

Ralf Klamma

Werkzeuge und Modelle für die übergreifende Untersuchung von Social Software

Tools and Models for Comprehensive Social Software Research

Social Software_Soziale Netzwerkanalyse_Modellierung_Musteranalyse_Visualisierung

Zusammenfassung. Für viele Unternehmen und Forscher ist die Netzwerk-Analyse von großen Datenbeständen immer noch mit hohen Kosten und vielerlei praktischen Hürden verbunden. In diesem Beitrag werden Werkzeuge und Modelle aus einer langjährigen Beschäftigung mit der Analyse von Web 2.0 Medien wie Blogs und Wikis in einem einheitlichen, multidisziplinären Rahmenwerk, der MediaBase, vorgestellt. Dabei kommen Methoden der sozialen Netzwerkanalyse zum Einsatz, die auf im Web gesammelten Datensätzen von sozialen Graphen in der MediaBase angewendet werden. Dabei werden automatisch Muster des sozialen Graphs auf der Ebene des Graphs, des Sub-Graphs und der Graph-Knoten erkannt und visualisiert. Der Ansatz wurde in vielen nationalen und europäischen Projekten, vor allem im Bereich des Technologie-gestützten Lernens entwickelt und evaluiert.

Summary. For researchers, practitioners as well as for many companies network analysis of huge data sets is costly and not practical at all. In this contribution we present a comprehensive unified framework for the collection and analysis of Web 2.0 media. This framework is the result of our experience with the analysis of media like blogs and wikis. In the framework we make use of social network analysis we apply on social graphs in the MediaBase. The MediaBase is a collection of crawled Web 2.0 media stored in the form of social graphs. On the social graphs, the included sub-graphs and the nodes we automatically detect patterns and visualize them. This framework was developed and evaluated in a number of national and European projects, in particular on technology enhanced learning.

1. Einleitung

Social Software ist Software, die menschliches Sozialverhalten nutzt oder verstärkt. Social Software konzentriert sich auf die Verknüpfung von Medienartefakten zwischen sozialen Entitäten (Klamma et al. 2006). Social Software ist im Grunde eine Weiterentwicklung traditioneller vernetzter Kommunikationsmedien (englisch: Computer Mediated Communication (CMC, Licklider et al., 1968)) und kollaborativer Informationssysteme für virtuelle Gemeinschaften und „Communities of Practice“ (Wenger, 1998). Eines der wichtigsten Ergebnisse der Benutzung von Social Software ist ein digitales soziales Netzwerk, manchmal auch ein so-

zialer Graph genannt, wobei die Knoten die möglichen Akteure abbilden und die Kanten in diesem Graph die realisierten Interaktions- oder Kommunikationsbeziehungen zwischen diesen Akteuren. Die Dichte der Kommunikationsbeziehungen in Teilgraphen kann höher als im Gesamtgraphen sein. Den dadurch ausgezeichneten Teilgraphen oder die dahinterliegende soziale Entität nennen wir hier Gemeinschaft (engl.: Community), da dieser Begriff sich in der Fachliteratur neutralisiert und durchgesetzt hat. Social Software ist innovativ, weil sie Brücken zwischen Medien und Gemeinschaften baut. Dabei konzentriert sich eine Social Software Instanz zumeist auf ein Medium, z.B. YouTube auf digitale Videos oder flickr.com auf digitale Bilder. Der Status quo

des Internets Entwicklungstrend ist „intelligent, simple, social“ (Bryant, 2003). Web 2.0-Technologien (O'Reilly, 2005) unterstützen und fördern die Interaktivität der Benutzer, ermöglichen damit eine Anpassung von Software an ihre Umwelt in Form digitaler Ökosysteme, anstatt zu verlangen, die Umwelt an die Software anzupassen. Oft wird der Erfolg von Social Software gemessen durch die Zahl der Benutzer. Obwohl dies einen guten Indikator für die Popularität einer bestimmten Plattform abgibt, liegt der Grund für die Wechsel existierender Gemeinschaften zu den neuen Medien oder das Aufkommen neuer Gemeinschaften immer noch in den Medien selbst. Die effizientere Form der Zusammenarbeit in den Gemeinschaften hat den „Long Tail“ (Anderson, 2005)

erzeugt, der charakteristisch für jede Art sozialer Portale ist und verdeutlicht, dass die wesentlichen Interaktionen in relativ kleinteiligen Gemeinschaften stattfindet, in denen Beziehungen existieren auf der Basis von Vertrauen und Leidenschaft für das gemeinsame Tun (Wegner, 1998). Die Kompetenz der Anwender im Umgang mit den neuen Medien führt zum letztendlichen Prozess der Aneignung der dieser Medien durch die Gemeinschaften. Bild 1 stellt die die Beziehung zwischen Medien und Community dar.

Wir ziehen hier zwei Schlussfolgerungen aus unseren Forschungen zu Social Software. Die erste Schlussfolgerung betrifft die Operationalisierung von Medientheorien in Gemeinschaften. Durch Beschreibungstechnologien wie XML (Brown et al., 2001) und RDF (Bray, 2003) sowie durch Syndizierungs-Technologien wie RSS (Cadenhead, 2006), ist die Erstellung, Nutzung und Wartung von verteilten, kollaborativen Metadaten für ein Medium in Social Software sehr einfach geworden. Allerdings gibt es für die durch Social Software erstellten Metadaten wie die durch „Tagging“ erzeugten „folksonomies“ noch keine Beschreibungsstandards. Im Gegensatz zu Standards in anderen IT-relevanten Bereichen fanden neue Standards nicht ihre operative Anwendung in Social Software, z.B. spielen Semantic Web Ontologie-Beschreibungen und deren Standardisierung nur eine winzige Rolle im Web und in der Social Software Welt. Wir können solche Probleme auf der Seite der Operationalisierung im Moment nicht lösen, im Bereich des Managements von Forschungsdaten schlagen wir ein vereinheitlichtes Datenmodell vor: die MediaBase. MediaBases ermöglichen es Gemeinschaften in einem interdisziplinär aufgebauten Rahmen-

werk trotz der verschiedenen Medialisierungsstrategien in Social Software ihre Daten zu verwalten, die durch sogenannte Crawler aus verschiedensten Quellen des Webs gewonnen werden. Die Gemeinschaften erzeugen und nutzen über verschiedene Social Software Instanzen hinweg Metadaten und haben ein Recht auf ihre eigenen sozialen „Metadaten“. Wie der „Long Tail“ in Bild 1 zeigt mag es Millionen von Gemeinschaften mit unterschiedlichsten Interessen geben. Eine große Herausforderung sind die unkalkulierbaren Folgen unserer Handlungen und die Handlungen anderer Menschen. Nutzer von Social Software agieren auf einer globalen Bühne. Jeder Weblog-Eintrag, jedes hochgeladene Video kann kommentiert, verändert und neu verteilt werden von jedem anderen Benutzer. Die Komplexität der Wechselwirkungen induziert durch dieses Universum von Möglichkeiten geht weit über die Bemühungen einer einzigen Person hinaus, diese Wechselwirkungen unter Kontrolle zu halten. Letztendlich sind es die Gemeinschaften, in denen Vertrauen herrscht und die durch gemeinsame Leidenschaften getragen werden, die die Privatheit und die Würde des Einzelnen bewahren.

