

Simone Braun, Andreas Schmidt und Valentin Zacharias

Mit Social Semantic Bookmarking zur nützlichen Ontologie

With Social Semantic Bookmarking towards useful ontologies

Kollaborative Ontologieentwicklung_SOBOLEO_SKOS_Social Bookmarking_Social Semantic Bookmarking_Ontologiereifung

Zusammenfassung. Dieser Artikel präsentiert das SOBOLEO-System und das zugrundeliegende Ontologiereifungsprozessmodell für die kollaborative Ontologieentwicklung. Man kann beobachten, dass die meisten aktuellen Ontologieentwicklungsprozesse und -werkzeuge von einer organisatorischen, personellen, technischen und zeitlichen Trennung zwischen Entwicklung und Nutzung der Ontologie ausgehen – eine Trennung, die wiederum häufig zu kostspieligen und nicht an ihre Nutzung angepasste Ontologien führt.

Unser Ansatz überwindet die durch diese Trennung verursachten Schwierigkeiten mittels Methoden und Werkzeugen, die die Nutzer der Ontologie in die Lage versetzen, diese selbst zu entwickeln, und zwar im gleichen System, das die Ontologie nutzt und zu dem Zeitpunkt und Umfang wie jeweils nötig (arbeitsintegriert)..

Summary. This paper presents the SOBOLEO system and the underlying Ontology Maturing process model of collaborative ontology development. We can observe that most current ontology development tools and methodologies stipulate a temporal, organizational and technical separation between development and use of the ontology – a separation that in turn can lead to expensive ontologies not adapted to their use.

Our approach overcomes the difficulties caused by this separation through the use of methods and tools that enable the users of the ontology themselves to continuously develop the ontology within the same system that uses the ontology, just in the point in time and to the extent that is needed (work integrated).

1. Einleitung

Ontologien versprechen eine Verbesserung des Informationsmanagements durch das Verbinden der Informationen mit maschinenverständlichem Hintergrundwissen. Durch dieses Hintergrundwissen, das den impliziten Kontext von Nutzern darstellt, sollen Computerprogramme in der Lage versetzt werden, Nutzerwünsche genauer zu verstehen und zu erfüllen. In der praktischen Nutzung sind Ontologien jedoch bisher der endgültige Durchbruch verwehrt geblieben. Eine mögliche Erklärung hierfür liegt in dem hohen Aufwand und der hohen

Komplexität der Ontologieentwicklung; eine hohe Komplexität, die nicht nur zu hohen Kosten führt, sondern auch Experten erfordert und damit meist eine organisatorische, personelle, technische und zeitliche Trennung zwischen Ontologieentwicklung und -nutzung. Diese Trennung zwischen Entwicklung und Nutzung wiederum kann zu Ontologien führen, die nicht auf Aufgabe und Nutzerkreis zugeschnitten sind, z. B. durch Überformalisierung oder durch das Fehlen von neuen, aber aktuell relevanten Begriffen.

In unserer Arbeit untersuchen wir mit dem hier vorgestellten SOBOLEO-System und dem Ontologiereifungsprozessmodell, inwieweit sich diese Probleme umgehen lassen, indem mit geeigneten

Werkzeugen und Formalismen die Nutzer der Ontologie in die Lage versetzt werden, diese selbst (weiter-) zu entwickeln; integriert in ihre Arbeit und die Nutzung der Ontologie und Stück für Stück, nur so weit wie gerade sinnvoll. Die dabei betrachtete Domäne ist die der gemeinschaftlichen Organisation von Internet-Informationsressourcen durch eine Ontologie; die Erweiterung von Social Bookmarking um Semantik zu *Social Semantic Bookmarking*.

Dieser Artikel beginnt mit einer Einführung in Social Semantic Bookmarking. Im darauf folgenden Abschnitt wird SOBOLEO als eine Implementierung des Social Semantic Bookmarking Ansatzes vorgestellt. Anschließend wird das Ontologiereifungsprozessmodell zur Beschrei-

bung einer solchen kollaborativen und arbeitsintegrierten Ontologieentwicklung vorgestellt. Abschließend werden verwandte Arbeiten diskutiert und zukünftige Entwicklungen skizziert.

2. Social Semantic Bookmarking

Social Semantic Bookmarking beschreibt ein Paradigma zur gemeinschaftlichen Strukturierung eines Informationsbestandes von Web-Ressourcen. Social Semantic Bookmarking stellt eine Kombination von Ideen aus dem Umfeld von Social Bookmarking und Semantischer Annotation dar und adressiert die jeweiligen Schwächen dieser Ansätze. Im Folgenden werden zuerst Social Bookmarking und Semantische Annotationsansätze kurz vorgestellt, bevor eine Beschreibung von Social Semantic Bookmarking am Ende dieses Abschnittes folgt.

2.1 Social Bookmarking Systeme

Social Bookmarking Systeme (z. B. delicious.com) gehören zu den erfolgreichsten Web-2.0-Anwendungen. Diese Anwendungen erlauben es Nutzern, Referenzen auf Web-Ressourcen zu sammeln, zu verwalten und auszutauschen. Zu jeder Ressource können die Nutzer ihrer Ansicht nach zur Beschreibung am besten geeigneten Schlagwörter, sog. *tags*, zuweisen. Dies erleichtert die Organisation, Navigation und Suche der Sammlung.

Der soziale Aspekt dieser Systeme zeigt sich darin, dass die verwendeten Tags von und für andere Teilnehmer sichtbar sind. So kann man eigene Tags mit anderen Teilnehmern teilen und deren Tags ebenfalls nutzen. Auch ist es möglich, über gemeinsam genutzte Tags oder gemeinsame annotierte Ressourcen andere Nutzer mit gleichen Interessen zu finden. Letztlich ist es auch möglich über die Nutzer zu deren Ressourcen zu gelangen und so neue interessante Themen zu entdecken.

Durch die Verwendung von Tags können Nutzer ohne Restriktion ihre eigene Sicht auf die Ressourcen und ihre Meinung vertreten. Auch müssen keine komplexen, vorgegebenen Schemata oder Syntax erlernt werden (s. McGregor & McCulloch 2006). Vermag dies auf den ersten Blick in einer chaotischen Samm-

lung enden, so zeigen Untersuchungen, dass sich in solchen Systemen aus den Tags (und ihrer Vergabe) ein gemeinsames Verständnis und Vokabular, auch als *folksonomy* bezeichnet, unter den Nutzern entwickelt (vgl. Golder & Huberman 2006, Marlow et al. 2006, Sen et al. 2006). Awareness-Mechanismen, wie z. B. Wortwolken (*tag clouds*) oder das Vorschlagen bereits verwendeter Tags, beeinflussen diesen Effekt zusätzlich positiv. Es entstehen somit Metadaten zu Inhalten ohne große Mühen und Kosten und dies flexibel und arbeitsintegriert, d. h. bei der Nutzung des Systems.

