templates in Wikis	11
2.2. Enterprise 2.0	12
2.2.1. Ursprung	12
2.2.2. Enterprise Wikis	13
2.3. Web 3.0	14
2.3.1. Ursprung	14
2.3.2. Semantic Web und semantische Technologien	14
2.3.3. Über Linked Data zum Web of Data	14
2.3.4. Ontologien	17
2.3.5. Semantische Wikis	18
2.3.5.1. Beispiel: semantische Annotation mit Semantic MediaWiki	19
2.3.6. Manuelle, unterstützte und automatisierte Strukturierung	20
2.4. Kollaborative Informationsstrukturierung mit Hybrid Wikis	20
2.4.1. Struktursystem	21
2.4.2. Eingabe	22
2.4.3. Darstellung, Navigation und Suche	23

2.4.4. Zusammenfassung	25
3. Literaturstudie	27
3.1. Eingrenzung	27
3.2. Vorgehensweise	28
3.3. Verwendete Möglichkeiten der Recherche	28
3.4. Ergebnisse der Suche	29
3.5. Beurteilung der Recherche und Eingrenzung der Auswahl	31
4. Analyse und Einordnung	33
4.1. Struktursystem	33
4.1.1. Extrahierte Eigenschaften	33
4.2. Eingabe	37
4.2.1. Extrahierte Eigenschaften	37
4.3. Darstellung, Navigation und Suche	40
4.3.1. Extrahierte Eigenschaften	40
4.4. Beurteilung von Hybrid Wikis im Vergleich	42
5. Gestaltungsempfehlungen für Hybrid Wikis	45
5.1. Struktursystem	45
5.1.1. Erweiterte Beschreibung von Link-Relationen	45
5.1.2. Gruppierte Attribute	46
5.1.3. Verschachtelte Attribute	47
5.1.4. Vorlagen für Wiki-Seiten	47
5.1.5. Vererbung auf Typ Tag-Ebene	48
5.1.6. Semantische Funktionalisierung von Attribut-Schlüsseln	48
5.1.7. Zugriffskontrolle auf Datenmodell-Ebene	49
5.2. Eingabe	50
5.2.1. Einfache Inferenzen für Typ Tags	50
5.2.2. Konsistenzprüfung während der Eingabe	51
5.2.3. Manuelle Markierung redundanter Inhalte	51
5.2.4. Kommentarfunktion für Belegungen von Attributen	52
5.2.5. Automatisierte Bildung von Einschränkungen für Datentypen	52
5.2.6. Import von Ontologien im RDF(S)/OWL-Format	53
5.3. Darstellung, Navigation und Suche	54
5.3.1. Vergleichssicht, Attribut-Sicht und kontextspezifische Visualisierung	54
5.3.2. Einbettung der Attribute von Zielseiten	55
5.3.3. Ausbau der Suche	56
6. Fazit und Ausblick	57
6.1. Zusammenfassung und Fazit	57
6.2. Zukünftige Forschungsfragen	57
Anhang	61
A. Einordnung: Struktursystem	61

B. Einordnung: Eingabe	67
C. Einordnung: Darstellung, Navigation und Suche	71
Literaturverzeichnis	75

Abkürzungsverzeichnis

ACE	Attempto Controlled English
AJAX	Asynchronous JavaScript and XML
API	Application Programming Interface
CRM	Customer Relationship Management
DTD	Document Type Definition
ERM	Entity Relationship Model
GUI	Graphical User Interface
EMF	Eclipse Model Framework
HR	Human Resource
HTML	Hyper Text Markup Language
HTTP	Hypertext Transfer Protocol
IP	Internet Protocol
KM	Knowledge Management
LACS	Lightweight Access Control Schema
LOD	Linked Open Data
ML	Markup Language
NER	Named Entity Recognition
NLP	Natural Language Processing
OWL	Web Ontology Language
RDF	Resource Description Framework
RDFS	RDF Vocabulary Definition Language
REST	Representational State Transfer
RSS	Really Simple Syndication
sebis	Software Engineering für betriebliche Anwendungen
SPARQL	SPARQL Protocol And RDF Query Language
SQL	Structured Query Language
SWRL	Semantic Web Rule Language
TUM	Technische Universität München
UGC	User Generated Content
UML	Unified Modeling Language
URI	Uniform Resource Identifier
URL	Uniform Resource Location
W3C	World Wide Web Consortium
WYSIWYG	What You See Is What You Get
XML	Extensible Markup Language

Abbildungsverzeichnis

1.1. Gegenmaßnahmen zu Wucherungen (engl. <i>proliferation</i>) in Wikis [HT08] . . .	2
2.1. Ein mit Tags annotiertes Bookmark (Del.icio.us)	7
2.2. Bottom-up Folksonomie (Del.icio.us - links) versus top-down Taxonomie (Yahoo Directory - rechts)	9
2.3. Entwicklung der Anzahl an englischen Artikeln auf Wikipedia (2001 bis heute) [WiXa]	10
2.4. Beispiel für den funktionalen Aufruf und die Darstellung einer Templateinstanz (MediaWiki) [Aum05]	12
2.5. Subjekt, Prädikat und Objekt als gerichteter Graph in Resource Description Framework (RDF) [RDF04c]	15
2.6. Schichtenmodell der Standards des Semantic Web (W3C) [Se007]	16
2.7. Semantische „Treppe“ [BT06]	17
2.8. Aus semantischer Annotation gesammelte Fakten des Wiki-Artikels über „Berlin“ in Semantic MediaWiki	19
2.9. Das Datenmodell von Hybrid Wikis [MNS11]	22
2.10. „Hybride“ Sicht auf eine Wiki-Seite in Hybrid Wikis/Tricia	24
2.11. Typ Tag-Sicht in Hybrid Wikis/Tricia	25
2.12. Sicht für Wiki-Seiten mit dem gleichen Typ Tag in Hybrid Wikis/Tricia . . .	26
2.13. Per Drill Down eingegrenzte Suche in Hybrid Wikis/Tricia	26
4.1. Der Unterschied zwischen implizitem und explizitem Schema (Eigene Darstellung)	36
4.2. Zusammenhang zwischen Freitext und Strukturteil in Wiki-Seiten (Eigene Darstellung)	37
4.3. Syntax-basierte Eingabe über kontrollierte natürliche Sprache in AceWiki [Kuh08b] (links oben), einer eigenen Notation in Kaukolu [Kie06] (links unten) und über eine einfache Doppelpunkt-Notation in WikSAR (SHAWN) [Aum05] (rechts)	38
4.4. Über kontrollierte natürliche Sprache formulierte Suchanfrage in AceWiki [Kuh08b]	42
5.1. Assoziationsattribute zweier Typ Tags in UML (Eigene Darstellung)	45
5.2. Konzeptuelles Mock-Up für Link-Attribute und optional angeheftete Typ Tags (Eigene Darstellung)	45
5.3. Templateinstanz von „Calcium“ mit gruppierten Eigenschaften in Wikipedia 46	
5.4. Konzeptuelles Mock-Up für verschachtelte Attribute in Form von Records (Eigene Darstellung)	47
5.5. Subtypen vom Typ „Business“ in Freebase	49

5.6. Festlegung von Synonymen und hierarchischen Aspekten eines Konzepts in MoKi [RGP ⁺ 09]	50
5.7. Statements zu Belegungen von Attributen in OntoWiki [ADR06]	52
5.8. Empfehlungen (Recommendations) in SnoopyDB [GZS11]	53
5.9. Vergleichssicht (Comparison View) in Social Infobox [HGT11]	54
5.10. Monatsansicht in HYENA [Rau10] (links) und Kartenansicht in OntoWiki [ADR06] (rechts)	55
5.11. Ausklappbare Attribute der Zielressource einer Link-Relation (Inline Reference View) in OntoWiki [ADR06]	56

Tabellenverzeichnis

3.1. Suchmöglichkeiten für die Schlagwortsuche	28
3.2. Ausgewählte Ansätze aus der Literaturstudie	31
A.1. Einordnung: Struktursystem 1	62
A.2. Einordnung: Struktursystem 2	63
A.3. Einordnung: Struktursystem 3	64
A.4. Einordnung: Struktursystem 4	65
B.1. Einordnung: Eingabe 1	68
B.2. Einordnung: Eingabe 2	69
C.1. Einordnung: Darstellung, Navigation und Suche 1	72
C.2. Einordnung: Darstellung, Navigation und Suche 2	73

1. Einleitung

Social Software wie Wikis findet nicht zuletzt dank des Erfolges von Wikipedia in Bezug auf das kollaborative Wissensmanagement zunehmend Gebrauch. Wikis haben ihren Ursprung im Web 2.0, in dem der Nutzer von Web Services die Erstellung der Inhalte übernimmt. Einerseits zeichnen sich Wikis durch ihre Offenheit und Einfachheit in Bezug auf die gemeinsame Erstellung von Inhalten aus, andererseits fehlt es vielen Wikis an erweiterten Möglichkeiten zur Strukturierung von Information.

Eine solche Einschränkung in der Mächtigkeit führt in Konsequenz zu einer Zunahme von Redundanzen und Inkonsistenzen. Das Suchen und Filtern von relevanten Inhalten („Nenne mir alle Projekte, die nach dem 15. Juni 2011 gestartet wurden und noch nicht abgeschlossen sind.“) ist über eine Titel- und Volltextsuche nicht möglich. Strukturiert vorliegende Daten (zum Beispiel im Excel-Format) verlieren durch den Import in ein Wiki gänzlich ihre Form. Inhalte werden nur visuell formatiert und wuchern unkontrolliert vor sich hin. Es entstehen auf Dauer sehr heterogene Strukturen, die von Maschinen nicht einheitlich interpretierbar sind.

In [LC01] werden entstehende Probleme und Konflikte durch die intensive Nutzung eines klassischen Wikis bei der New York Times Digital beschrieben. [DBEP⁺08] fanden heraus, dass nur knapp 36 Prozent aller Änderungen in Wikipedia mit dem inhaltlichen Umgestalten von Artikeln einher gingen, der Rest fällt auf das Optimieren struktureller Bestandteile wie Titel und Links, um Ausuferungen einzudämmen.

Jener Zustand stellt ein zunehmendes Problem dar, vor allem im Unternehmensbereich, in dem das effektive Wissensmanagement über Social Software inzwischen ein wichtiger Erfolgsfaktor ist. Aktuelle Studien befassen sich bereits mit den Ursachen, Konsequenzen und der Beschreibung von Gegenmaßnahmen [HT08].

1.1. Motivation

Da Ursachen und Konsequenzen anhand von Befragungen ausgiebig untersucht wurden [HT08], liegt der Fokus momentan auf der Entwicklung von Gegenmaßnahmen. Während der Begriff Gegenmaßnahmen sehr negativ klingt, eröffnet er neue Möglichkeiten zur Steigerung der Qualität des Informationsmanagements über Wikis. Möglichkeiten zur Kontrolle und Strukturierung von Inhalten über Wikis lassen sich prinzipiell in zwei Dimensionen einteilen (siehe Abbildung 1.1). Zum einen wird zwischen manuellen und automatisierten, zum anderen zwischen präemptiv und nachfolgenden Möglichkeiten unterschieden.

Bewährte Ansätze wie das Tagging zur (hierarchischen) Kategorisierung von Inhalten sowie die Wiederverwendung bestehender Konzepte anhand von semi-strukturierten Vorlagen (Beispiel: die Infobox in MediaWiki) oder die Bildung semantischer Verknüpfungen (Beispiel: die Factbox in Semantic MediaWiki) finden ihren Einsatz, um Informationsein-

	Preemptive	Subsequent
Manual	<ul style="list-style-type: none">• Employee training• Quality standards, rules of conduct• Tagging, categorization• Use of templates	<ul style="list-style-type: none">• Wiki Gardening
Automated	<ul style="list-style-type: none">• Rights management• Workspaces	<ul style="list-style-type: none">• Bots, Plug-ins, Scripts• Watchlists, Patrolling of edits

Abbildung 1.1.: Gegenmaßnahmen zu Wucherungen (engl. *proliferation*) in Wikis [HT08]

heiten geordnet zu beschreiben und in Relation zu setzen. Das Semantic Web und der Linked Data- bzw. Web of Data-Ansatz liefern grundlegende Prinzipien und Standards zur Gestaltung.

1.2. Forschungsfragen

Die beschriebenen Probleme könnten unter Einsatz der ersten beispielhaft genannten Konzepte und Methoden gemildert und sogar behoben werden. Aus dieser Motivation heraus ergeben sich diverse Forschungsfragen, die den Leitfaden für diese Bachelorarbeit bilden sollen.

- Wie hat sich das Web über die letzten Jahre hin verändert und welche allgemeinen Konzepte und Methoden zur Strukturierung und Repräsentation von Information in Wikis haben sich entwickelt?
- Welche Umsetzungen auf Wiki oder Wiki-ähnlicher Basis, die die Bildung von Strukturen erlauben und in wissenschaftlicher Literatur beschrieben werden, lassen sich finden?
- Welche konkreten Konzepte und Methoden verwenden die über eine Literaturstudie gesammelten Ansätze und wie lassen sich die beschreibenden Quellen einordnen?
- Wie können zukünftige Erweiterungen von Hybrid Wikis auf Basis der extrahierten Fakten aussehen?

1.3. Vorgehensweise

Am Anfang der Arbeit wird in einem theoretischen Teil auf die Entwicklung des Internets über das letzte Jahrzehnt hinweg eingegangen und entsprechenden Trends benannt und erläutert, um einen Einstieg in die Thematik zu schaffen. Jede „Ära“ des Internets hat eine eigene charakteristische Möglichkeit gefunden, mit der Datenflut umzugehen. Diese werden jeweils beispielhaft erläutert. Auch Wikis haben sich nach ihrer Entstehung weiterentwickelt und sich an neue Techniken und Einsatzgebiete angepasst. Hybrid Wikis wird

als ein erster fortschrittlicher Ansatz zur kollaborativen Strukturierung und Optimierung von Information vorgestellt.

Anschließend wird eine Literaturrecherche zum Thema (die Methodik beruht auf der Beschreibung in [OS10]) getätigt. Hier werden Vorgehensweise, Eingrenzungen und verwendete Möglichkeiten der Recherche erläutert, die Ergebnisse präsentiert als auch eine Bewertung der Studie sowie der Auswahl der Ergebnisse vorgenommen.

Nachdem die in einem ersten Schritt über eine Schlagwortsuche gesammelte und katalogisierte Literatur analysiert wurde, wird eine Einordnung der Quellen über die gefundenen Konzepte und Methoden durchgeführt, die extrahierten Eigenschaften und Erkenntnisse zusammenfassend beschrieben sowie eine Bewertung von Hybrid Wikis auf Basis des Vergleichs vorgenommen.

Die gewonnenen Erkenntnisse dienen zur Formulierung von Gestaltungsempfehlungen für Hybrid Wikis und es wird zum Abschluss die getätigte Arbeit zusammengefasst und vorausschauend auf zukünftige Forschungsfragen, die auf den Erkenntnissen dieser Arbeit aufbauen können, eingegangen.

2. Grundlagen - die Entwicklung des Webs

2.1. Web 2.0

2.1.1. Ursprung

Der Begriff Web 2.0 fand seine erste Erwähnung im Dezember 2003. In seinem Artikel „2004: The Year of Web Services“ spricht der Autor Erich Knorr von „[...] the Web 2.0, where the Web becomes a universal, standards-based integration platform. Web 1.0 (HTTP, TCP/IP and HTML) is the core of enterprise infrastructure.“ [Kno03]. Tim O'Reilly greift den Begriff Web 2.0 auf und versucht in seiner Veröffentlichung „What Is Web 2.0: Design Patterns and Business Models for the Next Generation of Software“ zu erklären, was das „neue“ Web vom alten unterscheidet. Außerdem gründet er auf Basis seiner gewonnenen Erkenntnis die Web 2.0 Konferenz. Seine Definition ähnelt der von Knorr. O'Reilly spricht u. a. von „The web as platform“ [O'R05]. Bereits eineinhalb Jahre nach der Beschreibung durch O'Reilly erzielte das neu geschaffene Buzzword 9,5 Millionen Suchtreffer bei Google [O'R05].

2.1.2. Definition

Eine genaue Definition des Begriffs Web 2.0 erweist sich als äußerst schwierig, zumal er sehr kontrovers diskutiert wird. Das Schlagwort beschreibt einerseits eine Reihe von neuen Technologien und konkreten Anwendungen, andererseits aber auch den Wandel der Verhaltensmuster von Internetnutzern und deren Rolle, die das Web bereits Anfang des 21. Jahrhunderts eingeschlagen hat [CHK07]. Inwieweit die Vergabe einer Versionsnummer Sinn macht, ist fraglich, da kein Vergleich zu einer klassischen Software gezogen werden kann, deren neue Releases durch ihre Entwickler gekennzeichnet werden. Eine objektive und allgemeingültige Nummerierung des Netzes ist aufgrund der hohen Dynamik und Unübersichtlichkeit faktisch unmöglich. Die Versionsnummer dient wohl der Markierung eines neuen Abschnitts oder einen neuen Ära des Internets, wobei deutlich gemacht werden soll, dass das Web in der Version 2.0 auf dem Vorgänger basiert und auf seinen Techniken aufbaut.

Es folgt eine kleine Auswahl an zufällig selektierten Definitionsversuchen aus dem Netz.

„Web 2.0 ist ein Schlagwort, das für eine Reihe interaktiver und kollaborativer Elemente des Internets, speziell des World Wide Webs, verwendet wird. Hierbei konsumiert der Nutzer nicht nur den Inhalt, er stellt als Prosumer¹ selbst Inhalt zur Verfügung.“ (Wikipedia, [Webb])

¹engl. *prosumer*; Mischform aus *producer* und *consumer*

„the internet viewed as a medium in which interactive experience, in the form of blogs, wikis, forums, etc, plays a more important role than simply accessing information.“ (Dictionary.com, [Webc])

„Bezeichnet ein Kooperationskonzept, das durch neue Technologien möglich wurde. Angebote gehören zum Web 2.0, wenn die Nutzer die Inhalte kommentieren, empfehlen oder mit anderen Angeboten verknüpfen können. Beispiele sind Videoseiten wie YouTube, deren Popularität vor allem auf dem problemlosen Bewerten, Kommentieren und Einbinden der Filme in eigene Web-Seiten beruht.“ (Harvard Business Manager, [Weba])

„Web 2.0 does not refer to a specific version of the Web, but rather a series of technological improvements.

Some examples of features considered to be part of Web 2.0 are listed below:

- Blogs - also known as Web logs, these allow users to post thoughts and updates about their life on the Web.
- Wikis - sites like Wikipedia and others enable users from around the world to add and update online content.
- Social networking - sites like Facebook and MySpace allow users to build and customize their own profile sand communicate with friends.
- Web applications - a broad range of new applications make it possible for users to run programs directly in a Web browser.“ (TechTerms.com, [Webd])

All diese Beschreibungen stimmen in ihren Grundtendenzen überein. Das Web 2.0 setzt auf User Generated Content (UGC) und Social Software, die zur Erstellung, Bearbeitung und möglichst uneingeschränkten Kommunikation der Inhalte dient. Der Nutzer wandelt sich von einer passiven, informationskonsumierenden hin zu einer aktiven, informationsproduzierenden und -verbreitenden Figur [SSM08]. Im Gegenzug hält sich der Bereitsteller eines Angebots zunehmend zurück und beobachtet, um neue Bedürfnisse der Anwender abzuleiten. Diese Entwicklung wird unterstützt durch altbewährte und neue Technik wie Really Simple Syndication (RSS)-Feeds zur Verteilung von Nachrichten auf XML-Ebene oder durch die Verwendung von Asynchronous JavaScript and XML (AJAX), um Webseiteninhalte dynamischer darstellen zu können und eine komfortablere Handhabung zu ermöglichen [O’R05].

Social Software befindet sich durch die ständige Optimierung und Verbesserung in einem nie endenden Betastadium. Ausprägungen sind Blogs (z. B. Wordpress²), Wikis (Wikipedia³), soziale Netzwerke (Facebook⁴, Google+⁵) und eine Reihe weiterer Applikationen wie das Videoportal YouTube⁶ oder der Microblogging-Dienst Twitter⁷.

²<http://de.wordpress.com/>

³<http://www.wikipedia.org/>

⁴<http://www.facebook.com/>

⁵<http://plus.google.com/>

⁶<http://www.youtube.com/>

⁷<http://www.twitter.com/>

2.1.3. Tags und Folksonomien

In einer Bibliothek machen sich Mitarbeiter die Arbeit einer sinnvollen Kategorisierung und Einordnung, im Web 2.0 wird diese Aufgabe an die Produzenten der Daten übergeben. Das sogenannte Collaborative oder Social Tagging erlaubt es dem Nutzer, manuell eigene (oder fremde) Inhalte gemeinsam mit frei wählbaren Tags bzw. Schlagwörtern zu versehen und Metadaten zu schaffen. Dieses Konzept macht dann Sinn, wenn es erstens eine enorme Menge an Inhalten gibt und zweitens keine Autoritätspersonen existieren (z. B. aus Kostengründen), die eine Schema diktieren [AGAH06]. „Getagged“ werden kann alles, was eine eindeutige Uniform Resource Location (URL) besitzt. Zur Festlegung von Tags wird dem Bearbeiter außerdem kein Spezialwissen abverlangt. Er gibt ein, was ihm einfällt.

Es existiert eine Vielzahl an Einsatzfeldern. In Archiven für wissenschaftliche Publikationen (CiteULike⁸), für Bookmarks (Del.icio.us⁹; vgl. Abbildung 2.1), für Fotos (Flickr¹⁰) oder aber auch im Blogging-Bereich (Blogger¹¹) kommen Tags zum Einsatz.



Abbildung 2.1.: Ein mit Tags annotiertes Bookmark (Del.icio.us)

Tags besitzen zwei besonders vorteilhafte Eigenschaften (nach [AGAH06]):

1. Sie bilden keine Hierarchie.
2. Die Ergebnismengen einzelner Tags sind nicht disjunkt.

Konkret bedeutet dies, dass man Inhalte mit beliebig vielen Tags versehen und eine Suche dadurch eingrenzen kann, indem man während der Suche mehrere Tags angibt und die Ergebnismenge eingrenzt. Dies bietet einen gewissen Vorteil in Bezug auf die Freiheit gegenüber klassischer Taxonomien wie man sie aus der Biologie kennt. Taxonomien¹² sind in der Regel streng hierarchisch aufgebaut und exklusiv.

Neben der Einfachheit und Freiheit der Strukturschaffung über Tags gibt es aber laut [AGAH06] auch einige Nachteile.

- Mehrdeutigkeit (Polysemie) von Tags führt mitunter zu ungewollten Suchergebnissen (Suche nach Pferd liefert Ergebnisse zum Tier und zur Schachfigur).
- Synonymie führt zu einer starken Streuung eines zentralen Tags und somit zu Inkonsistenzen (z. B. Verwendung des Wortes Apfelsine statt Orange).

⁸<http://www.citeulike.org/>

⁹<http://www.delicious.com/>

¹⁰<http://www.flickr.com/>

¹¹<http://www.blogger.com/>

¹²Die Anwendung einer Taxonomie zur Klassifizierung von Webseiten findet man u. a. im Yahoo Directory (<http://dir.yahoo.com/>).

- Pluralismen können dazu führen, dass keine Suchergebnisse ausgegeben werden, falls das System zu strikt ist (Suche nach Autos liefert keine Ergebnisse mit dem Tag Auto oder umgekehrt).
- Variation des Basislevels von Tags (ein Katzenliebhaber/-kenner würde den Tag Perser als Basistag für einen Artikel oder Link zu Perserkatzen verwenden, ein durchschnittlicher Tierbesitzer den Tag Katze)
- Sprachen schaffen verschiedene Tags mit der gleichen Bedeutung (black ↔ schwarz) und somit redundante Begriffe.