Die zweite Schlussfolgerung lautet, dass nur die Gemeinschaften letztendlich wissen können, was sie über die Gemeinschaft wissen wollen und wie sie damit umgehen wollen. Es sollten dazu Werkzeuge zur Verfügung stehen, die weitestgehend automatisiert aus den in den MediaBases gesammelten Daten Analysen erstellen und sie den Gemeinschaften zur Verfügung stellen. Damit verbunden können schon existierende Lösungsstrategien angeboten werden, die sich in anderen Fällen bewährt haben. Die letzte Entscheidung muss die Gemeinschaft selbst über-

nehmen. Die Gemeinschaft hat das letzte Wort über den Erfolg von Social Software. Nur wenn die Social Software in der Lage ist, den Kontext einer Gemeinschaft angemessen darzustellen und das Design sich den in der Gemeinschaft vereinbarten kollaborativen und ästhetischen Anforderungen anpasst, ist Social Software in der Lage, in einer großen Menge an Nutzern zu bestehen, die sich durch weitgehende Heterogenität auszeichnen.

Dieses Papier ist wie folgt gegliedert. Abschnitt 2 betrifft die wichtigsten Merkmale von Social Software und untersucht ihre Beziehungen auf einer Sammlung von Medien und Gemeinschaften, die auch die Forschung motiviert. Die damit verbundenen Arbeiten in den Medien für digitale soziale Netzwerke sowie Gemeinschaften werden in Abschnitt 3 eingeführt. In Abschnitt 4 schlagen wir unseren konzeptioneller Ansatz für ein allgemeines Modell von Social Software vor. Beispielhafte werden Ausprägungen innerhalb der Social Software Welt für Blogs und Podcasts diskutiert. Eine abgeleitete MediaBase und digitale soziale Netzwerk-Analyse-Werkzeuge für die Gemeinschaften sind Prototypen für die Anwendbarkeit unseres Ansatzes. Die Schlussfolgerungen und ein Ausblick auf die zukünftige Arbeit sind in Abschnitt 5 gegeben.

2. Lösungsansatz

Social Software ist evolutionär und die Entwicklung bisher weitestgehend von darwinistischen Prinzipien getrieben, die sich aus der Vielzahl der Anbieter und Nachfrager auf einem globalen Markt der Möglichkeiten ergeben. Die Theorie der sozialen Netzwerke wurde allerdings ursprünglich in den Bereichen Geistes- und Sozialwissenschaften erforscht (Wellman et al. 2002). Digitale soziale Netzwerke spiegeln die gesellschaftlichen Strukturen wieder, die während der Nutzung digitaler Medien entstehen. Während die soziale Interaktion in der professionellen Kommunikation immer existierte, ist die Persistenz bei weltweitem Zugriff neuartig, gefolgt von einem kulturellen Wandel des Internets durch die enorme Zahl der Breitbandanschlüsse in den westlichen Industrienationen. Darüber hinaus unterstützt jede Gemeinschaft bestimmte Arten von Medien. Sie beeinflussen, wie

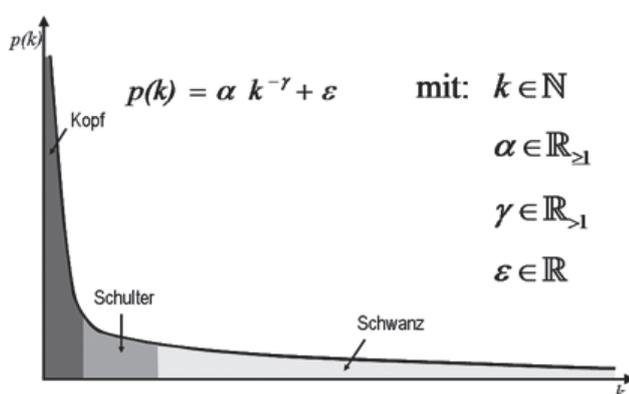


Bild 1: Der „Long Tail“ von Gemeinschaften und Medien

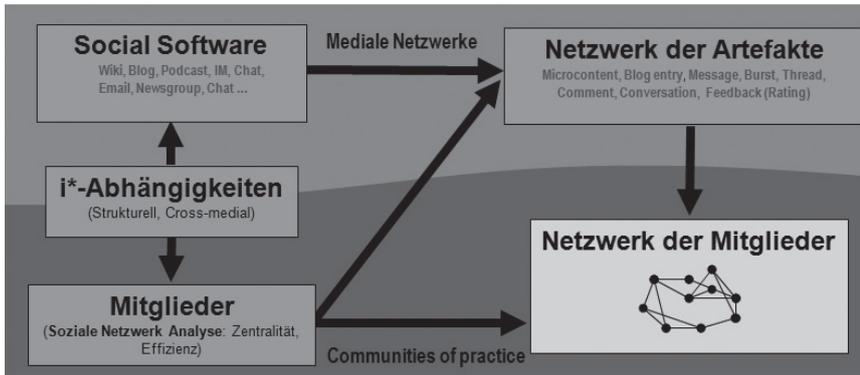


Bild 2: Digital Social Software

die Interaktion zwischen den Mitgliedern. In Bild 2 ist der grundsätzliche Lösungsansatz dargestellt, wie die Wechselwirkung zwischen Mediennetzwerken und Gemeinschaften (hier Communities of Practice) erforscht werden können. Der Trick ist, dass wir für die Analyse der medialen Netzwerke die gleichen Methoden anwenden, die wir auch bei der Analyse der sozialen Interaktionen zwischen den Mitgliedern einer Gemeinschaft verwenden würden. Aus der Analyse der Social Software und den dadurch erzeugten Medienartefakten schließen wir auf die Gemeinschaften, wobei wir auf grundlegenden Eigenschaften von Netzwerken, Gemeinschaften und Akteuren zurückgreifen, um die Kopplung zwischen Medien und Gemeinschaften zu gewährleisten. Diese Abhängigkeiten modellieren wir in einer speziellen Notation, die der i*-Methodik (Yu, 1995) zugrundeliegt.

