



DIPLOMARBEIT

Folksonomy

ausgeführt am Institut für
Gestaltungs- und Wirkungsforschung
der Technischen Universität Wien

unter der Anleitung von
ao.Univ.Prof. Peter Purgathofer

durch

Christine Albrecht
Matrikelnummer: 9825424

Gudrunstr. 121/46
A-1100 Wien

Wien, am 17. März 2006

Kurzfassung

Folksonomy ist eine neuartige Form des Datenmanagements, die auf dem Vorgang des Taggings basiert. Durch die sich immer stärker ausbreitende Vernetzung werden neue Arten der Organisation der dynamischen Inhalte des Webs notwendig. Folksonomy bietet diese neue Denkweise, die von herkömmlichen Ansätzen, wie Taxonomien und Ontologien, vollständig abkommt. Die Daten werden mit Tags belegt anstatt sie hierarchisch zu strukturieren. Durch das Fehlen dieser hierarchischen Abhängigkeiten sind Folksonomien weitaus flexibler und dynamischer als starre Klassifizierungssysteme und sind den gesteigerten Anforderungen des Webs gewachsen.

In dieser Arbeit wird auf die Unterschiede zwischen herkömmlichen hierarchischen Organisationsmodellen und Folksonomien eingegangen und ein Überblick über das Thema Folksonomy gegeben.

Danksagung

Ich danke meinen Eltern Radovan und Elfriede Albrecht dafür, dass sie mir durch ihre Unterstützung ermöglicht haben mein Studium erfolgreich zu vollenden.

Ich danke meinem Freund David Schmitt, weil er mir Mut gegeben hat und mich auch in schwierigen Zeiten ertragen hat.

Ich danke meinem Betreuer Peter Purgathofer dafür, dass er es mir möglich gemacht hat, meine Arbeit in kurzer Zeit zu einem positiven Abschluss zu bringen.

Inhaltsverzeichnis

1	Einleitung	1
2	Geschichte der Ontologie	3
2.1	Entstehung	5
2.2	Philosophische Ontologie	9
2.3	Ontologie in der Informatik	12
2.3.1	Ontologien in Datenbanken- und wissensbasierten Systemen	13
2.3.2	Ontologie in der Künstlichen Intelligenz	15
2.3.3	Weitere Anwendungen in der Informatik	16
2.4	Zusammenfassung	17
3	Geschichte und Theorie der Folksonomien	18
3.1	Grundlagen	18
3.2	Entwicklung	19
3.2.1	Definition nach Vander Wal	22
3.2.2	Mathematische Definition	23
3.2.3	Arten von Folksonomien	24
3.3	Power Laws	26
3.4	Zusammenfassung	29
4	Folksonomy vs. Ontologie	31
4.1	Vor- und Nachteile von Folksonomien	32
4.2	Vor- und Nachteile von Ontologien	35
4.3	Unterschiede zwischen Folksonomien und Ontologien	36

4.4	Zusammenfassung	37
5	Anwendungen von Folksonomien	38
5.1	Del.icio.us	38
5.2	Flickr	41
5.3	Furl	44
5.4	Technorati	46
5.5	CiteULike	47
5.6	43 Things	48
6	Visualisierung von Folksonomien	51
6.1	Tagwolken	52
6.2	Mindmaps	53
6.3	Del.icio.us Soup	57
6.4	Graph Del.icio.us Related Tags	59
6.5	Expialidocio.us	60
6.6	Flickr Related Tag Browser	61
6.7	Flickr Colors	63
6.8	Flickr Graph	64
6.9	Flickr Album	65
7	Zusammenfassung	66
	Abbildungsverzeichnis	68
	Literaturverzeichnis	70

Kapitel 1

Einleitung

Diese Arbeit beschäftigt sich mit dem Thema Folksonomy. Das Wort Folksonomy setzt sich aus den Wörtern „Folk“ – *Volk, Menschen* – und „nomia“ – *Management* – zusammen. Folksonomy bedeutet also in etwa von Menschen gemachtes Management. In der Literatur wird immer wieder der Fehler gemacht zu behaupten Folksonomy stamme von „Folk“ und „Taxonomie“ ab. Doch mit einer Taxonomie hat Folksonomy kaum etwas gemeinsam.

