



Hochschule für Angewandte Wissenschaften Hamburg  
*Hamburg University of Applied Sciences*

## Anwendung 2

Lars Mählmann

Deliever who I mean

Lars Mählmann  
Deliever who I mean

Anwendung 2 eingereicht im Rahmen der Diplomprüfung  
im Studiengang Informatik  
Studienrichtung Softwaretechnik  
am Fachbereich Elektrotechnik und Informatik  
der Hochschule für Angewandte Wissenschaften Hamburg

Betreuender Prüfer :  
Zweitgutachter : Prof. Dr. Kai von Luck

Abgegeben am 13. März 2008

## Inhaltsverzeichnis

<b>1</b>	<b>Einleitung</b>	<b>3</b>
<b>2</b>	<b>Deliever who I mean</b>	<b>4</b>
2.1	Social Networking . . . . .	4
2.2	Konzept . . . . .	5
<b>3</b>	<b>Anwendungen</b>	<b>6</b>
3.1	Backend . . . . .	6
3.1.1	Polyphonet, An Advanced Social Network Extraction System from the Web . . . . .	6
3.1.2	Augmenting employee profiles with people-tagging . . . . .	6
3.1.3	Towards Effective Browsing of Large Scale Social Annotations . . . . .	8
3.1.4	Weitere Ansätze . . . . .	8
3.2	Zusammenfassung . . . . .	9
3.3	Frontend . . . . .	10
3.3.1	Magnet: Supporting Navigation in Semistructured Data Environments . . . . .	10
3.3.2	Keywords and RDF Fragments: Integrating Meta-data and Full-Text Search in Beagle++ . . . . .	11
3.4	Zusammenfassung . . . . .	12
<b>4</b>	<b>Fazit</b>	<b>12</b>
	<b>Literatur</b>	<b>14</b>

## 1 Einleitung

Diese Ausarbeitung diskutiert die Erweiterung einer Intranetsuche, um die Sucherweiterung nach Personen. Es wird auf die Probleme eingegangen und mögliche Lösungsansätze aus ähnlichen Projekten vorgestellt.

Die Gliederung der Ausarbeitung ist wie folgt: Es wird vorgestellt, was eine Intranetsuche beschreibt und die Erweiterung durch eine Personensuche ( siehe 2). Es werden die entstehenden Schwierigkeiten beschrieben (siehe Kapitel 2.2). In dem darauffolgenden Abschnitten 3 werden einzelne Arbeiten aus dem Forschungsbereich vorgestellt, die sich direkt oder indirekt damit beschäftigen. Es folgt eine Bewertung der einzelnen Arbeiten in Kapitel 3.2. Abschliessend wird eine Abgrenzung zwischen der Problemstellung und den Projekten beschrieben. In dem Fazit 4 wird ein Ausblick auf noch bestehende Probleme der Arbeit (Aufgabenstellung) gegeben.

## 2 Deliver who I mean

In Firmen besteht meistens ein Computernetzwerk, das unabhängig von Ort und Zeit sämtliche Abteilungen miteinander verbindet. Zugänglich ist ein solches Netzwerk für Mitarbeiter über das Intranet. Im Intranet werden Dokumente, Anleitungen, Projekte, Schulungsunterlagen oder Benutzerverzeichnisse zur Verfügung gestellt. Bei größeren Firmen werden die Strukturen eines Intranets schnell groß, unübersichtlich und das Finden von Informationen schwierig. Die im Intranet angebotenen Suchfunktionen sind eine Hilfe, aber die wertvollen Information oft nicht verfügbar, indiziert oder schlicht nicht katalogisiert. Viele der Informationen sind nur verfügbar, wenn man die richtigen Personen kennt, bzw. fragt. Die Herausforderung besteht darin, die richtige Person zu finden, die unsere Frage beantworten kann. Das Finden der richtigen Person basiert meistens auf unserem "Social Network" ( [Yu und Singh \(2002\)](#)). Könnte man die normale Suche so erweitern, daß auch Personen gefunden werden, würde sich die Effektivität eines Intranets deutlich verbessern. Im Folgenden soll erklärt werden, was ein "Social Network" ist, wie es für eine Suche genutzt werden kann und welche Probleme bestehen.