Die Modellierungssprache i* ist dafür besonders geeignet, denn sie ermöglicht die Beschreibung strategischer Ziele und Abhängigkeiten zwischen Akteuren eines sozio-technischen Systems, die in Form von (menschlichen oder technischen) Agenten dargestellt werden. Hierbei sei erwähnt, dass der primäre Einsatzzweck von i* nicht die Prozessmodellierung ist, auch wenn die Modellierung von Abhängigkeiten oft Teile eines Prozesses andeutet. i* modelliert Agenten als benannte Kreise, die mit Darstellungen der Zielhierarchie der Agenten gefüllt sind. Zwischen den Agenten bestehen Abhängigkeiten verschiedenen Typs, die im Modell durch Rechtecke (Abhängigkeit bzgl. Erfüllung einer Aufgabe) bzw. sechseckige Rauten (Abhängigkeit bzgl. Verfügbarkeit einer Ressource) dargestellt sind. Dabei wird die Richtung der Abhängigkeit durch halbrunde Pfeilspitzen auf den Kanten

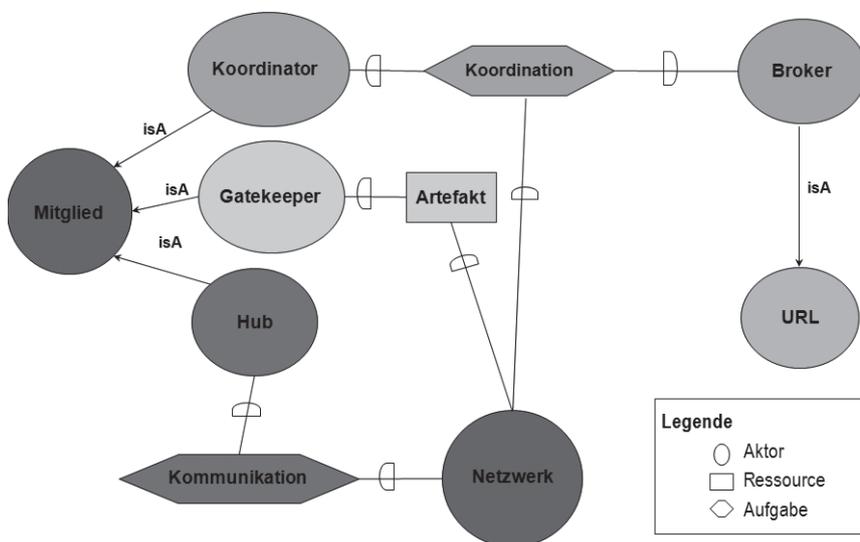


Bild 3: Abhängigkeiten zwischen Mitgliedern einer Gemeinschaft und Medienartefakten modelliert in i*

dargestellt, die zu Abhängigkeitsobjekten bzw. von ihnen weg auf andere Objekte zeigen. In Bild 3 sind mehrere Abhängigkeitsverhältnisse dargestellt, die sich teils aus der sozio-technischen Systementwicklung teils aus der sozialen Netzwerkanalyse ergeben.

Die soziale Netzwerkanalyse (englisch: Social Network Analysis (SNA)) (Brandes und Erlebach, 2005) hat sich gerade mit dem Boom der Web 2.0-Technologien neue Anwendungsgebiete erschlossen. Die wichtigsten Grundlagen und Konzepte der SNA für Social Software werden kurz erläutert. Wichtige Netzwerk-Eigenschaften sind die „Small World“ Eigenschaft (Watts und Strogatz, 1998) und die Skalenfreiheit von Netzwerken (Barabási, 2001). Ein „Small World“-Netzwerk ist ein Netzwerk, in dem jeder Knoten trotz hoher Anzahl von Knoten in einer sehr kleinen Anzahl von Schritten erreicht werden kann. In einem skalenfreien Netzwerk gehorchen dazukommende Knoten einem einfachen Gesetz. Diese Knoten orientieren sich bei der Vernetzung mit existierenden Knoten an der Anzahl der bereits existierenden Verbindungen dieser Knoten (englisch: Preferential Attachment). Der mathematische Zusammenhang ist in Bild 1 erläutert. Dieses Gesetz wird auch manchmal das Matthäus-Prinzip genannt. Dies impliziert, dass es eine Menge von Knoten mit nur ein paar Verbindungen (die Gemeinschaften im „Long Tail“) und eine kleine, aber signifikante Minderheit von Knoten, die Verbindungen zu einer sehr großen Anzahl von anderen Knoten haben, so genannte „Hubs“ (vgl. Bild 3), die eine zentrale Rolle in der Verbreitung von Information spielen. Weitere Rollen sind „Gatekeeper“, die den Zugang zu einem Artefakt einer Gemeinschaft (Netzwerk) kontrollieren, sowie die Rollen „Kordinator“ und „Broker“, die in Zusammenarbeit verschiedene Quellen, z.B. URL, für eine Aufgabe sammeln.

Akteure müssen nicht menschlich sein. Es gibt drei spezielle Arten von Akteuren, aus denen ein digitales soziales Netzwerk bestehen kann: Mitglied, Medium und Artefakt. Mitglied steht für eine Person oder ein Mitglied eines Netzwerkes (einer Gemeinschaft). Ein Medium ermöglicht die Speicherung und Übertragung von Informationen. Ein Artefakt ist ein Objekt, erstellt von Mitgliedern in einem

Medium. Die Mitglieder haben wie Netzwerke ebenfalls Eigenschaften, die aus der SNA abgeleitet werden. Zentralität ist eine wichtiges Maß für Eigenschaften von Mitgliedern. Drei wichtige Arten von Zentralität (Freeman, 1979) sind wie folgt aufgeführt.

- Degree-Zentralität ist der Grad der Vernetzung eines Mitglieds mit anderen Mitgliedern und deutet die Popularität oder das Prestige eines Mitglieds an.
- Die Closeness-Zentralität berücksichtigt die Nähe eines Mitglieds zu jedem anderen Mitglied
- Betweenness-Zentralität versucht für ein Mitglied eine Schätzung der Möglichkeit abzugeben, wie viel Einfluss auf die Kommunikation von anderen genommen werden kann.

Aus diesen Baseigenschaften lassen sich relativ einfach Muster ableiten, die komplexes Verhalten innerhalb von Gemeinschaften abbilden können. Ein Beispiel für so ein Muster ist das strukturelle Loch. Ein solches Mitglied ist eine nicht redundante Beziehung zwischen zwei Nachbarn dieses Mitglieds (Burt, 1992), d.j. für die Kommunikation zwischen den Nachbarn, ist dieses Mitglied besonders wertvoll. Es ist eine Funktion der Zeit und die Energie, die ein Mitglied investiert in den Beziehungen mit jedem seiner Nachbarn.

Neben den vorgestellten methodischen Aspekten der Modellierung mittels i^* und der Analyse mittels SNA verwenden wir in unserem Ansatz Theorien, auf die wir hier aus Platzgründen nicht näher eingehen können. Dazu gehört die Actor-Network Theory (ANT, Latour 1999), die uns erlaubt eine theoretische Fundierung der Auflösung von menschlichen und nicht-menschlichen Akteuren, sowie deren Aggregation zu Agenturen. In unserer Auffassung liefert die ANT Beiträge zu blinden Flecken der Community-of-Practice Theorie (Wenger, 1998) und der Activity Theorie (Engeström 1987), die vor allem den Lernprozess beziehungsweise den Umgang mit soziotechnischen Systemen in den Griff nehmen. Allen Theorien fehlt ein fundamentales Verständnis von der Operationalisierung von Medien, welche in der Transkriptivitäts-Theorie von Jäger (2002) geliefert wird. Die Entfaltung dieser Theorien in der Gestaltung von

informatischen Designtheorien wurden an anderer Stelle vorgestellt (Jarke und Klamma 2005).

