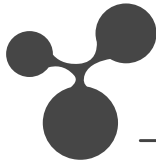


Technische Universität Dresden – Fakultät Informatik  
Professur für Multimediaetechnik, Privat-Dozentur für Angewandte Informatik

Prof. Dr.-Ing. Klaus Meißner  
PD Dr.-Ing. habil. Martin Englien  
(Hrsg.)



# GENE ME '11

---

GEMEINSCHAFTEN IN NEUEN MEDIEN

an der  
Fakultät Informatik der Technischen Universität Dresden

mit Unterstützung der

3m5. Media GmbH, Dresden  
Communardo Software GmbH, Dresden  
GI-Regionalgruppe, Dresden  
FERCHAU Engineering GmbH, Dresden  
IBM, Dresden  
itsax.de | pludoni GmbH, Dresden  
Kontext E GmbH, Dresden  
objectFab GmbH, Dresden  
queo GmbH, Dresden  
Robotron Datenbank-Software GmbH, Dresden  
SALT Solutions GmbH, Dresden  
SAP AG, Resarch Center Dresden  
Saxonia Systems AG, Dresden  
T-Systems Multimedia Solutions GmbH, Dresden  
Transinsight GmbH, Dresden  
xima media GmbH, Dresden

am 07. und 08. September 2011 in Dresden

[www.geneme.de](http://www.geneme.de)  
[info@geneme.de](mailto:info@geneme.de)

---

## B.5 Web 3L: Informationssuche und -verteilung mittels sozialer, semantischer Netze

*Manfred Langen<sup>1</sup>, Walter C. Kammergruber<sup>2</sup>, Karsten Ehms<sup>1</sup>*

<sup>1</sup> *Siemens AG, Corporate Research and Technology*

<sup>2</sup> *Technische Universität München*

### 1 Einleitung

Die Heterogenität von strukturierter, semi-strukturierter und unstrukturierter Information ist eine bekannte, jedoch bisher nicht gelöste Herausforderung an die Methoden der Informationsstrukturierung und des Information Retrieval.

Neue Ansätze betrachten den Bereich der nutzergenerierten Metadaten (Tagging) und verfolgen den Weg einer semantischen Anreicherung von Folksonomien zu einer leichtgewichtigen Ontologie. Die zunehmende Verbreitung von Social Software eröffnet weitere Möglichkeiten, mit der Zugehörigkeit von Personen zu sozialen Netzen zusätzliche Struktur-Information auszuwerten. Infolgedessen wird es im Information Retrieval möglich, durch eine Kombination aus semantischen und sozialen Relationen die Informationsversorgung treffsicherer zu machen. Neben der Suche als Information-Pull Interaktion kann auch die zielgerichtete Verteilung von Information (Information-Push) von diesem Ansatz profitieren. Diese beiden Anwendungsklassen sind in einem global agierenden Unternehmen von entscheidender Bedeutung für die Produktivität der Wissensarbeiter.

### 2 Konzeption – Das Web 3L Modell

Die hier ausgeführte Konzeption interpretiert das „Web 3.0“ [WaDe2006] nicht nur vage als Zusammenführung von Web 2.0 und Semantic Web, sondern konkret als Kombination von drei thematisch zunächst unabhängig betrachteten Netzwerken, die intelligent miteinander in Verbindung gebracht werden können. Es handelt sich um:

- ein Netz aus Hyperlinks, das „klassische“ WWW (L1)
- ein semantisches Netzwerk von Begriffen (L2)
- ein soziales Netzwerk von Akteuren (L3)

Aufgrund dieser drei logisch und in ihrer Realisierung technisch separierten Layer wählen wir für unser Modell die Bezeichnung Web 3L (Web of 3 Layers).