Die entstehenden Folksonomies sind jedoch nur schwach oder gänzlich unstrukturiert. Die fehlende Semantik erschwert eine effiziente Unterstützung beim Suchen und (Wieder-)Finden von Informationsressourcen, insbesondere in komplexen Domänen, durch Probleme wie Mehrsprachigkeit (Tags in unterschiedlichen Sprachen), Polysemie und Homonymie (Tags mit mehreren Bedeutungen), Synonyme und unterschiedliche Schreibweisen (verschiedene Tags mit gleicher Bedeutung) oder Tags auf unterschiedlichen Abstraktionsebenen (vgl. Golder & Huberman 2006, Guy & Tonkin 2006).

2.2 Semantische Annotationssysteme

Diese sich aus der Unstrukturiertheit von Folksonomies ergebenden Schwächen ließen sich durch die Verwendung von semantischen Annotationen und Ontologien beheben. Ontologien ermöglichen die Erstellung u. a. von semantischen Relationen, um etwa unterschiedliche Abstraktionsebenen abbilden bzw. die Verwandtheit einzelner Elemente ausdrücken zu können. Gleichzeitig ermöglichen Ontologien die Unterstützung für Synonyme, trennen Homonyme und bieten Hilfe für Mehrsprachigkeit.

Im Gegensatz zu einfach nutzbaren Tags sind Ontologien jedoch sehr viel schwieriger handzuhaben. Existierende Werkzeuge zur Ontologieerstellung (z. B. Protégé) sind für die meisten Endanwender viel zu kompliziert. Daher findet man bei ontologiegestützten Anwendungen eine strikte Trennung zwischen Ontologieentwicklung bzw. -pflege und der Nutzung vor: Wissensingenieure entwickeln und pflegen zusammen mit Domänenexperten die für eine Domäne relevante Ontologie im Voraus. Nutzer eines

semantischen Systems verwenden dann diese Ontologie zur Annotierung von Ressourcen. Dem Endanwender wird es hierdurch – im Gegensatz zu Social Tagging Systemen – nicht ermöglicht, selbst Domänenwissen einzubringen und zusammen mit anderen in einem Lernprozess ein gemeinsames Verständnis und Vokabular zu entwickeln. Die Folge davon sind Ontologien, die nicht ideal auf ihren Anwendungsbereich zugeschnitten sind. Aufgrund des oftmals nur eingeschränkten Verständnisses der Domäne durch Wissensingenieure entstehen fehlerhafte, unvollständige oder nicht für die Benutzer nachvollziehbare Ontologien (vgl. Barker et al. 2004). Statt dass Benutzer Änderungen selbst vornehmen können, müssen Änderungswünsche gesammelt, von Wissensingenieuren verstanden und in die Ontologie eingepflegt werden. Im Gegenzug müssen diese Änderungen an der Ontologie dann wieder den Anwendern erklärt werden (vgl. Hepp 2007). Dieser zeitaufwendige Prozess führt dazu, dass das Vokabular für die Beschreibung von aktuell relevanten Ressourcen, gerade in sich schnell ändernden Domänen, nicht rechtzeitig bereitsteht (vgl. Hepp 2007). Damit sinkt allerdings der Nutzen semantischer Anwendungen erheblich, und Anwender werden demotiviert, an der Weiterentwicklung der Ontologie mitzuwirken.

2.3 Von Social Bookmarking und Semantischer Annotation zu Social Semantic Bookmarking

Es zeigt sich also, dass: (1) Folksonomies durch ihre geringe Struktur inhärenten Limitierungen unterliegen und (2) Ontologien durch die Trennung von Erstellung und Anwendung in den derzeit vorherrschenden Methodiken problematisch und in vielen Domänen unbenutzbar sind. Demgegenüber haben (3) Folksonomies mit ihrem Erfolg im großen Stile die Nützlichkeit für große Nutzergruppen gezeigt und (4) Ontologien ein deutlich größeres Potential für weitergehende Anwendungsmöglichkeiten.

Daher ist unser Ziel, die Vorteile von Social Bookmarking mit den Vorteilen von Semantischen Annotationen zu verbinden und dabei die jeweiligen Nachteile zu überwinden. Ausgehend von einfachen Tags sollen alle Nutzer in die kollaborative Entwicklung von Ontologien einbezogen werden. Hierzu wird die

Entwicklung der Ontologie in deren Nutzung, z. B. für die Annotation von Web-Ressourcen, integriert. Alle Teilnehmer einer Gemeinschaft können aus der Nutzung heraus entstehende neue Konzeptideen in Form von Tags für die Weiterentwicklung von Ontologien einbringen. Diese werden in der Gemeinschaft aufgegriffen, zu Konzepten konsolidiert und erweitert und durch semantische Relationen zu leichtgewichtigen Ontologien formalisiert. Werden Änderungen in der Ontologie benötigt, können diese durch die Nutzer direkt durchgeführt werden – die zeitaufwendige Trennung von Ontologieentwicklung und -nutzung entfällt somit.

Das sich aus dieser Kombination ergebene Paradigma des Social Semantic Bookmakings erlaubt die Annotation von Web-Ressourcen durch Nutzer mit Konzepten aus einer Ontologie, deren Definition und Beschreibung kollaborativ im selben System und durch die gleichen Nutzer weiterentwickelt wird.

3. Das SOBOLEO System

Das SOBOLEO (Social BOOKmarking and Lightweight Engineering of Ontologies) System ist eine Implementierung des Social Semantic Bookmarking Paradigmas. Es unterstützt eine Gemeinschaft von Nutzern dabei, gemeinsam eine Sammlung von Internetressourcen zu erstellen und mithilfe einer leichtgewichtigen Ontologie zu verwalten. Bei der Erstellung des Systems wurde dabei besonderer Wert darauf gelegt, dass die Ontologie einfach zu erstellen ist und die Erstellung stückchenweise während der Nutzung erfolgen kann.

Im allgemeinen Verständnis beinhalten leichtgewichtige Ontologien Konzepte, Konzepttaxonomien, Beziehungen zwischen Konzepten und diesen beschreibende Eigenschaften (vgl. Gomez-Perez et al. 2004). Das SOBOLEO System erlaubt es, leichtgewichtige Ontologien entsprechend dem SKOS (Simple Knowledge Organization System) Vokabular zu erstellen (Miles & Bechhofer 2009). Das SKOS Vokabular erlaubt es, Konzepte mit Bezeichnern und Beschreibungen in verschiedenen Sprachen zu versehen und sie über die Relationen „enger“, „weiter“ und „ähnlich“ in Beziehung zu setzen.