3. Die MediaBase für das Management von Forschungsdaten

Basierend auf den Arbeiten im vorigen Abschnitt schlagen wir ein Modell für Social Software vor. Das allgemeine Modell dient der Vereinheitlichung der Daten aus verschiedenen Social Software Instanzen auf einer aggregierten Ebene. Auf Grundlage des allgemeinen Modells können spezialisierte Datenmodelle geschaffen werden, die in der Lage sind, mittels Crawlern, Daten aus Web-Quellen automatisiert in Datenbanken einzufügen. Im Gegensatz zu den reinen Webseiten-Crawlern, die in Suchmaschinen und Web-Archiven zum Einsatz kommen, erzeugen wir auf der Idee von Vorlagen spezialisierte Crawler, die auf die Besonderheiten verschiedener Social Software Instanzen schon bei der Abschöpfung der Daten achten, so dass es zu weniger Datenmüll kommen. Diese gespeicherten Daten werden dann in der MediaBase für die Gemeinschaften verwaltet und den Analysewerkzeugen der Gemeinschaften zur Verfügung gestellt.

3.1 Ein Modell für Gemeinschaften und ihre Mitglieder als Akteure in Social Software

Ausgehend von den oben genannten digitalen sozialen Netzwerk Theorien, skizzieren wir ein erweitertes Entity-Relationship-Modell, dargestellt in Bild 4. Wir nennen es ein ANT-Modell der Social Software.

Alles ist ein Akteur (Aktor). Spezielle Akteure sind (in Grün) Medium, Artefakt, Mitglied und als eigentlicher Besitzer der MediaBase (in Blau) Gemeinschaft als die zentralen Konzepte von Social Software. In jeder Art von Social Software, gibt es ein zentrales Medium, um das die angebotenen Dienste aufgebaut sind, z.B. kann man in einem Medium Blog Artefakte wie Blogeinträge, Kommentare, Tags und eine Liste mit lesenswerten Blogs anlegen (speichert). Medien sind in der Lage bestimmte Artefakte darzustellen, z.B. gibt es die Medien Email und Foren, die in der Lage sind Thread-Artefakte darzustellen, während dies in Blogs unüblich ist (hellgrau unterlegt). Die Trennung Medium/Artefakt ist dabei strikt operational. Was ein Medium und was ein Artefakt ist, entscheidet sich erst in der Verwendung und im Vergleich mit anderen Medien. Dienste erzeugen oder nutzen Artefakte innerhalb eines Mediums. In Tabelle 1 haben wir die möglichen Medien und Artefakte zusammengefasst.

Diese Dienste können sehr komplexe Medienoperationen sein, die entsprechend aufwendig in der Social Software zu implementieren sind. Ohne die Konzentration auf ein Medium würde Social Software nicht mehr nutzbar sein, wegen der hohen Komplexität der Kommuni-

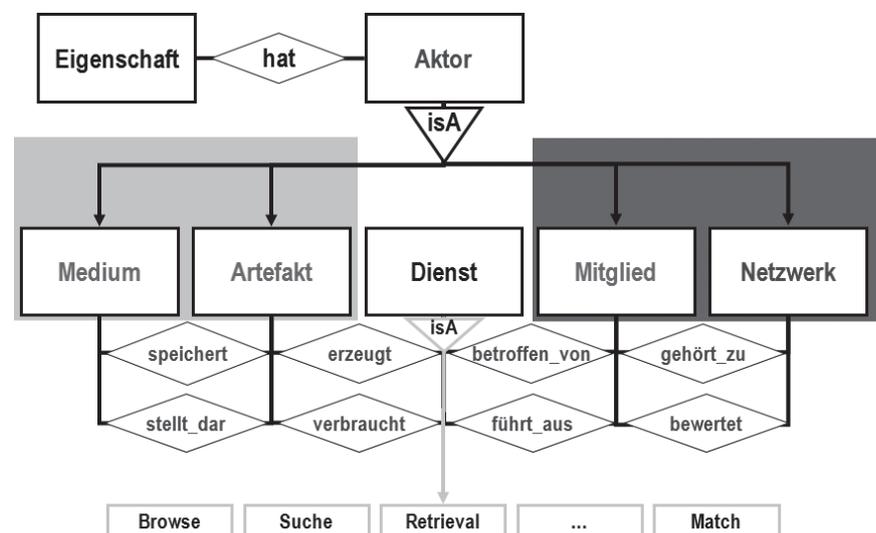


Bild 4: ANT Modell der MediaBase

kation und des Vertrauens in der Social Software. Entsprechend erfolgreich sind mono-mediale Social Software-Instanzen und Medien werden mittels Mash-ups gemischt. Aktivitäten der Gemeinschaft enthalten wichtige Erkenntnisse, z.B. der Gebrauch von Medien innerhalb der Gemeinschaft, der ist auf Medien gespeichert Zugehörigkeit zu. Mitglieder gehören auch zu einer Gemeinschaft (dunkelgrau unterlegt) und führen die Dienste aus, sind aber auch betroffen von der Ausführung von Diensten. Die Gemeinschaft bewertet die Mitglieder ständig, weisen ihnen einen sozialen Status zu, z.B. den Rang eines Experten. Alle Konzepte und ihre Beziehungen können verfeinert werden.

3.2 MediaBase Diskurse: Blogwatcher und Post-Cast-Watcher

Zur Illustration des Modells der MediaBase geben wir im Folgenden in aller Kürze zwei Beispiele für die Verfeinerung des Modells (Bild 5). Das erste Beispiel sind Podcasts, eine multimediale Social Software, die es Gemeinschaften ermöglicht, Ton- und/oder Videoaufnahmen zu erstellen und in Webseiten zu integrieren. Bei Podcasts handelt es sich um Mediendateien bzw. eine Serie von Medienbeiträgen in Form von Audio- oder Videoformaten, die bei Bedarf automatisch über Feeds und unter Zuhilfenahme eines Klienten bezogen werden können. Der Klient ruft die in einem Feedbeitrag enthaltenen Verweise ab, um die Dateien

zu laden und abzuspielen. Die Nachricht des Verfassers ist somit in den Dateien enthalten. Ein Ersteller von Video- und Audioaufnahmen kann diese mit relativ geringem Aufwand produzieren und den Bezieher automatisch über deren Existenz informieren. Podcasts wurden einige Zeit ernsthaft als Ablösung von klassischen Radio- und TV-Modellen diskutiert, haben heute aber enorm hohe praktische Bedeutung, z.B. in Lerngemeinschaften, da mittlerweile unzählige Podcasts mit Sprachlektionen existieren. In modernisierter Form haben sich Podcasts zu Web- oder Screencasts entwickelt, eine Form multimedialer Aufbereitung von Handbüchern zu komplexen IT Produkten, die immer weitere Verbreitung finden. Das zweite Beispiel sind Blogs. Ein Blog, kurz für „Weblog“, ist eine chronologisch sortierte Sammlung von Artikeln, genannt Blogbeiträge, die inhaltlich zum Thema des Blogs gehören und meist von einem einzelnen, manchmal auch von mehreren Autoren verfasst werden. Die Blogbeiträge tragen Titel, die enthaltenen Nachrichten können mit Verweisen und Multimediateilen versehen werden. Oft ist die Vergabe von Tags zu Einträgen durch den Verfasser sowie die Suche danach durch den Leser innerhalb der Blogsoftware möglich. Der Leser eines Blogs kann zum einen den Blog besuchen und dort direkt die Blogbeiträge lesen, zum anderen diese als Feed abonnieren und sich zustellen lassen. Blogs sind weltweit populär, aber insbesondere in den USA, wo sich mittlerweile Blogs als ernsthafte Mitspieler in den politischen Medien etabliert haben,

z.B. die Huffington-Post, aber auch riesige Gemeinschaften gebildet haben, so z.B. die Mommy-Bloggers, Millionen von Müttern, die Erfahrungen aus ihrem Leben teilen.