Homonymie ist nur bedingt ein Problem, da durch zusätzliche Tags während der Suche der Begriff eingegrenzt werden kann.

Das Ergebnis einer Strukturschaffung auf Basis von Tags ist immer eine Mischung aus allgemeingültigen, vom Publikum akzeptierten zentralen und aus sehr spezifischen dezentralen Begriffen. Diese Menge an Tags wird auch Folksonomie genannt. Der eingedeutschte Begriff stammt von dem englischen Wort *folksonomy*, das wiederum aus den Begriffen *folks* und *taxonomy* zusammengesetzt ist. Eine Folksonomie ist also eine Art Taxonomie, die von dem Volk der Internetnutzer in einem bottom-up Prozess geschaffen wird. Sie kann laut [O'R05] sogar als der Nachfolger der klassischen top-down Taxonomie gesehen werden, die den Nutzern durch Administratoren oder Experten eines Systems im Web 1.0 vorgegeben wurde. [Wei05] bekräftigt diese Einschätzung. Starre Taxonomien, die eine Baumstruktur schaffen, werden für die unkontrollierte und vernetzte Welt des Internets als unzureichend bewertet. Das sinnbildliche Zusammenrechnen einzelner Blätter hält er für zutreffender und optimaler.

Folksonomien lassen sich anhand von Tag Clouds (siehe Abbildung 2.2) visualisieren. Meist wird die Popularität bzw. Häufigkeit einzelner Tags dazu verwendet, um sie anhand ihrer Darstellung in Form der Größe voneinander abzugrenzen. Weiterhin werden die Begriffe oft alphabetisch sortiert.

Ein Schritt zur Unterstützung des meist manuellen Tagging-Prozesses sind automatisch abgeleitete Vorschläge auf Basis des vorhandenen Wissens. Diese können mitunter dazu dienen, die angesprochenen Problematiken zu mindern. Entsprechende Mechanismen wurden bereits entwickelt und erfolgreich eingesetzt (z. B. für das Fotoportal Flickr, wie in [SvZ08] beschrieben). Probleme wie Homonymie und Synonymie könnten durch die Verwendung von Thesauren¹³ verringert werden.

Tags lassen sich als ein erster semi-strukturierter Annotationsansatz zur Beschreibung von Inhalten bezeichnen. Neben den einfachen Tags, die zu einer, wie hier beschrieben, flachen Hierarchie führen, gibt es Erweiterungskonzepte wie strukturierte Tags, die bereits aus Eigenschaften-Werte Paare bestehen oder negative Tags, die über ein Ausschlussverfahren Inhalte voneinander abgrenzen [BEKAW09].

2.1.4. Wikis

Da der Fokus der im Hauptteil folgenden Literaturstudie auf Wiki-basierten Ansätzen liegt (siehe Abschnitt 3.1), wird nun eine kurze Einleitung zu Wikis und deren Prinzipien gegeben.

¹³Wortnetze, die Zusammenhänge von einzelnen Begriffen beschreiben

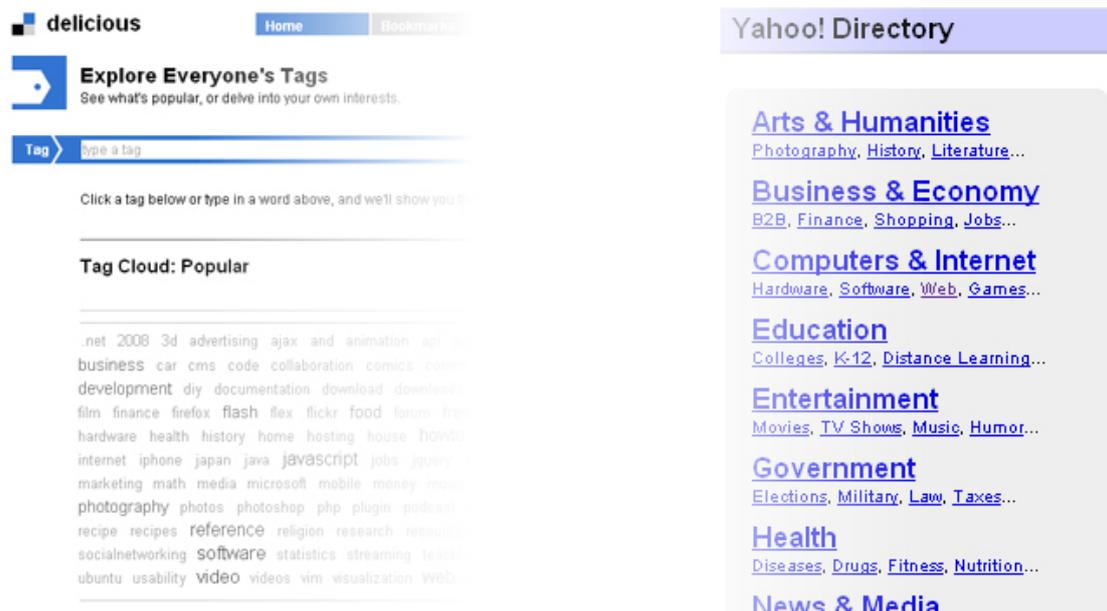


Abbildung 2.2.: Bottom-up Folksonomie (Del.icio.us - links) versus top-down Taxonomie (Yahoo Directory - rechts)

2.1.4.1. Ursprung

In dem Buch „The wiki way: quick collaboration on the web“ [LC01] wird die erste Wiki-Engine namens WikiWikiWeb¹⁴ beschrieben. Sie wurde bereits 1995 von Ward Cunningham für das Portland Pattern Repository¹⁵ entwickelt.

2.1.4.2. Die Wiki-Prinzipien

Cunningham hat auf seiner eigenen Wiki-Seite sämtliche Prinzipien gesammelt, auf deren Basis er das WikiWikiWeb entwickelt hat. Wikis sollten seiner Meinung nach mitunter einfach zu bedienen, offen, inkrementell, evolutiv, einheitlich und beobachtbar sein [Cun06]. Benutzer sollten Inhalte intuitiv und einfach erstellen, bearbeiten und verknüpfen können. Zusätzliche Syntax macht nur dann Sinn, wenn etablierte Markup Language (ML) wie die Hyper Text Markup Language (HTML) die gewünschte Funktionalität noch nicht bieten. Jeder sollte die Möglichkeit haben, suboptimal gegliederte oder inhaltlich falsche Artikel korrigieren zu können. Das Gesamtuniversum des Wikis soll inkrementell und evolutiv wachsen. Links zu noch nicht vorhandenen Seiten unterstützen ein stetiges Wachstum. Einheitliche Layouts sorgen für eine gute Interpretationsmöglichkeit der Inhalte. Überschriften bzw. Seitennamen sollten möglichst präzise sein, damit Konflikte vermieden werden. Änderungen müssen beobachtbar sein.

Essentielle Funktionen eines jeden Wikis sind der Bearbeitungsmodus, die Möglichkeit der internen Verlinkung, die Versionierung und die Suche. Jede Wiki-Seite lässt sich über einen „Bearbeiten“-Button von jedem editieren. In diesem Modus modifiziert man

¹⁴Der Begriff ist abgeleitet von dem hawaiianischen Wort *wiki* und bedeutet übersetzt *schnell*.

¹⁵<http://c2.com/cgi/wiki?PortlandPatternRepository>

2. Grundlagen - die Entwicklung des Webs

meist geringfügig strukturierten Text über einen Editor, der über eine Markup Language bzw. deren Syntax formatiert wird. Einzelne Seiten lassen sich untereinander über ihren Titel verlinken. Im WikiWikiWeb bildet der Titel ein WikiWord. WikiWords folgen dem CamelCase-Prinzip¹⁶, bei dem einzelnen Wörter aus dem Seitentitel zu einem Wort zusammengefügt werden, um Links zu bilden, wobei jedes einzelne Wort mit einem Großbuchstaben beginnt. Ein großer Nachteil dieser Methode ist die schwere Lesbarkeit mancher Titel (Beispiel: „SoftwareEngineeringForBusinessInformationSystems“). Das MediaWiki¹⁷, auf dem Wikipedia basiert, bietet CamelCase-Support. Die Funktion ist aber standardmäßig deaktiviert. Die Versionierung erlaubt eine Verfolgung aller vorangegangenen Änderungen einer Seite. Neben dem „Was“ erfährt man meist auch wer etwas geändert hat (z. B. anhand der Internet Protocol (IP)-Adresse oder einem Accountnamen bei registrierten Nutzern) und wann etwas geändert wurde. Eine Vergleichsmöglichkeit zwischen verschiedenen Versionen ist sinnvoll und auch oft vorhanden. Als Einstieg dient den Nutzern neben der Hauptseite in den meisten Fällen eine Titel- oder Volltextsuche [EG05] .

Enorme Popularität erlangte das „Wiki-Prinzip“ durch die 2001 an den Start gegangene Onlineenzyklopädie Wikipedia¹⁸. Nach rund zehn Jahren existieren in der englischen Ausgabe inzwischen mehr als 3,5 Millionen Einträge. Seit dem Jahr 2005 sind im Schnitt jährlich 500.000 neue Einträge hinzugekommen (siehe Abbildung 2.3).

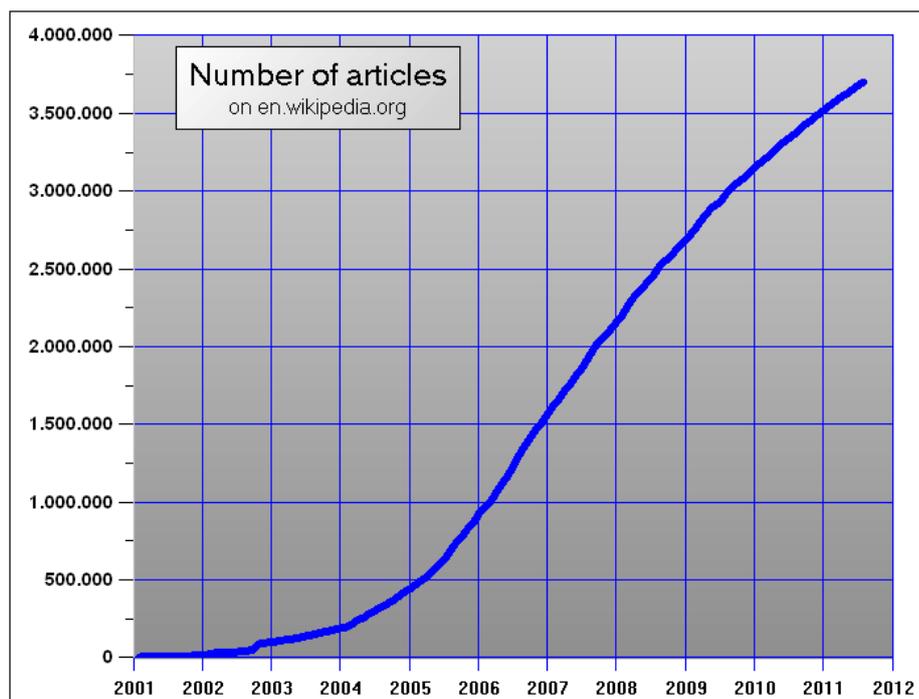


Abbildung 2.3.: Entwicklung der Anzahl an englischen Artikeln auf Wikipedia (2001 bis heute) [WiXa]

Die Einsatzmöglichkeiten von Wikis sind vielfältig: in Unternehmen zur Dokumentation oder Koordination von Projekten, im E-Learning Bereich an Schulen oder Universitäten

¹⁶vgl. <http://en.wikipedia.org/wiki/CamelCase>

¹⁷<http://www.mediawiki.org/wiki/MediaWiki/de>

¹⁸<http://de.wikipedia.org/wiki/Wikipedia:Hauptseite>

und als Wissenslexikon (z. B. in der Biologie oder Chemie).

2.1.4.3. Templates in Wikis

Vielen Wikis mangelt es neben dem Konzept der WikiPage und deren Verlinkungen an Möglichkeiten zur Organisation von Information. Inzwischen bieten viele Wiki-Engines neben der Zuweisung von Kategorien die Möglichkeit der Erstellung von Templates, die sich nach einmaliger Definition beliebig oft wiederverwenden lassen. In Wikipedia werden diese Vorlagen durch Infoboxen repräsentiert. Sie beinhalten ergänzende Information und kommen zumeist bei der Beschreibung von geographischen Objekten wie Städten oder Ländern zum Einsatz, ebenso wie zur biologischen Klassifizierung von Tieren oder bei Personen, um Eckdaten wie Geburtstag, Geburtsort oder Beruf zu sammeln. Sie sind Bestandteil der Formatvorlagen¹⁹ und bilden ein Konstrukt bestehend aus einer Benennung und Attributen, die wiederum über Überschriften untergliedert werden können. Einzelne Attribute sind durch einen String benannt und besitzen in der konkreten Instanz einen Wert der entweder wieder ein String oder ein Verweis auf eine interne Wiki-Seite sein kann. Attribute können in Bezug auf ihre Angabe optional oder verpflichtend sein²⁰.

Während Templates wie die Infobox generische Vorlagen bilden, sind Instanzen konkret verwendete und mit Information befüllte Templates. Es lassen sich außerdem (zunächst) zwei Arten von Vorlagen unterscheiden: Functional und Creational Templates [VZ07].

Functional Templates Ein Functional Template (für ein Beispiel siehe Abbildung 2.4) besteht aus einer Seite, die einen Satz von benannten Platzhaltern oder Parametern bereitstellt, definiert in einer Mischung aus HTML und Wiki-ML. Diese freien Plätze werden dann später, ähnlich dem Aufruf einer Funktion in Java, mit entsprechenden Informationen über die vorhandene Signatur befüllt. Ein großer Vorteil dieser Methode ist die Tatsache, dass alle Instanzen auf ein und dem selben, global definierten Template basieren. Ändert sich die Signatur der Vorlage, stehen die neuen Attribute allen Seiten zur Verfügung, die das Template momentan aufrufen. Dieses Muster sichert die vereinheitlichte Darstellung und lässt eine Abfrage über z. B. die automatisierte Informationsextraktion zu (siehe u. a. DBpedia²¹). Ein Nachteil ist mitunter die Tatsache, dass man sich zunächst durch das Vorlagenverzeichnis wühlen muss, falls man bestehende Templates verwenden möchte. Zusätzlich ist im Fall von MediaWiki die Aneignung einer spezifischen Syntax notwendig, um eigene Vorlagen zu definieren. Templates werden aufgrund ihrer hoher Barrieren deshalb von wenigen Personen erstellt und verwaltet und von vielen genutzt.

Creational Templates Im Gegensatz zu der eben genannten Form stehen kreative Templates. Es besteht nur eine schwache Bindung zwischen Template und Instanz, da die Templates nur als „Copy & Paste“-Vorlage dienen. Änderungen am Template zu einem späteren Zeitpunkt haben keinen Einfluss auf konkrete Instanzen und umgekehrt. Während dieser Ansatz zum einen eine höhere Flexibilität dem Strukturierenden gegenüber ermöglicht, leidet darunter die gesamte Vereinheitlichung des Template-Systems.

Lightly Constrained Templates Diese sind ein dritter Ansatz (vorgeschlagen durch [VZ07]), der auf den Creational Templates aufbaut und sie um Integritätsbedingungen er-

¹⁹siehe <http://de.wikipedia.org/wiki/Hilfe:Formatvorlagen>

²⁰Eine Übersicht zur Verfügung stehender Infoboxen ist auf <http://de.wikipedia.org/wiki/Kategorie:Vorlage:Infobox> zu finden.

²¹<http://dbpedia.org/>

```
1  {{Infobox Town AT |
2   name = Innsbruck |
3   image_coa = InnsbruckWappen.png |
4   image_map = Karte-tirol-I.png |
5   state = [[Tyrol]] |
6   regbzg = [[Statutory city]] |
7   population = 117,342 |
8   population_as_of = 2006 |
9   pop_dens = 1,119 |
10  area = 104.91 |
11  elevation = 574 |
12  lat_deg = 47 |
13  lat_min = 16 |
14  lat_hem = N |
15  lon_deg = 11 |
16  lon_min = 23 |
17  lon_hem = E |
18  postal_code = 6010-6080 |
19  area_code = 0512 |
20  licence = I |
21  mayor = Hilde Zach |
22  website = [http://innsbruck.at] |
23  }}
```



Country	Austria
State	Tyrol
Administrative region	Statutory city
Population	117,342 ⁽²⁰⁰⁶⁾
Area	104.91 km²
Population density	1,119 /km²
Elevation	574 m
Coordinates	47°16′N 11°23′E﻿ / ﻿47.267°N 11.383°E﻿ / 47.267; 11.383
Postal code	6010-6080
Area code	0512
Licence plate code	I
Mayor	Hilde Zach
Website	www.innsbruck.at ^{}

Abbildung 2.4.: Beispiel für den funktionalen Aufruf und die Darstellung einer Templateinstanz (MediaWiki) [Aum05]

weitert. Validatoren überprüfen vorab festgelegte Einschränkungen und teilen in Instanzen über Warnungen mit, ob diese eingehalten wurden oder nicht. So kann ein gewisser Grad an Konsistenz gewährleistet werden, ohne dabei die Flexibilität des Nutzers einzuschränken.

2.2. Enterprise 2.0

Die zunehmende Globalisierung der Wirtschaft, kurze Produktlebenszyklen und steigende Anforderungen an die Flexibilität der Mitarbeiter lässt die IT-gestützte Zusammenarbeit zunehmend an Bedeutung gewinnen [RCB10]. Mitarbeiter müssen sowohl in räumlicher als auch zeitlicher Unabhängigkeit effektiv zusammenarbeiten können, um diesem Trend gerecht zu werden. Es gibt zwar klassische Groupware und Knowledge Management (KM)-Systeme, um diese Art des Arbeitens zu unterstützen, jedoch setzen sich Web 2.0 Technologien wie Wikis oder Blogs zunehmend im Unternehmensumfeld durch, nachdem sie sich im Privatgebrauch bewährt haben. Es wird mehr auf einzelne Module gesetzt als auf Komplettlösungen. Der Sammelbegriff für diese Entwicklung nennt sich Enterprise 2.0.

2.2.1. Ursprung

Andrew P. McAfee spricht im Frühjahr 2006 erstmals von Enterprise 2.0. Er stellt fest: „Wikis, blogs, group-messaging software and the like can make a corporate intranet into a constantly changing structure built by distributed, autonomous peers — a collaborative platform that reflects the way work really gets done.“ [McA06]. Er sieht die genannten Tools im Vorteil gegenüber den alten Kommunikationsmöglichkeiten wie E-Mail oder

Instant-Messengern, die er als Kanäle bezeichnet, und Plattformen wie Intranets, firmeneigene Webseiten oder Informationsportale. Auch KM-Systeme, die beide Aspekte der Kommunikation vereinen, werden als unzureichend gesehen. Das Problem der Kommunikationskanäle ist, dass sie keine öffentliche Einsicht erlauben. Wissen wird immer nur bidirektional geteilt. Plattformen hingegen haben das Problem, dass keiner weiß wie und wann sie genutzt werden.

2.2.2. Enterprise Wikis

Auf Unternehmensbedürfnisse ausgelegte Wikis lösen diese Probleme unter anderem über Versionierung und Aufzeichnung von Aktivitäten, Bewertung und Diskussion der Inhalte und der Vergabe von Zugriffsrechten. Schnittstellen sorgen für die nahtlose Integration in die vorhandene Systemlandschaft. [Cho06]

Einer Umfrage zufolge, bei der 168 Nutzer von Wikis in Unternehmen teilnahmen [MWY06], kristallisierten sich drei positive Effekte von Wikis heraus: Verbesserung des Rufes (der Organisation und einzelner Individuen), Vereinfachung des Arbeitsalltages und Unterstützung zur Optimierung von Prozessen.

Enterprise Wikis sind in ihrem Einsatz extrem flexibel. Eine Auflistung der Verwendungsmöglichkeiten in verschiedenen Unternehmensbereichen soll dies verdeutlichen:

- Softwareentwicklung (inklusive Dokumentation, Kundenabnahme, Bugtracking, Qualitäts- und Prozessmanagement, Installationshinweise, Anforderungsbeschreibung)
- E-Learning (Tests, Organisation von Trainingseinheiten)
- Projektmanagement (zeitliche Planung, Festlegung und Verteilung von Deliverables, Statusreports, Standards und Praktiken)
- Wissensmanagement (Urlaubskalender, Tutorials, persönliche Blogs, Unternehmensinformationen, Best Practices, Human Resource (HR)-Informationen, Versicherungshinweise)
- Organisation von Arbeits- oder Nutzergruppen
- technischer Support (Downloads, Tutorials, Troubleshooting)
- Marketing und Customer Relationship Management (CRM) (Beobachten und Festhalten von Trends, Datensammlungen, Marketingmaterial, Informationen zu Partnerschaften)
- Ressourcenmanagement (z. B. die Verwaltung eines Nutzungsplans für Maschinen)
- Forschung und Entwicklung (Anforderungen an neue Produkte, Produktinformationen)

Außerdem wurde festgestellt, dass sich Wikis vor allem dann bewährt haben, wenn es um die Handhabung von Aufgaben oder Problemen geht, denen keine routinemäßigen Vorgänge zugrunde liegen oder das Ergebnis vorab noch nicht feststeht (also vor allem im Projektmanagement) [MWY06].

Populäre Beispiele für Wikis, die sich im Unternehmenskontext bewährt haben, sind Atlassian Confluence, DokuWiki, DrupalWiki, FosWiki, TikiWiki und TWiki.²²

2.3. Web 3.0

2.3.1. Ursprung

Das Web 3.0 erweitert das Web 2.0 um die Eigenschaften und Funktionalitäten mehrerer semantischer Netze (zentraler engl. Begriff: *Semantic Web*). Während die Inhalte des Web 2.0 größtenteils von Menschen für Menschen geschaffen wurden, zielt das Web 3.0 darauf ab, diese vorhandene Information maschinenlesbar zu machen und legt den Fokus auf die Bedeutungs- und Beziehungsebene der Inhalte, um Computer als Helfer oder Agenten in das Web einzubinden. So sollen über Ansätze aus dem Bereich der künstlichen Intelligenz (engl. *artificial intelligence*) Maschinen dazu in der Lage sein, Information miteinander zu verknüpfen, neue Zusammenhänge abzuleiten und komplexe Suchabfragen zu erledigen. Die „Semantic Web Road map“ von Tim Berners-Lee [BL98] beschreibt bereits 1998, wie das semantische Web aussehen könnte. John Markoff, Autor der New York Times, stellt 2006 fest: „[...] commercial interest in Web 3.0 — or the ‘semantic Web’, for the idea of adding meaning — is only now emerging.“ [Mar06].

2.3.2. Semantic Web und semantische Technologien

Zunächst muss differenziert werden zwischen dem Konzept des Semantic Web, dessen Aufbau und Standards durch das World Wide Web Consortium (W3C) beschrieben werden (siehe Abbildung 2.6), und semantischen Technologien zur Schaffung von semantischen Netzen. Letztere können auch innerhalb von Organisationsgrenzen im Wissensmanagement zum Einsatz kommen, während das Semantic Web ein ganzheitliches Konzept des Internets darstellt, ähnlich dem Web 2.0 [BT06].