### 2.1 Social Networking

Bei der Suche nach einer Person oder einer Personengruppe wird versucht Gemeinsamkeiten zwischen diesen Personen zu finden. Die Gemeinsamkeiten können verschiedene Relationen sein, wie Freundschaft, Arbeitskollegen, gleiche Interessensgebiete, Visionen, Geschäftsbeziehung, dieselbe Universität, Konferenzen, etc . . . ([Matsuo u. a. \(2006\)](#)).

Die Struktur wird in Graphen abgebildet, wobei Personen die Knoten darstellen und Beziehungen im Netzwerk als Kanten bezeichnet und beschrieben werden Granovetter ([Granovetter \(1973\)](#)). Diese Darstellung wird in den verschiedenen vorgestellten Arbeiten benutzt und weiter ausgebaut

Solche Beziehungen oder Verbindungen bezeichnet man als "Social Networking". Als Begründer des Begriffes "Soziale Netzwerke" wird J.S. Barnes ([Barnes \(1954\)](#)) betrachtet, er führt in das Netzwerk Konzept soziale, produktive und individuelle Beziehungsfelder ein. Diese Beziehungsfelder sind lockere, indirekte, teilweise unüberschaubare Geflechte sozialer Beziehungen, an denen die Mitglieder sich mehrheitlich und wechselseitig beteiligen. Dabei können Netzwerke in Größe, Dichte, Stabilität und Qualität unterschiedlich sein. Innerhalb der Netzwerke ist der soziale Rückhalt ([J.S. \(1981\)](#)) durch Informationen instrumentiert und durch die faktische Verfügbarkeit von Hilfeleistung ([Siegrist \(1995\)](#)) gegeben.

Wie können diese Begriffe und Darstellung in einem Computersystem abgebildet und nutzbar gemacht werden? Der nächste Abschnitt beschreibt die Probleme beim Aufbau eines

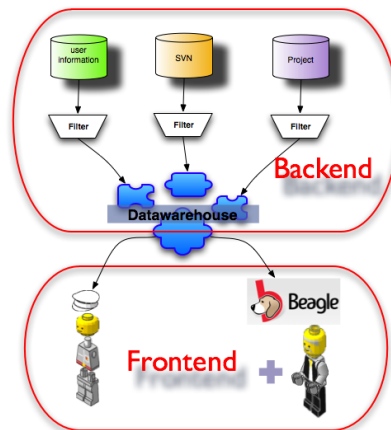


Abbildung 1: Abbildung eines Konzeptes zur Personensuchen

solchen und zeigt ein Konzept an dem die verschiedenen Projekte beschrieben werden sollen, die sich mit einzelnen Bereichen beschäftigen.

## 2.2 Konzept

Für den Aufbau eines "Social Networks" und die Suche nach Personen gilt es verschiedene Schwierigkeiten zu bewältigen.

1. In einem ersten Schritt müssen Daten zur Verfügung stehen, um ein "Social Network" aufzubauen. Diese Daten müssen aus dem Intranet extrahiert werden.
2. Sind die Daten extrahiert, müssen die Daten ausgewertet werden. Welche Daten passen auf welche Personen und welche Personen sind hilfreich eine Person zu beschreiben und Relationen zwischen Personen aufzubauen.
3. Stehen diese Informationen zur Verfügung müssen diese in einer maschinenlesbaren Form hinterlegt sein, um sie mit einer Suchmaschine zu finden.

Die Grafik<sup>1</sup> zeigt wie so eine Struktur aussehen könnte. Man kann es aufteilen in ein Backend und ein Frontend. In dem Frontend sind die Personenbeschreibungen in einem Profil abgelegt, welches mit einer Suchmaschine auffindbar ist. Im Backend werden die Daten aus den einzelnen Strukturen des Intranets extrahiert und im zweiten Schritt in einem Datawarehouse abgelegt und mit Data Mining Algorithmen bewertet und in regelmäßigen Abständen aufbereitet um die Aktualität zu behalten.