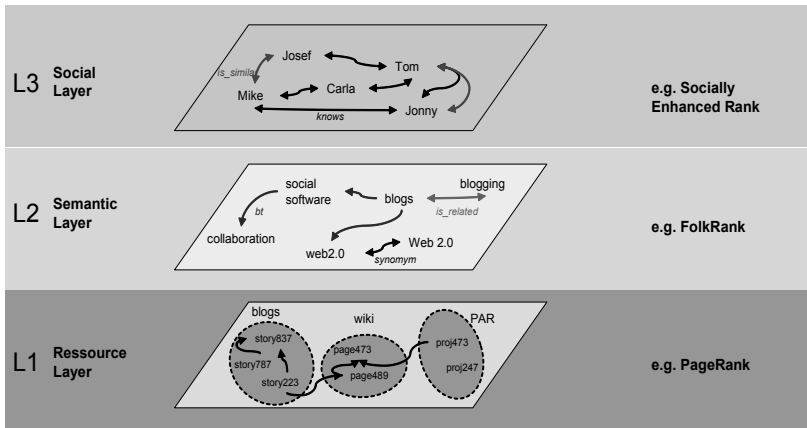


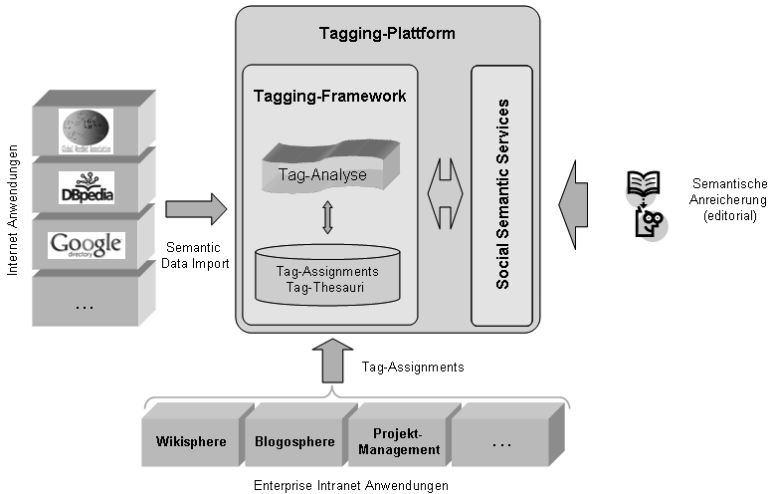
Abbildung 1: Die drei Layer des Web 3L Modells

### 3 Vom Tagging zu leichtgewichtigen Ontologien

Analog zum Grundgedanken unseres Forschungsansatzes [Thes2007, LaKa2008], eine sinnvolle Verknüpfung von Web 2.0 und Semantic Web Technologien zu finden, werden Folksonomien [Pete2009] aus dem Web 2.0 schrittweise zu „semantischen Folksonomien“ ausgebaut. Andere Arbeiten in diesem Umfeld kommen von Cattuto et al. [CaBe2008], die verschiedene Methoden zum automatischen Auffinden von Relationen zwischen Tags untersuchen oder Angeletou et al. [AnSa2007], die das Verfahren FLOR zum Mapping von Tags auf bestehende Semantic Web Entitäten beschreiben. Anders als die genannten Arbeiten verfolgen wir das Ziel, eine konzeptionell fundierte Architektur für Geschäftsanwendungen einzusetzen, um so einen Mehrwert für die Aufgaben eines Wissensarbeiters zu realisieren.

#### 3.1 Architektur einer Tagging-Plattform

Tagging ist eine Funktionalität, die heute in vielen Web 2.0 Applikationen angeboten wird. Synergieeffekte zwischen verschiedenen Web 2.0 Applikationen werden jedoch kaum genutzt, da die Tag-Zuordnungen verteilt gespeichert sind. Unser Ansatz eines applikationsübergreifenden Tagging Framework kann die Quervernetzung von Ressourcen, Personen und Metadaten herstellen und so einen Beitrag zur Lösung des „Orchestrierungsproblems“ [Ehms2010] eines Wissensarbeiters leisten.



**Abbildung 2: Architektur der Tagging-Plattform**

Das in Abbildung 2 dargestellte Tagging-Framework ist das Kernelement einer Tagging-Plattform. Es erhält seine Daten durch den Export von Tags aus unterschiedlichen Social Software-Applikationen (z.B. Siemens Wikisphere und Siemens Blogosphere). Eine semantische Anreicherung der Tag-Daten erfolgt durch statistische Verfahren (Text-mining, Co-Occurrence-Analyse, etc.), Import von semantischen Strukturen im Internet (z.B. DBpedia) und durch die explizite Ergänzung von Relationen (z.B. Synonyme) mittels Thesaurus Editor. Dadurch können, neben den automatisch abgeleiteten Zusammenhängen, semantische Beziehungen einer leichtgewichtigen Ontologie explizit hergestellt werden [LaKa2010].