2.3.3. Über Linked Data zum Web of Data

Das Internet unterlag in den letzten Jahren einem Wandel von nebeneinander existierender Information hin zu verlinkten Daten. Linked Data bezieht sich auf Daten im (semantischen) Web, die maschinenlesbar sind, deren Bedeutung explizit formuliert wurde, die auf andere Daten verweisen und auf die wiederum verwiesen werden kann. Um dies zu realisieren, werden RDF-Statements (prinzipiell ohne feste Repräsentationsform, in der Regel aber über das Extensible Markup Language (XML)-Format ausgedrückt) dazu verwendet, um strukturierte Subjekt-Prädikat-Objekt Relationen zur Beschreibung und Verknüpfung beliebiger Ressourcen (real existierende Einheiten unserer Welt wie Personen, Gebäude, geographische Objekte, ...) und ihren Beschreibungen zu formen. Das letztendliche Ergebnis dieser Vorgehensweise wird auch als Web of Data bezeichnet [BHBL09].

Berners-Lee [BL06] hat vier Regeln festgelegt („Linked Data Principles“), die als Orientierungshilfe dienen sollen:

1. Uniform Resource Identifier (URI) formen Bezeichner für Objekte.

²²vergleiche [GFG11]; Auf die Angabe weiterführender Links wird an dieser Stelle verzichtet.

2. Hypertext Transfer Protocol (HTTP) URI verlinken die festgelegten Bezeichner mit beschreibenden Inhalten.
3. Wenn jemand einen URI aufruft, muss sinnvolle, auf Standards basierende (RDF, SPARQL²³) Information zur Verfügung gestellt werden.
4. Verknüpfungen zu weiteren URI sorgen dafür, dass man zusätzliche Dinge findet.

Die Eigenschaften des Web of Data lassen sich wie folgt zusammenfassen [BHBL09]:

1. Das Web of Data ist generisch und kann jede Art von Daten enthalten.
2. Jedermann darf Inhalte veröffentlichen.
3. Autoren sind in der Wahl des verwendeten Vokabulars zur Beschreibung ihrer Daten in keiner Weise eingeschränkt.
4. Entitäten sind über RDF-Links verknüpft, es entsteht ein Netz aus verbundenen Daten, über das neue Zusammenhänge entdeckt werden können.

URI und das HTTP bilden die beiden Basistechnologien des Web of Data. URI sind dabei generischer als URL. Während URL Dokumente oder HTML-Seiten im Netz markieren, dienen URI als verallgemeinerte Variante zur Lokalisation von Inhalten über einen Bezeichner. Über HTTP lassen sich URI abfragen.

RDF ist ein Graph-basiertes Datenmodell (vgl. Abbildung 2.5) zur Beschreibung von Ressourcen und ihren Zusammenhängen. Eine Ressource (engl. *resource*), die über eine URI identifiziert wird, bildet das Subjekt. Diese besitzt eine oder mehrere benannte Eigenschaften oder Prädikate (engl. *property*), die wiederum benannte Ressourcen sein können. Eigenschaften haben eine Belegung mit einem Wert (engl. *property value*), der neben einer weiteren Ressource auch ein Text-Literal sein kann.



Abbildung 2.5.: Subjekt, Prädikat und Objekt als gerichteter Graph in RDF [RDF04c]

Das W3C²⁴ definiert auf der offiziellen Seite den Standard so:

„RDF is a standard model for data interchange on the Web.

[...] RDF extends the linking structure of the Web to use URIs to name the relationship between things as well as the two ends of the link (this is usually referred to as a ‘triple’).

[...] This linking structure forms a directed, labeled graph, where the edges represent the named link between two resources, represented by the graph nodes. This graph view is the easiest possible mental model for RDF and is often used in easy-to-understand visual explanations.“ ([RDF04b])

²³RDF Query Language; siehe <http://www.w3.org/TR/rdf-sparql-query/>

²⁴Gründer: Tim Berners-Lee

2. Grundlagen - die Entwicklung des Webs

URI und RDF-Statements, die diese verwenden, bilden die beiden untersten Ebenen des Semantic Web-Stacks (veranschaulicht in Abbildung 2.6).

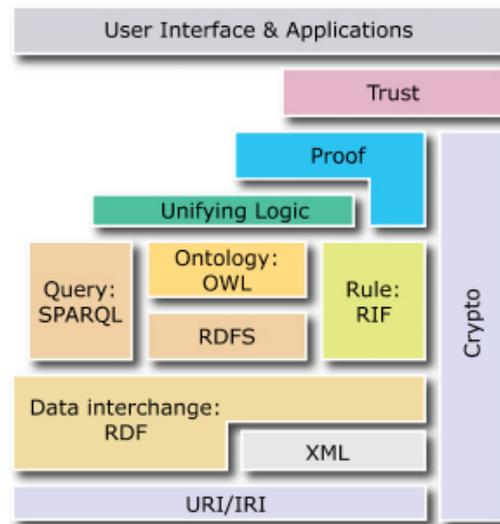


Abbildung 2.6.: Schichtenmodell der Standards des Semantic Web (W3C) [Se007]

RDF Vocabulary Definition Language (RDFS)²⁵ erlaubt die Definition eines grundlegenden Vokabulars zur Beschreibung von Eigenschaften und Klassen zu RDF-Ressourcen [OWL04] und wird ähnlich wie bei normalen XML-Dokumenten über Document Type Definition (DTD) am Anfang eines Dokuments importiert.

Strukturierte Abfragen finden über die SPARQL Protocol And RDF Query Language (SPARQL)²⁶ statt. Es ähnelt von seiner Syntax her anderen Query-Languages wie z. B. der Structured Query Language (SQL). Die Ergebnisse einer Suchabfrage werden entweder in XML oder als RDF-Graph zurückgegeben. Ein sehr einfaches, schematisches Beispiel soll das Konzept verdeutlichen²⁷.

Vorhandenes Tripel:

```
Subject: <http://example.org/book/book1>
Predicate: <http://purl.org/dc/elements/1.1/title>
Object: "SPARQL Tutorial"
```

Abfrage in SPARQL:

```
SELECT ?title
WHERE
{
<http://example.org/book/book1>
```

²⁵<http://www.w3.org/TR/rdf-schema/>

²⁶<http://www.w3.org/TR/rdf-sparql-query/>

²⁷vgl. <http://www.w3.org/TR/rdf-sparql-query/#WritingSimpleQueries>; abgerufen am 31.08.2011

```
<http://purl.org/dc/elements/1.1/title>
?title
}
```

Ergebnis der Abfrage:

```
?title: "SPARQL Tutorial"
```

2.3.4. Ontologien

Die klassische Ontologie bezeichnet die Lehre des Seins in der Philosophie. In der Informatik ist eine Ontologie „[...] an explicit specification of a conceptualization.“ [RG94], also eine explizite Spezifikation einer Konzeptionalisierung. Spezifisches Wissen kann über die Verwendung eines Vokabulars formal modelliert und repräsentiert werden. Wichtige Elemente sind Klassen (z. B. „Stadt“), Typen („geographischer Punkt“ für „Stadt“), Instanzen („München“), Relationen („München“ „ist Landeshauptstadt von“ „Bayern“), Vererbung und Axiome (wahre Aussagen, die sich nicht über Begriffe beschreiben lassen) (vgl. [WiXb]). Ontologien besitzen die höchste semantische Reichhaltigkeit (siehe Abbildung 2.7).

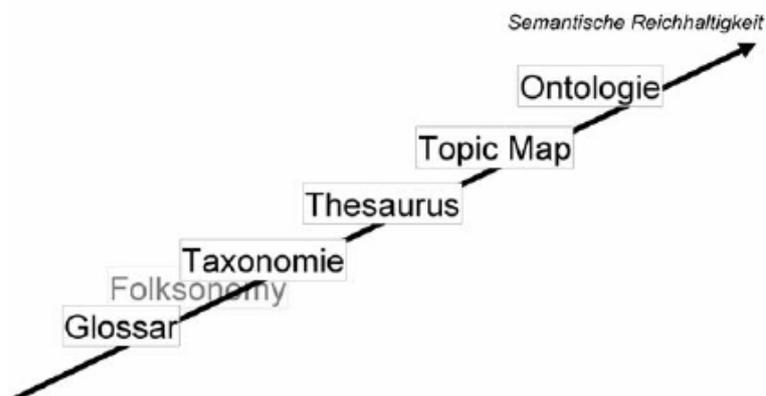


Abbildung 2.7.: Semantische „Treppe“ [BT06]

[BT06] teilt Ontologien in vier verschiedene Kategorien ein:

1. domänenspezifische Ontologien (Eingrenzung auf einen anwendungsspezifischen Bereich wie z. B. Banken, Universitäten, etc.)
2. Metadaten-Ontologien (Dateigröße, Format, Änderungsdaten; dienen zur Beschreibung von Informationsquellen)
3. Upper Level-Ontologien (auch: generische Ontologien; dienen zur Darstellung allgemeinen Wissens; Grundlage spezifischerer Ontologien)
4. Light Weight-Ontologien (Festlegung von Konzepten/Begriffen und Eigenschaften sowie Bildung von Relationen)

Schwergewichtige Ontologien erweitern die leichtgewichtigen um Axiome, Einschränkungen und Inferenzen durch die Festlegung von Regeln [BT06].

Ontologien bilden die nächste Ebene des Semantic Web-Stacks und werden über die Web Ontology Language (OWL)²⁸ beschrieben, die RDF(S) um zusätzliche Konstrukte zur Erhöhung der Ausdrucksmächtigkeit erweitert. Basiselemente sind Klassen, Eigenschaften und Instanzen von Klassen [WiXc]. Klassen lassen sich hierarchisch in Relation zueinander setzen sowie mit Kardinalitäten versehen. Eigenschaften können ob ihrer Belegung eingeschränkt und Charakteristiken wie Transitivität oder Symmetrie festgelegt werden. Außerdem lassen sich sowohl Klassen als auch Eigenschaften auf ihre Gleichheit hin definieren. [OWL04][RDF04a]

2.3.5. Semantische Wikis

Semantische Wikis ergänzen die Funktionalitäten eines normalen Wikis mit ausgewählten semantischen Mitteln und einem einfachen aber mächtigen Modell, in dem jede Seite ein Konzept einer Domäne darstellt, typisierte Links Beziehungen zwischen Seiten ausdrücken und Eigenschaften jedes Konzeptes über Attribute ausgedrückt werden [DIVZ09]. Intern werden die geschaffenen Strukturdaten oft in RDF(S) oder OWL repräsentiert (siehe Abschnitt 2.3.3).

Darüber hinaus bieten semantische Wikis nach [SBBK08] charakteristische Eigenschaften:

- einen Formalismus zur Annotation von Inhalten eines Wiki-Artikels (v. a. Links)
- eine semantische Suche, die eine detaillierte Abfrage auf Basis der geschaffenen Struktur ermöglicht (z. B. eine Suche über Facetten)
- zusätzlich möglicherweise eine (halb-)automatische Extraktion von Metadaten aus Wiki-Artikeln zur Unterstützung der Annotation

In ihrer Zielsetzung lassen sich zwei Typen von semantischen Wikis unterscheiden. Die erste Gruppe möchte über den Einsatz semantischer Technologien die Navigation und generell das Arbeiten mit dem Wiki erleichtern (Semantik für Wikis), die andere setzt gezielt auf die kollaborative Entwicklung semantisch reichhaltiger Strukturen (Wikis für Semantik) [SBBK08].

Semantische Wikis abstrahieren oft nicht von der internen Repräsentation der Daten. Da selbst noch relativ einfache Standards wie RDF Nichtexperten in der Regel fremd sind und kompliziert erscheinen, kommt es hier oft zu Verständnisproblemen.

Punkte, die momentan im Hinblick auf semantische Wikis in Übereinstimmung mit [SBBK08] im Blickpunkt stehen, sind ...

- die Flexibilität (Spannungsfeld zwischen unstrukturierten Daten, die einfach einzugeben sind, versus vorgegebene Schemata, die einen höheren Aufwand der Pflege nach sich ziehen),
- der gezielte Einsatz von Abfragesprachen (SPARQL), die neue Strukturen zu nutzen wissen,

²⁸<http://www.w3.org/TR/owl-features/>

- der Prozess der Extraktion der gewünschten Daten (Algorithmen und weitere Hilfsmechanismen aus den Bereichen Ontologie-Lernen und Text-Mining),
- das Reasoning, zu Deutsch die Deduktion (Anwendung einer einfachen Regelsprache zur Ableitung von Schlussfolgerungen),
- die Visualisierung (unterschiedliche, auch anwendungsspezifische Sichten auf verschiedene Aspekte der Struktur, z. B. in Form von Tabellen und Diagrammen),
- sowie die Adaption (nach z. B. Themenbereich oder Präferenzen der Nutzer) der Inhalte.

2.3.5.1. Beispiel: semantische Annotation mit Semantic MediaWiki

Ein erstes Beispiel für ein semantisches Wiki ist das Semantic MediaWiki [KVV06], das die Factbox als eine verbesserte Möglichkeit zur Infobox (siehe Abschnitt 2.1.4.3) anbietet, um Information aus Artikeln zu ontologisieren.

Betrachtet man Artikel im MediaWiki als ontologische Elemente (Klassen und Eigenschaften), dann lassen sich über die erweiterte Markup Language des Semantic MediaWikis Aussagen über eben diese Elemente in Form von semantischen Annotationen formulieren.

Existiert in einem Artikel mit dem Titel „München“ die Annotation [[Kategorie:Stadt]], wird aus „München“ automatisch eine Instanz der OWL-Klasse „Stadt“. Somit dienen Kategorien zur Bildung von Klassen.

Angenommen, es existiert ein Link im Artikel zu „München“ zur Partnerstadt „Verona“ in Form von [[Verona]], dann kann man ihn erweitern auf [[Partnerstadt::Verona]]. Dadurch entsteht eine Relation zwischen zwei Links.

Relationen, die zu keinem Artikel in Form eines Links führen, nennt man Attribute. [[Vorwahl:=089]] sagt, ohne auf einen Link zu verweisen, aus, dass „München“ die Vorwahl 089 besitzt. Da „Vorwahl“ ebenso ein ontologisches Element darstellt, lässt sich ein Typ für dessen Belegung mit Werten definieren, hier also zum Beispiel Integer.

All diese Statements innerhalb eines Artikels führen zur automatischen Generierung einer Factbox (für ein Beispiel siehe Abbildung 2.8). Sie wird ergänzend unterhalb eines Wiki-Artikels angezeigt, um die Fakten gesammelt darzustellen.

Facts about Berlin ⓘ	
Area	891,690,000 m² (891.69 km², 89,169 ha, 344.283 miles²) + 🔍
Capital of	Germany + 🔍
Coordinates	52°30′60″N, 13°24′0″E + 🔍 ⓘ
Homepage	http://www.berlin.de/english/index.html ⓘ + 🔍
Population	3,391,407 + 🔍
Population (on date)	3405000 (2006-11) + 🔍
State	Berlin (state) + 🔍

Abbildung 2.8.: Aus semantischer Annotation gesammelte Fakten des Wiki-Artikels über „Berlin“ in Semantic MediaWiki

2.3.6. Manuelle, unterstützte und automatisierte Strukturierung

Während Konstrukte wie Ontologien früher von Experten entworfen und formalisiert worden sind, ist es in Zeiten von Social Software dazu gekommen, diese Aufgabe immer stärker auf die Nutzer eines dafür ausgerichteten Systems zu übergeben. Die kollaborative Ontologieentwicklung profitiert von den Wiki-Prinzipien um über vereinfachende Graphical User Interface (GUI)s oder eine spezielle Syntax Inhalte manuell zu annotieren und dadurch in Konsequenz im besten Fall formale, leichtgewichtige Ontologien zu erstellen (vgl. unter anderem [ADR06]). Dies sind grob umrissen die Eigenschaften der zu sammelnden Ansätze und dienen somit als Ausgangspunkt der Recherche (mehr dazu in Abschnitt 3.1).

Dem gegenüber stehen Prozesse, die auf Basis bestehender unstrukturierter bis semi-strukturierter Wiki-Artikel über den Einsatz unterschiedlichster Methoden, wie der Verwendung von Algorithmen²⁹, die automatisierte Extraktion verlinkter Daten zu vollziehen, die im Anschluss dazu dienen können, über die Named Entity Recognition (NER) als Teilbereich des Natural Language Processing (NLP) neue unstrukturierte Inhalte auf Vergleichsbasis automatisiert zu annotieren. Auch werden so gewonnene Linked Open Data (LOD) Sets immer öfter in neu entstehende Plattformen eingespielt, um einen unterstützenden Einstiegspunkt für die Community zu schaffen (bestes Beispiel ist die Open Graph-Plattform Freebase³⁰, die aufgrund ihres Erfolges inzwischen von Google aufgekauft wurde³¹).

2.4. Kollaborative Informationsstrukturierung mit Hybrid Wikis

Hybrid Wikis³² ist eine am sebis Lehrstuhl der TUM entworfene, leichtgewichtige Erweiterung für die ebenfalls dort entwickelte Enterprise Collaboration-Plattform Tricia. Während Tricia als Basiskomponente ein traditionelles Wiki-System mit Standardfeatures wie einem What You See Is What You Get (WYSIWYG)-Editor zum Bearbeiten von Inhalten, der Möglichkeit der Zuweisung von Schlagwörtern zu Seiten sowie dem Aufbau von Seitenhierarchien stellt, reichert Hybrid Wikis dieses mit ausgewählten Features an, um Nutzern eine Möglichkeit zu bieten, ihre Inhalte mit zusätzlicher Struktur zu versehen. Die Zusammenführung eines unstrukturierten Textabschnitts, der über einfache Formatierungen, ähnlich wie in einem Word-Dokument, gegliedert wird, und einem Strukturteil mit ausgewählten Elementen zur erweiterten Beschreibung führt zu einer hybriden Darstellungsform der Inhalte.

²⁹Oft in Bezug auf Wikipedia und unter Verwendung der Gliederung der Infoboxen von MediaWiki. Eine solche Vorgehensweise ließe sich aber prinzipiell auf jedes Wiki anwenden, das eine ähnliche Regelmäßigkeit in der Darstellung aufweist.

³⁰<http://www.freebase.com>

³¹vgl. <http://www.cloudave.com/140/google-buys-freebase-this-is-huge/>; abgerufen am 04.10.2011

³²<http://www.matthes.in.tum.de/wikis/sebis/hybrid-wiki>; abgerufen am 15.10.2011

2.4.1. Struktursystem

Die von Hybrid Wikis bereitgestellten Elemente können zur Bildung von Modellen verwendet werden. Einzelne Wiki-Seiten werden über eine URL identifiziert und besitzen ihre Gültigkeit innerhalb eines (Namens-)Raums (Space), der einen in sich geschlossenen Teilbereich des Systems bildet. In den Seiten selbst finden neben klassischen Tags Attribute und Typ Tags, denen Constraints hinzugefügt werden können, ihre Verwendung.

- **Attribute** Attribute sind Eigenschaften-Werte-Paare. Jedes Attribut besitzt einen Namen oder Schlüssel sowie einen oder mehrere zugewiesene Werte, die neben typisierten Text-Literalen auch Links sein können [MNS11]. Seiten können eine beliebige Anzahl an Attributen angeheftet werden. Die Attribut-Liste wird alphabetisch sortiert angezeigt, bei Mehrfachzuweisung von Werten lassen sich auch diese geordnet darstellen (z. B. nach Relevanz) [MNS11]
- **Typ Tags** Neben Attributen können sogenannte Typ Tags angegeben werden, die den Typ der zu beschreibenden Wiki-Seite bestimmen [MNS11]. Auch Typ Tags können in beliebiger Menge pro Seite hinzugefügt werden. Zusätzlich besteht ein Zusammenhang zwischen Typ Tags und Attributen. Werden Typ Tags mit Attributen gemischt, kommt es zur impliziten Bildung von Klassen, die ein Set von Attributen unter einem oder mehreren Typ Tags zusammenfassen. Beide Elemente bleiben aber lose aneinander gekoppelt, so dass einzelne Attribute mit anderen Typ Tags kombiniert werden können, genauso wie zusätzliche Attribute zu einem Typ Tag annotiert werden können, die bisher noch nicht im Zusammenhang mit diesem standen.

Vererbung in Form von Typ-Subtyp-Relationen ist explizit nicht möglich. Sollte eine Typ Tag die Generalisierung eines spezifischeren Typ Tags darstellen, so muss bei Verwendung des spezifischeren Typ Tags stets manuell der generalisierende Typ Tag mit angegeben werden, um Integritätsverletzungen zu verhindern [MNS11].

- **Integritätsbedingungen (Constraints)** Um ein gewisses Maß an Konsistenz gewährleisten zu können, gibt es die Möglichkeit der Verwendung von Constraints für Definitionen von Attributen und Typ Tags, die durch Validatoren geprüft werden.

Typ Tag-Definitionen beinhalten ein Set von Attributen, die in Verbindung mit dem entsprechenden Typ Tag auftreten werden sollen. Zusätzlich lässt sich auch der Wertebereich, der Typ und die Anzahl der Werte der einzelnen Attribute festlegen. Man könnte also dem Typ Tag „Buch“ ein Attribut „Autor“ zuweisen und bestimmen, dass der Wert mindestens einen Link haben muss, der wiederum auf eine Wiki-Seite verweist, die vom Typ „Person“ sein muss.

Für Attribute, die per Definition zu einem Typ Tag gehören, lassen sich ebenfalls Integritätsbedingungen festlegen. Es gibt unter anderem die Überprüfung der Verwendung von Enumerationen (z. B. „hoch“, „mittel“ und „niedrig“), von Hyperlinks (und welche(n) Typ Tag(s) er besitzen soll) sowie der Anzahl von Werten (z. B. mindestens ein Wert, genau ein Wert oder höchstens ein Wert).

Sollte durch die Validatoren eine Verletzung der Bedingungen festgestellt werden, wird ein Hinweis neben dem entsprechenden Attribut angezeigt. Dieser Hinweis dient als Empfehlung, weshalb man auch von „weicher“ Validierung spricht. [MNS11]

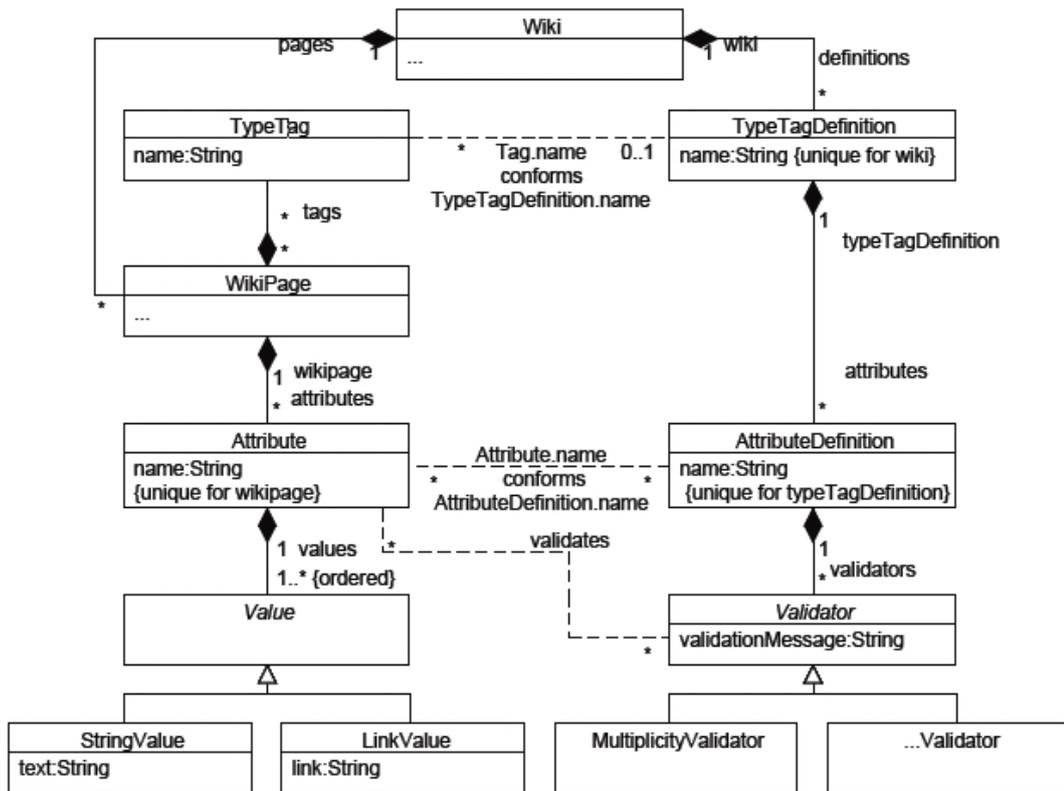


Abbildung 2.9.: Das Datenmodell von Hybrid Wikis [MNS11]

Abbildung 2.9 fasst die genannten Konzepte in einem Unified Modeling Language (UML)-Klassendiagramm zusammen.