## 3 Anwendungen

In diesem Kapitel werden mehrere Arbeiten vorgestellt, die ein "Social Network" aufgebaut haben oder in dem Kontext des "Social Networks" Untersuchungen vorstellen. Die einzelnen Projekte werden kurz dargestellt und an Hand der in Kapitel 2.2 beschriebenen Schritte eingeteilt und beurteilt.

### 3.1 Backend

Die ersten Projekte beziehen sich auf die in Kapitel 2.2 vorgestellte Abbildung 1 des Backends. In dem es um die Informationsgewinnung aus dem Netzwerk geht, d.h. die Datenextrahierung und die Auswertung mit Hilfe von Data Mining Algorithmen.

#### 3.1.1 Polyphonet, An Advanced Social Network Extraction System from the Web

Polyphonet 2 ist ein Projekt aus der ein Social Network aufgebaut werden kann. Das System findet Beziehungen zwischen Personen, Beziehungen zwischen verschiedenen Gruppen von Personen und ordnet Schlagwörter einer Person zu. Das Ergebnis des Projektes wurde als eine "Super Social Network Mining" Architektur vorgestellt und auf mehreren Konferenzen erfolgreich getestet (Matsuo u. a. (2006)). Zunächst wurden verschiedene Algorithmen entwickelt, um Informationen aus dem jeweiligen System zu extrahieren. Im zweiten Schritt wurden mehrere Data-Mining Algorithmen verwendet, die eine Einteilung von verschiedenen Relationen in Kategorien vornimmt. Durch die Skalierung der Daten verbesserte man die Möglichkeiten Beziehungen zwischen Personen und Begriffen zu erstellen. Es zeigt die Beziehung von Personen zwischen einzelnen Fachgebieten.

In Polyphonet werden verschiedene Algorithmen zum Extrahieren von Daten beschrieben, die Algorithmen basieren auf Google Hacks Calishain und Dornfest (2003) und wurden angepasst auf die Umgebung. Es wird auf die Skalierung der Daten eingegangen und erklärt wie man mit der Menge von verschiedenen Beziehungen zwischen Personen und Dokumenten arbeitet und diese verarbeitet. Es werden weitere Algorithmen vorgestellt, mit denen die Metadaten von Personen an Hand von Personen - Wort - Matrizen besser sortiert werden können.

#### 3.1.2 Augmenting employee profiles with people-tagging

Ziel der Arbeit war die Erweiterung von bestehenden Intranetprofilen 3. Es wurde versucht automatisch Inhalte aus anderen Systemen, z. B. "Web 2.0" Technologie ( Blogs oder Social



Abbildung 2: Screenshot von einer Person aus Polyphoner



Abbildung 3: Abbildung eines Benutzerprofils mit der Erweiterung von Fringe

Bookmarks) in die Profile zu integrieren. Benutzer wurden befähigt an Hand von Schlüsselwörtern andere Benutzer zu "taggen" bzw. zu klassifizieren. Durch das "Taggen" mit "Social Bootmarks" wird ein Standardbenutzerprofil erweitert [Farrell u. a. \(2007\)](#).

Es wurde untersucht inwieweit und wie Benutzer diese Anwendung verwenden und es die Qualität der bestehenden Profile verbessert und ob man diese Technologie des "taggens" auf mehrere Anwendungen erweitern kann.

Mit Hilfe von verschiedenen Algorithmen (Data mining, text processing) wurde ermittelt, womit eine Person "getaggt" wurde. Dabei wurde versucht ähnliche Begriffe zu einem zusammenzufassen und gezählt wie häufig eine Person mit verschiedenen Thematiken in Zusammenhang gebracht wurde. Durch Algorithmen, ähnlich dem Google' PageRank, <sup>1</sup> wurde ermittelt welche Person zu den jeweiligen Begriffen die "beste Bewertung" hat.