SOBOLEO besteht aus vier Hauptteilen: (1) einem kollaborativen Echtzeit-Editor zur Änderung der Ontologie, (2) einem Werkzeug zur Annotation von Webressourcen, (3) einer semantischen Suchmaschine zur Suche nach annotierten Ressourcen und Personen und (4) einem Ontologie-Browser zum Durchstöbern der Ontologie und der Bookmark-Sammlung. Jeder dieser Teile wird im Folgenden detailliert beschrieben.

SOBOLEO ist als eine Web-Anwendung realisiert und läuft vollständig – unter der starken Verwendung von AJAX-Technologien – im Browser. Eine In-

stallation ist nicht notwendig, allerdings kann der Nutzer das Annotationswerkzeug über ein Bookmarklet (eine Zeile Javascript, die wie ein Browser-Lesezeichen gehandhabt wird) in den Browser integrieren. Die einzelnen Hauptteile werden im Folgenden beschrieben.

3.1 Annotation von Webressourcen

Mittels des Annotationswerkzeugs (Bild 1), das über ein im Browser gespeichertes Bookmarklet zugänglich ist, kann der Nutzer eine momentan im Browser betrachtete Internetressource annotieren

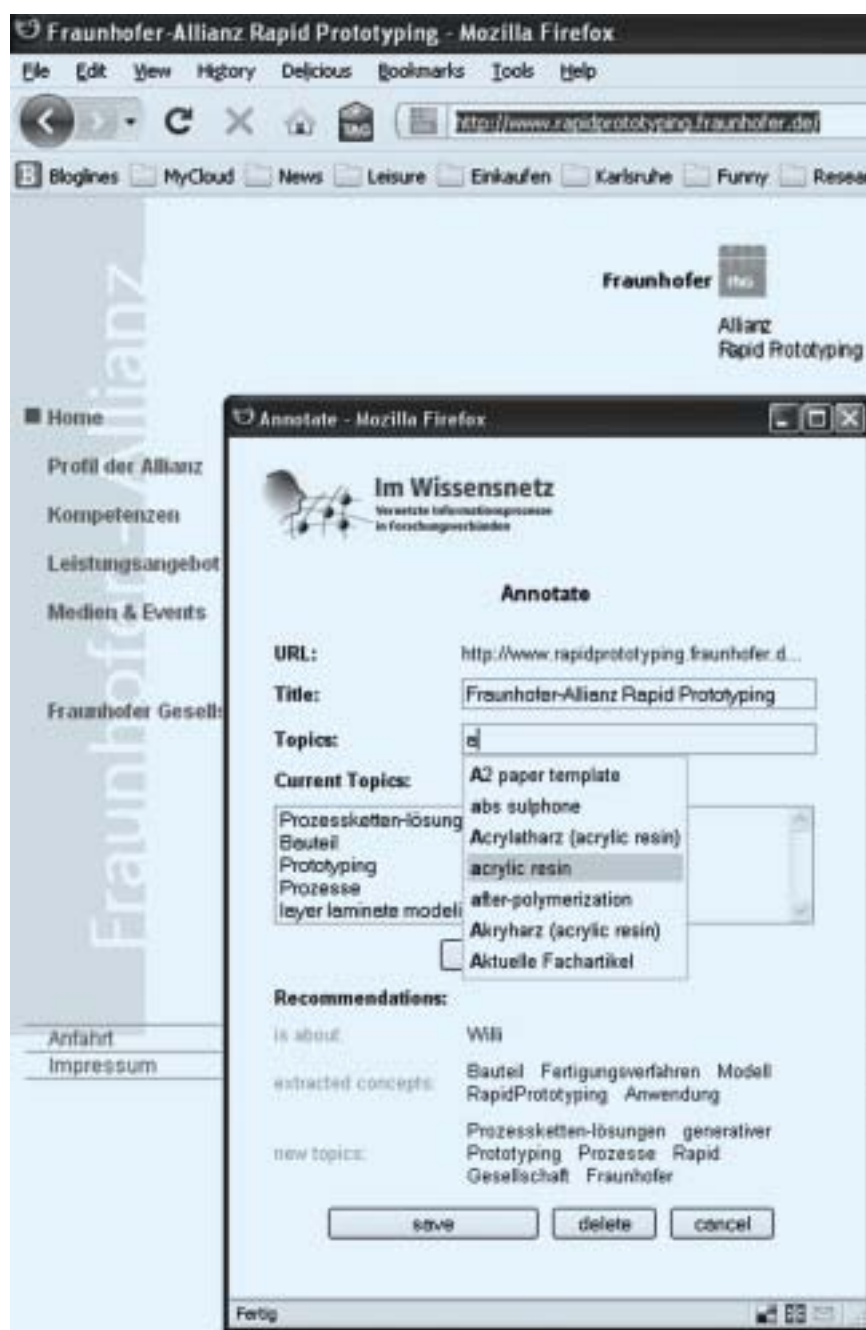


Bild 1: Annotation einer Webseite mit AutoComplete-Assistenzfunktion und Annotationsempfehlungen

und den Link der gemeinsamen Bookmark-Sammlung hinzufügen. Es öffnet sich hierfür ein Popup-Fenster und das URL- und Titelfeld werden mit den Angaben der gerade geöffneten Internetseite ausgefüllt. Ist die angegebene URL bereits in der Sammlung enthalten, so werden auch die bereits existierenden Annotationen angezeigt. Um die Ressource zu annotieren, kann der Nutzer Konzepte aus der Ontologie oder beliebige Schlagwörter verwenden; neue Schlagwörter werden dabei automatisch als „prototypische Konzepte“ der Ontologie hinzugefügt. Die AutoComplete-Assistenzfunktionalität bei der Annotation erleichtert dem Nutzer die Eingabe von Begriffen zur Annotation. Gleichzeitig schafft sie Awareness über bereits existierende Konzepte in der Ontologie und ermuntert dadurch den Nutzer, möglichst diese Konzepte wiederzuverwenden. Weitere Unterstützung für den Nutzer bietet die Annotationsempfehlungskomponente. Mit dieser werden automatisch aus der zu annotierenden Internetressource Konzepte der gemeinsamen Ontologie und neue (noch nicht in der Ontologie enthaltene), aber für die Ressource charakteristische Begriffe extrahiert sowie die Ressource klassifiziert und dem Benutzer als Empfehlung im Annotationsprozess vorgeschlagen. Diese Assistenzfunktionalität hilft dem Benutzer neben der Empfehlung zur Annotation auch, sich einen schnellen Überblick bezüglich des Inhalts eines fremden Dokuments zu verschaffen. Schließlich wird beim Speichern einer Annotation noch der Inhalt der Seite über einen Crawler ermittelt und zusammen mit den zur Annotation verwendeten Konzepten in den Index aufgenommen.