In Tricia werden Änderungen an Inhaltsobjekten wie Wiki-Seiten versioniert. Es wird festgehalten, wer welches Objekt zu welcher Zeit und aus welchem Grund geändert hat. In Hybrid Wikis werden zusätzlich Änderungen von Strukturelementen festgehalten. Ein Notification-Dienst erlaubt Benachrichtigungen (zum Beispiel per E-Mail) zu allen aktuellen Änderungen, die Freitext und Strukturelemente betreffen. [MN11]

Strukturierte Inhalte lassen sich über spreadsheetartige Formate importieren und exportieren. Der Import von Inhalten über beispielsweise das gängige Excel-Format ist dank der weichen Handhabung von Integritätsbedingungen möglich, da nach Verletzungen direkt nach der Integration gezielt gesucht werden und diese dann gezielt korrigiert werden können. Daten aus Hybrid Wikis lassen sich außerdem auf das Eclipse Model Framework (EMF)-Austauschformat abbilden, wobei Modelle als .ecore-Dateien und Instanzdaten als .xmi-Dateien gespeichert werden. [MNS11]

2.4.2. Eingabe

Die Eingabe der Strukturelemente erfolgt getrennt vom eigentlichen Inhalt einer Seite anhand eines UI mit formularbasierten Feldern, was schlussendlich zu einer tabellenähnlichen Darstellung der geschaffenen Struktur führt. Der Nutzer wird während der Eingabe

be durch einen Mechanismus zur Autovervollständigung sowie durch Vorschläge unterstützt.

- **Autovervollständigung** Die Eingabefelder für Namen und Werte von Attributen zeigen während dem Tippen kontextbezogene Angaben in Form einer Drop Down-Liste, sortiert nach Relevanz, an. Wird zum Beispiel der Wert eines Attributs eingegeben, erscheinen abhängig vom zugehörigen Attribut-Namen Vorschläge zur Belegung, wobei solche, die existierende Constraints erfüllen, höher gelistet werden [MNS11].
- **Vorschläge** Durch eine statistische Analyse von oft verwendeten Kombinationen von Tags, Typ Tags und Attributen in Hybrid Wikis können für Typen und Attribute Vorschläge gegeben werden [MN11]. Besitzen mehrere Seiten die gleichen (Typ) Tags, so besteht eine Ähnlichkeit zwischen den Seiten und die Tags können als Indikatoren für deren Nähe gesehen werden. Aus dieser Tatsache heraus lassen sich bei der nutzergesteuerten Eingabe eines Typ Tags Vorschläge für die im Zusammenhang am häufigsten verwendeten oder vordefinierten Attribute generieren. Sollten keine Typ Tags sondern nur Attribute existieren, wird die Ähnlichkeit anhand der Attribute analysiert [MNS11].

Dadurch, dass Autovervollständigung und Vorschläge sowohl im ad-hoc Strukturierungsprozess während der Erstellung des Strukturteils einer neuen Seite als auch im Definitionsprozess für Strukturelemente ihre Anwendung finden, können Modellspezialisten als Ontologieentwickler und Domänenexperten als Wiki-Autoren in Dialog treten [MN11].

2.4.3. Darstellung, Navigation und Suche

Das auf jeder Seite präsente Suchfeld (vgl. Punkt 1 in Abbildung 2.10) ist verallgemeinert der zentrale Navigationspunkt. Zusätzliche Sichten und das Konzept der Referenzierung erleichtern eine Bewegung über Links zwischen einzelnen Seiten.

Hybrid Wikis kombinieren klassische Wikis und ihre Artikel mit den bereits erläuterten Strukturelementen. Beide Ansichten existieren dabei zusammen in einer zweigeteilten Darstellungsform. Diese Art der Darstellung ist an die Integration von Infoboxen in MediaWiki und die Factbox in Semantic MediaWiki angelehnt. Es existieren aber weder starr definierte Templates im Hintergrund, noch ist die tabellarische Ansicht eine Zusammenfassung aus im eigentlichen Text annotierter Elemente.

Typ Tags und Attribute bilden zusammen flexible Schemata und lassen sich zur Erhöhung der Übersicht in einer Sammelansicht aller existierenden Typ Tags pro Space und zusätzlich zu einem Vergleich von Wiki-Seiten mit dem gleichen Typ Tag in Matrixform darstellen (siehe Abbildung 2.11 und 2.12). Standardmäßig ist die Tabelle nach Häufigkeit von nicht-leeren Attributen beziehungsweise Werten sortiert [MN11], prinzipiell kann aber über alle Spalten hinweg sortiert werden. Erreichbar ist diese Darstellung über einen einfachen Klick auf ein beliebiges Typ Tag in einer Seite (vgl. Punkt 2 in Abbildung 2.10) [MNS11].

Eine Visualisierung von Hybrid Wiki Inhalten ist per Representational State Transfer

2. Grundlagen - die Entwicklung des Webs

The screenshot shows the Hybrid Wikis/Tricia interface. At the top, there is a search bar with the text "What are you looking for?" and a magnifying glass icon. Below the search bar is a navigation menu with links for TEAM, RESEARCH, PUBLICATIONS, EVENTS, COURSES, STUDENT PROJECTS, PARTNERS, CAREER, and CONTACT. The main content area displays the page title "Hybrid Wikis" and a list of tags: [eam](#), [hybrid/wiki](#), [research](#), [semantic](#), [social software](#), [wiki](#), and [wiki-based-approach](#). A "News" section contains a short article snippet dated 20th of July 2011. An "Objective" section provides a detailed description of the Hybrid Wikis approach. On the right side, there is a metadata table with the following entries:

Types:	project research project
Acronym	HyWi
Contact	Christian Neubert
Project start	2008
Research area	Social Software EAM
Status	active
Team members	Christian Neubert Alexander Steinhoff Dr. Christian M. Schweda Prof. Dr. Florian Matthes
Incoming Links	
Project of	Bachelorthesis Martin

Abbildung 2.10.: „Hybride“ Sicht auf eine Wiki-Seite in Hybrid Wikis/Tricia

(REST)ful Application Programming Interface (API) Abfragen über den externen Systemkartographie-Service SyCa³³ möglich.

Hybrid Wikis verwenden zur erweiterten Navigation und Verknüpfung von Inhalten sogenannte uni-direktionale (vgl. [MBB⁺08]) Links, die aus einem gerichteten, aus- und eingehenden Link bestehen. Es gibt zwei Möglichkeiten, ausgehende Links zu definieren. Einmal aus dem Freitext einer Seite selbst heraus und zum Zweiten über die Zuweisung als Wert einer Eigenschaft. „Freie“ eingehende Links werden als Referenzierung auf der Zielseite angezeigt, Wertverknüpfungen über ein automatisch erzeugtes Attribut mit der ursprünglichen Benennung und einem „of“-Postfix samt Backlink zur Ursprungsseite als Wert (vgl. Punkt 3 in Abbildung 2.10). Das System bildet also automatisch inverse Verknüpfungen zwischen Seiten. Dadurch, dass einer Seite der Link „gehört“ [MNS11], muss vorab die Richtung des Verweises bestimmt werden. Ein guter Indikator hierfür sind in der Regel die Multiplizitäten (z. B. das Bundesland „Bayern“ besitzt viele Städte, somit sollten die Städte Richtung Land verweisen, um die eingehenden Links beim Land zu sammeln).

Neben der Volltextsuche existiert eine facettierte Suche. Diese erlaubt eine schrittweises Eingrenzen (Drill Down) der Suchergebnisse anhand der vorhandenen Strukturelemente (vgl. Abbildung 2.13). Der Verzicht auf eine komplexe Abfragesprache erleichtert den Nutzern die Verwendung. [MNS11]

Suchergebnisse lassen sich im Freitextbereich von Wiki-Seiten einbetten. [MN11]

³³<http://www.matthes.in.tum.de/wikis/sebis/syca>; abgerufen am 15.10.2011

All Type Tags in sebis Public Website

Type Tag	Type Definition
bachelorthesis (1)	Create Type Definition for bachelorthesis Student (1) ; Status (1) ; Title (de) (1) ; Supervisor (1) ; Advisor (1) ; Start Date (1)
course (17)	All courses (lectures, lab courses, seminars, ...) offered by sebis currently or in the past. Module No. (17) ; Type (17) ; Lecturer (17) ; SWS (17) ; Language (17) ; Level (17) ; Acronym (17) ; Rhythm (17) ; ECTS (15) ; Organizer (10) ;

Abbildung 2.11.: Typ Tag-Sicht in Hybrid Wikis/Tricia

2.4.4. Zusammenfassung

Neben der sehr fundamentalen Kombination und Nutzung von unstrukturiertem und strukturiertem Inhalt bedienen sich Hybrid Wikis verallgemeinert zweier Methoden: der kollaborativen Ontologieentwicklung unter Einsatz von Wiki-Autoren zur bottom-up Schaffung impliziter Struktur und einem den in Abschnitt 2.1.4.3 beschriebenen Lightly Constrained Templates ähnlichen Konzept, das top-down Definitionen durch Domänenexperten erlaubt (vgl. [MN11]). Durch diese Synergie, den Einsatz von Vorschlägen und die weiche Prüfung der Einhaltung von Bedingungen wird die Qualität der entstehenden Datenstruktur erhöht. Zusätzlich fördern die Designprinzipien den Anwender zur ausgiebigen Beschreibung seiner Inhalte, da er passende Attribute und (Typ) Tags vorgeschlagen bekommt. Klassische Tagging-Probleme wie Pluralismen oder Mehrdeutigkeit (vgl. Abschnitt 2.1.3) werden durch die beschriebenen Mechanismen schon im Vorfeld eingedämmt.

2. Grundlagen - die Entwicklung des Webs

Wiki Pages with Type Tag **student project** in sebis Public Website

Showing 1 to 10 of 69 entries

Search:

First Previous 1 2 3 4 5 Next Last

	Advisor (69)	Supervisor (67)	Type (66)	Student (59)	Title (de) (58)	Project (43)	Status (40)	Stu dat (39)
Master's Thesis	...							
Master's Thesis	...							
Diplomarbeit	...							
Diplomarbeit	...							

Abbildung 2.12.: Sicht für Wiki-Seiten mit dem gleichen Typ Tag in Hybrid Wikis/Tricia

- ▼ Content type
- Wiki Page (8)
- ▼ Space
- sebis Public Website (8)
- ▼ Type
- student project (8)**
- ▼ Attribute
- Start Date (8)
- ▼ Start Date
- 05/15/2011 (1)
- 15.03.2011 (1)
- 2010/02/15 (1)
- 2010/10/15 (1)
- 2011/04/15 (1)
- 2011/05/15 (1)
- 2011/06/15 (1)
- 2011/09/15 (1)
- Special

2011 bachelorthesis thesis tricia

Results 1 - 8 of 8

Search

- [Bachelor's Thesis - ...](#)
- [bachelorarbeit bachelorthesis data migration mergers and acqu](#)
- [Master's Thesis ...](#)
- [2011 diplomathesis masterthesis studentproject teaching](#)
- [Bachelor Thesis ...](#)

Abbildung 2.13.: Per Drill Down eingegrenzte Suche in Hybrid Wikis/Tricia

3. Literaturstudie

Nachdem eine theoretische Einführung in das Thema vorgenommen und Hybrid Wikis als Referenzansatz vorgestellt wurde, wird nun recherchiert, um Ansätze zu finden, die sich dieser oder ähnlicher Grundlagen bedienen. Die Literaturstudie orientiert sich in großen Teilen an dem in [OS10] beschriebenen Leitfaden für strukturierte Literaturrecherchen, der im speziellen Empfehlungen zur Gestaltung bei der Betrachtung von Informationssystemen gibt.

Diese Methoden finden gänzlich oder abgewandelt ihre Verwendung:

- Festlegung der Forschungsfrage sowie der Ziele der Studie („Purpose of the literature review“) - siehe Kapitel 1
- Beschreibung der Vorgehensweise, Eingrenzung und Suche („Searching for the literature“) - dieses Kapitel
- Beurteilung sowie begründete In- und Exklusion verschiedener Funde („Practical screen“) - dieses Kapitel
- Katalogisierung der Ansätze und anschließende Analyse und Einordnung der Quellen („Data extraction“) - siehe Kapitel 4
- Formulierung von Folgerungen auf Basis der vorangegangenen Arbeit („Synthesis of studies“) - siehe Kapitel 5

3.1. Eingrenzung

Da es im Rahmen der Bachelorarbeit schwierig wird, eine ganzheitliche Literaturanalyse im Bereich der Web-basierten Tools oder Ansätze zur Strukturierung von Information durchzuführen, wurde die Breitensuche auf Basis mehrere Kriterien eingeschränkt. Zunächst wird der Fokus auf das Wiki-Universum und verwandte Bereiche gelegt. Weiterhin werden nur möglichst neuartige Ideen und Umsetzungen aus dem universitären bzw. wissenschaftlichen Umfeld betrachtet, um einen guten Vergleich mit Hybrid Wikis zu ermöglichen, auf das dieses Kriterium ebenfalls zutrifft. Aus der letzten Einschränkung folgt, dass die Tools oder Ansätze wenn möglich nicht älter als fünf bis maximal sieben Jahre alt sein sollen, um einen Blick auf aktuelle Tendenzen in diesem Bereich zu erhalten. Außerdem müssen Publikationen in Form von Konferenz-, Journal-, oder Workshop-Beiträgen vorliegen, die über wissenschaftliche Suchmaschinen gefunden werden können. Mengemäßig wird die Zahl auf nicht mehr als 30 ausgewählte Ansätze beschränkt werden, um ein gewisses Maß an Übersicht gewährleisten zu können.

3.2. Vorgehensweise

Die konkrete Vorgehensweise gestaltet sich wie folgt: Zunächst wurden über die in Abschnitt 2.4 vorgenommene Charakterisierung von Hybrid Wikis sowie der vorangegangenen Theorie Schlagwörter bestimmt (die da lauten: *Wiki Information Structuring*, *Wiki Structuring Concepts*, *Wiki Content Structuring*, *Wiki Information Management*, *Web Collaboration*, *Knowledge Wiki*, *Wiki Structured Data*, *Collaborative Data Structuring*, *Collaborative Information Structuring*) die als Startpunkt für die Suche über Datenbanken verwendet werden. Die gefundenen Ansätze werden auf Querverweise und Nennungen von ähnlichen Umsetzungen hin untersucht, um sukzessive eine solide Basis aufbauen zu können.

Abschließend werden die initialen Suchergebnisse unter Angabe zusätzlicher Informationen aufgelistet und eine Beurteilung der Recherche abgegeben.

3.3. Verwendete Möglichkeiten der Recherche

Die Schlagwortsuche findet über Suchmaschinen zu wissenschaftlicher Literatur und einigen, per TUM eAccess¹ zugänglichen, Datenbanken statt. Es wurden mehrere wissenschaftliche Suchmaschinen verwendet, um einen hohen Deckungsgrad für die entsprechenden Schlagwörter und die Quersuche zu erhalten.

Name	URL
Google Scholar	http://scholar.google.com/
Microsoft Academic Search	http://academic.research.microsoft.com/
arXiv.org	http://www.arxiv.org/
SpringerLink	http://www.springerlink.com/
ACM Digital Library	http://portal.acm.org/
IEEE Xplore Digital Library	http://ieeexplore.ieee.org/

Tabelle 3.1.: Suchmöglichkeiten für die Schlagwortsuche

¹<http://www.ub.tum.de/eaccess/>

3.4. Ergebnisse der Suche

Es folgt eine alphabetisch sortierte Auflistung der auf Basis der vorangegangenen Festlegungen gesammelten Literatur mit Quellenangabe, Titel, Schlagwörtern und einer Kurzbeschreibung.

Quelle(n)	Ansatz ²	Schlagwörter ³	Kurzbeschreibung
[Kuh08a], [Kuh08b]	AceWiki	<i>Controlled Natural Language, Attempto Controlled English, Semantic Wiki, Semantic Web, Ontology, Usability, User Interfaces, Knowledge Representation</i>	Semantisches Wiki mit einem eingeschränkten Wortschatz (Attempto Controlled English (ACE)) zur Repräsentation von Information über grammatikalisch korrekte Sätze.
[LZ04]	Artificial Memory	<i>Personal Knowledge Management, Distributed Web Knowledge Management, Ontology Visualization</i>	Prototypisches System zur Organisation von Wissen aus Dokumenten.
[BSB ⁺ 09]	CNL Wiki	<i>Controlled Natural Language, Rabbit, Large Scale, Meta Model, Ontology Development, Semantic Wiki, Text Generation, User Interaction, Semantic Annotation, Data Management, Attempto Controlled English</i>	Controlled Natural Language (CNL)-Plugin für Semantic MediaWiki.
[BTPC07], [DBPCT07], [DBEP ⁺ 08]	Freebase	<i>Distributed Data Structure, Information Integration, User Interface, Collaborative System, Object Oriented, Query Language, Semantic Network, Reading and Writing</i>	Kollaborativ erstellte Open Graph-Datenbank menschlichen Wissens.
[Rau10]	HYENA	<i>Wiki, Structured Data, RDF</i>	Hybride Plattform bestehend aus einem Wiki und einem RDF-Editor.
[MNS11], [MN11]	Hybrid Wikis	<i>Enterprise Wikis, Semantic Wikis, Enterprise Information Management, Enterprise 2.0, Web Collaboration</i>	Erweiterung der Enterprise-Plattform Tricia um Attribute, Typen und Integritätsbedingungen zur Ergänzung von Freitext in Wiki-Seiten um Struktur.
[Kie06]	Kaukolu	<i>Entry Barrier, Knowledge Management, Open Source, Semantic Wiki</i>	Prototyp eines semantischen Wikis zum Einsatz für das Wissensmanagement in Unternehmen.

²Das Wort „Ansatz“ bezieht sich in der Regel auf prototypische bis ausgereifte Umsetzungen unter Beschreibung der zugrundeliegenden theoretischen Ideen, die in einem konkreten Informationssystem münden.

³Die Schlagwörter stammen entweder aus der Quelle selbst, aus den angegebenen Suchmöglichkeiten und/oder wurden manuell auf Basis der Analyse ergänzt.

3. Literaturstudie

[Sch06], [SWG06], [SEG ⁺ 09]	KiWi (IkeWiki)	<i>Ease of Use, Formal Ontology, Interactive User Interface, Knowledge Engineering, Knowledge Management, Ontology Engineering, Semantic Wiki</i>	Semantisches Wiki mit Fokus auf der Usability zur Unterstützung der kollaborativen Wissensentwicklung.
[RGP ⁺ 09], [GKNL ⁺ 09]	MoKi	<i>Enterprise Modelling, Semantic Wiki, Enterprise Modelling Wiki</i>	Wiki zur Modellierung von Konzepten wie Aktivitäten und Prozesse im Unternehmenskontext.
[ADR06], [DAR06]	OntoWiki	<i>Ease of Use, Knowledge Base, Knowledge Engineering, Knowledge Representation, Open Source Software, Search Strategy, Semantic Web, Tool Support, Real Time</i>	Wiki-ähnliches Tool zur flexiblen, verteilten Wissensentwicklung unter Verwendung vieler verschiedener Sichten.
[Sou06]	Rhizome	<i>Application Framework, Open Source, Semantic Web, Semantic Wiki</i>	Application Framework zum Erstellen semantischer Wikis auf Basis von RDF und einer eigenen Syntax zur Annotation.
[Ore05]	SemperWiki	<i>Collaborative Authoring, Semistructured Data, Information Space, Semantic Annotation, Semantic Wiki</i>	Kombiniert die Usability von persönlichen mit den Möglichkeiten von semantischen Wikis in einem neuen System.
[BSZ07]	SOBOLEO	<i>Ontology Engineering, Social Bookmarking, Social Software</i>	System, das kollaborativ und Web-basiert die Erstellung, Erweiterung und Pflege von Ontologien und gemeinsamer Lesezeichensammlung ermöglicht.
[KVV06]	Semantic MediaWiki	<i>Wiki, Semantic Wiki, Semantic Technologies, MediaWiki</i>	Erweitert die Markup Language des MediaWikis zur semantischen Annotation von Kategorien, Attributen und Relationen in Wiki-Artikeln.
[GZTS10], [ZGS10], [GZS11]	SnoopyDB	<i>Semistructured Data, Recommendations, Human Interaction, Ranking, Semantic Web, RDF, Heterogeneity, Schema Proliferation, Solution, Collaboration</i>	Kollaboratives, Web-basiertes Informationssystem zur semistrukturierten Repräsentation von Resources über Empfehlungen.
[HGT11]	Social Infobox	<i>Collaborative Knowledge Construction, Collaborative Tagging, Data Semantics, Data Structure, Information Management, Knowledge Engineering, Semantic Web, Social Tagging, Structured Data</i>	Kollaborative Wissensentwicklung über Social Property Tagging von Ressourcen.

[BGLP06], [BLGE ⁺ 08]	SweetWiki	<i>Wiki, Semantic Web, Social Tagging, Ontology, Web 2.0</i>	Semantisches Wiki zur RDF-basierten Annotation von Artikeln mit Schlagwörtern und Eigenschaften.
[BH04]	wiki@nt	<i>Knowledge Integration, Ontology Development, Semantic Heterogeneity, Team Work, Version Control, Agent Based, Partial Order</i>	Umgebung zur Ontologieentwicklung mit Wiki-ähnlichem Charakter.
[Aum05], [Aum10]	WikSAR (SHAWN)	<i>Semantic Web, Semantic Web Technology, Semantic Wiki</i>	Semantisches Wiki, mit dem sich über eine einfache Doppelpunkt-Syntax per Social Tagging Inhalte Tripel-basiert annotieren lassen.

Tabelle 3.2.: Ausgewählte Ansätze aus der Literaturstudie

3.5. Beurteilung der Recherche und Eingrenzung der Auswahl

Ein erster Überblick führt zur Erkenntnis, dass sich der Großteil der Ansätze Semantic Web-Standards zur internen Repräsentation bedient. Dies ist nicht sehr verwunderlich, da wir uns gerade in einer Entwicklungsphase des Internets befinden, in der die Verwendung dieser Standards zur Beschreibung von Relationen bis hin zu ganzen Ontologien im Trend liegt (vgl. Abschnitt 2.3).

Auch eine Differenzierung der Ansätze nach solchen, die jene Technologien zur Optimierung der Repräsentation, Navigation und Suche in Wikis einsetzen (Semantik für Wikis) und solche, die nur dazu dienen, semantische Strukturen zu schaffen (Wikis für Semantik), lässt sich, wie in Abschnitt 2.3.5 beschrieben, feststellen.