<sup>1</sup>Ermittelt wird der PageRank eines Dokuments rekursiv anhand von Verweisen auf dieses Dokument. Je mehr solcher Verweise existieren und je höher jener PageRank der hierauf verweisenden Dokumente ist, desto höher fällt entsprechend der Wert des Dokuments aus, auf das verwiesen wird, <http://de.wikipedia.org/wiki/Google>, <http://www.google.com/technology/>

### 3.1.3 Towards Effective Browsing of Large Scale Social Annotations

Es gibt viele "Services", die den unterschiedlichen Benutzern die Möglichkeiten bieten ihre favorisierten Daten mit anderen Interessierten zu teilen und zu verwalten. Durch die steigende Vielzahl der sozialen Annotationen ergibt sich das Problem der effizienten Suchgestaltung. In kleineren Suchgebieten helfen z. B. TagClouds bei der Verwaltung, bei steigender Größe ist dies jedoch kompliziert. Eine Lösung für dieses Dilemma der Anwender und Serviceanbieter stellt ELSABer [Li u. a. \(2007\)](#). ELSABer kann größere Datenmengen effektiv, hierarchisch und semantisch durchsuchen und Verbindungen herstellen.

Dies erfolgt durch

- Erstellen semantischer Beziehung von ähnlichen Begriffen (z. B. Film, Movie) und dazugehörigen Begriffen
- Erstellung von hierarchischen Beziehungen in einer geordneten Abfolge (z. B. Tier, Katze, Siamkatze)
- Die Verbreitung "Sozialen Annotationen" wird für eine effektive Suche weiter untersucht.

Effective Large Scale Annotation Browser (ELSABer) kann, durch Einbeziehung diverser Komponenten, wie eine Suchfunktion bzw. Suchmaschine hier Lucene<sup>2</sup> und Clusteranalyse<sup>3</sup>, zu einer zeitlichen und personellen Suche erweitert werden. Der Nachteil bei dieser Anwendung ist, daß es keine vollständige Implementierung gibt und der Aufwand nicht genau beschrieben ist, da mit steigender Anzahl von "Tags" auch dessen Sichtung mehr Aufwand bedeutet.

### 3.1.4 Weitere Ansätze

Es gibt weitere Arbeiten, wie "Learning Social Networks from Web Documents Using Support Vector Classifiers" von [Makrehchi und Kamel \(2006\)](#) oder "Data Mining Through Fuzzy Social Network Analysis" von [Nair und Sarasamma \(2007\)](#). Es wurde im erst genannten Artikel versucht mit Support Vector Maschinen halbstrukturierte Datenmengen zu klassifizieren und um die fehlenden Daten zu erweitern. Probleme bei dieser Technik entstehen durch die Voraussetzungen des Trainingsets und dadurch, daß das Daten nicht vollständig, automatisch bearbeitet werden können. Im zweiten Fall wurde versucht mit Hilfe von Fuzzy Logik ein

---

<sup>2</sup>Apache Lucene is a high-performance, full-featured text search engine library, <http://lucene.apache.org>

<sup>3</sup>Unter Clusteranalyse versteht man ein strukturentdeckendes, multivariates Analyseverfahren zur Ermittlung von Gruppen (Clustern) von Objekten, deren Eigenschaften oder ihre Ausprägungen bestimmte Ähnlichkeiten oder Unähnlichkeiten aufweisen, <http://de.wikipedia.org/wiki/Clusteranalyse>



Data Mining auf schon bestehenden Daten vorzunehmen und darüber ein "Social Network" aufzubauen.

Ein einfach gehaltenen Bericht von [Matsumura u. a. \(2005\)](#) untersucht, wie mit Hilfe eines "Message Boards" und Data-Mining ein "Social Network" aufgebaut und Aussagen über die Teilnehmer getroffen werden können. Es wurde ein Distance Vektor aufgebaut in der Form, dass gesagt wurde umso mehr Nachrichten hin und her gehen und umso direkter die Nachrichten verschickt wurden, desto höher war die Gewichtung.