3.2 Editieren der Ontologie

Zum kollaborativen Erstellen und Verändern der Ontologie steht ein Editor zur Verfügung (Bild 2). Dieser Editor erlaubt es mehreren Benutzern gleichzeitig die Ontologie zu editieren. Änderungen werden sofort für alle Nutzer im Editor sichtbar und in der Nutzung der Ontologie wirksam, z. B. bei der Annotation von Webressourcen oder in der Suche in der Bookmark-Sammlung. Die Benutzeroberfläche des Editors gliedert sich in drei Teile: die Baumansicht der Ontologie, die Detailansicht des momentanen Konzeptes in der Mitte und ein Nachrichtenfenster rechts (hier nicht zu sehen). In der



Bild 2: Echtzeit-Ontologieeditor zum kollaborativen Erstellen und Verändern der Ontologie

Baumansicht der Ontologie werden die Konzepte mit ihren bevorzugten Bezeichnungen (skos:prefLabel) und ihren Relationen zu engeren und weiteren Konzepten dargestellt (skos:narrower/skos:broader). Der Baum kann durch Drag-and-Drop, Menübefehle oder Tastenkombinationen verändert werden. Tasten erlauben das Hinzufügen und Entfernen von Konzepten. Unter „prototypical concepts“ werden alle durch die Annotation neu hinzu gekommenen Begriffe gesammelt. Diese können dann von den Nutzern in die Ontologie einsortiert und konsolidiert werden. Eine Suchkomponente erleichtert dem Nutzer die Navigation im Ontologiebaum. Direkt über der Baumansicht kann der Nutzer über ein Eingabefeld nach Konzepten in der Ontologie suchen. Dabei wird der Nutzer ebenfalls wieder durch eine automatische Vervollständigung unterstützt.

In den drei oberen Eingabefeldern des „Edit Concept“-Bereichs kann der Nutzer den bevorzugten bzw. die alternativen Bezeichner und die Beschreibung des ausgewählten Konzepts erstellen und ändern. Darunter kann der Benutzer die Relationen „engeres Konzept“, „weiteres

Konzept“ und „ähnliches Konzept“ bearbeiten (skos:narrower/skos:broader/skos:related). Auch hier wird der Nutzer wieder durch eine AutoComplete-Komponente unterstützt.

Änderungen an der Ontologie erfolgen in Echtzeit und sind sogleich für alle Nutzer sichtbar. Um den Nutzer über Aktivitäten an der Ontologie auf dem Laufenden zu halten, erzeugt das System für jede Veränderung automatisch eine Nachricht, die genau darüber informiert, welcher Nutzer welche Änderung vorgenommen hat. Diese Nachrichten werden in einer separaten Anzeige im rechten Bildschirmbereich dargestellt. Diese dient neben der Information über Modifikationen auch als Chatfenster zum Austausch mit anderen Nutzern. Ein Eingabefeld unterhalb der Anzeige erlaubt es, Nachrichten an alle anderen Nutzer, die zur gleichen Zeit die Ontologie bearbeiten, zu senden und so ggf. konkurrierende Änderungen zu diskutieren.

3.3 Navigation durch die Ontologie und Bookmark-Sammlung

Im Browse-Bereich kann der Nutzer durch die Ontologie und den Inhalt der

Results: 18 documents and 5 people found
 I understood that you searched for: [sintering anwendung sintern](#)
 Broaden your query (also show documents annotated with): [laser sintering](#) [selective laser sintering](#) [RP technique](#)
 Narrow your search (only show documents annotated with): [laser sintering](#) [rapid tooling](#) [3D-Printer](#)

RTejournal - Forum für Rapid Technologie - Laser-Sintern. Die SchiÄsselltech...
[rapid tooling rapid manufacturing laser sintering](#)
 RTejournal - Forum für Rapid Technologie - Laser-Sintern. Die SchiÄ... Beschreibung lasergesinterter Teile. Obwohl der Einsatz von Laser-Sintern für Rapid Prototyping und ... Rapid Technologie Direkt zum Inhalt
<http://www.rtejournal.de/archivausgabe/1/14/gd/>

Lasersintern ? Wikipedia
[laser sintering](#)
Anwendung beim Sintern von Kunststoff und teilweise beim Sintern von Metall mit Spezialinterpulver möglich ... Selektives Lasersintern (SLB) ist ein Verfahren, um räumliche Strukturen durch Sintern aus ... komplizierter Teile verwendet. Der Trend geht allerdings dahin, die Technologie auch als Rapid
http://de.wikipedia.org/wiki/Selbstlaser_Sintern

Selective-Laser-Sintering
[selective laser sintering](#)
 ... die sich aufschmelzen lassen. Verwendete Werkstoffe für das Kunststoff-Laser Sintern sind z.B. ... ModelTechnik Rapid Prototyping GmbH | Diegelsb: 3b
http://www.modeltechnik.de/selective_laser_sintering_4103.html?productid=7379387&id=891&id=76510...

RTejournal - Forum für Rapid Technologie - Rapid Tooling für den Spritzguss...
[rapid tooling laser sintering Spritzguss](#)
 1. Einleitung Die Anwendung von Rapid-Prototyping (RP) Technologien hat sich in den letzten ... RTejournal - Forum für Rapid Technologie - Rapid Tooling für den Spritzguss ... Rapid Technologie Direkt zum Inhalt
<http://www.rtejournal.de/archivausgabe/1/20/gd/>

RTejournal - Forum für Rapid Technologie - Grundlagen des Rapid Prototyping
[3D printing based deposition modeling laminated object manufacturing laser sintering Extrusion](#)
 Verfahren und insbesondere der Materialien wird mehr und mehr das Feld der direkten Anwendung der Rapid ... RTejournal - Forum für Rapid Technologie - Grundlagen des Rapid Prototyping ... für Rapid Technologie
<http://www.rtejournal.de/archivausgabe/1/15/gd/>

Dr. Mark Heffe
 used narrower concept 1x [selective laser sintering](#) 3x [laser sintering](#) and looked at or searched for one of the concepts.
[contact Dr. Mark Heffe](#)

Frithjof Baumann
 used narrower concept 2x [laser sintering](#) and looked at or searched for one of the concepts.
[contact Frithjof Baumann](#)

Dr. Hans-Jürgen Richter
 used narrower concept 1x [selective laser sintering](#) 1x [3D-Printer](#) and looked at or searched for one of the concepts.
[contact Dr. Hans-Jürgen Richter](#)

Neyir Sevilmis
 used narrower concept 1x [selective laser sintering](#) and looked at or searched for one of the concepts.
[contact Neyir Sevilmis](#)