Ansätze, die für eine bestimmte Domäne oder einen bestimmten Anwendungsbereich konzipiert wurden (Beispiele: die Annotation von Genomen in der Biologie [HOG⁺08], die Sammlung mathematischer Formeln [Lan08], der Einsatz von Wikis als virtuelle Klassenzimmer [Cum08] oder die gezielte Extraktion und Manipulation von Datensätzen aus Wikipedia [ABK⁺07]) wurden ausgeschlossen.

4. Analyse und Einordnung

Die Auswertung des Ergebniskatalogs der Breitensuche hat zur Extraktion verschiedener Konzepte und Methoden (im Folgenden als Eigenschaften betitelt) der betrachteten Systeme geführt („Data extraction“), denen die Quellen nun entsprechend zugeordnet werden. „Struktursystem“, „Eingabe“ und „Darstellung, Navigation und Suche“ bilden dabei die festgelegten Oberkategorien. Die Erkenntnisse aus der Einordnung der Quellen in Tabellenform werden durch zusammenfassende Beschreibungen erläutert¹.

4.1. Struktursystem

Die zugehörige Einordnung der Quellen in Tabellenform befindet sich in Anhang A².

4.1.1. Extrahierte Eigenschaften

- **Datenmodell** Jeder Ansatz bedient sich einem grundlegenden Datenmodell. Dabei wurden drei verschiedene identifiziert:
 - ein eigenes Datenmodell (zwei Ansätze; vgl. Hybrid Wikis, Freebase)
 - in einem Einzelfall ein an das Entity Relationship Model (ERM)³ angelehnte Modell (vgl. Artificial Memory)
 - das Graph-basierte RDF(S)/OWL-Modell, welches bei 79 Prozent der betrachteten Ansätze zum Einsatz kommt (vgl. u. a. Kaukolu, Rhizome, SweetWiki)
- **Import/Export** Der Im- und Export von Daten und Ontologien spielt vor allem im Unternehmensbereich eine wichtige Rolle, da man die im Wiki geschaffenen Informationen in anderen Programmen weiterverwenden kann, ohne die Struktur zu verlieren.
 - Circa 60 Prozent der RDF(S)/OWL-zentrischen Ansätze erwähnen die Möglichkeit des Austauschs der über diese Standards geschaffenen Konzepte und konkreten Daten, um diese dann z. B. in Ontologieeditoren weiter zu optimieren

¹Anmerkung 1.: In den folgenden Unterabschnitten werden immer wieder beispielhaft Ansätze genannt, auf die zuvor formulierte Aussagen zutreffen. Aus Gründen der Übersicht wurden diesen nicht zusätzlich noch die entsprechenden Quellen angehängt. Der Rückschluss auf die verwendeten Quellen ist aber ganz einfach über die Tabellen möglich, indem man über den Namen des Ansatzes (Zeile) und die aktuell betrachtete Eigenschaft (Spalte) zur entsprechenden Stelle der Einordnung navigiert.

²Anmerkung 2.: Die in drei Oberkategorien aufgeteilten Tabellen mussten aufgrund von Platzproblemen gestückelt und auf mehrere Seiten verteilt werden. Die Gesamtheit der Tabellen in Anhang A gehören zum Teil „Struktursystem“, die von Anhang B zu „Eingabe“ und die von Anhang C zu „Darstellung, Navigation und Suche“. Es wird jeweils, wie an dieser Stelle bereits geschehen, nur einmal zu Beginn auf den entsprechenden Abschnitt im Anhang zur Orientierung verwiesen und nicht einzeln referenziert.

³dient zur semantischen Datenmodellierung

oder bereits vorhandene domänenspezifische Abbildungen aus dem Wissensmanagement in ein Wiki zu importieren.

- In drei Ansätzen (vgl. Hybrid Wikis, MoKi, OntoWiki) werden spreadsheetartige Formate (z. B. das .xlsx-Format von Microsoft Excel oder das .csv-Format, in dem Daten über Kommas separiert werden) zum Einsatz gebracht, weil in ihnen strukturierte Datensätze weitestgehend erhalten bleiben (Links gehen allerdings verloren).
- Hybrid Wikis besitzt als einziger Ansatz zusätzlich die Möglichkeit des Austauschs der Inhalte (Konzepte und Daten) über das Eclipse Modeling Framework (EMF).
- **Strukturelemente** Prinzipiell besitzt jeder Ansatz mehrere oder alle der folgenden Strukturelemente.

- **Attribute und Relationen** Etwa 90 Prozent der Ansätze beschreiben im weitesten Sinne eine Verwendung von Attributen und Relationen zur Darstellung von strukturierten Daten. Attribute sind Schlüssel-Werte-Paare, Prädikat-Objekt-Statements in RDF(S)/OWL-zentrischen Ansätzen (vgl. SnoopyDB) oder Prädikate und Objekte im grammatikalischen Sinne in Bezug auf natürliche Sprache (vgl. AceWiki, CNLWiki). Schlüssel sind in der Regel einfache Text-Literale. Im RDF-Sinne können sie aber auch wieder Links zu Wiki-Seiten bilden, in denen sie näher beschrieben werden. Werte sind zunächst einmal wie ihre Schlüssel einfache Text-Literale oder Links. Sind Werte Links zu internen oder externen Seiten, spricht man von Relationen (vgl. Semantic MediaWiki), die durch den Schlüssel beschrieben (oder „typisiert“; vgl. KiWi (IkeWiki)) werden.

Eine weitere wichtige Funktionalität ist die Mehrfachzuweisung von Werten zu Attributen (erwähnt in drei Ansätzen), um die redundante Verwendung von Schlüsseln zu vermeiden (so können z. B. mehrere Autoren zu einem Buch hinzugefügt werden, ohne jedes mal den gleichen Attribut-Namen pro Autor verwenden zu müssen; vgl. SweetWiki, OntoWiki, Social Infobox, Hybrid Wikis).

Eine Sortierungsfunktion für Werte und generell Attribute anhand ihrer Schlüssel kann dazu dienen, die Wichtigkeit einzelner Angaben durch die Festlegung einer Reihenfolge hervorzuheben (vgl. Hybrid Wikis, SnoopyDB) und für mehr Übersicht zu sorgen (alphabetische Sortierung; vgl. Hybrid Wikis).

Text-Literale bei Werten können zusätzlich als bestimmter Datentyp interpretiert werden, zum Beispiel als boolesche Variablen (wahr oder falsch), Währungen oder Datumsangaben.

Speziell dafür ausgerichtete Schlüssel können semantische Funktionen erfüllen. Verwendung finden unter anderem:

- * „is-a“ als Vererbungsbeziehung (vgl. u. a. WikSAR (SHAWN))
- * „partOf“ als Aggregationsbeziehung (vgl. u. a. WikSAR(SHAWN))
- * „seeAlso“ um inhaltliche Nähe zu anderen Inhalten auszudrücken (vgl. u. a. SweetWiki)

- * Alias-Definitionen als Synonyme für eindeutige Bezeichner von Wiki-Seiten (vgl. Freebase) und auch Attribute (vgl. Kaukolu)
- * (Akronyme als Kürzel für Seiten-Titel)

Hierarchien und Aggregationen dienen zur verbesserten Navigation, Synonyme und Abkürzungen zur Begriffsklärung während der Eingabe in Form von Vorschlägen, falls ein Autovervollständigungsmechanismus vorhanden ist.

- **Typen** Klassische Tags werden durch den Nutzer frei vergeben, bilden eine Folksonomie und beschreiben den Inhalt, dem sie zugewiesen werden (vgl. Abschnitt 2.1.3). Typen sind eine Sonderform von Tags, da sie zusätzliche Aufgaben erfüllen. Sie dienen zur Schemabildung und können u. a. über Vererbungsbeziehungen in Relation zueinander gesetzt werden (Übergeordnete Typen; vgl. Freebase). OWL als Semantic Web-Standard bietet noch weitaus mehr Optionen zur Beschreibung von Beziehungen zwischen Typen bzw. deren Klassen wie Äquivalenz und Disjunktion⁴. Da anhand der Quellen nie sicher gesagt werden kann, inwieweit diese Möglichkeiten voll ausgereizt werden, wurden die Quellen nur auf das Vorhandensein von Vererbungsbeziehungen für Typen geprüft. Etwa 74 Prozent der Ansätze, die Typen prinzipiell verwenden, erlauben auch eine Hierarchisierung dieser.

Eine zusätzliche Option ist die Mehrfachtypisierung von Wiki-Seiten. Der Gebrauch dieser Methode wird in drei Ansätzen (vgl. Hybrid Wikis, Freebase, KiWi (IkeWiki)) beschrieben.

- **implizite und explizite Schemata** Einzelne Typen dienen als Bezeichner für eine Klasse oder ein Schema, das eine Reihe von Attributen bündelt. Ist eine feste Definition eines solchen Bündels vorhanden, spricht man von einem expliziten Schema (Schemata; vgl. Freebase oder OWL-Klassen in RDF(S)/OWL-zentrischen Ansätzen; vgl. u. a. Semantic MediaWiki), das bei 53 Prozent der Ansätze thematisiert wird. Fehlen Typen zur festen Kopplung, werden Wiki-Seiten aber trotzdem Attribute angeheftet, kommt es auf Dauer zu einer impliziten Formung von Schemata, bei denen die Häufigkeit der in Kombination verwendeten Attribute ausschlaggebend zur Bildung beitragen (vgl. SnoopyDB, Social Infobox). Je mehr Attribute einer Wiki-Seiten zugewiesen und je mehr Wiki-Seiten zur Beschreibung von real existierenden Einheiten (z. B. eine Beschreibung zum „Eiffelturm“) vorhanden sind, desto genauer kann man das implizite Schema, das sich dahinter verbirgt, erkennen und voraussagen. Diese impliziten Schemata werden dann über Empfehlungen und Vergleiche kommuniziert (vgl. SnoopyDB, Social Infobox). Werden Tags und Attribute zu einem Schema kombiniert, kann diese Vorgabe in Form einer festen Definition verpflichtend sein. Ein Sonderfall, der nur bei Hybrid Wikis seine Verwendung findet, ist die Kombination aus impliziter und expliziter Schemabildung. Hier existieren sowohl fest definierte Schemata als auch über einen Vergleich verwandter Seiten abgeleitete Vorschläge für zusätzliche Attribute. Hinzu kommt die einzigartige Tatsache, dass alle Schemata flexibel sind und Typen mit Attributen nur lose gekoppelt. Dies bedeutet, dass zur Laufzeit Attribute frei mit Ty-

⁴für genauere Informationen zur Mächtigkeit von OWL vgl. [OWL04]

pen, unabhängig von festgelegten Schemata, kombiniert werden können, was bei Ansätzen, die eine explizite Definition der Schemata vorsehen, in der Regel nicht möglich ist. Man kann zwar nicht verwendete Attribute leer lassen, aber keine neuen zur Laufzeit hinzufügen. Außerdem werden bei der Mehrfachtypisierung von Seiten Attribute streng nach Vorgabe pro Typ gegliedert (vgl. Freebase). Abbildung 4.1 fasst diese Unterschiede noch einmal veranschaulicht zusammen.

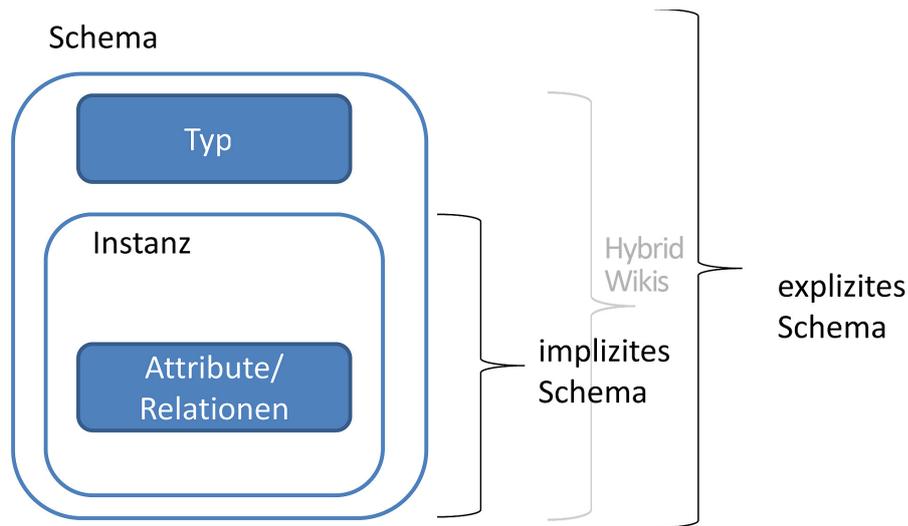


Abbildung 4.1.: Der Unterschied zwischen implizitem und explizitem Schema (Eigene Darstellung)

- **Freitext und/oder Struktur in Wiki-Seiten** In klassischen Wikis erfolgt die Beschreibung einer Einheit durch einen Freitext, der nur über einfache Formatierungsoptionen wie Paragraphen gegliedert ist. Die hier betrachteten Ansätze verzichten entweder ganz auf einen solchen Freitext und verwenden nur einen Strukturteil (42 Prozent) oder kombinieren beides (42 Prozent). Die Kombination aus Freitext und Struktur vereint die Ausdrucksstärke beider Konzepte: der Freitext kann über eine Schlagwortsuche erschlossen werden, Strukturen erlauben eine gezielte Abfrage anhand von Details. Außerdem kann auf diese Art und Weise die Kollaboration zweier Zielgruppen von Nutzern sichergestellt werden: Wiki-Autoren und Domänenexperten, die es gewohnt sind, über einen WYSIWYG-Editor im Wikipedia-Stil Beschreibungen zu verfassen und Ontologieexperten, die ihren Fokus auf eine möglichst starke Präzisierung ihrer Inhalte legen.

Existieren sowohl Freitext als auch Strukturteil, gibt es zwei Möglichkeiten wie diese miteinander zusammenhängen können. Entweder wurden über semantische Annotationen im Freitext Attribute und/oder Relationen geschaffen (zu 75 Prozent), die dann zusammengefasst dargestellt werden (z. B. über eine Factbox; vgl. Semantic MediaWiki) oder der Strukturteil wurde separat und unabhängig zum Beschreibungstext erstellt (zu 25 Prozent). Abbildung 4.2 fasst den erkannten Zusammenhang noch einmal zusammen.

- **Integritätsbedingungen** Um Inkonsistenzen einzudämmen setzen 42 Prozent der

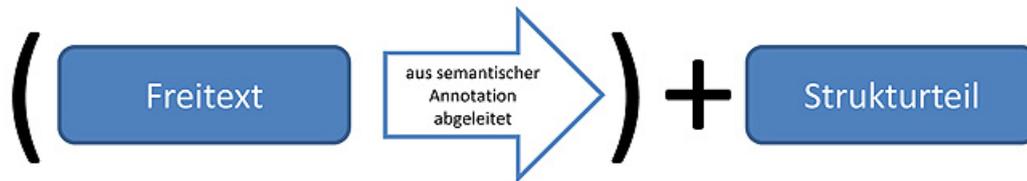


Abbildung 4.2.: Zusammenhang zwischen Freitext und Strukturteil in Wiki-Seiten (Eigene Darstellung)

Ansätze gezielt Integritätsbedingungen ein. Diese sind eng verzahnt mit Definitionen von Attributen und Typen bzw. ihren Schemata und tragen zur Gesamtkonsistenz der Strukturen innerhalb eines Systems bei.

Die am häufigsten verwendeten Integritätsbedingungen sind:

- Festlegung der Multiplizität von Attributen und Relationen (vgl. Multiplizitäten in Hybrid Wikis und Kardinalitäten in RDF(S)/OWL-zentrischen Ansätzen))
- Beschränkung auf Wertebereiche (z. B. bei numerischen Werten) und Enumerationen (Aufzählungen mit Auswahlmöglichkeit aus n Elementen; vgl. Artificial Memory, Hybrid Wikis) für die Auswahl bei Werten
- Beschränkung auf einen bestimmten Datentyp für Werte (vgl. SnoopyDB)
- Beschränkung auf ein bestimmtes Format zu einem Datentyp (Beispiel: dd.mm.yyyy für Datumsangaben; vgl. SnoopyDB)
- erwartete Typen für Zielseiten als Sonderfall der Typisierung von Links als Datentyp (vgl. Freebase, Artificial Memory, Hybrid Wikis)

In einem Fall wird die Generierung von Einschränkungen für erlaubte Wertetypen und die zugehörige Formatierung über eine Häufigkeitsanalyse der Verwendungen automatisch durch das System erzeugt und propagiert (vgl. SnoopyDB).

Wie Verletzungen gehandhabt werden, wird nur im Fall von Hybrid Wikis beschrieben, nämlich „weich“ in Form von Hinweisen.

4.2. Eingabe

Die zugehörige Einordnung der Quellen in Tabellenform befindet sich in Anhang B.

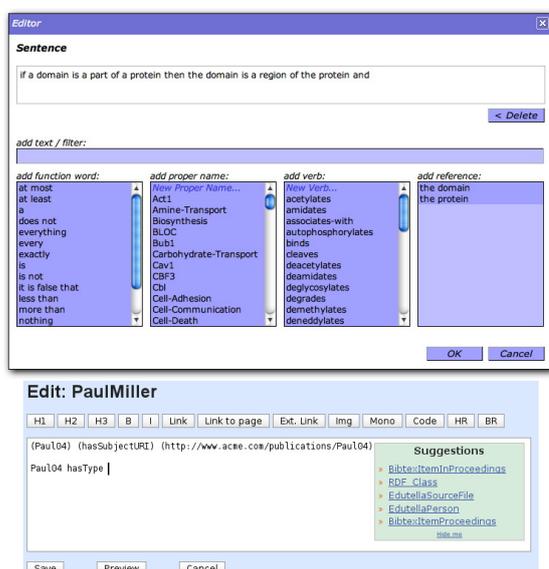
4.2.1. Extrahierte Eigenschaften

- **Arten der Eingabe**

- **UI-basiert** Mit rund 47 Prozent verzichten fast die Hälfte der Ansätze auf eine Syntax oder spezielle Notation zur Eingabe für Strukturelemente zugunsten der Usability und bietet stattdessen vordefinierte Formularfelder, kombiniert mit verschiedenen Unterstützungsmechanismen, zur Eingabe an. Diese Methode ermöglicht eine Abstraktion von komplexeren Grundmodellen wie RDF(S) und

OWL und erlaubt den Nutzern dadurch eine einfachere Schaffung von Strukturen.

- **Syntax-basiert** Eine eigene Sprache zur Formung von Strukturdaten setzen 32 Prozent der Ansätze ein. Eine weitere, sehr unkomplizierte Form der Syntax-basierten Eingabe, die zweimal beschrieben wurde, ist die Verwendung einer Doppelpunkt-Notation (z. B. „liegtIn: Bayern“; vgl. WikSAR (SHAWN)) zur Bildung von Schlüssel-Werte-Paaren. Die dritte identifizierte Form der Syntax-basierten Eingabe funktioniert über die Verwendung eines begrenzten Wortschatzes (z. B. Attempto Controlled English (ACE); vgl. AceWiki), um über grammatikalisch korrekte Sätze Tatsachen auszudrücken (z. B. „München ist Landeshauptstadt von Bayern und liegt nicht in Österreich.“). Diese Form ist sehr mächtig, da sie die gesamte Vielfalt einer natürlichen Sprache berücksichtigt (z. B. Verneinung, Symmetrie, Transitivität, ...; vgl. CNL Wiki). Abbildung 4.3 liefert einen Überblick über die abgegrenzten Eingabeformen.



HomerSimpson

- `Something > LivingThing > Someone` : `HomerSimpson`
- `PartOf` : `TheSimpsons`
- `WorksAt` : `PowerPlant`
- `LivesIn` : `SpringField`
- `InterestsIn` : `DuffBeer`

Homer Jay Simpson (voiced by Dan Castellaneta), is one of the main characters in the animated television series; `TheSimpsons`.

Homer works in the Springfield Nuclear `PowerPlant`, although "working" in this case refers to dozing and eating doughnuts. He spends a great deal of his time at Moe's Tavern with his lifelong friends `BarneyGumble`, Carl, Lenny, and `MoeSzyzslak`, the bartender.

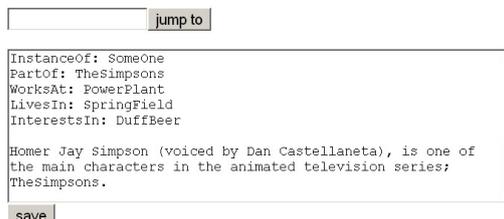


Abbildung 4.3.: Syntax-basierte Eingabe über kontrollierte natürliche Sprache in AceWiki [Kuh08b] (links oben), einer eigenen Notation in Kaukolu [Kie06] (links unten) und über eine einfache Doppelpunkt-Notation in WikSAR (SHAWN) [Aum05] (rechts)

- **Unterstützungsmechanismen** Der Bereich der Unterstützungsmechanismen ist in der Betrachtung mit am interessantesten, da hier über verschiedene Mechanismen präventiv oder nachträglich Nutzereingaben und daraus entstehende Strukturen optimiert werden.
 - **Autovervollständigung** Die Autovervollständigung unterstützt den Nutzer während der Eingabe für ein bestimmtes Feld durch Vorschläge (in Hybrid Wikis als Empfehlungen bezeichnet) und versucht die geplante Eingabe anhand vorhandener Daten vorauszusagen. Zusätzlich dient die Autovervollständigung zur Begriffsklärung, da sie bestenfalls die Verwendung synonyme oder homonyme Begriffe und von Pluralismen verhindert. Die Vorschlagsliste enthält nach Relevanz sortierte Voraussagen für die Eingabe. Außerdem findet eine

Priorisierung der zur Verfügung stehenden Begriffe anhand von häufiger Verwendung im Kontext, von Vorabdefinitionen oder erwartete Typen statt. Ungefähr 78 Prozent der Ansätze, die auf eine UI-basierte Eingabe setzen, bieten auch einen Autovervollständigungsmechanismus an.

- **Empfehlungen** Empfehlungen (in Hybrid Wikis als Vorschläge bezeichnet) unterstützen den Strukturierungsprozess durch die Auflistung zusätzlicher Strukturelemente, die bisher noch nicht auf der aktuellen Wiki-Seite verwendet wurden, diese aber über einen Kombinationsvergleich auf anderen Seiten als zutreffend erfasst wurden. Lediglich drei der betrachteten Ansätze bieten Empfehlungen als Form der Unterstützung an.
- **Auflistung ähnlicher/verwandter Wiki-Seiten** Neben Empfehlungen für zusätzliche Strukturelemente gibt es in Social Infobox eine Auflistung verwandter Wiki-Seiten, die über einen Klick zu einem direkten Vergleich als Hilfe gegenübergestellt werden können. Diese Liste wird umso genauer, je mehr Attribute der aktuellen Seite hinzugefügt werden, da sich dadurch das implizite Schema genauer voraussagen lässt.
- **Inferenzen** Inferenzen sind einfach gesagt logische Schlussfolgerungen, die anhand bereits gebildeter Aussagen abgeleitet werden. Angenommen, es existieren zwei Städte „A“ und „B“. Zusätzlich steht fest: „A liegt in Amerika“ und „B liegt in Europa“. Außerdem ist durch eine entsprechende Angabe bekannt: „Es existiert keine Zugverbindung zwischen Amerika und Europa“. Eine Schlussfolgerung wäre nun, dass es zwischen den Städten „A“ und „B“ auch keine Zugverbindung geben kann. Zur Ableitung solcher Folgerungen gibt es spezielle Reasoning-Engines. Umfassendes Reasoning und inferenzierte Schlussfolgerungen setzen nur die beiden Ansätze ein, die auch kontrollierte natürliche Sprache verwenden (vgl. AceWiki, CNL-Wiki). Einfache Inferenzen oder Schlussfolgerungen kommen hingegen zum Einsatz, um Typen anhand von Relationen zwischen Wiki-Seiten abzuleiten (vgl. Freebase, KiWi (IkeWiki)).
- **automatisierte Informationsextraktion** Für die automatisierte Informationsextraktion gilt im Prinzip das gleiche wie für Inferenzen. Bisher noch wenig verwendet (zwei Ansätze), könnte sie in Zukunft für etliche Aufgaben in Wikis eingesetzt werden, zum Beispiel für das Extrahieren und Projizieren von Entitäten aus einem Freitext auf Konzepte in einer Ontologie, das automatisierte Erstellen von Zusammenfassungen für Freitexte und die Bildung von Links über eine Analyse im System vorhandener Wiki-Seiten anhand ihrer Titel. Da die zugrundeliegenden Algorithmen ähnlich wie beim Reasoning sehr komplex sind und ein eigenes Forschungsgebiet bilden, wird an dieser Stelle auf den Beitrag [SS09] der Macher von KiWi (IkeWiki) verwiesen, der den gezielten Einsatz der automatisierten Informationsextraktion in Wikis genauer ausführt.
- **Gezielte Zurverfügungstellung einer Struktur-/Datenbasis** Der Import und die Zurverfügungstellung von mitunter domänenspezifischen Typen, Attributen und vorhandenen Daten kann dem Nutzer den Einstieg in den Strukturierungsprozess erleichtern. Diese können als Grundlage kollaborativ optimiert

und erweitert werden (vgl. Freebase).