### 3.2 Zusammenfassung

Die hier vorgestellten Arbeiten sind ein kleiner Ausschnitt aus der Forschung die in den letzten beiden Jahren zu dem Thema verfasst worden sind. Die Schwierigkeit die aussagekräftigen Daten aus dem System herauszufiltern und richtig zu bewerten sind allen Arbeiten gemein. Die verschiedenen Berichte zeigen alle einen guten Ansatz, um Information aus einem Netzwerk heraus zu filtern. Bei den Arbeiten ([3.1.1](#), [3.1.3](#)) ist ein Problem gewesen, dass die Dateien und Dokumente nicht bekannt waren und keine exakten Aussagen darüber getroffen werden konnten, mit welchen Datentypen man gerade arbeitet. Eine weitere Schwierigkeit ist die Menge an Daten und der Aufwand der damit verbunden ist. Bei der Arbeit [3.1.3](#) entwickelte sich der Vergleich von Eigenschaften und Personen mit einem  $O(n^2)$ . Bei der Towards Effective Browsing of Large Scale Social Annotations ([3.1.3](#)) konnte über den Aufwand keine Aussage gemacht werden, da nur exemplarisch ein Prototyp entwickelt wurde. Bei Fringe in [3.1.2](#) vertraute man mit dem Tag Mechanismus darauf, dass die Benutzer nur korrekte Angaben machen. Es gibt keine direkte Überprüfung der Eingaben. Eine Thematik die bei den Projekten nicht benannt wurde, ist die Zeit, wann und wie lange die Informationen gültig sind oder welche wichtige Informationen enthalten.

Die Probleme bei der Informationsgewinnung im "Backendbereich" in den vorgestellten Arbeiten waren: Die aufwendigen Vergleiche von Benutzerinformationen und die Bestimmungen der Relationen bzw. Zuordnung von Personen, Gruppen und Schlüsselworten (Vgl. [Matsuo u. a. \(2006\)](#)). Bei Fringe [3.1.2](#) muß man den "taggen" der Benutzer vertrauen und hat nur eine beschränkte Kontrollstruktur, womit Profilbeschreibung nicht vollständig oder einfach falsch sein können. Weitere Arbeiten sind nur bedingt nutzbar, weil diese nicht eine komplette Suche abbilden (siehe [Matsumura u. a. \(2005\)](#)) oder nicht automatisiert ausführbar sind (siehe [Makrehchi und Kamel \(2006\)](#)).

### 3.3 Frontend

Die folgenden Arbeiten beziehen sich auf die in Kapitel 2.2 vorgestellte Abbildung 1 des Frontends. Es werden Projekte vorgestellt, die sich mit einer strukturierten Speicherung von Daten beschäftigen. Das heißt die Daten werden in einer Struktur abgespeichert, die maschinenlesbar ist, erweiterbar und geeignet sein könnten, die gewonnenen Informationen aus dem Data Mining zu speichern und die Daten für eine Suchmaschine indizierbar zu halten. Diese Aufbereitung ist nötig, um eine für den Computer lesbare und interpretierbare Darstellung zu besitzen. Es müssen die gewonnenen Informationen in einer Struktur zur Verfügung gestellt werden, die eine Person repräsentieren und durch die Suchmaschine auffindbar sind.

#### 3.3.1 Magnet: Supporting Navigation in Semistructured Data Environments

Die Notwendigkeit der Unterstützung von suchenden Benutzern ist mit der steigenden Bedeutung von Systemen, die willkürlich semistrukturierte Beziehungen beinhalten, gestiegen. Bei der derzeitigen Unterstützung der Benutzer benötigt dieser entweder Kenntnisse über Domainspezifische Benutzerinterfaces oder muss "schema" Experte sein [Sinha und Karger \(2005\)](#). In dieser Arbeit wurde ein allgemeinnützlich Tool (Magnet, siehe Abbildung 4) entwickelt, das dem Benutzer in der Orientierung hilfreich ist und Verfeinerungsmöglichkeiten bei der Suche innerhalb der semistrukturierten Informationen bietet. Es wird gezeigt, wie ein solches Tool hergestellt werden kann, ohne dass domainspezifische Voraussetzungen benötigt werden. Zusätzlich zu der Beschreibung der generellen Annäherung an das Problem, wird eine Anzahl von natürlichen, allgemeinnützlichen Verfeinerungstaktiken vorgestellt, wobei viele Taktiken aus älteren Arbeiten über eine textuelle Informationssuche stammen.