Valentin Zacharias
 used narrower concept 1x [laser sintering](#)
[contact Valentin Zacharias](#)

Bild 3: Semantische Suche zum Suchen und Wiederfinden von annotierten Internetressourcen und Kompetenzträgern

Bookmark-Sammlung navigieren. Ausgehend von den Wurzelkonzepten („root concepts“), kann sich der Nutzer durch die Konzepte der Ontologie klicken. Dabei wird das aktuell ausgewählte Konzept mit seinen bevorzugten und alternativen Bezeichnungen sowie seiner Beschreibung angezeigt. Zusätzlich werden alle zugehörigen engeren, weiteren und ähnlichen Konzepte als Links zur weiterführenden Navigation eingeblendet. Nach den Konzeptdetails folgen, als „Newest Documents“ bezeichnet, aufgelistet alle Ressourcen, welche mit dem aktuell ausgewählten oder einem engeren Konzept annotiert sind. Die Ressourcen werden mit ihrem Titel und den Konzepten, mit denen sie annotiert sind, angezeigt. Die Reihenfolge wird durch den Zeitpunkt, zu dem sie in die Sammlung aufgenommen wurden, bestimmt. So erscheinen die neuesten Ressourcen zuoberst. Über einen Klick auf den Titel gelangt man direkt zur Webseite der Ressource (vgl. auch Bild 3).

Über die annotierten Internetressourcen hinausgehend bietet der Browse-Bereich auch eine themenbezogene Bereitstellung von Personen – als „Interesting People“ dargestellt. Dies erfolgt nutzungsintegriert und mittels des Hintergrundwissens über die eingebundene Domänenontologie. Das bedeutet, dass

die personenbezogenen Profildaten u. a. automatisch aus den Nutzungsaktivitäten der Anwender gewonnen werden. Grundannahme hierbei ist, dass die Nutzung und insbesondere Nutzungshäufigkeit Indikatoren für die Expertise des Benutzers sind.

Es werden Personen aufgelistet, welche das aktuell ausgewählte oder ein engeres Konzept bereits zur Annotation von Web-Ressourcen verwendet haben oder danach in SOBOLEO gesucht bzw. es sich im Browse-Bereich angesehen haben. Ausgehend von der Häufigkeit dieser Aktivitäten erfolgt dann das Ranking der Personen, wobei die Verwendung zur Annotation und des aktuell ausgewählten Konzepts höher gewichtet werden. Zusätzlich macht das System auf der Startseite des Browse-Bereichs mit den Wurzelkonzepten der Ontologie auf andere Nutzer aufmerksam, die ein ähnliches Profil wie man selbst aufweisen. Auf diese Art und Weise wird eine Ambient Awareness über vorhandene Fähigkeiten und Interessen geschaffen, die Vernetzung der Nutzer untereinander und der Aufbau von Teams und Communities gefördert.

3.4 Ontologiebasierte Suche nach Dokumenten und Personen

Die semantische Suchmaschine in SOBOLEO erlaubt das Suchen und Wieder-

finden von Internetressourcen in der gemeinsamen Bookmark-Sammlung (Bild 3). Ähnlich gängiger Suchmaschinen können auch in SOBOLEO Suchbegriffe in ein Textfeld eingegeben werden. Die eingegebene Zeichenkette wird dabei jedoch auf Bezeichner von Konzepten hin analysiert. Wird ein Bezeichner eines Konzeptes erkannt, wird dies vom System als Referenz auf dieses Konzept verstanden, das heißt es sucht nach Internetressourcen, welche mit diesem Konzept oder dessen engeren Konzepten annotiert sind. Die semantische Suche ist mit der Volltextsuche in den Inhalten aller annotierten Internetseiten integriert. Die Reihenfolge der Ergebnisse ergibt sich durch die Kombination aus Volltext- und semantischer Suche. Auf der Suchergebnisseite werden alle gefundenen Ressourcen mit ihrem Titel (als Link auf die ursprüngliche Internetseite), ihren annotierten Konzepten, einem kurzen Auszug des Seiteninhalts (mit Hervorhebung der Suchbegriffe) und der genauen URL aufgelistet. Über einen Editier-Link zu jeder gefundenen Ressource, kann der Nutzer deren Annotationen einfach und schnell modifizieren oder löschen. Hierfür öffnet sich das bekannte Annotations-Popup-Interface. Darüber hinaus werden oberhalb der Ergebnisliste alle von der SOBOLEO-Suchmaschine aus der Suchanfrage

verstandenen Konzepte angezeigt. Abhängig von der Eingabe, den Ergebnissen und der Ontologie schlägt das System weitere Möglichkeiten zur Suchanfragenverfeinerung bzw. -ausweitung vor.

Die Suche ermöglicht dem Benutzer, auch Personen zu einem bestimmten Thema zu finden. Die Suche nach Personen erfolgt hierbei integriert mit der Suche nach annotierten Internetressourcen mit der Eingabe in das Suchfeld. Ebenso werden nicht nur die erkannten Konzepte, sondern auch deren Unterkonzepte unter Ausnutzung der broader-narrower-Relation in der Ontologie mit in die Personensuche einbezogen. So wird etwa im vorliegenden Beispiel keine Person gefunden, die das von der Suchmaschine aus der Eingabe erkannte Konzept „sintering“ verwendet hat. Aber fünf Personen haben z. B. mit einem Unterkonzept dessen annotiert. Entsprechend der Häufigkeit und Verteilung in der Nutzung der identifizierten Konzepte bzw. Unterkonzepte erfolgt dann das Ranking der gefundenen Personen. So wird die Verwendung bei der Annotation höher bewertet als bei der Suche oder im Browse-Bereich. Ebenso erhält die Verwendung des erkannten Konzepts eine höhere Gewichtung als die Verwendung eines Unterkonzepts. Dieses Ranking wird für den Benutzer durch die Anzeige der Ergebnisse, die in die Suche mit eingeflossen sind, transparent gemacht. Über einen Klick auf eine der gefundenen Personen kann der Benutzer direkt mit diesen z. B. über Email in Kontakt treten.

4. Der Weg zur Ontologie – Ontologiereifungsprozessmodell

4.1 Der Weg zur Ontologie – Theorie

Auch wenn die Verbindung von den als schwerfällig geltenden Ontologien mit freien Tagging-Systemen sinnvoll erscheint, um die Vorteile beider Welten miteinander zu verbinden, so ist hier ein Neudenken des Ontologieverständnisses an sich und der Entwicklungspfade erforderlich. Klassischerweise wird Ontologieentwicklung als ein einmaliger bzw. allenfalls sich in größeren Zeitabständen wiederholender Prozess gesehen, bei dem es darum geht, ein bestehendes Domänenverständnis zu formalisieren und

dabei unterschiedliche Perspektiven zum Konsens zu führen. Diese Sichtweise ist allerdings nur dann sinnvoll, wenn man davon ausgehen kann, dass die Ontologie ein stabiles, bereits etabliertes Strukturierungswissen abbildet, weil auch eine stabile Ontologie als kontrolliertes Vokabular angestrebt wird. Dies begründet den Ruf der „Schwerfälligkeit“, weil in langen Aktualisierungszyklen gedacht wird, und führt oft dazu, dass am Modellierungszweck „vorbei“ modelliert wird.