- **manuelle Konsistenzprüfung** Um Redundanzen zu verhindern, können sich ähnelnde Konzepte oder Inhalte einer Wiki-Seite über ein besonderes Attribut zur Bildung einer Relation verknüpft und als solche markiert werden. Die Gültigkeit der Markierung muss im Anschluss manuell geprüft und behoben werden (vgl. Freebase). Ein zweiter Weg ist die Anheftung von Kommentaren an einzelne Attribute. Diese Methode dient zur Hinweisgebung bei Fehlern in der Belegung mit Werten. Solche Kommentare können ebenfalls wieder auf ihre Existenz und Gültigkeit hin geprüft und behoben werden (vgl. OntoWiki).
- **kontextspezifische Widgets als Eingabehelfer** Je nach Datentyp für die Belegung eines Wertes können Widgets dazu dienen, die Eingabe zu erleichtern und die Einhaltung von Formatierungen, die später bei einer Suche von Bedeutung sind, zu forcieren. Die einzigen beiden genannten Hilfen sind Kalender-Pop Ups und Karten zur Markierung von Orten (vgl. OntoWiki, HYENA).
- **Inline-Editierung von Schlüssel-Werte-Paaren** Die Inline-Editierung ist eine AJAX-basierte Variante zur Bearbeitung von Werten und den Schlüsseln mehrerer Wiki-Seiten. Werden über entsprechende Sichten oder die Suche Seiten des gleichen Typs und mit gleichen Attributen herausgefiltert, kann man durch einen einfachen Klick in das entsprechende Wertefeld die Belegung horizontal nacheinander ändern, ohne dabei die jeweilige Wiki-Seite einzeln aufrufen zu müssen. Dies ist eine bequeme Methode, um schnell und effizient Belegungen von Attributen mit dem gleichen Schlüssel anzupassen, falls sich Gegebenheiten geändert haben (z. B. zeitliche Verschiebungen haben die Raumzuweisungen zu vielen Vorlesungen beeinflusst und diese müssen nun alle angepasst werden; vgl. hierzu OntoWiki, Hybrid Wikis).

4.3. Darstellung, Navigation und Suche

Die zugehörige Einordnung der Quellen in Tabellenform befindet sich in Anhang C.

4.3.1. Extrahierte Eigenschaften

- **Sichten**

- **Attribut-Sicht** Die Attribut-Sicht gibt einen Überblick der im System vorhandenen Attribute und in welchem Zusammenhang sie verwendet werden. In der Regel findet man spaltenweise separiert Wiki-Seiten, denen das Attribut zugewiesen wurde, sowie die konkreten Belegungen mit Werten.
- **Typ-Sicht** Ähnlich wie bei der Attribut-Sicht gibt es Übersichten für Typen. Falls entsprechende explizite Schemata existieren werden die zugehörigen Attribute pro Typ aufgelistet, ebenso wie die Wiki-Seiten, die diese verwenden.
- **Vergleichssicht** Zur Gegenüberstellung zweier, über eine Wiki-Seite strukturierte, Ressourcen kann die Vergleichssicht dienen. Diese können anhand ihrer

Eigenschaften und deren Belegungen direkt gegenübergestellt werden, um Gemeinsamkeiten und Unterschiede festzustellen und Angaben zu erweitern und optimieren (vgl. Social Infobox).

- **kontextspezifische Sichten** Alle Wiki-Seiten können gesammelt und abhängig vom Kontext über einen verwendeten Datentyp in einer Sicht dargestellt werden. Wie schon bei den Widgets wurde eine Kalenderansicht zur Aggregation von Datumsangaben (vgl. HYENA, OntoWiki) sowie eine Kartenansicht zur Sammlung von Positionsangaben (vgl. OntoWiki) aus Wiki-Seiten heraus gefunden.
- **Einbettung von Attributen** In OntoWiki werden Attribute an Links angeheftet und angezeigt (Inline Reference View), falls diese auf eine interne Wiki-Seite zeigen, die Attribute besitzt. So erhält man im Voraus schon zusätzliche Information zur verlinkten Seite, ohne diese konkret aufrufen zu müssen.
- **Navigation** Die Navigation erfolgt in erster Linie über Links oder zentrale Suchfelder. Davon abgesehen gibt es noch zwei weitere Möglichkeiten, die identifiziert wurden.
 - **Inverse Links** Links sind normalerweise uni-direktional, sprich sie zeigen von einer Wiki-Seite auf eine andere und existieren nur als ausgehende Links. Durch die zusätzliche Anzeige eingehender Links als Referenzierungen auf der Zielseite erfolgt eine bidirektionale Verknüpfung. 37 Prozent der Ansätze setzen diese Methode ein, um zusätzlich eine Rückwärts-Navigation zu erlauben.
 - **Typenhierarchien** Voraussetzung für eine Navigation über hierarchisierte Typen ist das Vorhandensein von Vererbungsbeziehungen zwischen eben diesen in Form von Typ-Subtyp-Relationen. Schafft man es, alle Typen auf einige wenige Basistypen zu verallgemeinern, können diese als zentrale Einstiegspunkte („Kategorien“) der Navigation dienen (vgl. Freebase).
- **Suche**
 - **Arten der Suche** Für die Arten von Suchen gilt das gleiche wie für die Eingabe: entweder ist sie UI-basiert oder man formuliert die Abfragen über die Syntax einer Sprache.
 - * **UI-basiert** Die UI-basierte Suche wird durch ein Suchfeld realisiert, über das per Schlagwörter eine Durchsuchung von Freitexten (Volltextsuche) und den Text-Literalen des Strukturteils stattfinden kann. Weiterhin gibt es als Zusatz eine facettierte Suche, die über einzelne vorhandene Strukturelemente die Filterung beziehungsweise Eingrenzung der Ergebnismenge erlaubt (vgl. z. B. OntoWiki, Rhizome). Orientiert sich diese Eingrenzung zusätzlich an der Abstufung der Strukturelemente (vgl. z. B. Hybrid Wikis: ... → Typ Tag → Attribut-Name → Wert) spricht man von einem Drill Down.
 - * **Syntax-basiert** Die SPARQL ist mit der Nennung in den Quellen von vier Ansätzen die beliebteste Syntax-basierte Abfragesprache. Sie ergänzt die Semantic Web-Standards RDF(S)/OWL. Die JSON/HTTP-basierte API von

Freebase, aufgeführt als eigene Abfragesprache, dient neben der Erstellung und Bearbeitung gleichzeitig auch zur Formulierung von Queries auf den Inhalten. Kontrollierte natürliche Sprache kann neben der Schaffung von Fakten auch zur Abfrage dieser dienen (vgl. Abbildung 4.4).

```
▶ What borders Spain?  
- Andorra  
- Atlantic-Ocean  
- France  
- Mediterranean-Sea  
- Portugal
```

Abbildung 4.4.: Über kontrollierte natürliche Sprache formulierte Suchanfrage in AceWiki [Kuh08b]

- **Einbettung von Suchergebnissen** Drei Ansätze bieten die Option der Einbettung von Suchergebnissen in Wiki-Seiten (vgl. Hybrid Wikis, HYENA, Semantic MediaWiki).

4.4. Beurteilung von Hybrid Wikis im Vergleich

Durch den gegenüberstellenden Vergleich anhand der Eigenschaften der einzelnen Ansätze mit Hybrid Wikis lässt sich erkennen, welche gängigen Konzepte und Methoden im Allgemeinen angewandt werden und welche Hybrid Wikis noch fehlen. Diese Lücken in der Tabelle wurden mit Farbindikatoren versehen, um dadurch verschiedene Sachverhalte auszudrücken. Eine Legende erklärt die verwendeten Indikatoren, der Vollständigkeit halber werden sie aber hier noch einmal ausführlich erklärt.

- **Alleinstellungsmerkmale** Alleinstellungsmerkmale sind in der Tabelle durch eine grüne Umrandung hervorgehoben. Sie bilden Eigenschaften, die im direkten Vergleich so nur bei Hybrid Wikis zu finden waren.

Diese sind:

- der Im- und Export von Konzepten und Daten über Formate des Eclipse Modeling Framework (EMF)
- die lose Kopplung von Typen und Attributen und die daraus entstehende Möglichkeit zur freien Kombination zur Laufzeit

- **Nicht vorhandene Eigenschaften** Nicht vorhandene Eigenschaften besitzen einen roten Hintergrund, fehlen ganz und können vom Autor auch nicht als inzwischen implementiert bestätigt werden. Sie dienen teilweise als Grundlage für die in Kapitel 5 formulierten Gestaltungsempfehlungen.

Diese sind:

- (kein ERM- oder RDF(S)/OWL-zentrisches Datenmodell)
- fehlende Möglichkeiten zum Im- und Export in Semantic Web-Standards wie RDF(S)/OWL
- keine semantisch-funktionalen Attribute (nur einfache Text-Literale)
- keine Vererbung zwischen Typen

- keine Möglichkeit der semantischen Annotation von Links oder Bildung von Attributen im Freitext
 - keine automatisierte Bildung von Integritätsbedingungen durch Häufigkeitsanalysen
 - (keine harten Fehlermeldungen für Integritätsverletzungen)
 - keine Syntax-basierte Eingabe (korrespondiert mit der fehlenden Möglichkeit zur semantischen Annotation)
 - keine zusätzliche Auflistung ähnlicher/verwandter Wiki-Seiten zur Verbesserung der Navigation oder zu einem Vergleich
 - keine einfachen logischen Inferenzen
 - keine automatisierte Informationsextraktion aus dem Freitext
 - keine Möglichkeit der Hinweisgebung auf Inkonsistenzen durch den Nutzer bei der Verwendung von Strukturelementen
 - keine Vergleichssicht zur direkten Gegenüberstellung zweier Wiki-Seiten und ihrer Strukturelemente zur Optimierung
 - keine kontextspezifischen Sichten/Visualisierungen von bestimmten Datentypen wie Datumsangaben oder Positionsangaben
 - keine Einbettung von Attributen zu Links, die die Zielseite beschreiben
 - keine Navigationsmöglichkeit über Typenhierarchien als Quasi-Kategorien (korrespondiert mit der fehlenden Vererbung)
 - keine Syntax-basierten Abfragen möglich
- **Teilweise vorhandene Eigenschaften** Teilweise vorhandene Eigenschaften (markiert mit einem gelben Hintergrund) sind Eigenschaften, die seit der Veröffentlichung der Quellen zu Hybrid Wikis implementiert und deshalb in ihnen noch nicht beschrieben wurden. Ihr Vorhandensein in einer grundlegenden aber noch ausbaufähigen Form kann aber vom Autor bestätigt werden.

Diese sind:

 - Überprüfung des Formats von Datentypen (für eine Basismenge vorhanden)
 - Widgets als Helfer bei der Eingabe (für Datumsangaben vorhanden)
 - **Vorhandene Eigenschaften** Vorhandene Eigenschaften - grün hinterlegt - wurden dem Wissen des Autors nach in der Zwischenzeit komplett implementiert oder fanden keine Erwähnung in den vorliegenden Quellen.

Diese sind:

 - Inline-Editierung von Schlüssel-Werte-Paaren (über Typ-Sicht)
 - Einführung einer zusätzlichen Attribut-Sicht (noch in Entwicklung)

5. Gestaltungsempfehlungen für Hybrid Wikis

Die Katalogisierung, Analyse und Extraktion von Konzepten und Methoden der gesammelten Ansätze hat zusätzlich zur Entwicklung von Gestaltungsempfehlungen für Hybrid Wikis geführt („Synthesis of studies“). Die im Folgenden ausgeführten Vorschläge haben ihren Ursprung in den betrachteten Quellen oder sind eigene Ideen, die auf Basis des gewonnenen Wissens formuliert wurden und über die das Strukturieren von Inhalten weiter gefördert und optimiert werden kann.

5.1. Struktursystem

5.1.1. Erweiterte Beschreibung von Link-Relationen

Schlüssel sind in Hybrid Wikis bisher nur einfache Text-Literale ohne weitere Bedeutung. Ist der zugehörige Wert aber ein Link, wird über den Schlüssel eine gerichtete Relation beschrieben, der man in manchen Fällen zusätzliche Attribute zur Beschreibung anheften könnte, ähnlich wie bei dem aus der UML-Modellierung bekannten Konzept der Bildung von Assoziationsklassen (vgl. Abbildung 5.1).

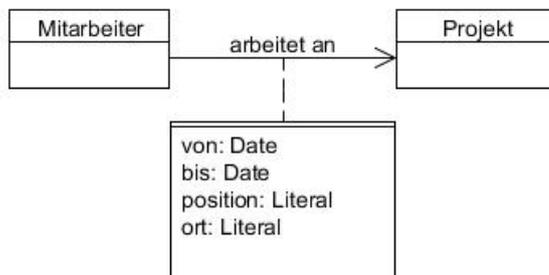


Abbildung 5.1.: Assoziationsattribute zweier Typ Tags in UML (Eigene Darstellung)

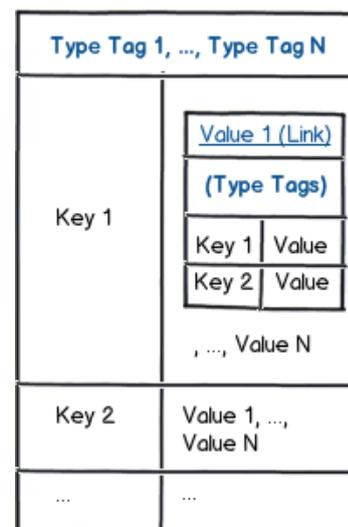


Abbildung 5.2.: Konzeptuelles Mock-Up für Link-Attribute und optional angeheftete Typ Tags (Eigene Darstellung)

Angenommen, es existiert eine Definition für das Attribut „Teilnahme“. Dieses wird dem Typ Tag „Mitarbeiter“ zugewiesen, der Wert soll ein Link zu einer Zielseite vom Typ „Projekt“ sein. Durch diese Festlegung entsteht eine Relation zwischen zwei Typ Tags, die durch das Attribut beschrieben wird. Dem Schlüssel oder dem Link der gerichteten Relation könnte man in diesem Zusammenhang jetzt zusätzliche Attribute (und Typ Tags) anheften, die diese näher beschreiben (z. B. den Ort und den Zeitraum der Teilnahme). In einer Seite, über die ein Mitarbeiter und seine Teilnahme an einem Projekt beschrieben wird, werden per Mouse Over-Tooltip oder in gewohnter zweisepaltiger Ansicht (siehe Abbildung 5.2) dann zusätzliche Informationen zur Teilnahme an dem bestimmten Projekt angezeigt, falls diese angegeben wurden.

Vorteile/Nutzen:

- Möglichkeit der detaillierten, zusätzlichen Beschreibung einer Relation zwischen zwei Wiki-Seiten eines bestimmten Typs

5.1.2. Gruppierete Attribute

Besitzt eine zu annotierende Wiki-Seite viele verschiedene Eigenschaften, wird die Liste der Attribute schnell sehr lang und unübersichtlich. Hinzu kommen eingehende Referenzen, die die Tabelle zusätzlich verlängern.

Abhilfe könnte die Gruppierung einzelner zusammengehöriger Attribute „on-the-fly“ während der Eingabe oder vorab, ähnlich wie bei der Definition von Typ Tags, stattfinden. „Straße“, Hausnummer“, „Postleitzahl“ und „Ort“ könnte man zu einer Attribut-Gruppe „Anschrift“ kombinieren. Typ Tag-Definitionen sind zur Kapselung von Attributen unzureichend, da sowohl eine Person als auch ein Unternehmen eine Anschrift besitzt kann, man aber eine Instanz von beiden nie mit dem Typ Tag „Anschrift“ versehen würde, da diese ja nicht die Entität direkt beschreiben.

Die Kapselung oder Gruppierung von Attributen findet man in den Templates von MediaWiki (siehe Abbildung 5.1.2). Eine zusätzliche Option zum Ein- und Ausklappen einzelner Gruppen wäre wünschenswert.

Vorteile/Nutzen:

- Kapselung zusammengehöriger Attribute, die nicht über ein Typ Tag gegliedert werden können/sollten

Eigenschaften	
Allgemein	
Name, Symbol, Ordnungszahl	Calcium, Ca, 20
Serie	Erdalkalimetalle
Gruppe, Periode, Block	2, 4, s
Aussehen	silbrig weiß
CAS-Nummer	7440-70-2
Massenanteil an der Erdhülle	3,39 % ^[1]
Atomat ^[2]	
Atommasse	40,078 u
Atomradius (berechnet)	180 (194) pm
Kovalenter Radius	176 pm
Van-der-Waals-Radius	231 ^[3] pm
Elektronenkonfiguration	[Ar] 4s ²
Austrittsarbeit	2,87 eV ^[4]
1. Ionisierungsenergie	589,8 kJ/mol
2. Ionisierungsenergie	1145,4 kJ/mol
Physikalisch ^[5]	
Aggregatzustand	fest
Kristallstruktur	kubisch flächenzentriert
Dichte	1,55 g/cm ³ (20 °C) ^[6]

Abbildung 5.3.: Templateinstanz von „Calcium“ mit gruppierten Eigenschaften in Wikipedia

- Verbesserung der Übersicht

5.1.3. Verschachtelte Attribute

Eine Idee zur „horizontalen“ Kapselung von Attributen wäre die Verschachtelung einzelner Attribute zu einem Gesamtattribut.

Das Konzept ist relativ einfach. Man verwendet bei der Attribut-Definition neben den Basistypen Boolean, Date, usw. zusätzlich bereits vorhandene Attribut-Namen als Wertetypen und kombiniert diese zu einem Record. Im Gegensatz zum Vorschlag in Abschnitt 5.1.1 gibt es also keinen Wert in Form eines Links, an den weitere Attribute und optionale Typ Tags angeheftet werden. Die Attribute werden stattdessen direkt als Belegung eines Schlüssels verwendet. Das Mock-Up in Abbildung 5.4 verdeutlicht den erläuterten Unterschied. Greift man das obige Beispiel auf, könnte man also ein Attribut „Anschrift“ definieren, das mit Werten der Attribute „Straße“, Hausnummer“, „Postleitzahl“ und „Ort“ belegt wird (Anschrift:= {Straße, Hausnummer, Postleitzahl, Ort}).

Vorteile/Nutzen:

- Verbesserung der Übersicht
- Bildung anonymer Attribute

5.1.4. Vorlagen für Wiki-Seiten

Die Einführung von benutzerdefinierten Vorlagen für Wiki-Seite würde unerfahrenen Nutzern eine Einstiegshilfe zum Verstehen von (Typ) Tags und Attributen und Experten eine Option zur Festlegung oft verwendeter Kombinationen der Strukturelemente bieten. Das eigentliche Konzept einer Wiki-Seite bleibt erhalten, jedoch lassen sich kreative Templates (vgl. Abschnitt 2.1.4.3) vorab definieren, die einen Namen bekommen und eine leere Wiki-Seite mit vordefinierten Tags, Typ Tags und Attributen bilden (z. B. ein Template „Mitarbeiterprofil“ mit den Typ Tags „Person“ und „Mitarbeiter“ und damit verbundene Attribute wie „Name“, „Alter“, „Position“, „Büro“, usw.). In der Instanz müssten diese dann nur noch ausgefüllt werden. Diese Vorabtypisierung von Seiten würde den gesamten Erstellungsprozess beschleunigen und Wiki-Autoren indirekt den Zusammenhang von Inhalt und Strukturelementen einer Seite verdeutlichen.

Der momentane Ablauf bei Erstellung einer neuen Wiki-Seite gestaltet sich in der Regel wie folgt: (1) neue Wiki-Seite erstellen, (2) Titel eingeben, (3) Tags hinzufügen, (4) Typ Tags hinzufügen, (5) Artikel formulieren und abspeichern, (6) Attribute hinzufügen und mit Werten belegen

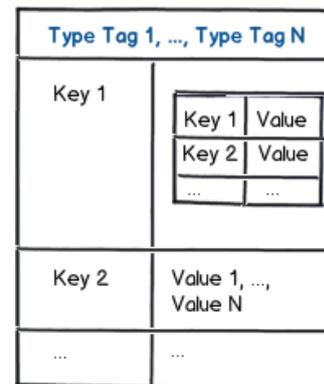


Abbildung 5.4.: Konzeptuelles Mock-Up für verschachtelte Attribute in Form von Records (Eigene Darstellung)

Der neue Ablauf wäre dann: (1) neue Wiki-Seite erstellen, (2) Wiki-Seiten-Template auswählen, (3) Titel eingeben, (4) Artikel formulieren und abspeichern (5) Attribute mit Werten belegen (und optional (6) Strukturelemente hinzufügen/entfernen)

Vorteile/Nutzen:

- Schnelle Auswahl und Erstellung einer Wiki-Seite anhand einer Vorlage (für oft verwendete Objekte wie Projekte oder Termine)
- Homogenisierung (Templates werden in der Regel häufig verwendet und vermeiden Redundanz)
- Geringerer Aufwand, da manuelles Hinzufügen von Typ Tags und Attributen entfällt

5.1.5. Vererbung auf Typ Tag-Ebene

Vererbung beziehungsweise Generalisierung findet ihre Anwendung in Programmiersprachen wie Java, aber auch in der Darstellung von Ontologien über OWL. Betrachtet man Typ Tag-Definitionen als Klassen, wäre eine Generalisierung auf Typ Tag-Ebene möglich. Vererbung wird mitunter in Freebase in Form von Parent Types zur Modellierung eingesetzt.

Vorteile/Nutzen:

- schnelle Ableitung von Typ Tag-Definitionen von übergeordneten Klassen
- Verbesserte Navigation durch Bildung von verallgemeinernden Kategorien als Einstiegspunkt der Navigation (siehe Abbildung 5.5)
- Vermeidung von Redundanz und doppelter manueller Vergabe von Typ Tags

5.1.6. Semantische Funktionalisierung von Attribut-Schlüsseln

Neben der bereits angesprochenen Vererbung gibt es mannigfaltige Szenarien zum semantisch-funktionalen Einsatz von Attribut-Schlüsseln.