### 3.2 Repräsentation leichtgewichtiger, semantischer Beziehungen

Der in Abbildung 1 dargestellte semantische Layer (L2) wird in Abbildung 3 detailliert aufgeschlüsselt. Es veranschaulicht, wie die vorgestellten verschiedenen Qualitäten von Termrelationen zunächst unabhängig voneinander in drei verschiedenen Sublayern repräsentiert werden.

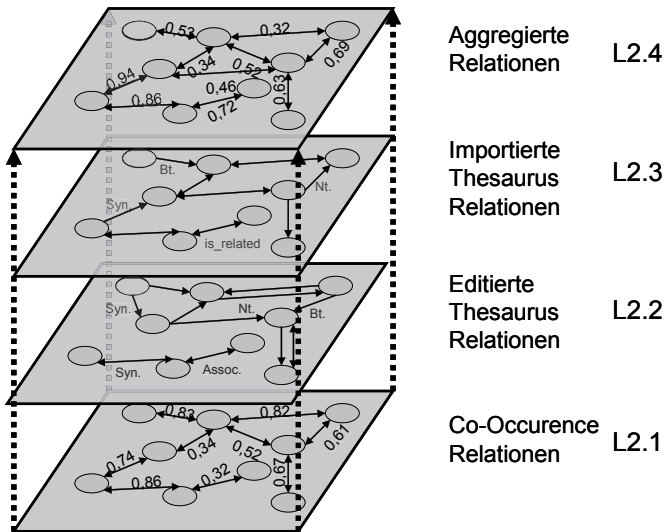


Abbildung 3: Multilayer Architektur des semantischen Layers L2

Die Tag-Daten werden in einem RDF-Modell beschrieben und in einem Triplestore gespeichert. Statistische Analysen, beispielsweise die Co-Occurrence-Analyse, liefern Hinweise auf Relationen, wobei diese weder vollständig noch unbedingt treffend sind. Sie bilden den Layer L2.1. Eine sinnvolle Ergänzung ist es, explizit Beziehungen zwischen Tags definieren zu können, analog zu den Beziehungen zwischen Konzepten bei Thesauri (siehe Abschnitt 3.3). Das Ergebnis wird in Layer L2.2 repräsentiert. Auf diese Weise wird eine Brücke geschlagen zwischen leicht zu gewinnenden Folksonomien und aufwändiger zu erstellenden Ontologien. Die Relationen von L2.1, L2.2, L2.3 können mittels spezieller mapping Verfahren<sup>1</sup> und individueller Gewichtungen zu einer Gesamtdarstellung L2.4 aggregiert werden.

### Tag-Thesaurus Editor

Abbildung 4 zeigt den im Projekt entwickelten Tag-Thesaurus Editor. Er wurde als Web-Applikation so gestaltet, dass er im Gegensatz zu Ontologie-Editoren wie Protégé<sup>2</sup> leicht bedienbar ist und keinerlei Einarbeitung oder Schulung erfordert. Um die Semantik eines bestimmten Tags zu bearbeiten, wird es aus der Menge aller Tags über eine komfortabel filterbare Liste ausgewählt (Abbildung 4, linke Seite).

<sup>1</sup> Die mapping Verfahren sind Gegenstand einer Patentanmeldung.

<sup>2</sup> <http://protege.stanford.edu/>

Ein Doppelklick auf das Tag „web2.0“ übernimmt dieses Tag als aktuellen Hauptterm in den Editor (in der Abbildung mit 1 gekennzeichnet) und zeigt die aktuell definierten semantischen Beziehungen (unter „Relations“).

Zur Unterstützung der semantischen Anreicherung für einen gegebenen Term werden im Tag-Thesaurus Editor Termvorschläge (unter „Suggestions“) in mehreren Kategorien angeboten, z.B. Synonym, Related Terms und weitere Relationen. Die Vorschläge in diesen Kategorien können mittels drag and drop in jede Tag-Thesaurus Relation (Synonym, Related, Broader und Narrower) übernommen werden. Im Beispiel ist das exemplarisch für den Term „Web 3.0“ gezeigt (siehe Pfeil 2 in der Abbildung). Auf diese Weise kann jeder Nutzer seine Terme in einem individuellen Tag-Thesaurus strukturieren. Infolgedessen entsteht eine Menge von individuellen Tag-Thesauri, die als Grundlage für einen gemeinsamen, gewichteten Tag-Thesaurus dienen können.