Eine Folksonomy ist ein soziales Netzwerk, dass sich aus der Tätigkeit des *Free Tagging* ergibt. *Tagging* ist das Benennen oder Beschlagworten von Objekten, wie zum Beispiel Bookmarks, Bildern oder Zitaten, mit Tags. Diese sind – im Fall von Free Tagging frei gewählte – Wörter, die die Objekte umschreiben. Der Benutzer sucht sich selbst Wörter als Tags aus, die für ihn persönlich die Objekte bestmöglich beschreiben, sodass er die Objekte anhand dieser Tags wiederfinden kann.

Mit Webrecherche und dem Testen bestehender Folksonomy-Systeme habe ich mir einen Überblick über das Thema verschafft.

Immer wieder wird Folksonomy in der Literatur mit Ontologien verglichen. Deshalb beginne ich meine Arbeit mit einer Einführung in die Geschichte der Ontologie von Plato und Aristoteles bis zu modernen Anwendungen aus dem Bereich der Informatik in Kapitel 2. Im Kapitel 3 erläutere ich die Geschichte der Folksonomien, die nicht annähernd so lange ist, wie die der Ontologie. Die ersten Wurzeln und Ideen von Folksonomien sind im Jahre 2001 mit dem Projekt Annotea entstanden und erst 2004 wurde das Wort „Folksonomy“ von Thomas Vander Wal geprägt. Die Statistik zeigt, dass viele Aspekte von Folksonomien der Power Law

Verteilung folgen. Deshalb gehe ich im letzten Teil dieses Kapitels noch auf diese Art der Verteilung ein.

Im weiteren Verlauf erläutere ich dann in Kapitel 4 die Unterschiede zwischen Folksonomy und Ontologie sowie deren Vor- und Nachteile, bevor ich in Kapitel 5 auf die Anwendungen von Folksonomien und in Kapitel 6 auf diverse Visualisierungen näher eingehe.

Kapitel 2

Geschichte der Ontologie

Die Ontologie im philosophischen Sinn ist die Wissenschaft des Seins, der Strukturen von Objekten, Prozessen und Relationen in jedem Bereich der Realität. Sie ist der Teil der Metaphysik¹, die die Prinzipien und Gründe des Seins und die Natur untersucht und erklärt.

Ontology:

- 1. a branch of metaphysics concerned with the nature and relations of being*
- 2. a particular theory about the nature of being or the kinds of existents*

[38]

Ontologie ist die Behauptung einer Logik des Seins. Sie postuliert eine Logik, die allem Sein vorausgesetzt ist. Alles soll aus der Seinslogik heraus allgemein und absolut begründet sein.

In Abbildung 2.1 sieht man eine beispielhafte Ontologie, wie sie in der Biologie zur Einteilung der Tierwelt verwendet wird. Diese Einteilung ist streng hierarchisch gegliedert und soll nur einen Ansatz einer Ontologie darstellen. Es wurden in dem Beispiel nur drei Detailstufen ausgeführt, die vollständige Ontologie kann den meisten Biologie Lehrbüchern entnommen werden.

¹Metaphysik (»metaphysica«, *meta ta physika*) ist die Wissenschaft von den Grundbegriffen (Prinzipien) des Erkennens und der Einzelwissenschaften in ihrem letzten für uns erreichbaren Sinne und in ihrem Zusammenhang untereinander und mit den Forderungen des nach Einheit und Geschlossenheit (Harmonie) der Weltanschauung strebenden Denkens.[21]