Mögliche Optionen wären:

- Alias-Definitionen zur Begriffsklärung (engl. *disambiguation*) für Akronyme, Homonyme und Synonyme von Wiki-Seiten-Titeln
- Einführung eines Vererbungsattributs („is-a“) zur Umsetzung von Typ Tag-Vererbung (Bildung von Taxonomien auf Definitions- und Instanzebene)
- Beschreibung von Teil-Ganzes-Beziehungen (Partonomien) über ein entsprechendes Attribut („partOf“) - gerade im Unternehmenskontext sinnvoll zur Aggregation von z. B. Einzelteilen zu einem Gesamtprodukt

In MoKi können beispielsweise direkt bei der Erstellung eines Konzepts (Concept) synonyme Begriffe sowie hierarchische Eigenschaften mit angegeben werden (vgl. Abbildung 5.6).

Vorteile/Nutzen:

Business

Business contains types related to business such as 'Company', 'Industry' and 'Retail Location'. Note, this category doesn't include types specific to Finance (stock exchange and currency) Still to do: populate example company types capture more companies [More](#)

[Edit Description](#)

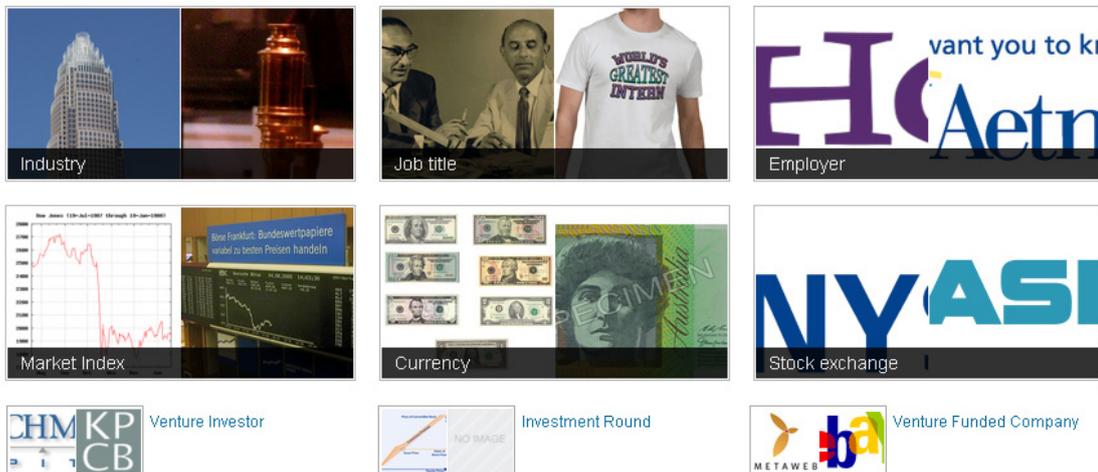


Abbildung 5.5.: Subtypen vom Typ „Business“ in Freebase

- Erhöhung der semantischen Reichhaltigkeit des Strukturmodells
- Begriffsklärung und Erhöhung der Konsistenz durch Vermeidung von Mehrdeutigkeiten
- Formung zusätzlicher Navigationsstrukturen

5.1.7. Zugriffskontrolle auf Datenmodell-Ebene

Das Thema Zugriffskontrolle beziehungsweise die Rechtevergabe auf Ontologie-Ebene und noch differenzierter auf Typen-, Instanz- und Attribut-Ebene wird in der betrachteten Literatur kaum behandelt. Lediglich die Macher von OntoWiki haben sich diesem Thema gewidmet. Sie beschreiben ein Modell zur Zugriffskontrolle auf einen RDF-Tripel Speicher [DA06].

In ihrer Anforderungsanalyse wird festgestellt, dass eine Berechtigungskontrolle auf Attribut- und Instanz-Ebene bestimmter Klassen notwendig ist. Die entsprechenden Berechtigungen sollen über Account-Metadaten wie Gruppenzugehörigkeit und Unternehmensinformationen wie Delegationsstrukturen vergeben werden. Auf der anderen Seite muss der Zugang auch über einzelne Ressourcen selbst geregelt werden können. Die Abfrage, das Hinzufügen und das Entfernen von Tripeln müssen als atomare Aktionen getrennt voneinander behandelt werden.

Es wird bei der Darstellung auf ein virtuelles Modell gesetzt, das durch festgelegte Regeln die Darstellung eines angeforderten Modells einschränkt. Zugriffsregeln werden über die Semantic Web Rule Language (SWRL) oder einer ähnlichen Sprache definiert, eine Query-Engine (zum Beispiel unter Verwendung der SPARQL) ermöglicht die Anwendung der Regeln als Filter für Abfragen auf dem Tripel-Speicher. Das minimalistische

Modify concept: Conference

Annotations

Description:

Synonyms:

Hierarchical Structure

Is a:

Is part of:

Properties

Property:	<input type="text" value="has program chair"/>
Property target:	<input type="text" value="Person"/>

Abbildung 5.6.: Festlegung von Synonymen und hierarchischen Aspekten eines Konzepts in MoKi [RGP⁺09]

RDF-Schema Lightweight Access Control Schema (LACS) stellt ein Basisvokabular zur Formulierung von Regeln und Abfragen.

Leider ist das beschriebene Framework sehr spezifisch. „[...] it is not intended to be a general access control framework for the Semantic Web. Instead it is designed to be a fast, reliable and easy to implement as part of an RDF triple store.“ [DA06] und kann somit nur schwer zur allgemeinen Gestaltungsempfehlung herangezogen werden.

Vorteile/Nutzen:

- fraglich, da es dem Wiki-Prinzip der Offenheit widerspricht
- denkbar beim Einsatz von Wikis im Unternehmensumfeld
- Zugriffsrechte wenn dann pro Seite geregelt (individuelle Erlaubnis zum Lesen, Ändern und Hinzufügen von Typ Tags und Attributen zu einer Wiki-Seite)

5.2. Eingabe

5.2.1. Einfache Inferenzen für Typ Tags

Wird in IkeWiki ein ausgehender Link mit „key“+ „Of“ typisiert, leitet das System für die Zielseite „Key“ als neuen Typ für die Seite ab.

„[...] for example, if a link from 'Mozart' to 'Die Zauberflöte' is annotated by 'composerOf', the system will automatically associate the type 'Composer' with the page describing 'Mozart' [...]“ (IkeWiki, [Sch06])

In Hybrid Wikis wäre eine solche Umsetzung zur Erweiterung der Empfehlungen für das Hinzufügen von Typ Tags denkbar, allerdings in einer etwas abgewandelten Form. Der

Attribut-Namen eingehender Links wird in Hybrid Wikis automatisch mit einem „Of“-Postfix versehen. Der vorgeordnete Schlüssel typisiert die aktuelle Wiki-Seite (im obigen Beispiel „Composer“) und könnte deshalb als Typ Tag-Empfehlung für die entsprechende Seite verwendet werden.

Vorteile/Nutzen:

- zusätzliche Vorschläge für Typ Tags über logische Schlussfolgerungen

5.2.2. Konsistenzprüfung während der Eingabe

In Freebase werden Schemata starr festgelegt. Jeder Wert einer Eigenschaft hat zusätzlich einen erwarteten Typ (Expected Type) des Links.

„[...] the entity at the end of this link is automatically typed with the expected type if needed.“ (Freebase, [BTPC07])

Wird während der Annotation ein Topic (entspricht einer Wiki-Seite als Entität bei Hybrid Wikis und wird über ihren Titel in einem Namensraum identifiziert) eingegeben, das nicht dem erwarteten Typ entspricht, weist das System darauf hin und ordnet der Zielseite bei Bedarf automatisch den erwarteten Typ zu.

Für diesen Fall könnte man in Hybrid Wikis eine Pop Up-Meldung einführen, falls die Zielseite laut Einschränkung nicht den erwarteten Typ Tag besitzt (in Form von: „Die ausgewählte Zielseite besitzt den erwarteten Typ Tag nicht. Jetzt hinzufügen oder anderen Link wählen?“).

- Aufdecken von möglichen Inkonsistenzen in Bezug auf fehlende Typ Tags

5.2.3. Manuelle Markierung redundanter Inhalte

Um Redundanz zu verringern, verwendet Freebase die Methode des manuellen Markings.

„[...] a user can mark a topic as being a duplicate entity of another, creating a relationship between those topics having the meaning 'should be merged with'.“ (Freebase, [BTPC07])

Zwei sich ähnelnde Beschreibungen können somit über ein spezielles Attribut als solche markiert werden. Andere Nutzer oder Systemadministratoren können diese Markierungen über eine Suche abfragen, prüfen und entscheiden, ob zwei Artikel verschmolzen werden sollen. Die Konsistenz würde so auf Dauer manuell erhöht werden.

Für Hybrid Wikis wäre ein solches Konzept theoretisch auch erweitert für Strukturelemente möglich (z. B. Markierung über ein Attribut, ob zwei Typ Tag-Definitionen sich zu ähnlich sind). Dessen Umsetzung würde aber erst Sinn machen, wenn Vererbung auf Typ Tag-Ebene eingeführt wurde, da sich unter den momentanen Umständen viele Typ Tag-Definitionen ähneln.

Vorteile/Nutzen:

- Erhöhung von Konsistenz
- Einbezug von Nutzermeinungen

5.2.4. Kommentarfunktion für Belegungen von Attributen

Um Ersteller und Eigner einer Wiki-Seite auf mögliche Fehler in seinen Angaben hinzuweisen, gibt es eine allgemeine Kommentarfunktion. Diese könnte man auf eine Kommentarfunktion für Attributwerte erweitern. OntoWiki erlaubt es den Anwendern, zu jeder Belegung kleine Hinweise (Statements) zu verfassen, falls diese Fehler aufweisen. Prinzipiell sollte es jedem registrierten Nutzer offenstehen, solche Notizen anzubringen.

Vorteile/Nutzen:

- Möglichkeit zur Hinweisgebung bei Fehlern in Angaben (z. B. bei falschem Datum oder inkorrekt hinterlegter Telefonnummer)
- Manuelle Konsistenzprüfung



Abbildung 5.7.: Statements zu Belegungen von Attributen in OntoWiki [ADR06]

5.2.5. Automatisierte Bildung von Einschränkungen für Datentypen

In SnoopyDB ist der User während des Hinzufügens von Schlüssel-Werte-Paaren dazu angehalten, stets einen Datentyp für seinen Wert auszuwählen oder einer automatischer Erkennung zuzustimmen. Der am meisten verwendete Datentyp für den Wert im Zusammenhang mit der Eigenschaft führt auf Dauer zur automatisierten Bildung eines Constraints. Der Nutzer erhält in Zukunft während der Eingabe dann einen Hinweis, dass er den korrekten Datentyp (samt Formatierung) verwenden soll.

„[...] the concept proposes a validation process, which includes determining a data type for each newly entered value[...] [...] if a property is added that already exists, has a data type assigned and is used by the majority of property instances, the user is prompted to enter values according to this data type.“ (SnoopyDB, [GZS11])

Die strikte Verwendung von Datentypen ist vor allem in Hinblick auf eine mögliche Einführung einer Expertensuche oder die kontextbezogene Darstellung von kleinen Widgets als Annotationshilfen (Kalender bei Datumsangaben, Karten bei Positionsangaben, ...) wichtig.

„The detection of data types, especially in the case of numeric types, is very important as it is crucial for queries based on numeric evaluation.“ (SnoopyDB, [GZS11])

Um Nutzer zusätzlich dazu anzuregen das richtige Format für die Belegung einer Eigenschaft zu wählen, wäre eine beispielhafte Hinterlegung des noch nicht ausgefüllten Feldes mit einem Wert sinnvoll. Das Schlüssel-Werte-Paar mit dem Schlüssel „employees“ gibt als Beispiel für seine Belegung den Wert 120 an (vgl. Abbildung 5.8). Durch diese Präsentation wird dem Nutzer verdeutlicht, dass ein numerisch ganzzahliger Wert erwartet wird.

Vorteile/Nutzen:

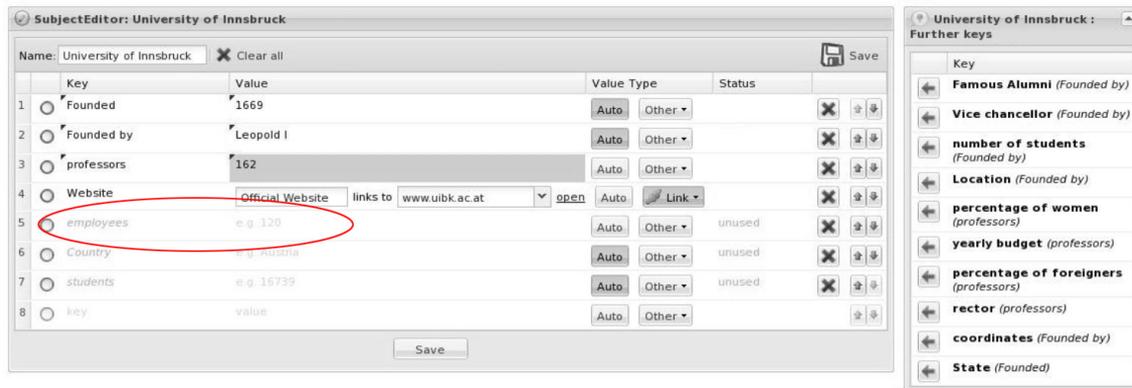


Abbildung 5.8.: Empfehlungen (Recommendations) in SnoopyDB [GZS11]

- Erhöhung von Konsistenz
- präventive Einhaltung von Constraints
- Grundlage für eine komplexere Suche

5.2.6. Import von Ontologien im RDF(S)/OWL-Format

Um den Anwender eines neu aufgesetzten Tricia-Systems mit der Hybrid Wikis-Erweiterung eine Basis an Attribut- und Typ Tag-Definitionen bereitstellen zu können, wäre ein Import vorhandener Datensets, variierend je nach Branche/Einsatzgebiet, denkbar. Freebase verwendet zum Start sogenannte Data Set Seedings.

„The current data in Freebase consists of millions of concepts (topics) and tens of millions of relationships between those topics. Areas of initial seeding include popular culture (e.g. films, music, books, sports, television), location information (restaurants, geolocations, businesses), scholarly information (linguistics, biology, astronomy), and general knowledge (Wikipedia). (Freebase, [BTPC07])

Da Hybrid Wikis bereits über eine Importfunktion für die Bildung von Definitionen und Constraints über .ecore-Dateien besitzt, wäre eine Ausweitung auf weitere Formate denkbar. Die Literaturstudie hat gezeigt, dass ein Großteil der Ansätze auf die Semantic Web-Formate RDF(S)/OWL aufbaut. So wäre die Übertragung von in Ontologieeditoren modellierten Konzepten möglich.

Vorteile/Nutzen:

- zusätzliche Unterstützung der Nutzer, um nicht bei „Null“ beginnen zu müssen
- präventive Vermeidung der Entstehung heterogener Strukturen
- Import von Ontologien aus Tools, die RDF(S)/OWL zur internen Repräsentation der Daten verwenden

5.3. Darstellung, Navigation und Suche

5.3.1. Vergleichssicht, Attribut-Sicht und kontextspezifische Visualisierung

Möglichen Erweiterungen im Hinblick auf Sichten sind im Prinzip keine Grenzen gesetzt. Dennoch sind manche im Hinblick auf die Förderung von Strukturierung sinnvoll und notwendig.

Vergleichssicht Diese Sicht (vgl. Abbildung 5.9) soll zum Vergleich zweier Wiki-Seiten anhand ihrer Struktur dienen. Die „Top 10“ der sich am meisten ähnelnden Seiten (Comparison Suggest) könnten zusätzlich zur direkten Gegenüberstellung angezeigt werden („Ähnliche/Verwandte Seiten“).

„Comparison suggest is the system suggests resources to compare based on similarity of properties and their values Because role of a property is describing characteristic to identify from others A user can find properties and values from comparison of a target and others[.]“ (Social Infobox, [HGT11])

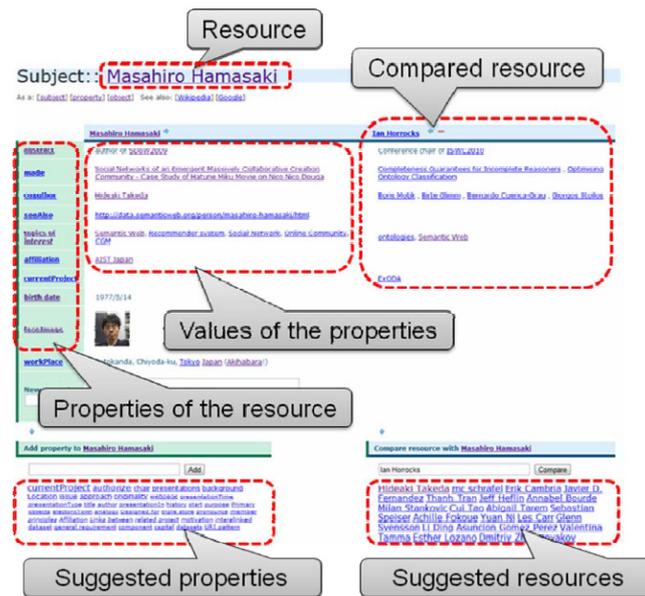


Abbildung 5.9.: Vergleichssicht (Comparison View) in Social Infobox [HGT11]

Vorteile/Nutzen:

- Entdecken zusätzlicher Typ Tags und Attribute
- nachträgliche Korrektur oder Erweiterung durch Erkenntnisgewinn über den Vergleich
- Optimierung von Strukturen

Attribut-Sicht Eine logische Ergänzung zu der bereits vorhandenen Typ Tag-Sicht ist eine Attribut-zentrische Sicht. Anzeigen sollte man neben dem Attribut-Namen auch per Definition deklarierte Zusammenhänge zu Typ Tags, Seiten samt Titel, die das Attribut verwenden, ebenso wie den Wert, mit dem es belegt wurde.

kontextspezifische Sichten (siehe Abbildung 5.10) HYENA bietet zum Beispiel eine Kalenderansicht für Datumsangaben. Gleiches könnte man für Positionsangaben über eine Karte realisieren, ähnlich umgesetzt wie in OntoWiki. Würde man bei Datumsangaben außerdem zwischen Start- und Enddatum differenzieren, könnte man Balkendiagramme, ähnlich wie Gantt-Diagramme in Microsoft Project, realisieren. Gerade im Unternehmenskontext wäre das eine hilfreiche Darstellungsform zur Abstimmung von zeitkritischen Projekten oder Aktivitäten, die im Wiki beschrieben werden.

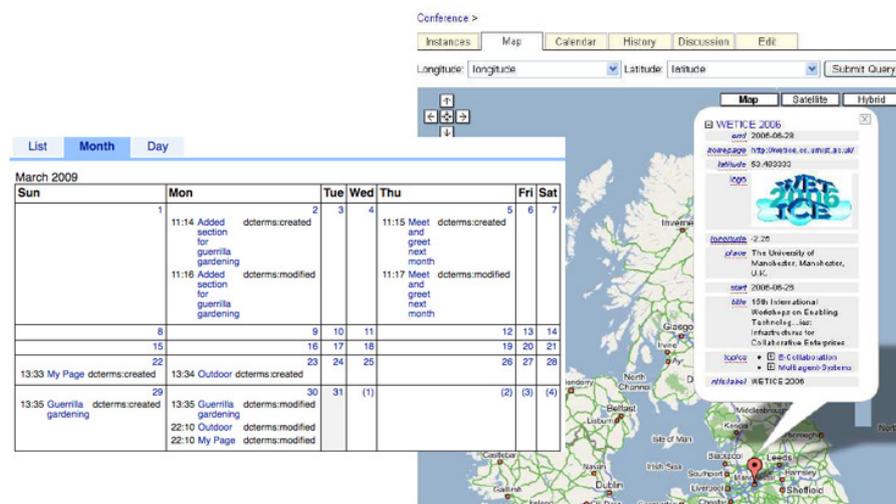


Abbildung 5.10.: Monatsansicht in HYENA [Rau10] (links) und Kartenansicht in OntoWiki [ADR06] (rechts)

Vorteile/Nutzen:

- Verbesserung der Übersicht
- kontextspezifische Visualisierung von Wertetypen

5.3.2. Einbettung der Attribute von Zielseiten

Entstanden ist diese Idee bei der Betrachtung von internen Links zu Ressourcen in OntoWiki (vgl. Abbildung 5.11). Diesen werden automatisch die Attribute der Zielseite angeheftet. Zur Erhöhung der Übersicht sollte man diese Liste ausklappbar gestalten und eventuell maximal fünf Attribute angezeigt bekommen.

Vorteile/Nutzen:

- Einsparung eines Navigationsschritts, da die wichtigsten Informationen der Zielseite (z. B. Telefonnummer oder Raum) bereits angezeigt werden
- oder: bestimmte aufgeführte Attribute regen zum Anklicken des Links an, um detailliertere Informationen zu erlangen



Abbildung 5.11.: Ausklappbare Attribute der Zielressource einer Link-Relation (Inline Reference View) in On-toWiki [ADR06]

5.3.3. Ausbau der Suche

Ein sehr wichtiger Punkt, dessen Voraussetzungen schon beschrieben wurden (siehe Abschnitt 5.2.5), ist die Erweiterung der Suche in Form eines „Expertenmodus“ als Option für erfahrene Nutzer.

Essentielle Funktionalitäten sind dabei:

- einfache Mehrfachauswahl und freie Kombination von Typ Tags, Werten und Schlüsseln
- Einführung von Operatoren (neben dem bereits vorhandenen AND OR und NOT) zur Verknüpfung der Mehrfachauswahl
- abhängig vom Typ des Wertes zusätzliche Eingrenzungsmöglichkeiten anbieten (z. B. Vergleichsoperatoren bei numerischen Werten, von-bis-Angaben für Zeiträume, ...)

Diese Thematik wurde bereits aufgegriffen und als Erweiterung diskutiert:

„From a technical point of view, attribute (and type tag) filters could be flexibly combined using boolean operators.“ (Hybrid Wikis, [MNS11])

Des Weiteren sollte man sich auch Gedanken zur Einführung einer Syntax-basierten Suche machen, um JOIN- und GROUP BY-Statements, ähnlich wie über SQL, realisieren zu können.

Vorteile/Nutzen:

- komplexere Abfragen ermöglichen präzisere Suchanfragen und eine Erhöhung der Güte der Ergebnismenge

6. Fazit und Ausblick

6.1. Zusammenfassung und Fazit

In dieser Bachelorarbeit wurde einführend auf die historische Entwicklung des Internets und die Evolutionsgeschichte von Wikis im Verlauf eingegangen. Grundlegende Ansätze zur Strukturierung wurden vom Tagging über das Templating bis hin zur Ontologieentwicklung über semantische Wikis beschrieben. Anschließend wurde Hybrid Wikis als ein erster Referenzansatz zur emergenten Schaffung von Struktur über Wikis vorgestellt. Danach wurde eine Literaturrecherche zur Sammlung und Katalogisierung weiterer Ansätze durchgeführt, die die Web-basierte Strukturierung von Information auf ähnliche Art und Weise erlauben. Die Quellen wurden analysiert und anhand extrahierter Eigenschaften in Tabellenform zugeordnet. Eine zusammenfassende Erläuterung von Erkenntnissen ergänzt die Einordnung. Außerdem wurden das gewonnene Wissen dazu genutzt, um Hybrid Wikis im Kontext zu analysieren, fehlende Eigenschaften aufzudecken und unterschiedliche Verbesserungsvorschläge initial anzusprechen und ihren Bedarf zu begründen.

6.2. Zukünftige Forschungsfragen

Für auf den Ergebnissen und Erkenntnissen aufbauende, zukünftige Forschungsarbeiten ergeben sich diverse Problem- und Aufgabenstellungen, die zur Aufarbeitung bereitstehen.