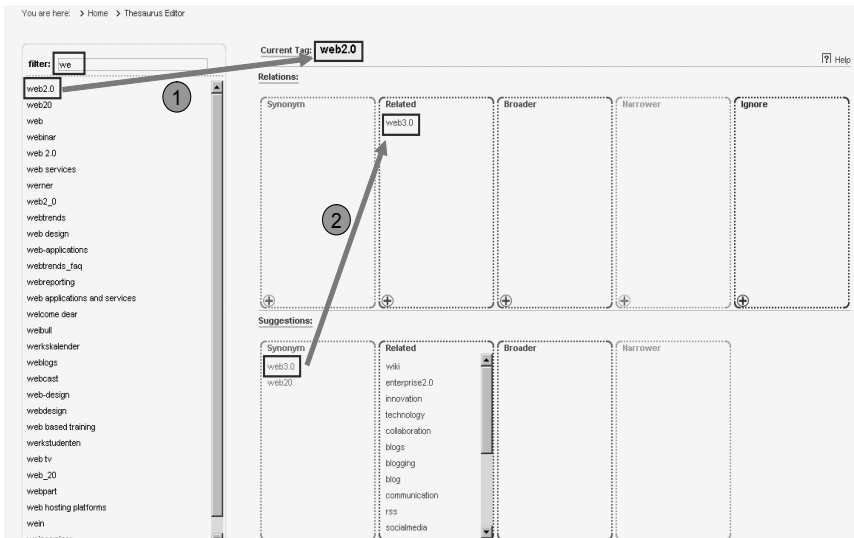


Abbildung 4: Tag-Thesaurus Editor

### 3.3 Import semantischer Strukturen

Die in Layer L2.3 abgebildeten semantischen Strukturen werden aus externen, frei verfügbaren, strukturierten Datenquellen importiert. Hierbei wurde in ersten Erprobungen sowohl DBpedia als auch DMOZ (dmoz.org) verwendet.

#### 4 Repräsentation sozialer Beziehungen

Ein weiterer wichtiger Baustein im Web 3L Modell ist der Social Layer (L3). Der Social Layer enthält als Knoten alle Nutzer, die in Social Software-Applikationen präsent sind. Dieser Layer beschreibt unterschiedliche Qualitäten an sozialen Beziehungen zwischen Personen, wobei sowohl explizite als auch implizite Vorgehensweisen zur Bestimmung der Relationen Anwendung finden. Eine explizite Beziehung entsteht z.B. durch sogenanntes „Following“ oder die Definition von „Friends“. Dies ist in Abbildung 5 im Layer L3.3 repräsentiert.

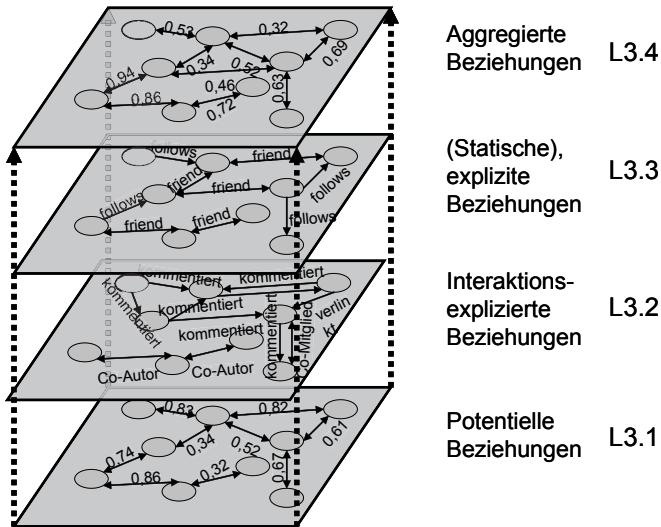


Abbildung 5: Multilayer Architektur des sozialen Layers L3

Implizite soziale Beziehungen können z.B. aus der Kommentierung oder Annotation von Inhalten abgeleitet werden. Diese interaktionsorientierten Relationen sind im Layer L3.2 repräsentiert. Desweiteren kann ein Vernetzungspotenzial von Personen aus ihren generierten Inhalten ermittelt werden, das sich für die Bildung von Communities verwenden lässt [KaLa2009]. Da hier keine explizite oder durch Interaktion implizierte Beziehung definiert ist, sprechen wir in Layer L3.1 von potentiellen Beziehungen. In Anlehnung an Granovetter's Konzept der verschiedenen Intensitäten von Beziehungen, korrespondiert L3.3 zu „strong ties“, L3.2 zu „weak ties“ und L3.1 zu „potential ties“ [Gran1973].