Magnet präsentiert Benutzern abfragende Verfeinerungen und Orientierungsmöglichkeiten, die sie anwenden können, um die gesuchte Information zu finden. Dabei zeigt Magnet, dass einige Suchtechniken für "fuzzy" Informationen, die für textuelle Suche effektiv sind, auch für die Suche in semistrukturierten Daten nützlich sind. Es wurden annähernde Evaluationen für unterschiedliche Datenquellen, mit variierenden Informationsschemata, durchgeführt, die Flexibilität des Browsing Interface auf Abfragen von Suchenden XML Initiativen examiniert und eine Benutzerstudie mit einem System, das die Arbeit mit Metadata Browsing erklärt, durchführt.

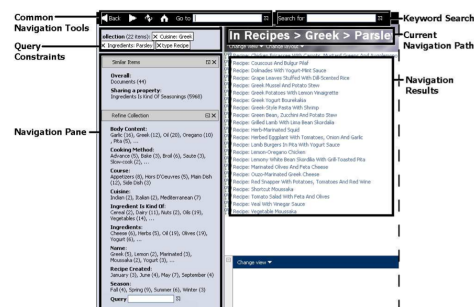


Abbildung 4: Abbildung von Magnet

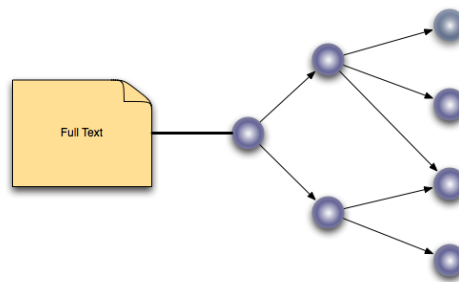


Abbildung 5: Konzept von Beagle++

### 3.3.2 Keywords and RDF Fragments: Integrating Meta-data and Full-Text Search in Beagle++

Der Untersuchungsschwerpunkt war es eine Suche zu entwickeln, die aus einer Volltextsuche und einer Metasuche auf einem Desktopcomputer besteht. Es wurde um die zusätzlich die Funktionalität eines "result rankings" und um die Möglichkeit einer unscharfen Suche [lofcu u. a. \(2005\)](#) erweitert. Die Volltextsuche wird mit Hilfe von Beagle<sup>4</sup> einer Desktopsuchmaschine basierend auf Lucene (siehe auch [3.1.3](#)) umgesetzt. Für die Metasuche wurden die Informationen aus den einzelnen Dokumenten in Resource Description Format (RDF, für mehr Details [Mählmann \(2008\)](#)) abgespeichert (siehe [Abbildung 5](#)). RDF ist ein Teil von Semantic Web und wird als Graph in einer XML Datei abgespeichert. Diese XML Beschreibung kann mit Hilfe eines Vector Space Models, wie in [Carmel u. a. \(2003\)](#) beschrieben, durchsucht werden und eine bessere Bewertung der Suchergebnisse erstellen.

<sup>4</sup>[www.beagle-project.org](http://www.beagle-project.org)

### 3.4 Zusammenfassung

In den vorgestellten Arbeiten wurde jeweils das Resource Description Framework eingesetzt, um eine Metabeschreibung von Dateien und Informationen anzulegen. In Magnet (3.3.1) wurde ausserdem gezeigt, dass man gut durch die Strukturen suchen und zusätzlich Verweise zu weiteren Beschreibungen Dokumenten einbinden kann.

RDF ist eine der Spezifikationen von Semantik Web.

RDF along with OWL are Semantic Web specifications <sup>5</sup>

Es ist daher ausgelegt für eine maschinenlesbare Beschreibung. Die Spezifikation lässt zusätzlich Erweiterung für die speziellen Bedürfnisse in einer Beschreibung zu (Vgl. <sup>6</sup>). Ein Beispiel dafür ist Beagle++ 3.3.2 bei der eine Desktopsuche mit einer Metabeschreibung durch RDF kominiert wurde. Die Kombination wurde mit einer grossen Menge unterschiedlicher Datentypen erfolgreich getestet.