In dem Modell für Blogs, ist der Akteur ist der Blogger, der zu einer oder mehreren Gemeinschaften gehört. Blogs als spezielles Medium gehört auch zu einer oder mehreren Gemeinschaften. Die Klassifizierung von Blog-Einträgen kann durch die Gemeinschaft mittels „Tagging“ gemacht werden. „Tagging“ ist eine Instanz des Dienst-Konzepts. Andere Elemente, die zu einem Blog gehören, sind Blogrolls, Benutzerprofile usw. Dieses Modell kann in einer Datenbank realisiert werden, z.B. in einer relationalen Datenbank. Jeder Blog besteht aus einer Anzahl von Einträgen, die von Autoren geschaffen sind, von einem Autor oder einer Gruppe von Autoren. Autoren schaffen auch Kommentare, Tags, Multimedia, Referenzen und Trackbacks, die gehören aber alle zu einem Blog-Eintrag und nicht dem Blog als Ganzes.

Neben Blogs, haben wir ein Modell für Podcasts vorgeschlagen. Neben den spezifischen Unterschieden, die darin bestehen, dass wir sogenannte Pod-Hosts haben und Podcasts aus Episoden bestehen (vgl. Bild 5 links), haben wir auch Gemeinsamkeiten, wie die Existenz von Tags. Somit können wir das Tagverhalten in Gemeinschaften miteinander vergleichen. Während wir schon aus anderen Arbeiten (Bischoff et al. 2008) wissen, dass Tagging sehr medienspezifisch ist, können wir mittels unseres Modelles auch

Tabelle 1: Medien und mögliche Artefakte

	E-Mail	Mailing List	Blog	Transaktionsbasierte Web-Site	Wiki	Chat Room	URL
Nachricht	+	+	-	-	-	-	-
Thread	-	+	-	-	+	-	-
Burst	+	+	+	+	+	-	-
Gespräch	-	-	-	-	-	+	-
Blog-Eintrag	-	-	+	-	-	-	-
Kommentieren	-	-	+	+	+	-	-
Web Page	-	-	-	-	+	-	+
Transaction	-	-	-	+	-	-	-
Ihre Meinung	-	-	-	+	-	-	-

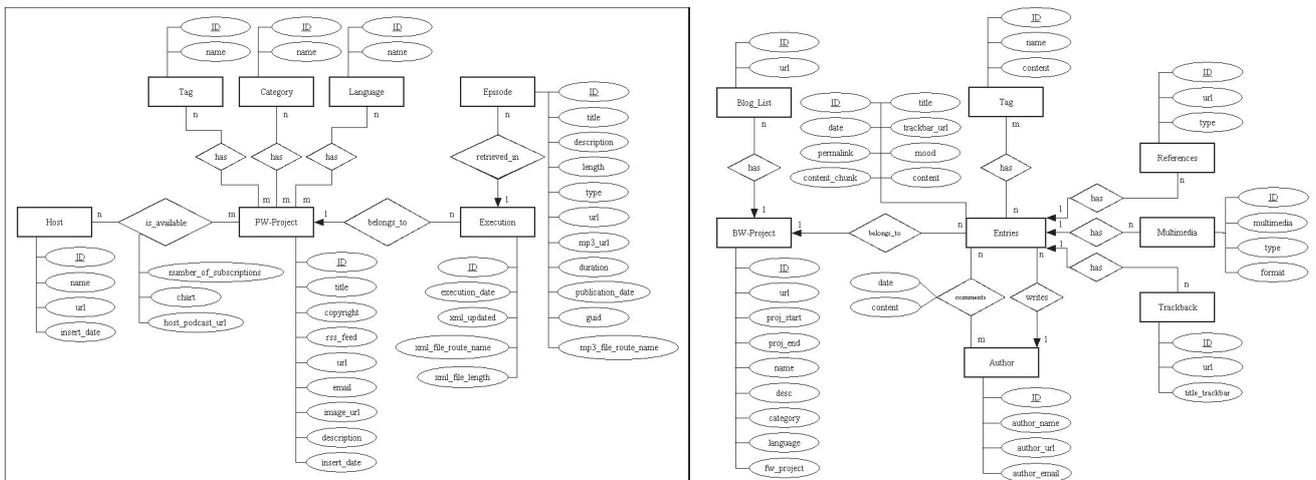


Bild 5: ERD Modell von Podcasts (links) und Blogs (rechts)

das Tagging innerhalb einer Gemeinschaft untersuchen und so z.B. zeigen, dass Experten anders als Amateure taggen (Klamma et al. 2009).

Es existieren nicht nur Datenmodelle zu Blog und Podcasts in der MediaBase, sondern auch zu vielen anderen Medien wie Email-Listen, Foren, Wikis, Feeds und News (Klamma et al., 2007, Klamma und Haasler 2008). Alle diese Modelle sind in einer relationalen Datenbank realisiert. Zu jedem Medium gibt es Crawler, die mittels Vorlagen (englisch: Templates) konfiguriert und parametrisiert werden. Diese Crawler bevölkern die Modelle für eine Gemeinschaft. Die Gemeinschaften haben die Möglichkeit über ein Verwaltungswerkzeug (MediaBase Commander) die Medien auszuwählen, die für ihre Belange von Bedeutung sind. Der MediaBase Commander ist selbst als Web 2.0 Plattform gedacht, so können Gemeinschaften, deren Sammlungen usw. getaggt, kommentiert und klassifiziert werden.

3.3 Pattern-basierte Media-Bases

Als Folge der Darstellung in einem gemeinsamen Modell von Social Software, können wir nun Muster des sozialen Graphs einer Gemeinschaft untersuchen, um eine angemessene und vollständige Betrachtung von Social Software durchzuführen. Unsere Ausgangshypothese war ja, das Social Software erst in der Aneignung durch eine Gemeinschaft ihr volles Potential entfalten kann. Muster sind wesentliche Indikatoren in digitalen

sozialen Netzwerken, weil sich verändernde Muster in Netzen im Laufe der Zeit zeigen, wie sich Netzwerk bilden, wie sie wachsen und auch wieder schwinden. Durch das Verständnis solcher Muster in verschiedenen Netzwerk-Typen, können wir auch mögliche Ursachen und Folgen des Wandels und Netzwerk ableiten und vorhersagen. Wir können die Evolution dieser digitalen Ökosysteme verstehen und gegebenenfalls verschiedenen Interventionen durchführen (Churhill und Halverson, 2005).