Wir sehen stattdessen die Ontologieentwicklung als kontinuierlichen Prozess an, der eingebettet und verwoben ist mit dem Nutzungsprozess, also in unserem Fall: dem Verwenden für Verschlagwortung und Suche. Dieser Prozess ist als Lernprozess zu verstehen, innerhalb dessen nicht nur eine Formalisierung stattfindet, sondern auch das Domänenverständnis an sich vertieft wird, so dass letztendlich als Qualitätsmerkmal ein Gleichgewicht entsteht zwischen:

- *Angemessenheit/Nützlichkeit* der Repräsentation der Domäne bzgl. des Verwendungszwecks und Nutzerkreises der Ontologie. Eine Ontologie kann nicht nutzungskontextfrei modelliert werden, sondern muss eine nützliche Abstraktion der Wirklichkeit darstellen. Dies ist nur zu erreichen durch eine enge Kopplung von Modellierung und Nutzung sowie direktes Feedback bei Änderungen an der Ontologie.
- *Sozialer Übereinkunft* bzgl. der Repräsentation des gemeinsamen Verständnisses unter den Nutzern. Eine Ontologie kann nicht außerhalb eines sozialen Kontextes modelliert werden, innerhalb dessen eine Übereinstimmung und Akzeptanz erzielt wird. Die Ontologie ist also Ergebnis eines kollaborativen Lern- und Aushandlungsprozesses. Die Nutzer vertiefen nach und nach ihr gemeinsames Verständnis über die reale Welt und erarbeiten ein (angemessenes) Vokabular zu deren Beschreibung.
- *Formalität/Struktur* als Voraussetzung für die Anwendung automatisierter Verfahren. Der wesentliche Vorteil von formalen Repräsentationen ist die Möglichkeit, sie adäquat automatisiert verarbeiten zu können. Formalität ist also nicht Ziel an sich für eine Ontologieentwicklung, sondern ein notwendiger Schritt, wenn automa-

tische Verfahren in der Lage sein sollen, das in der Ontologie enthaltene Wissen zu nutzen. Schaut man sich reale Anwendungsfälle an, so ist das regelmäßig nicht für alle Teile einer Ontologie möglich bzw. erforderlich, so dass mit einer Ko-Existenz unterschiedlicher Formalitätsgrade innerhalb einer Ontologie zu rechnen ist: unstrukturierter Text, informelle Tags und formale Konzepte. Betrachtet man insbesondere gerade aufkommende Ideen und Konzepte, so können diese nicht sogleich in eine Ontologie integriert werden, da sie noch nicht klar definiert sind. Dies bedeutet, dass der Aufbau einer Ontologie einem ständigen Weiterentwicklungsprozess unterliegt.

4.2 Das Ontologiereifungsprozessmodell: vom Social Tagging zur Ontologie

Diese Qualität oder auch „Reife“ der Ontologie entwickelt sich im Rahmen eines kollaborativen Lernprozesses, der sich auf der Basis des Wissensreifungsmodells (Maier & Schmidt 2007) beschreiben lässt:

- *Entstehen von Ideen:* Neu aufkommende Ideen werden von einzelnen Nutzern als Konzeptideen oder informelle Tags eingeführt; z. B. wird bei der Annotierung oder beim Suchen nach Ressourcen erkannt, dass es einen benötigten Tag bisher noch nicht gibt und dieser wird hinzugefügt. Diese Tags werden individuell benutzt und informell kommuniziert. Sie sind primär individuell (danach bemisst sich auch ausschließlich die Nützlichkeit).
- *Konsolidierung in Communities:* In der kollaborativen (Wieder-)Verwendung der Tags entwickelt sich ein gemeinsames Vokabular (Folksonomy) in der Community. Ziel ist hier, Vorteile durch Angleichung der Begrifflichkeiten zu erzielen, zu denen das vereinfachte Auffinden der Ressourcen anderer, die Nutzung der Urteile anderer etc. zählen. Hierbei werden Konzeptideen (die in Form von Tags vorliegen) verfeinert, andere verworfen; z. B. werden Ähnlichkeiten oder Unterschiede zwischen Tags festgestellt; etwa wenn bemerkt wird, dass die Verwendung eines synonymen Begriffs die Suche verbessert und damit die Nützlichkeit erhöht.

Dadurch wird auch die Formalität erhöht: aus unstrukturierten Tags werden einfache Synonymmengen.

- **Formalisierung:** Die Community beginnt die Konzepte in hierarchischen oder auch beliebigen Ad-hoc-Relationen zu strukturieren, z.B. im Zuge dessen die Nutzer unterschiedlicher Abstraktionsebenen gewahrt werden, was in sog. leichtgewichtigen Ontologien resultiert. Die soziale Übereinkunft erstreckt sich nun nicht mehr nur auf Begriffe und ihre (potentiell synonyme) Benennung, sondern auch auf die Unterordnung. Dort wo das Problem der unterschiedlichen Abstraktionsebenen zu beobachten ist, wird die Nützlichkeit gesteigert, da die damit verbundenen Suchprobleme gelöst werden.
- **Axiomatisierung:** Will man weitergehende Zusammenhänge für komplexere Fragestellungen repräsentieren und nutzbar machen, können Experten differenziertes Hintergrundwissen einbringen, das einen (oft nicht reflektierten, weil stabilen), Konsens in der Domäne repräsentiert. So können z.B. verbesserte logische Schlussfolgerungen ermöglicht werden. Durch den hohen Formalisierungsgrad ist eine spezielle Modellierungskompetenz erforderlich.

Die Ontologieentwicklung verläuft dabei (in Analogie zu Lernprozessen) nicht zwingend von Grund auf und auch nicht strikt linear, sondern vielmehr in verschiedenen Iterationen zwischen den unterschiedlichen Phasen. Auch bewegen sich die Konzepte nicht gleichzeitig durch den Reifungsprozess. So gibt es üblicherweise zu jedem Zeitpunkt individuell genutzte Tags, gemeinsame aber noch nicht formale Terminologie wie auch formal beschriebene Konzepte nebeneinander.