Betrachtet man das Speichern und Erstellen von Benutzerprofilen im "Frontend" 3.3 zeigt sich mit dem Resource Description Framework (RDF) eine gute Lösung. Die vorgestellten Projekte konnten verschiedene Daten erfolgreich in einer RDF Metabeschreibung ablegen. Bei dem Umgang mit den unterschiedlichen Datentypen oder Datenmengen stellte sich dieses Format als zuverlässig heraus (siehe Makrehchi und Kamel (2006)). Die in den verschiedenen Projekten benutzte Suche "Lucene" kann ebenso wie die darauf basierende Suchmaschine "Beagle" mit den Daten und Metadaten umgehen und scheint geeignet für die Durchsuchung von RDF Profilen.

## 4 Fazit

Die hier vorgestellten Projekte zeigen, daß es grundsätzlich möglich ist eine Personensuche zu implementieren. Allerdings ist die Funktionalität auf Grund der Analyse von den Datenmengen beschränkt. Die Suche und die Auswertung der Informationen wurde auf unbekannt Daten ausgeführt und nur syntaktisch vorgenommen eine weitere Schwierigkeit war dass ein sehr grosser und unbekannter Bereich nach Daten durchsucht wurde. Die Vertrauenswürdigkeit der Informationen ist nicht direkt erkennbar oder überprüfbar. Vergleicht man das mit dem Eingangs erwähnten Beispiel einer Suche in einem Intranet lassen sich die genannten Probleme verringern.

---

<sup>5</sup><http://www.w3.org/RDF/>

<sup>6</sup><http://www.xml.com/pub/a/98/06/rdf.html>

In einem Intranet kann man von der Vertrauenswürdigkeit der Informationen ausgehen. Es sind die Informationsquellen bekannt. Die Quellen können unterschiedlich betrachtet werden. Die Auswertung wird somit semantisch angereichert mit Informationen über den Ursprung der Daten. Die Daten in einem Intranet sind zudem teilweise strukturiert, wie z. B. in einem Benutzerverzeichniss (z. B. LDAP Bäume).

Die Suche ist in einem Intranet eingeschränkt auf bestimmte Bereiche und somit die Informationsgewinnung vereinfacht. Ein nächster Schritt wäre demnach in einem Intranet sich verschiedene Quellen auszusuchen und diese mit den Algorithmen und vorgestellten Anwendungen zu untersuchen und zu ermitteln, wie hilfreich eine so erstellte Suche sein kann.

## Literatur

- [Barnes 1954] BARNES, J.: Class and Committees. In: *Human Relations* (1954)
- [Calishain und Dornfest 2003] CALISHAIN, Tara ; DORNFEST, Rael: *Google Hacks: 100 Industrial-Strength Tips and Tools*. Sebastopol, CA, USA : O'Reilly & Associates, Inc., 2003. – ISBN 0596004478
- [Carmel u. a. 2003] CARMEL, David ; MAAREK, Yoelle S. ; MANDELBROD, Matan ; MASS, Yosi ; SOFFER, Aya: Searching XML documents via XML fragments. In: *SIGIR '03: Proceedings of the 26th annual international ACM SIGIR conference on Research and development in information retrieval*. New York, NY, USA : ACM, 2003, S. 151–158. – ISBN 1-58113-646-3
- [Farrell u. a. 2007] FARRELL, Stephen ; LAU, Tessa ; NUSSER, Stefan ; WILCOX, Eric ; MULLER, Michael: Socially augmenting employee profiles with people-tagging. In: *UIST '07: Proceedings of the 20th annual ACM symposium on User interface software and technology*. New York, NY, USA : ACM, 2007, S. 91–100. – ISBN 978-1-59593-679-2
- [Friedman 2003] FRIEDMAN, Nat: Dashboard. In: *In Proceedings of the Linux Symposium Conference on the Linux kernel and major OS infrastructure and research projects, 2003*
- [Granovetter 1973] GRANOVETTER, Mark: The strength of weak ties. In: *American Journal of Sociology* (1973)
- [Iofciu u. a. 2005] IOFCIU, Tereza ; KOHLUETTER, Christian ; PAIU, Raluca ; NEJDL, Wolfgang: Keywords and RDF Fragments: Integrating Metadata and Full-Text Search in Beagle++. In: *Workshop on The Semantic Desktop - Next Generation Personal Information Management and Collaboration Infrastructure at the International Semantic Web Conference, 2005*
- [J.S. 1981] J.S., House: *Work Stress and social support*. Reading, Mass : Addison-Wesley, 1981
- [Li u. a. 2007] LI, Rui ; BAO, Shenghua ; YU, Yong ; FEI, Ben ; SU, Zhong: Towards effective browsing of large scale social annotations. In: *WWW '07: Proceedings of the 16th international conference on World Wide Web*. New York, NY, USA : ACM, 2007, S. 943–952. – ISBN 978-1-59593-654-7
- [Mählmann 2008] MÄHLMANN, Lars: Entwicklung eines Systems zur Erstellung und Suche von Benutzerprofilen, 2008