Ein Muster enthält die Beschreibung des Problems, das gerichtet ist, den Kontext, wo das Muster angewendet wird, welche Kräfte ins Spiel kommen, die empfohlene Lösung und die Gründe dahinter (Alexander, 1978). In unserem Modell haben wir ein Netzwerk gebildet aus den Beziehungen von Akteuren und in den Mustern gibt es einen Zusammenhang aus verschiedenen Kräften. Die Hauptgedanken der ANT-Modell und im Muster sind die Akteure und die Kräfte (Latour, 1999), die eine Umverteilung der Aktion zur Folge haben. Wenn ein Muster in i^* modelliert und angewendet wird, um ein digitales Netzwerk (Gemeinschaft) zu analysieren, können wir für jede Kraft menschliche und nicht-menschliche Akteure benennen und für jede Relation zwischen den Kräften gibt es eine entsprechende Beziehung zwischen den Akteuren. Als Endziel sollen die Gemeinschaften selbst in die Lage versetzt werden, diese Modelle zu schaffen und sie mittels der Muster zu analysieren. Dazu haben wir in der MediaBase ein Repo-

sitory von Mustern geschaffen, welche die Gemeinschaften auswählen und auf ihre eigenen sozialen Graphen anwenden können. Dabei entspricht der soziale Graph, der aus der MediaBase kommt, dem was die Gemeinschaft tut oder getan hat, während die Modelle ausdrücken können, was die Gemeinschaft tun will. Aus dem Vergleich von Soll und Ist können die Gemeinschaften dann Handlungen ableiten, die schon in den Musterbeschreibungen angelegt sind oder über völlig neue Strategien nachdenken, die ihre Handlungsfähigkeit in digitalen sozialen Netzwerken erhalten oder verbessern können.

Basierend auf der Modellierung in i^* und der Muster Sprache PALADIN (Pattern Language für Störungen in digitalen sozialen Netzen), die wir in XML definiert haben, haben wir die PALADIN Anwendung entwickelt. PALADIN ist dazu geeignet, um zu einem Muster einen ersten visuellen Einblick in eine Reihe von digitalen sozialen Netzwerken zu erhalten. Es bietet ein Web-Interface, um Muster zu definieren, zu durchsuchen und speichern.. Das Repository selbst ist ein XML-Datenbank-Management-System. PALADIN enthält einen Algorithmus, der auf Grundlage der gegebenen sozialen digitalen Netzwerke, die ausgewählten Muster instanziiert und parametrisiert. Wie in Abschnitt 2 hervorgehoben, konzentrieren wir uns auf die strukturellen Eigenschaften der Gemeinschaften, z.B. Zentralitäten von Mitgliedern. Aufgrund dieser strukturellen Eigenschaften werden komplexere Musterausdrücke erzeugt, die mittels

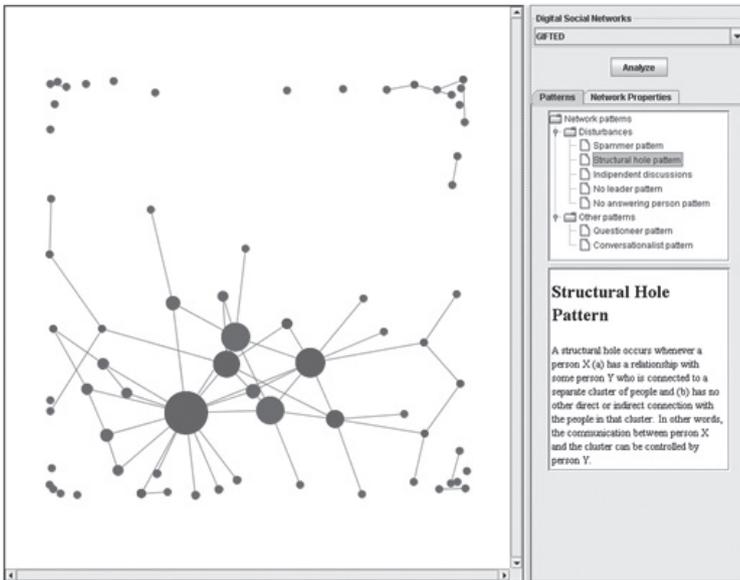


Bild 6: Strukturelle Löcher in einer Mailingliste (Bildschirmabzug der Software PALADIN)

der PALADIN Sprache definiert werden können und die später dann von dem Algorithmus ausgewertet werden. Realisiert werden die Mustersausdrücke als SQL-Anfragen an die gespeicherten sozialen digitalen Netzwerken im relationalen Teil der MediaBase.

Die Kommunikation zwischen den Mitgliedern besteht im Austausch von Artefakten. Nach der Aktivität der Mitglieder und deren Muster der Kommunikation, können ihnen unterschiedliche Rollen zugeschrieben (al. Cheung 2005; Fisher et al., 2006; Madanmohan und Navelkar, 2004; Turner, 2005):

- Ein Fragender sucht Hilfe und Informationen.

- Eine Beantwortung Person gibt Antworten auf die aufgeworfenen Fragen ohne Diskussionen beteiligt zu. Ein Gesprächspartner beteiligt sich an Diskussionen aktiv.
- Ein Troll zielt darauf ab, Aufmerksamkeit und ab nutzlos Diskussionen.
- Ein Spammer sendet Nachrichten irrelevant für die Gemeinschaft, z. B. Anzeigen oder Binärdateien.

Diese und weitere Muster wurden mittels PALADIN definiert und auf digitalen sozialen Netzwerken. Bild 6 zeigt ein strukturelles Loch in digitalen sozialen Netzwerken. Die Muster spiegeln die Existenz von Trollen, Spammern, Gesprächspartner,

Fragesteller, usw. Ob ein Muster eine positive oder negative Störung des digitalen sozialen Netzwerkes darstellt, ob die Gemeinschaft damit leben kann, sie das Auftreten dieses Musters in Zukunft verhindern, fördern oder ignorieren will, liegt allein in der Gemeinschaft selbst begründet. PALADIN liefert zu jedem Muster mögliche Reaktionen, ohne zu überprüfen, ob eine solche Intervention auch zu den gewünschten Resultaten führt. Interventionen in digitalen sozialen Netzwerken sind ein überaus komplexes Geschehen, da sie wie Störungen erstens sehr schwer zu erkennen sind und zweitens nicht nur auf Änderungen auf der Software-Ebene beruhen, sondern auch in Schulungen oder in geänderten Verhaltensweisen beruhen können.