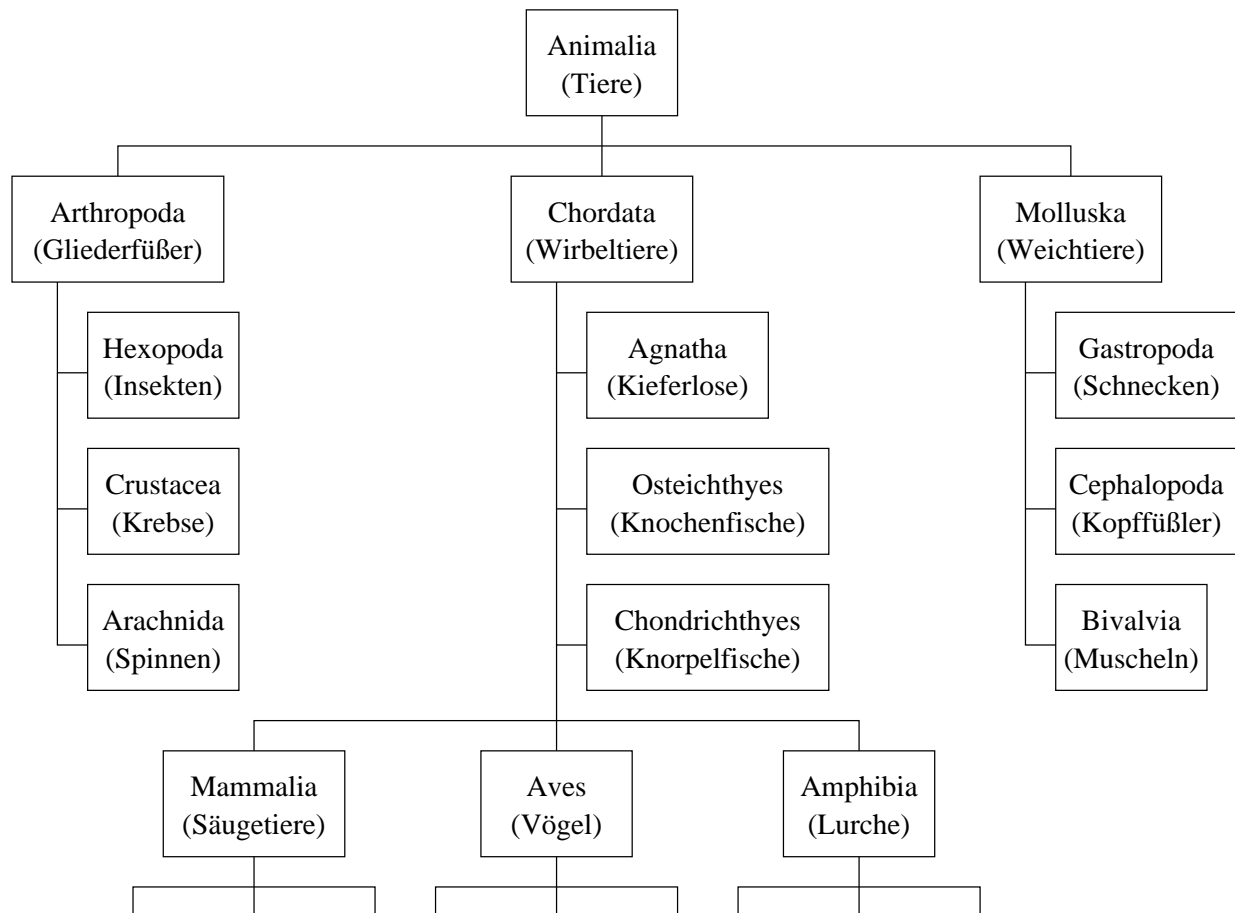


Abbildung 2.1: Ontologie der Tierwelt

Die Ontologie hat im Laufe der Zeit weitere Bedeutungen hinzugewonnen. Im Bereich der Informatik kann man im Wesentlichen zwischen zwei Arten unterscheiden – Ontologie in der Künstlichen Intelligenz und Ontologie im Bereich der Informatik allgemein.