5. Verwandte Arbeiten

Im folgenden Abschnitt werden ähnliche Anwendungen vorgestellt, die Social Bookmarking um mehr semantische Unterstützung erweitern. Das heißt diese Anwendungen erlauben es ihren Nutzern, ihre Folksonomy über semantische Definitionen und Beschreibungen zu erweitern, z.B. durch die Ergänzung von Tags mit Hintergrundwissen über Syno-

nyme oder Unter-/Oberbegriffsrelationen.

Hauptunterscheidungsmerkmal ist dabei, neben Geltungsbereich und Ausnutzungsgrad der Semantik, insbesondere der Freiheitsgrad, der Nutzern in dem zugestanden wird, was sie selbst beitragen können und was automatisch durch das System ermittelt wird. So basiert das Fuzzy-System, ähnlich wie SOBOLEO, in hohem Maße auf der Nutzergemeinschaft und deren Beiträge. Mit dem Fuzzy-System (<http://www.fuzzy.com>), ein PhD-Projekt an der Universität Oslo, können Anwender neben Bookmarks für Webseiten auch Bookmarks u.a. für Bücher und Events organisieren. Fuzzy basiert auf Topic Maps zur Beschreibung der semantischen Strukturen. Ein wichtiges Ziel ist, die kollaborative Entwicklung von Ontologien zu unterstützen. Daher bietet es seinen Nutzern großen Freiraum, Tags und Relationen frei zu definieren und zu editieren. So können die Nutzer neben hierarchischen und ähnlichen Relationen aus 22 spezifischen, vordefinierten Assoziationstypen auswählen, um Tags miteinander zu verknüpfen. Diese Tag-Relationen gelten für das ganze System und sind auch von anderen Nutzern editierbar. Ein weiteres, im Mittelpunkt stehendes Konzept sind Abstimmungen und Bewertungen für Pflege- und Aufräumaktivitäten. So können die Nutzer Bookmarks, Tags mit denen ein Bookmark annotiert ist, Relationen zwischen Tags und andere Nutzer bewerten und abstimmen.

Die Anwendungen BibSonomy und GroupMe! hingegen limitieren ihre Nutzer auf einen bestimmten Typ von Relationen, den sie hinzufügen können. BibSonomy (<http://www.bibsonomy.org>) ist ein Forschungsprojekt der Knowledge and Data Engineering Group der Universität Kassel. Es bietet seinen Nutzern neben der Verwaltung von Internetressourcen auch Funktionalitäten zur Organisation von bibliographischen Daten, wie z.B. die Unterstützung zur Importierung und Exportierung bibliographischer Informationen im Bibtex-Format. Im Kern unterscheidet sich BibSonomy von gängigen Social Bookmarking Diensten dadurch, dass es seinen Nutzern die Möglichkeit bietet, Ober- und Unterbegriffsrelationen zwischen Tags zu erstellen. Jedoch sind diese Relationen nur lokal eingeschränkt wirk-

sam. Das heißt jeder Nutzer kann für sich selbst seine eigenen Relationen erstellen und pflegen, aber diesbezüglich nicht von den Beiträgen anderer Nutzern profitieren.

GroupMe! (<http://www.groupme.org>) wird von der Semantic Web Gruppe der Universität Hannover entwickelt. Die im Mittelpunkt des Systems stehende Funktionalität ist die Erweiterung der Tagging-Idee um das Konzept der „Gruppe“. Das heißt, alle annotierten Internetressourcen (z.B. Webseiten, Musik, Videos, Photos und Newsfeeds) können in Gruppen zusammengefasst und organisiert werden. Dadurch wird eine weitere Informationsebene eingeführt, die zur Suche und Navigation genutzt werden kann.

Die kommerziellen Systeme Twine und ZigTag zielen insbesondere auf hohe automatische Unterstützung. Die Semantik wird bei diesen Systemen intern abgeleitet und vordefiniert. Die Nutzer haben darauf nur eingeschränkt Einfluss. Twine (<http://twine.com>) unterstützt die Organisation und den Austausch von Bookmarks für Webseiten, Photos, Videos, Bücher und sonstige Produktartikel. Das Hauptprinzip zur Organisation der Bookmarks stellen so genannte „twines“, eine Art von Nutzerinteressensgruppen, dar. Die Nutzer können einer solchen Interessensgruppe beitreten oder selbst eine gründen, um über diese Bookmarks auszutauschen. Über die Bookmarks, die in derselben Twine-Gruppe gruppiert sind, werden Tags (mit denen diese Bookmarks annotiert sind) miteinander in Beziehung gesetzt. Eine weitere Funktionalität von Twine ist die facetiierte Filterung von Suchergebnissen. Das heißt die Nutzer können über sieben Kategorien wie Tags, Personen, Orte oder Elementtyp ihre Suche filtern.

Mit ZigTag (<http://zigtag.com>) können die Anwender Bookmarks für Internetressourcen annotieren und organisieren. Im Gegensatz zu anderen Anwendungen, können sie ihre Bookmarks mit vordefinierten Begriffen annotieren. Das bedeutet, sie wählen ihre Tags bei der Annotierung aus einer vorgegebenen Liste von Tags mit einer bestimmten Bedeutung aus. Zusätzlich können die Anwender für den privaten Gebrauch eigene Tags hinzufügen. Es ist aber für den Anwender nicht möglich, weitere Semantik hinzuzufügen. Relationen zwischen Tags werden von dem System gesetzt.

Die Anwendung Faviki (<http://faviki.com>) baut auf die Standardisierung basierend auf Wikipedia/Dbpedia. Das heißt, die Nutzer können Bookmarks mit Tags, die Titel von Wikipedia-Artikel sind, annotieren. Es ist nicht möglich weitere semantische Relationen zwischen Tags hinzuzufügen. Für eine detaillierte Betrachtung der Systeme sei der interessierte Leser auf Braun et al. 2009 verwiesen.

6. Ausblick

Die zahlreichen Evaluierungsstudien des SOBOLEO-Systems haben gezeigt, dass die Balance zwischen den Möglichkeiten durch eine ontologiebasierte semantische Annotierung und der Zugänglichkeit für eine möglichst große Nutzergruppe eine schwierige Gratwanderung ist, bei der sich der verwendete Ontologieformalismus SKOS als sehr hilfreich herausgestellt hat. Umgekehrt zeigten die Studien, dass die Verbindung der Welt des Tagging und der Welt der Ontologien gelungen scheint: Notwendigkeit und Nützlichkeit von leichtgewichtigen Ontologien wird in der Benutzung des Systems relativ schnell erkannt, ohne dass ihre Entwicklung zuviel Aufwand bedeutet.