## **5 Verknüpfungen zwischen sozialem und semantischen Layer**

Das Web 3L Modell (Abbildung 1) stellt den sozialen und semantischen Layer zunächst separiert dar. Tatsächlich sind beide leicht miteinander verknüpfbar. In Social Software-Applikationen werden nutzergenerierte Inhalte von Personen mit Tags versehen. Damit besteht vom Tag (semantischer Layer L2) eine Verbindung zu einer Person (social Layer L3), die das Tag vergeben hat. Aufbauend auf dieser inhärenten Begebenheit lassen sich nachfolgende Zusammenhänge zusätzlich herstellen.

### **5.1 Ähnlichkeit von Interessensprofilen**

Durch die Verwendung von Tags an nutzer-generierten Inhalten entsteht für jeden Nutzer eine persönliche Tag-Cloud. Die Ähnlichkeit von Tag-Clouds lässt sich durch die Cosinus-Distanz bestimmen, so dass implizit eine Ähnlichkeit der dazugehörigen Personen in Bezug auf ein thematisches Interesse geschlossen werden kann. Auf dem sozialen Layer lässt sich überprüfen, ob bereits eine Beziehung zwischen diesen „ähnlichen“ Personen besteht (L3.2 oder L3.3). Ist dies nicht der Fall, kann eine Social Software-Applikation eine explizite Vernetzung vorschlagen [ViKa2009].

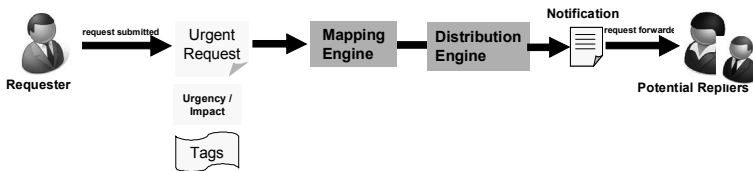
### **5.2 Personen folgen Themen**

Auf dem Social Layer L3.3 ist berücksichtigt, dass Personen anderen Personen „folgen“ können. Weiterhin ist denkbar, dass Personen einem oder mehreren Tags folgen. Durch diese Funktion werden Personen direkt mit Tags verbunden, ohne dass sie nutzergenerierte Inhalte erzeugen müssen. Dies ermöglicht eine Profilbildung auch für Nutzer, die „nur“ lesend auf Social Software-Anwendungen zugreifen.

## **6 Nutzenpotenzial für Urgent Requests**

Sogenannte Urgent Request Systeme – eine Abwandlung klassischer Q&A Systeme – werden in Firmen zur gegenseitigen Hilfe in geschäftsrelevanten Aufgaben verwendet. Der Fragende wendet sich dabei an eine zunächst nicht genau spezifizierte Gruppe von potentiellen Antwortgebern. Eine derartige Funktion stellt bei Siemens das TechnoWeb 2.0 bereit, ein Tool zur Expertenvernetzung, das seit 2010 für Technologie-Experten der gesamten Siemens AG angeboten wird [KäHe2009]. Breite Akzeptanz und ein positiver ROI lassen sich jedoch nur erreichen, wenn nicht zu viele (Wissens-)Mitarbeiter durch die Anfrage in ihrer Tätigkeit unterbrochen werden und gleichzeitig genügend geeignete Mitarbeiter erreicht werden, um das Problem des Fragenden zu lösen. Die Zielgenauigkeit der Verteilung von Urgent Requests sollte daher durch das vorgestellte Web 3L Modell verbessert werden.





**Abbildung 6: Smarte Verteilung von Urgent Requests**

Abbildung 6 zeigt den prinzipiellen Ablauf beim Urgent Request Prozess, wobei die „Mapping Engine“ auf Module des Web 3L Modells zurückgreift. Ein Lösungsweg verwendet die Tags eines Requests und expandiert die Begriffe auf dem Semantischen Layer (L2.1, L2.2, L2.3). Daraufhin wird ein Abgleich mit den Tags der Netzwerke in TechnoWeb vorgenommen und damit eine Verbindung zu Personen (Social Layer L3.3) hergestellt. Erscheint die Anzahl der Personen zu groß, kann eine Filterung dadurch vorgenommen werden, dass nur Personen berücksichtigt werden, die eines oder mehrere der Tags bereits selbst verwendet haben. Insgesamt sind verschiedene Varianten anwendbar, mit Hilfe derer die Zahl der potentiellen Antwortgeber vergrößert oder eingeschränkt werden kann, um so eine optimale Menge an Adressaten zu finden.