In einem erhobenen Datensatz mit 156 Mailinglisten kulturwissenschaftlicher Provenienz haben wir diese Muster untersucht und auch gefunden. Der Datensatz enthält ca. 250.000 Emails von mehr als 25.000 Akteuren in 180.000 einzelnen Diskursen). Weiterhin haben wir uns mit den eher alltäglich anfallenden Kommunikationsmustern festgestellt, dabei aber multimodale Netzwerke zugelassen, d.h. Medien als Knoten im Netzwerk im Verhältnis zu den Teilnehmern untersucht (vgl. Bild 7). Dadurch konnten z.B. Fragen nach der Konsistenz der Teilnahme oder die Peripherialität von Teilnehmern über verschiedene Medien hinweg untersucht werden. Dabei haben wir bei aller Vorläufigkeit solcher Studien folgende Eindrücke gewonnen. 1) Die meisten Netzwerke in unserem Datensatz waren monologischer Natur, d.h. Diskurse in Mailinglisten sind nicht sehr häufig. 2) Eine kleine Reihe von Aktivitäten führt die tägliche Diskussion in den Listen an. 3) Kulturwissenschaftliche Mailinglisten werden von peripherer Teilnahme beherrscht. 4) Teilnehmer in kulturwissenschaftlichen Mailinglisten zeigen ein konsistentes Verhalten über mehrere Mailinglisten hinweg.

Um diese Form von Analyse zu unterstützen, ist die Visualisierung von Störungen von erheblicher Bedeutung. Wir wenden Cross-Media-Visualisierungstechniken in der MediaBase an (Klamma et al. 2006), z.B. Treemaps unter Verwendung der squarified Treemap Algorithmus (Bruls, 2000). In PALADIN stehen darüber hinaus vier verschiedene Layouts für die Visualisierung von digitalen sozialen

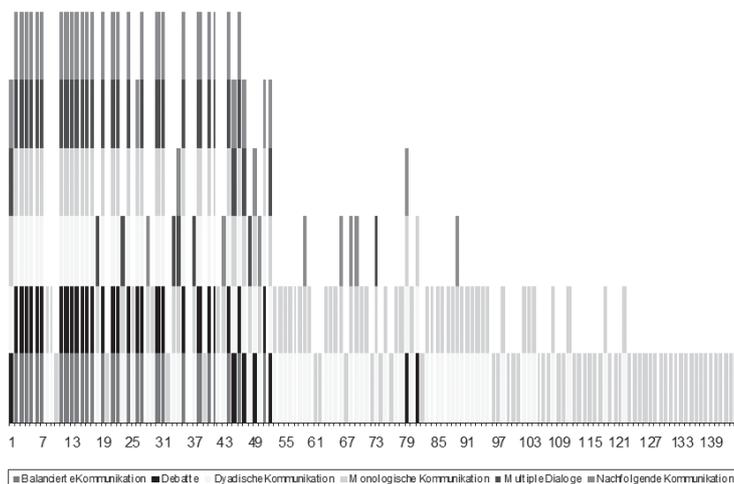


Bild 7: Kommunikationsmuster über verschiedene Medien von Gemeinschaften

Netzwerken zur Verfügung. Damit wird es möglich, großes und unübersichtliche digitale soziale Netzwerke (die gefürchteten Wollknäuel) interaktiv zu erkunden, z.B. über die Änderung der Größe von Knoten und die Dicke von Ecken aufgrund der SN, z.B. basierend auf der Edge-Betweenness Zentralität (EBC) (Freeman et al., 1991). Solche Funktionalitäten werden durch Tasten und durch Schieberegler realisiert. Die Anzeige kann in verschiedenen Farben erfolgen.

4. Schlussfolgerungen und künftige Arbeit

In diesem Papier haben wir ein allgemeines Modell für die Forschung im Bereich der Social Software vorgestellt. Zudem haben wir Werkzeuge zur Analyse und Visualisierung der aus der Interaktion der Mitglieder in digitalen sozialen Netzwerken entstehenden Graphen vorgestellt und anhand von Beispielen und Datensätzen erläutert. Medien und Gemeinschaften sind mittels Social Software verbunden. Das erfolgreiche Arbeiten, Forschen und Lernen hängt von der richtigen Mischung von Medien, deren Nutzung und den damit notwendigen Kompetenzen innerhalb der Gemeinschaft ab. Die Vielzahl der existierenden Ansätze unterschätzen dabei häufig die mit den Medien verbundenen Operationen, die Heterogenität und Nutzung der damit verbundenen Metadaten und die tiefen sozialen Strukturen, die innerhalb der Gemeinschaften wesentlichen Einfluss auf die Nutzung der Medienoperationen Einfluss haben. Der Ansatz wurde in verschiedenen nationalen und internationalen Projekten entwickelt und getestet und hat sich mittlerweile im Bereich des Technologie-gestützten Lernens verstetigt. Die PROLEARN MediaBase und PALADIN sind integriert in das PROLEARN Academy Portal www.prolearn-academy.org. MediaBases werden weiterhin in den EU Projekten TELLNET (Teacher's Life Long Learning Networks) und TELMAP (Future gazing Technology Enhanced Learning – The Roadmap for the unknown Learning Landscape) eingesetzt. In TELLNET wird vor allem die Wechselwirkung von Kompetenzgewinn im Bereich der Lesbarkeit von sozialen digitalen Netzwerken und die Kollaborationsqualität in einem Netz-

werk von tausenden von europäischen Schulen untersucht. Die Interventionen bestehen aus Workshops mit Lehrern, die Kooperationen zwischen europäischen Schulen aus verschiedenen Ländern erfolgreich durchgeführt haben oder dies vorhaben. Für diese Lehrer werden Visualisierungen ihrer bisherigen oder möglichen Kollaborationsaktivitäten erzeugt und diskutiert. In einem angemessenen Zeitraum wird dann untersucht, ob diese Interventionen Einfluss auf das Kollaborationsgeschehen im Schulnetzwerk hatte. In TELMAP geht es um die Unterstützung Methoden der Technologievorhersage, hier das Roadmapping, durch die Weisheit der Vielen. Es wird untersucht, ob sich bestimmte Trends durch schwache Signale (englisch: Weak Signals) in Web 2.0 Medien bestätigen oder verwerfen lassen. Beide Projekte sind 2010 gestartet und werden ihre Ergebnisse in den nächsten Jahren veröffentlichen.

Auch bleibt die Frage zu klären, ob der Einsatz von Social Software im Technologie-gestützten Lernen lohnenswert ist (Klamma et al. 2007). Dalsgaard argumentiert, dass Social Software wie del.icio.us und Wikipedia nützliche Werkzeuge zur Erstellung persönliche Lernumgebungen sind, die zusammen mit traditionellen Learning Management Systemen (LMS) eingesetzt werden können (Dalsgaard, 2006), d.h. Social Software hat noch den Status von Hilfswerkzeugen. Die weitgehende Transformation von Lernszenarien durch Social Software ist noch nicht gut verstanden. Das EU Projekt ROLE (Responsive Open Learning Environments) untersucht die systematische Verbindung verschiedenster Lernmaterialien und Werkzeuge in einer Vielzahl von Lerngemeinschaften mittels standardisierter Web 2.0 Technologien. Die große Herausforderung hier ist es die gewaltigen Möglichkeiten der Technologie mit angemessenen pädagogischen Modellen für neue Formen des Selbst-regulierten Lernens zu verbinden, so dass für die Lerngemeinschaften die Möglichkeiten besteht, sich auf ihre Lernziele zu konzentrieren und trotzdem an der technischen Weiterentwicklung zu partizipieren. Im Kern beruhen solche Ansätze auf die intelligente Nutzung von Empfehlungen für Lernmaterialien, Werkzeuge und Lernpartner auf der Grundlage des sozialen Graphs. Allerdings sind auch hier die ethischen und

datenschutzrechtlichen Grundsätze zu berücksichtigen, die letztes Endes zu der Einsicht führen, dass die Gemeinschaften über ihre Daten selbst bestimmen müssen und können.