- **Qualitative Überprüfung der verwendeten Quellen** Während der Analyse und Einordnung hat sich herausgestellt, dass viele Quellen die verwendeten Konzepte und Methoden nur sehr vage beschreiben oder es schwierig war, Eigenschaften trennscharf zu identifizieren und eine Einordnung überhaupt vorzunehmen, ohne sich dabei in Details der Ontologieentwicklung zu verlieren. Eine qualitative Überprüfung und Bewertung der Quellen könnte nachträglich dazu beitragen, den Ergebniskatalog und die Einordnung zu optimieren und mit weiteren Quellen zu ergänzen (nachträgliche Durchführung des in [OS10] beschriebenen Punktes „Quality appraisal“).
- **Tiefensuche zur Verdichtung der gewonnenen Erkenntnisse** Die durchgeführte, ansatzzentrische Breitensuche über einzelne Schlagwörter und die Analyse der Systeme hat zur Extraktion von Eigenschaften geführt, die in Kapitel 4 zusammenfassend beschrieben und die Quellen entsprechend zugeordnet wurden. Eine auf den Erkenntnissen aufbauende Tiefensuche könnte detailliertere Erkenntnisse zu den zusammenfassend beschriebenen Eigenschaften liefern.
- **Detaillierter Vergleich einzelner Ansätze** Viele der Ansätze beschreiben ihre Konzepte und Methoden, wie bereits erläutert, nur sehr vage, besitzen maximal eine

prototypische Implementierung oder wurden inzwischen wieder aufgegeben. Freebase als Plattform und OntoWiki als Wiki-Software haben sich hingegen sehr stark weiterentwickelt und bieten inzwischen neue Funktionen. Sinnvoll wäre deshalb eine genauere Systemanalyse durch praktischen Einsatz und Einbezug von Dokumentation zur detaillierten Gegenüberstellung etablierter Systeme, um weitere, für Hybrid Wikis interessante, Features aufdecken zu können.

- **Umfragen zu Gestaltungsempfehlungen** Die im letzten Abschnitt des Hauptteils getätigten Empfehlungen sind nur sehr kurz und knapp als grundlegende Ideen für Erweiterungen beschrieben, um der zukünftigen Planung und Umsetzung möglichst viel Spielraum zu lassen. Bevor dies jedoch geschehen kann, sollte man Bedarfsanalysen in Form von Umfragen durchführen.
- **Prototypische Implementierung** Sollte bestätigter Bedarf für eines oder mehrere „Features“ bestehen, wäre der prototypische Entwurf und die Implementierung, je nach Aufwand, in Form einer zukünftigen Abschlussarbeit denkbar.

Anhang

Schemabildung		Wiki-Seiten		Freitext und/oder Struktur ?		Zusammenhang Freitext/Struktur (falls beides vorhanden)	
Implizit	explizit (Typen als Bezeichner)	lose Kopplung von Typen und Attributen	Freitext	Struktur	Im Freitext annotiert und daraus abgeleitet	separat und unabhängig erstellt	
[MNS11], [MNI11]	[MNS11], [MNI11]	[MNS11], [MNI11]	[MNS11], [MNI11]	[MNS11], [MNI11]	[MNS11]	[MNS11]	
[ZGS10], [GZTS10], [GZS11]	[BTPC07], [DBEP+08]	[ZGS10], [GZTS10], [GZS11]	[BTPC07]	[ZGS10], [GZTS10], [GZS11]	[BTPC07], [DBEP+08], [DBPCT07]	[BTPC07], [DBEP+08]	
[HGT11]	[L204]	[L204]		[HGT11]			
	[ADR06]	[ADR06]		[ADR06], [DAR06]			
	[RGP+09], [GKNL+09]	[RGP+09], [GKNL+09]		[RGP+09], [GKNL+09]			
	[Rau10]	[Rau10]	[Sch06] [Aum05] [BGLP06] [Rau10] [KVV06]	[Sch06] [Aum05] [BGLP06] [Rau10] [KVV06]	[Sch06] [Aum05] [BGLP06] [Rau10] [KVV06]	[Sch06] [Aum05] [BGLP06] [Rau10] [KVV06]	
	[Kun08a], [Kun08a]	[Kun08a], [Kun08a]		[Kun08a], [Kun08a]			
	[BSB+09] [Kle06]	[BSB+09] [Kle06]		[BSB+09] [Kle06]			
[Ore05]	[BH04]	[Ore05]	[Ore05]	[Ore05]	[Ore05]	[Ore05]	

Tabelle A.3.: Einordnung: Struktursystem 3

Integritätsbedingungen				Validierung				
Verwendung	Multiplizitäten (Werte)	Wertebereich/Enumerationen (Werte)	Datentypen (Werte)	Format von Datentypen (Werte)	Erwartete Typen für Links (Werte)	automatisierte Bildung von Integritätsbedingungen (über Häufigkeit der Verwendung)	weich (Hinweise)	hart (Fehlermeldungen)
[MNS1:1], [MNI:1]	[MNS1:1], [MNI:1]	[MNS:1], [MNI:1]	[MNS1:1], [MNI:1]	[GZS1:1]	[MNS1:1], [MNI:1]	[GZS1:1]	[MNS1:1], [MNI:1]	
[GZS1:1]			[GZS1:1]	[GZS1:1]				
[BTP:C07]	[LZ04]				[BTP:C07]			
[LZ04]	[LZ04]				[LZ04]			
[Rau:0]		[Rau:0]						
[Kun08b]		[Kun08b]					[Kun08b], [Kun08a]	
[BSB+09]	[BSB+09]	[BSB+09]			[BSB+09]		[Kle09]	
[BH+04]	[BH+04]		[BH+04]					

Tabelle A.4.: Einordnung: Struktursystem 4

Auflistung ähnlicherwörter Seiten (z. B. für einen direkten Vergleich)	Reasoninglogische Inferenzen	automatisierte Informationsextraktion	Gezielte Zuverfügungstellung einer Struktur-/Datenbasis [MNS11]	manuelle Konsistenzprüfung (Kommentare zu Attributen, Marking)	Widgers als Helfer abhängig vom Datentyp (Kalender, ...)	Inline-Bearbeitung von Schlüssel -Werte-Paaren
[HGT11] [ADR06]	[BTPC07]	[RCP+09], [GKNL+09] [SEG+09]	[BTPC07], [DBEP+08], [DBPCT07]	[BTPC07]	[ADR06], [DAR06]	[ADR06]
	[Sch06], [SWG06]					
	[kun08a], [kun08a]					
	[BSB+09]					
	[BH04]					

Tabelle B.2.: Einordnung: Eingabe 2

C. Einordnung: Darstellung, Navigation und Suche

Darstellung, Navigation und Suche					
Sichten					
	Attribut-Sicht	Typ-Sicht	Vergleichs-sicht (zweier-Wiki-Seiten)	kontextspezifische Sichten (Kalendersicht, Karte, ...)	Eingebettete Attribute von Zielseiten
Hybrid Wikis		[MNS11], [MN11]			
SnoopyDB					
Freebase					
Artificial Memory					
Social Infblox					
OrdoWiki	[ADP06], [DAP06]	[ADP06], [DAP06]	[HGT11]	[ADR06], [DAR06]	[ADR06]
Moki					
KIM (IkeWiki)					
WIKSAR (SHAWN)	[Purn05]				
SweetWiki					
HYENA		[Raur10]		[Raur10]	
Semantic MediaWiki					
AceWiki					
CNL Wiki					
Katkolli					
Rizzone					
SeipertWiki					
SOBOLEO					
wiki@unt					
	Legende:				
				Alleinstellungsmerkmal von Hybrid Wikis	
				Nicht vorhanden in Hybrid Wikis	
				inzwischen teilweise vorhanden in Hybrid Wikis (nicht durch Quellen belegbar)	
				Vorhanden in Hybrid Wikis (nicht durch Quellen belegbar)	

Tabelle C.1.: Einordnung: Darstellung, Navigation und Suche 1

Suche		Syntax-basiert				eingebettete Suchabfragen in Wiki-Seiten
Art der Suche	UI-basiert	facettedrill Down	SPARQL	eigene Abfragesprache	kontrollierter Wortschatznatürliche Sprache	
		[MNS11]				[MNS11]
Volltext-/Schlagwortsuche						
		[BTPC07]		[BTPC07], [DBPCT07]		
		[LZ04]				
		[ADR06], [DAR06]				
		[SEG+09]				
		[Rau10]	[BGLP06] [Rau10] [KVV06]	[KVV06]		[Rau10] [KVV06]
			[BSB+09]		[kuh08b], [kuh08a]	
		[Sou06]				
		[Ore05]				
		[BSZ07]				

Tabelle C.2.: Einordnung: Darstellung, Navigation und Suche 2

Literaturverzeichnis

- [ABK⁺07] S. Auer, B. Bizer, G. Kobilarov, J. Lehmann, and Z. Ives. DBpedia: A nucleus for a web of open data. In *In 6th Int'l Semantic Web Conference, Busan, Korea*, pages 11–15. Springer, 2007.
- [ADR06] S. Auer, S. Dietzold, and T. Riechert. OntoWiki - a tool for social, semantic collaboration. In *International Semantic Web Conference*, pages 736–749, 2006.
- [AGAH06] S. A. Golder and B. A. Huberman. Usage patterns of collaborative tagging systems. *Journal of Information Science*, 32:198–208, 2006.
- [Aum05] D. Aumueller. SHAWN: Structure helps a wiki navigate. In *Proc. of the BTW-Workshop WebDB Meets IR*, 2005.
- [Aum10] D. Aumueller. Usability meets instant gratification on the semantic web. *Computing Research Repository*, abs/1011.2, 2010.
- [BEKAW09] F. Bry, M. Eckert, J. Kotowski, and K. A. Weiand. What the user interacts with: Reflections on conceptual models for semantic wikis. In *Semantic Wikis*, 2009.
- [BGLP06] M. Buffa, F. Gandon, C. Lecompte, and J. Passeron. SweetWiki: semantic web enabled technologies in wiki. In *International Symposium on Wikis*, pages 69–78, 2006.
- [BH04] J. Bao and V. Honavar. Collaborative ontology building with wiki@nt - a multi-agent based ontology building environment. In *Workshop on Evaluation of Ontology-based Tools*, 2004.
- [BHBL09] C. Bizer, T. Heath, and T. Berners-Lee. Linked Data - the story so far. *International Journal on Semantic Web and Information Systems*, 5:1–22, 2009.
- [BL98] T. Berners-Lee. Semantic Web Road Map. <http://www.w3.org/DesignIssues/Semantic.html>, September 1998. (abgerufen am 25.08.2011).
- [BL06] T. Berners-Lee. Linked Data. <http://www.w3.org/DesignIssues/LinkedData.html>, 2006. (abgerufen am 30.08.2011).
- [BLGE⁺08] M. Buffa, F. L. Gandon, G. Ereteo, P. Sander, and C. Faron. SweetWiki: A semantic wiki. *Journal of Web Semantics*, 6:84–97, 2008.
- [BSB⁺09] J. Bao, P. R. Smart, D. Braines, G. Jones, and N. R. Shadbolt. A controlled natural language interface for Semantic MediaWiki. In *2nd Annual Conference of the International Technology Alliance*, <http://tw.rpi.edu/wiki/TW-2009>, volume 5. Citeseer, 2009.

- [BSZ07] Simone Braun, Andreas Schmidt, and Valentin Zacharias. SOBOLEO: vom kollaborativen Tagging zur leichtgewichtigen Ontologie. In Tom Gross, editor, *Mensch & Computer - 7. Fachübergreifende Konferenz - M&C 2007*, pages 209–218, München, 2007. Oldenbourg Verlag.
- [BT06] A. Blumauer and P. Tassilo. Semantic Web und semantische Technologien: Zentrale Begriffe und Unterscheidungen. In A. Blumauer and P. Tassilo, editors, *Semantic Web. Wege zur vernetzten Wissensgesellschaft*, pages 9–25. Springer Verlag, Berlin, Heidelberg, 2006.
- [BTPC07] K. Bollacker, P. Tufts, T. Pierce, and R. Cook. A platform for scalable, collaborative, structured information integration. 2007.
- [CHK07] P. Cyganski, B. Hass, and Th. Kilian. Potenziale sozialer Netzwerke für Unternehmen. In *Web 2.0. Neue Perspektiven für Marketing und Medien*, pages 101–120. Springer Verlag, 2007.
- [Cho06] M. Choate. What makes an enterprise wiki? <http://www.realstorygroup.com/Feature/145-Wikis-in-the-enterprise>, April 2006. (abgerufen am 07.08.2011).
- [Cum08] R. E. Cummings. *Wiki writing: Collaborative learning in the college classroom*. Univ of Michigan Pr, 2008.
- [Cun06] W. Cunningham. Wiki design principles. <http://c2.com/cgi/wiki?WikiDesignPrinciples>, 2006. (abgerufen am 05.08.2011).
- [DA06] S. Dietzold and S. Auer. Access control on RDF triple stores from a semantic wiki perspective. In C. Bizer, S. Auer, and L. Miller, editors, *Proc. of 2nd Workshop on Scripting for the Semantic Web at ESWC, Budva, Montenegro, June 12, 2006*, volume 183 of *CEUR Workshop Proceedings ISSN 1613-0073*, June 2006.
- [DAR06] S. Dietzold, S. Auer, and T. Riechert. Kollaborative Wissensarbeit mit Onto-Wiki. In *Proceedings of the INFORMATIK 2006 Workshop: Bildung von Sozialen Netzwerken in Anwendungen der Social Software*, 2006.
- [DBEP+08] K. D. Bollacker, C. Evans, P. Paritosh, T. Sturge, and J. Taylor. Freebase: a collaboratively created graph database for structuring human knowledge. In *International Conference on Management of Data*, pages 1247–1250, 2008.
- [DBPCT07] K. D. Bollacker, R. P. Cook, and P. Tufts. Freebase: A shared database of structured general human knowledge. In *National Conference on Artificial Intelligence*, pages 1962–1963, 2007.
- [DIVZ09] A. Di Iorio, F. Vitali, and S. Zacchiroli. Wiki semantics via wiki templating. *Computer*, pages 1–20, 2009.
- [EG05] A. Ebersbach and M. Glaser. Wiki. *Informatik-Spektrum*, 28:131–135, 2005. 10.1007/s00287-005-0480-7.

- [GFG11] K. Grzeganeck, I. Frost, and D. Gross. Spoilt for choice - wiki software for knowledge management in organisations. <http://oreilly.com/web2/archive/what-is-web-20.html>, 2011. (abgerufen am 31.08.2011).
- [GKNL⁺09] C. Ghidini, B. Kump, S. N. Lindstaedt, N. Mahbub, V. Pammer, M. Rospocher, and L. Serafini. MoKi: The enterprise modelling wiki. In *European Semantic Web Symposium / Conference*, pages 831–835, 2009.
- [GZS11] W. Gassler, E. Zangerle, and G. Specht. The Snoopy Concept: Fighting heterogeneity in semistructured and collaborative information systems by using recommendations. In *Collaboration Technologies and Systems (CTS), 2011 International Conference on*, pages 61–68, may 2011.
- [GZTS10] W. Gassler, E. Zangerle, M. Tschuggnall, and G. Specht. SnoopyDB: narrowing the gap between structured and unstructured information using recommendations. In *ACM Conference on Hypertext*, pages 271–272, 2010.
- [HGT11] M. Hamasaki, M. Goto, and H. Takeda. Social Infobox: collaborative knowledge construction by social property tagging. In *Conference on Computer Supported Cooperative Work*, pages 641–644, 2011.
- [HOG⁺08] J.W. Huss, C. Orozco, J. Goodale, C. Wu, S. Batalov, T.J. Vickers, F. Valafar, and A.I. Su. A gene wiki for community annotation of gene function. *PLoS biology*, 6(7):e175, 2008.
- [HT08] H.-J. Happel and M. Treitz. Proliferation in enterprise wikis. In *Proc. 8th Intl. Conf. on the Design of Cooperative Systems (COOP08)*, pages 123–129, 2008.
- [Kie06] M. Kiesel. Kaukolu: Hub of the Semantic Corporate Intranet. In M. Völkel and S. Schaffert, editors, *SemWiki*, volume 206 of *CEUR Workshop Proceedings*. CEUR-WS.org, 2006.
- [Kno03] E. Knorr. 2004: The year of web services. http://www.cio.com/article/32050/2004_The_Year_of_Web_Services?page=1&taxonomyId=3044, December 2003. (abgerufen am 03.08.2011).
- [Kuh08a] T. Kuhn. AceWiki: A natural and expressive semantic wiki. *Computing Research Repository*, abs/0807.4, 2008.
- [Kuh08b] T. Kuhn. AceWiki: Collaborative ontology management in controlled natural language. *Computing Research Repository*, abs/0807.4, 2008.
- [KVV06] Markus Krötzsch, Denny Vrandečić, and Max Völkel. *Semantic MediaWiki*. 2006.
- [Lan08] C. Lange. SWiM—a semantic wiki for mathematical knowledge management. *The Semantic Web: Research and Applications*, pages 832–837, 2008.
- [LC01] B. Leuf and W. Cunningham. The wiki way: quick collaboration on the web. 2001.

- [LZ04] L. Ludwig and X. Zhou. Artificial Memory prototype for personal semantic subdocument knowledge management. In *3rd International Semantic Web Conference (ISWC2004)*, 2004.
- [Mar06] J. Markoff. Entrepreneurs see a web guided by common sense. <http://www.nytimes.com/2006/11/12/business/12web.html?pagewanted=1>, November 2006. (abgerufen am 25.08.2011).
- [MBB⁺08] D. E. Millard, C. P. Bailey, P. Boulain, S. Chennupati, H. C. Davis, Y. Howard, and G. Wills. Semantics on demand: Can a semantic wiki replace a knowledge base? *New Rev. Hypermedia Multimedia*, 14:95–120, January 2008.
- [McA06] A. McAfee. Enterprise 2.0: the dawn of emergent collaboration. *IEEE Engineering Management Review*, 34:38–38, 2006.
- [MN11] F. Matthes and C. Neubert. Enabling knowledge workers to collaboratively add structure to enterprise wikis. In *12th European Conference on Knowledge Management - ECKM 2011*. 2011.
- [MNS11] F. Matthes, C. Neubert, and A. Steinhoff. Hybrid Wikis: Empowering users to collaboratively structure information. In *6th International Conference on Software and Data Technologies (ICSOFT)*, 2011.
- [MWY06] A. Majchrzak, C. Wagner, and D. Yates. Corporate wiki users: results of a survey. In *International Symposium on Wikis*, pages 99–104, 2006.
- [O’R05] T. O’Reilly. What is Web 2.0, design patterns and business models for the next generation of software. <http://oreilly.com/web2/archive/what-is-web-20.html>, September 2005. (abgerufen am 03.08.2011).
- [Ore05] E. Oren. SemperWiki: a semantic personal wiki. 2005.
- [OS10] C. Okoli and K. Schabram. A guide to conducting a systematic literature review of information systems research. <http://sprouts.aisnet.org/10-26/>, 2010. (abgerufen am 05.07.2011).
- [OWL04] OWL Web Ontology Language. <http://www.w3.org/TR/owl-features/>, 2004. (abgerufen am 28.09.2011).
- [Rau10] A. Rauschmayer. Structure your wiki: Improving support for structured data in wikis. 2010.
- [RCB10] A. Richter and A. C. Bullinger. Enterprise 2.0 – gegenwart und zukunft. vorschlag einer forschungsagenda. In *Tagungsband der Multikonferenz Wirtschaftsinformatik*, 2010.
- [RDF04a] Rdf Vocabulary Description Language 1.0: RDF Schema. <http://www.w3.org/TR/rdf-schema/>, 2004. (abgerufen am 28.09.2011).
- [RDF04b] Resource Description Framework (RDF). <http://www.w3.org/RDF/>, 2004. (abgerufen am 25.08.2011).

- [RDF04c] Resource Description Framework (RDF): Concepts and Abstract Syntax. <http://www.w3.org/TR/rdf-concepts/>, 2004. (abgerufen am 25.08.2011).
- [RG94] T. R. Gruber. A translation approach to portable ontology specifications. 1994.
- [RGP⁺09] M. Rospocher, C. Ghidini, V. Pammer, L. Serafini, and S. N. Lindstaedt. MoKi: the modelling wiki. In *Semantic Wikis*, 2009.
- [SBBK08] S. Schaffert, F. Bry, J. Baumeister, and M. Kiesel. Semantic wikis. *IEEE Software*, 25:8–11, 2008.
- [Sch06] S. Schaffert. IkeWiki: A semantic wiki for collaborative knowledge management. In *Workshop on Enabling Technologies: Infrastructure for Collaborative Enterprises*, pages 388–396, 2006.
- [Se007] Semantic Web Layer Cake. <http://www.w3.org/2007/03/layerCake-small.png>, 2007. (abgerufen am 25.08.2011).
- [SEG⁺09] S. Schaffert, J. Eder, S. Grünwald, T. Kurz, M. Radulescu, R. Sint, and S. Stroka. KiWi - a platform for semantic social software. In *Semantic Wikis*, 2009.
- [Sou06] A. Souzis. Bringing the Wiki-Way to the Semantic Web with Rhizome. In *Semantic Wikis*, 2006.
- [SS09] P. Smrz and M. Schmidt. Information extraction in semantic wikis. In *Proceedings of the 4th Workshop on Semantic Wikis, European Semantic Web Conference*. Citeseer, 2009.
- [SSM08] K. Stanoevska-Slabeva and M. Meckel. Web 2.0 – Grundlagen, Auswirkungen und zukünftige Trends. In *Web 2.0. Die nächste Generation Internet*, pages 13–38. Nomos Verlagsgesellschaft, 2008.
- [SvZ08] B. Sigurbjörnsson and R. van Zwol. Flickr tag recommendation based on collective knowledge. In *WWW '08: Proceeding of the 17th international conference on World Wide Web*, pages 327–336, New York, NY, USA, 2008. ACM.
- [SWG06] S. Schaffert, R. Westenthaler, and A. Gruber. IkeWiki: A user-friendly semantic wiki. In *European Semantic Web Symposium / Conference*, 2006.
- [VZ07] F. Vitali and S. Zacchiroli. Templating wiki content for fun and profit. 2007.
- [Weba] A-Z: Glossar - Harvard Business Manager. <http://www.harvardbusinessmanager.de/extra/glossar/a-599776.html>. (abgerufen am 04.08.2011).
- [Webb] Web 2.0 - Wikipedia. http://de.wikipedia.org/wiki/Web_2.0. (abgerufen am 04.08.2011).
- [Webc] Web 2.0 | Define Web 2.0 at Dictionary.com. <http://dictionary.reference.com/browse/web+2.0>. (abgerufen am 04.08.2011).

- [Webd] Web 2.0 Definition. <http://www.techterms.com/definition/web20>. (abgerufen am 04.08.2011).
- [Wei05] D. Weinberger. Trees vs. leaves. <http://www.hyperorg.com/backissues/joho-jan28-05.html>, January 2005. (abgerufen am 04.08.2011).
- [WiXa] Number of articles on en.wikipedia.org. <http://en.wikipedia.org/wiki/File:EnwikipediaArt.PNG>. (abgerufen am 04.08.2011).
- [WiXb] Ontologie (Informatik). [http://de.wikipedia.org/wiki/Ontologie_\(Informatik\)](http://de.wikipedia.org/wiki/Ontologie_(Informatik)). (abgerufen am 06.09.2011).
- [WiXc] Web Ontology Language. http://de.wikipedia.org/wiki/Web_Ontology_Language. (abgerufen am 18.09.2011).
- [ZGS10] E. Zangerle, W. Gassler, and G. Specht. Recommending structure in collaborative semistructured information systems. In *Proceedings of the fourth ACM conference on Recommender systems, RecSys '10*, pages 261–264, New York, NY, USA, 2010. ACM.