## 7 Nutzenpotenzial für soziale, semantische Suche

Die Nutzenpotenziale bei der Suche reichen von einer Suchstring-Expansion bis hin zum Auflösen von Ambiguitäten (Mehrdeutigkeiten). Bei der Suche nach Informationsobjekten mit Hilfe von Schlüsselwörtern können dem Benutzer verwandte Begriffe als mögliche Alternativen für die Verfeinerung oder Erweiterung der Suche grafisch aufbereitet angezeigt werden.

Am Beispiel einer Tag-basierten Suche wird exemplarisch gezeigt, wie ein Suchergebnis mittels semantischem und sozialen Kontext angereichert werden kann (Abbildung 7).

**Advanced Tag Search**

search

filter: innovation x

Results 1-9 < 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 >

**Test\_SM**  
2010-12-20T10:08:43  
https://technweb.siemens.com/web/test\_sm/overview  
entry-level innovation emerging markets smart

**Invitation for "Be an inventor opportunity"- International Student Competition at Siemens Healthcare**  
2010-12-17T13:37:42  
https://technweb.siemens.com/c/blogs/find\_entry?entryId=8732597  
students think tank healthcare innovation

**Das neue Stromzähler**  
2010-12-16T11:43:42  
https://blogs.eps.siemens.com/sebs/paul.heller/stories/33661  
cc communication energy erlangen innovation löscher siemensstv strategy technology video

**Community Treffen in Erlangen**  
2010-12-15T13:59:51  
https://technweb.siemens.com/c/blogs/find\_entry?entryId=3375331  
innovation event events

**Knowledge and Innovation**  
2010-12-14T04:03:46  
https://technweb.siemens.com/c/blogs/find\_entry?entryId=8341825  
innovation knowledge management process

**Happy Birthday (Fröhliche Geburtstag), Werner!**  
2010-12-13T17:55:58  
https://blogs.eps.siemens.com/sebs/fauna.schachle/stories/33615  
3 birthday cary celebration customer excellence innovation stammisch video werner

**Project Management Webinar**  
2010-12-09T12:36:02  
https://blogs.eps.siemens.com/sebs/otto.werner/stories/30539  
pnews innovation project management web2.0


**State Visit - Factory Visits: Christian Wulff and Vladimir Putin at Siemens**  
2010-12-08T17:30:49


**Filter your search results**  
 Technweb  Blogosphere  Wikisphere


**Your relations: (1)**  
jerm  
Less ▲


**Everybodys relations: (8)**  
innovation management  
research  
open\_innovation  
collaboration  
change\_management  
invention  
inventors  
jerm  
Less ▲


**Related People: (15)**

 **Martin Hopf Erlangen** 47 (> SCD)

 **Paul Hübner** 41 (> SCD)

 **Hans Buthopp** 20 (> SCD)

 **Michael Beckmann** 18 (> SCD)

 **Thomas Weyer** 14 (> SCD)

Less ▲

**Co-Tags: (20)**  
More ▼

**DMOZ: (0)**  
More ▼

**Abbildung 7: Tag-basierte soziale, semantische Suche**

Im Einzelnen zeichnet sich die soziale, semantische Suche durch folgende Funktionalitäten aus:

- **Anzeige von Co-Tags im Suchergebnis**  
Im Suchergebnis wird angezeigt, welche Tags zusätzlich zum angegebenen Suchterm die jeweilige Ressource kennzeichnen.
- **Angabe von verwandten Termen zu den Suchtermen**  
In der rechten Spalte werden mit dem Suchterm (semantisch) verbundene Terme angezeigt. Dies umfasst mit dem Thesaurus Editor erstellte eigene Relationen, semantische Relationen aus der Community sowie Co-Tags und aus dem Kategoriensystem von DMOZ importierte Relationen.
- **Angabe von relevanten Personen**  
Es wird angezeigt, welche Personen das Tag (zuletzt) häufig verwendet haben.

Durch Aktivieren eines der aufgeführten Terme kann die Suche fokussiert (UND-Verknüpfung) oder durch eine ODER-Verknüpfung expandiert werden. Dies ermöglicht ein dynamisches Navigieren durch den Informationsraum, ohne dass weitere Begriffe eingetippt werden müssen.