Danksagung

Diese Arbeit wurde von den EU Projekten ROLE, TELLNET und TELMAP unterstützt. Ich bedanke mich bei den vielen Kollegen in diesen Projekten für die inspirierenden Diskussionen und die vielen praktischen Beiträge zur MediaBase.

Literatur

- Alexander, C.: A Pattern Language: Towns, Buildings, Construction, Oxford University Press, Oxford, 1978.
- Anderson, C.: The long tail, Wired, Ausgabe 12.10, 2004 – <http://www.wired.com/wired/archive/12.10/tail.html> (letzter Zugriff November 2010).
- Barabási, A.-L.: The Physics of the Web, Physics World, 2001.
- Bischoff, K.; Firan, C. S.; Nejdil, W.; Paiu, R.: Can all tags be used for search? CIKM 2008: 193–202.
- Brandes, U.; Erlebach, T.: Network Analysis – Methodological Foundations, Springer-Verlag, Berlin Heidelberg, 2005.
- Bruls, M.; Huizing, K.; van Wijk, J. J.: Squarified Treemaps, Proc. Joint Eurographics and IEEE TCVG Symposium on Visualization, S. 33–42, 2000.
- Bryant, L.: Smarter, Simpler, Social, Tech. Report, Headshift.com, 2003.
- Burt, R.: Structural Holes: The Social Structure of Competition, Harvard University Press, Boston, MA, 1992.
- Cheung, K.; Lee, F.; Ip, R.; Wagner, C.: The development of a successful online community. International Journal of the Computer, the Internet and Management, Vol. 13, No. 1, Jan.–Apr, 2005.
- Churchill, E. F.; Halverson, C. A.: Guest Editors' Introduction: Social Networks and Social Networking, IEEE Internet Computing, 9(5):4–9, 2005.
- Dalsgaard, C.: Social software: E-Learning beyond learning management systems, Europ. Journal of Open, Distance and E-Learning, 2006, http://www.eurodl.org/materials/contrib/2006/Christian_Dalsgaard.htm (letzter Zugriff November 2010).
- Engeström, Y.: Learning by Expanding: An Activity – Theoretical Approach to Developmental Research, 1987.
- Fisher, D.; Smith, M.; Welsler, H.: You are who You Talk to: Detecting Roles in Usenet Newsgroups, Proceedings of 39th HICSS, 2006.

- Freeman, L.: Centrality in Social Networks, Conceptual Clarification, *Social Networks* Vol. 1, S. 215–239, 1979.
- Freeman, L. C.; Borgatti, S. P.; White, D. R.: Centrality in valued graphs: A measure of betweenness based on network flow, *Social Networks*, Vol. 13, S. 141–154, 1991.
- Jäger, L.: Transkriptivität. Zur medialen Logik der kulturellen Semantik, Jäger, L.; Stanitzek, G. (eds.): *Transkribieren – Medien/Lektüre*, Fink, München, S. 19–41, 2002.
- Jarke, M.; Klamma R.: Transkriptivität als informatives Designprinzip. Mediale Spuren in rechnergestützten Entwicklungsprozessen, G. Fehrmann; E. Linz; C. Epping-Jäger (eds.): *Spuren Lektüren. Praktiken des Symbolischen*, München: Fink, S. 105–120, 2005.
- Klamma, R.; Haasler, C.: Wikis as Social Networks: Evolution and Dynamics. *Proceedings of the Second KDD Workshop on Social Network Mining and Analysis*, August 24, 2008, Las Vegas, Nevada, USA, 2008.
- Klamma, R.; Cao, Y.; Spaniol, M.: Watching the Blogosphere: Knowledge Sharing in the Web 2.0, N. Nicolov, N. Glance, E. Adar, M. Hurst, M. Liberman, J. H. Martin and F. Salvetti (Eds.): *International Conference on Weblogs and Social Media*, Boulder, Colorado, USA, March 26–28, S. 105–112, Klamma, R.; Spaniol, M.; Cao, Y.; Jarke, M.: *Pattern-Based Cross Media Social Network Analysis for Technology Enhanced Learning in Europe*, *Proceedings of the 1st European Conference on Technology Enhanced Learning*, Crete, Greece, October 3–5, Springer LNCS, 2006.
- Klamma, R., et al.: Social Software for Life-long Learning, *Journal of Educational Technology & Society*, 10(3):72–83, 2007.
- Klamma, R.; Cao, Y.; Jarke, M.: Storytelling on the Web 2.0 as a New Means of Creating Arts, in Borko Furht (Ed.): *Handbook of Multimedia for Digital Entertainment and Arts*, Springer, S. 623–650, 2009.
- Latour, B.: On Recalling ANT, J. Law/J. Hassard (Eds): *Actor-Network Theory and After*, Oxford, S. 15–25, 1999.
- Licklider, J.C.R.; Taylor, R.; Herbert, E.: The Computer as a Communication Device, *International Science and Technology*, April 1968.
- O'Reilly, T.: *What Is Web 2.0 – Design Patterns and Business Models for the Next Generation of Software*, 2005, www.oreilly.com, <http://www.oreillynet.com/pub/a/oreilly/tim/news/2005/09/30/what-is-web-20.html> (letzter Zugriff November 2010).
- Shirky, C.: Social Software: A New Generation of Tools, *Esther Dyson's Monthly Report* (10), 2003.
- Turner, T.; Smith, M.; Fisher, D.; Welsler, H.: Picturing USENET: Mapping Computer-Mediated Collective Action, *Journal of Computer-Mediated Communication*, 10(4), 2005.
- Yu, E.: Towards Modelling and Reasoning Support for Early-Phase Requirements Engineering. In: *Proceedings 3rd IEEE Int. Symp. on Requirements Engineering (RE'97)* Jan. 6–8, 1997, Washington D.C., USA, S. 226–235, 1997.
- Watts, D.; Strogatz, S.: Collective Dynamics of 'Small World' Networks, *Nature* 393, S. 440–442, 1998.
- Wellman, B.; Boase, J.; Chen, W.: The networked nature of community on and off the Internet, *IT and Society*, vol. 1, No. 1, S. 151–165, 2002.
- Wenger, E.: *Communities of Practice: Learning, Meaning and Identity*, Cambridge University Press, UK, 1998.



Ralf Klamma ist Privatdozent und akademischer Oberarzt am Lehrstuhl Informatik 5 (Informationssysteme und Datenbanken) der RWTH Aachen. Seine Forschungsinteressen

umfassen unter anderem Community Informationssysteme, Metadaten, Social Software, die Analyse von Netzwerken, Technologie-gestütztes Lernen, Mobile Multimedia und Web 2.0 Anwendungen.

E-Mail: klamma@dbis.rwth-aachen.de

<http://dbis.rwth-aachen.de/cms/staff/klamma>