Über das reine Social Bookmarking Paradigma (also die kollaborative Annotation von Web-Ressourcen) hinaus lässt sich das SOBOLEO-System auch für andere Anwendungsfälle einsetzen, wo kollaborativ Strukturierungswissen aufgebaut werden soll und klassische Ansätze mit zentral gepflegten kontrollierten Vokabularen fehlgeschlagen sind. Ein Beispiel ist das Kompetenzmanagement, wo Ansätze mit Kompetenzkatalogen regelmäßig durch veraltete Inhalte, falsches Detaillierungsniveau und überbordende Aufwände zur Modellierung u. ä. nur beschränkten Erfolg haben. Deshalb haben wir im Rahmen des Projektes MATURE (<http://mature-ip.eu>) begonnen, ein auf dem Prinzip des „People Tagging“ basierendes leichtgewichtiges Kompetenzmanagement „von unten“ umzusetzen. Dabei werden Mitarbeiterseiten in offiziellen Mitarbeiterverzeichnissen oder aber private Webseiten oder Seiten in sozialen Netzwerken mit Begriffen annotiert, die beschreiben, mit was man die jeweiligen Personen in Verbindung bringt. Im Rahmen des kollektiven Aufräumens (gardening) können diese Begriffe schrittweise

formalisiert und aggregiert werden, so dass dort, wo z.B. für organisationale Auswertungen erforderlich, Kompetenzen definiert werden können. Die ersten Ergebnisse sind erfolversprechend, allerdings zeigt sich bei solchen Anwendungsfällen auch, dass neben den rein technischen Aspekten auch das gesamte sozio-technische System gesehen werden muss. Hierzu gehören beispielsweise auch Gestaltungsmaßnahmen wie z.B. wer kann vergebene Tags zu einer Person sehen? Müssen sie von der annotierten Person freigegeben werden bzw. kann sie unerwünschte Tags löschen? Im Rahmen eines Piloten sollen diese Fragen systematisch im Zusammenhang mit dem Organisationskontext geklärt werden.

Literatur

- Barker, K.; Chaudhri, V.K.; Chaw, S.Y.; Clark, P.; Fan, J.; Israel, D.; Mishra, S.; Porter, B.W.; Romero, P.; Tecuci, D.; Yeh, P.Z.: A Question-Answering System for AP Chemistry: Assessing KR&R Technologies. In: *Ninth International Conference of Knowledge Representation and Reasoning 2004*. (Hrsg. Dubois, D.; Welty, C.; Williams, M.-A.) Menlo Park: AAAI Press, 2004, 488–497.
- Braun, S.; Schora, C.; Zacharias, V.: Semantics to the Bookmarks: A Review of Social Semantic Bookmarking Systems. In: *International Conference on Semantic Systems*. In: *Fifth International Conference on Semantic Systems 2009*. (Hrsg. Paschke, A.; Weigand, H.; Behrendt, W.; Tochtermann, K.; Pellegrini, T.) Graz: Verlag der Technischen Universität Graz, 2009, 445–454.
- Golder, S.; Huberman, B.A.: The Structure of Collaborative Tagging Systems. *Journal of Information Sciences* 2 (32) (2006) 198–208.
- Gomez-Perez, A.; Fernandez-Lopez, M.; Corcho, O. (Hrsg.): *Ontological Engineering*. London: Springer.
- Guy, M., Tonkin, E.: Folksonomies: Tidying Up Tags? *D-Lib Magazine* 1 (12) (2006), <http://www.dlib.org/dlib/january06/guy/01guy.html> (Letzter Zugriff 05.10.2009).
- Hepp, H.: Possible Ontologies: How Reality Constraints Building Relevant Ontologies. *IEEE Internet Computing* 1 (11) (2007) 90–96.
- Macgregor, G.; McCulloch, E.: Collaborative Tagging as a Knowledge Organisation and Resource Discovery Tool. *Library Review* 5 (55) 2006, 291–300.
- Maier, R.; Schmidt, A.: Characterizing Ontology Maturing. In: *Fourth Conference Professional Knowledge Management – Experiences and Visions*. (Hrsg. Gronau, N.) Potsdam: GITO, 2007, 325–334.
- Marlow, C.; Naaman, N.; Boyd, D.; Davis, M.: Position Paper, Tagging, Taxonomy, Flickr, Article, ToRead. In: *Collaborative Web Tagging Workshop at the 15th International World Wide Web Conference 2006*. (Hrsg. Smadja, F.; Tomkins, A.; Golder, S.) 2006.

Miles, A.; Bechhofer, S.: SKOS Simple Knowledge Organization System Reference. *W3C Proposed Recommendation 15 June 2009*, 2009.

Sen, S.; Lam, S.K.; Rashid A.M.; Cosley, D.; Frankowski, D.; Osterhouse, J.; Harper, F.M.; Riedl, J.: tagging, communities, vocabulary, evolution. In: *20th Conference on Computer Supported Cooperative Work 2006*. (Hrsg. Hinds, P.; Martin, D.) New York: ACM Press, 2006, 181–190.



1



2



3

1 Dipl.-Mediensystemwiss. Simone Braun ist seit Dezember 2005 wissenschaftliche Mitarbeiterin am Forschungszentrum Informatik (FZI) in Karlsruhe. Dort verfolgt sie ihre Promotion über kollaborative Ontologieentwicklung mit Web 2.0-Ansätzen, die von Prof. Rudi Studer betreut wird. Weiterhin beschäftigt sie sich mit agilen Methoden für das Kompetenzmanagement und der Rolle sozialer Strukturen in der beruflichen Zusammenarbeit.
E-Mail: Simone.Braun@fzi.de
www.fzi.de/mitarbeiter/braun

2 Dr. Andreas Schmidt ist seit Februar 2000 wissenschaftlicher Mitarbeiter am Forschungszentrum Informatik (FZI) in Karlsruhe. Hier hat er sich in einer Vielzahl von Projekten insbesondere mit Wissensreife, Unterstützung des arbeitsbegleitenden Lernens durch Ausnutzung des Arbeitskontextes, und Ambient Assisted Living – dem technologieunterstützten Wohnen für ältere Menschen – beschäftigt. Seit September 2005 ist er Abteilungsleiter und verantwortlich für das Themenfeld Wissen & Lernen.
E-Mail: Andreas.Schmidt@fzi.de
<http://andreas.schmidt.name/>

3 Dr. Valentin Zacharias ist wissenschaftlicher Mitarbeiter am Forschungszentrum Informatik (FZI) in Karlsruhe. Promoviert hat er in der Semantik Web-Forschungsgruppe von Professor Studer an der Universität Karlsruhe über Verifikation und Validierung von wissensbasierten Systemen. Sein momentanes Forschungsinteresse ist die Nutzung von leichtgewichtigen, minimal-semantischen und heuristischen Methoden für das Semantik Web und das Daten Web.
E-Mail: zach@fzi.de
<http://vzach.de/blog>