## **8 Fazit**

Das beschriebene Web 3L Modell ist die konzeptionelle Basis für die Implementierung eines Tagging Frameworks, welches sowohl semantische als auch soziale Verknüpfungen zwischen nutzergenerierten Inhalten unterstützt.

Es wurde damit ein Lösungsansatz beschrieben, der Web 2.0 Metadaten (Tagging) mit semantischen Technologien (Thesaurus-Editor) verbindet. Konkrete Nutzenpotenziale wurden in Kapitel 5 sowie den beiden Anwendungsszenarien (Kapitel 6 und 7) dargestellt. Diese können sowohl die Informationsversorgung als auch die Vernetzung von Wissensarbeitern in Unternehmen durch eine geeignete Aufmerksamkeitssteuerung erheblich verbessern.

## **Forschungsrahmen**

Die beschriebenen Arbeiten erfolgen im Kontext des Förderprojekts Theseus Alexandria. THESEUS ist ein vom Bundesministerium für Wirtschaft und Technologie (BMWi) initiiertes Forschungsprogramm mit dem Ziel, eine neue internetbasierte Wissensinfrastruktur zu entwickeln, um das Wissen im Internet besser zu nutzen und zu verwerten. Im Rahmen des Anwendungsszenarios Alexandria soll in Anlehnung an die gleichnamige antike Bibliothek eine Endnutzerorientierte Wissensplattform aufgebaut werden [Thes2007].

## **Literatur**

- [AnSa2007] Angeletou, S., Sabou, M., Specia, L., Motta, E., Bridging the gap between folksonomies and the semantic web: An experience report. In Workshop: Bridging the Gap between SemanticWeb andWeb, volume 2, 2007.
- [CaBe2008] Cattuto, C., Benz, D., Hotho, A., Stumme, G., Semantic grounding of tag relatedness in social bookmarking systems. The Semantic Web - ISWC 2008, pages 615–631, 2008.
- [Ehms2010] Ehms, K., Persönliche Weblogs in Organisationen - Spielzeug oder Werkzeug für ein zeitgemäßes Wissensmanagement? Doktorarbeit, Universität Augsburg, 2010
- [Gran1973] Granovetter, Mark S. (May 1973), „The Strength of Weak Ties“, American Journal of Sociology 78 (6): 1360–1380

- 
- [KäHe2009] Käfer, G., Heiss, M.: Wissensnetze als Basis für Enterprise 2.0 - Ein Erfahrungsbericht der Siemens AG aus 10 Jahren Wissensvernetzung als Basis für die Einführung von Enterprise 2.0. In: "Geteiltes Wissen ist doppeltes Wissen!" M. Bentele, R. Hochreiter, H. Krcmar, M. Weber (Hrsg.), KnowTech 2009, 201-205
- [KaLa2009] Kammergruber, W., Langen, M., Tagging als soziales Bindeglied für Communities. In: Virtuelle Organisation und Neue Medien 2009 GeNeMe 2009 Dresden: (2009), p. 35--43.
- [LaKa2008] Langen, M., Kammergruber, W.: Tagging vs. Ontologien? In: Bentele, M., Gronau, N., Hochreiter, R., Schütt, P., Weber, M. (Hrsg.) KnowTech 2008 Erhöhen Sie Ihren BQ!, 10. Kongress zum IT-gestützten Wissensmanagement, Frankfurt am Main, KnowTech 2008
- [LaKa2010] Langen, M., Kammergruber, W., Vom Tagging zu leichtgewichtigen Ontologien. In: Bentele, M., Gronau, N., Schütt, P., Weber, M. (Hrsg.) KnowTech 2010, 12. Kongress zum IT-gestützten Wissensmanagement
- [Pete2009] Peters, I.: Folksonomies: Indexing and Retrieval in Web 2.0., De Gruyter, Saur, München, 2009.
- [Thes2007] <http://theseus-programm.de/scenarios/de/alexandria> Förderkennzeichen 01MQ07008 des BMWi
- [ViKa2009] Viermetz, M., Kammergruber, W., Ziegler, C., Discovering Communities of Interest in a Tagged Online Environment. CASoN, 2009.
- [WaDe2006] Wahlster, W., Dengel, A., Web 3.0: Convergence of Web 2.0 and the Semantic Web. Technology Radar Feature Paper Edition II/2006