- [Makrehchi und Kamel 2006] MAKREHCHI, Masoud ; KAMEL, Mohamed S.: Learning Social Networks from Web Documents Using Support Vector Classifiers. In: *WI '06: Proceedings of the 2006 IEEE/WIC/ACM International Conference on Web Intelligence*. Washington, DC, USA : IEEE Computer Society, 2006, S. 88–94. – ISBN 0-7695-2747-7
- [Matsumura u. a. 2005] MATSUMURA, Naohiro ; GOLDBERG, David E. ; LLORÀ, Xavier: Mining directed social network from message board. In: *WWW '05: Special interest tracks and posters of the 14th international conference on World Wide Web*. New York, NY, USA : ACM, 2005, S. 1092–1093. – ISBN 1-59593-051-5
- [Matsuo u. a. 2006] MATSUO, Yutaka ; MORI, Junichiro ; HAMASAKI, Masahiro ; ISHIDA, Keisuke ; NISHIMURA, Takuichi ; TAKEDA, Hideaki ; HASIDA, Koiti ; ISHIZUKA, Mitsuru: POLYPHONET: an advanced social network extraction system from the web. In: *WWW '06: Proceedings of the 15th international conference on World Wide Web*. New York, NY, USA : ACM, 2006, S. 397–406. – ISBN 1-59593-323-9
- [Nair und Sarasamma 2007] NAIR, Premchand S. ; SARASAMMA, Suseela T.: Data Mining Through Fuzzy Social Network Analysis. In: *North American Fuzzy Information Processing Society, 2007*, S. 251 – 255. – ISBN 1-4244-1214-5
- [Siegrist 1995] SIEGRIST, Johannes: *Medizinische Soziologie : mit 13 Tabellen*. München, Wien, Baltimore : Urban und Schwarzenberg, 1995. – ISBN 3-541-06385-8
- [Sinha und Karger 2005] SINHA, Vineet ; KARGER, David R.: Magnet: supporting navigation in semistructured data environments. In: *SIGMOD '05: Proceedings of the 2005 ACM SIGMOD international conference on Management of data*. New York, NY, USA : ACM, 2005, S. 97–106. – ISBN 1-59593-060-4
- [Tang u. a. 2007] TANG, John C. ; LIN, James ; PIERCE, Jeffrey ; WHITTAKER, Steve ; DREWS, Clemens: Recent shortcuts: using recent interactions to support shared activities. In: *CHI '07: Proceedings of the SIGCHI conference on Human factors in computing systems*. New York, NY, USA : ACM, 2007, S. 1263–1272. – ISBN 978-1-59593-593-9
- [Yu und Singh 2002] YU, Bin ; SINGH, Munindar P.: Searching Social Networks. Raleigh, NC, USA : North Carolina State University at Raleigh, 2002. – Forschungsbericht