Ontologie im Sinne der Künstlichen Intelligenz ist eine explizite formale Spezifikation wie Objekte, Konzepte und andere Entitäten und ihre Beziehungen zueinander repräsentiert werden. Diese Bedeutung leitet sich direkt von der philosophischen Ontologie ab. Das „Sein“ in der Künstlichen Intelligenz umfasst alles Repräsentierbare. Wenn das Wissen über einen spezifischen Bereich in einer deklarativen Sprache beschrieben wird, dann sind alle inkludierten Objekte in diesem Kontext enthalten. Die Namen der Entitäten im Universum werden in Text übersetzt, der für Menschen verständlich ist und beschreibt, was diese Namen bedeuten. For-

male Axiome beschränken die Interpretation auf eine wohlformulierte Verwendung der Terme. Formal gesehen bildet die Ontologie in der Künstlichen Intelligenz eine logische Theorie.

Die zweite, allgemeinere Bedeutung von Ontologie im Bereich der Informatik ist das hierarchische Strukturieren von Wissen über Objekte. Diese werden nach ihren essenziellen Qualitäten beurteilt und kategorisiert.[27]

Die Geschichte der Ontologie vereint eine beschreibende und eine erzeugende Seite. Die beschreibende Seite macht sich über die Aspekte der Realität, die unabhängig von jeglicher Taxonomie existieren, Gedanken, während die erzeugende Seite neue Kategorien schafft, um die Realität zu beschreiben. Die erzeugende Seite gibt der Ontologie Kraft sich über neue Bereiche von Entitäten auszubreiten, die beschreibende Seite bindet die Ontologie an die zugrundeliegende Realität.

2.1 Entstehung

Lange bevor das Wort Ontologie überhaupt existierte, stellte der Philosoph Plato (427 v. Chr. - 347 v. Chr.) die Frage nach dem „Sein“ und der Existenz im Allgemeinen und entwickelte hierarchische Einteilungen, die wir heute als Ontologien bezeichnen würden. In seinem Werk „Sophist“[43] findet ein Dialog über ein mögliches hierarchisches Modell der Welt statt, ausgehend von der Frage, ob Sophist, Staatsmann und Philosoph das gleiche oder drei verschiedene Entitäten sind.

Auch Aristoteles (384 v. Chr. - 322 v. Chr.) machte sich Gedanken über eine mögliche Einteilung der Welt und versuchte eine allgemeine Kategorisierung zu schaffen. Er teilte die Realität in zehn Kategorien ein, von denen jede Einzelne wiederum aus mehreren Unterkategorien besteht:

Von den ohne Verbindung gesprochenen Worten bezeichnen die einzelnen entweder ein Ding, oder eine Größe, oder eine Beschaffenheit oder eine Beziehung, oder einen Ort, oder eine Zeit, oder einen Zustand, oder ein Haben, oder ein Thun, oder ein Leiden.

Ein Ding ist, um es im Umriss anzudeuten, z.B. der Mensch, das Pferd; eine Größe ist z.B. das Zweieilige, oder Dreieilige; eine Beschaffenheit ist z.B. weiß, sprachge-

lehrt; eine Beziehung ist z.B. doppelt, halb, größer; ein Ort ist z.B. im Lykeion, auf dem Markte; eine Zeit ist z.B. Gestern, vorm Jahre; ein Zustand z.B. das Liegen, Sitzen; ein Haben z.B. Schuhe anhaben, bewaffnet sein; ein Thun z.B. er schneidet, er brennt; ein Leiden z.B. er wird geschnitten, er wird gebrannt. [16]

Das Wort *Ontologie* (lat.: *ontologia*) wurde Anfang des 17. Jahrhunderts unabhängig von den zwei Philosophen Rudolf Goclenius (1547-1628), in seinem „*Lexicon philosophicum*“, und Jacobus Lorhardus (1561-1609), in „*Theatrum philosophicum*“, geschaffen.

Jacobus Lorhardus (Jakob Lorhard) war Professor an der Universität St. Gallen (Schweiz). Seine Werke waren lange Zeit unbekannt. Das „*Theatrum philosophicum*“ – eigentlich „*Ogdoas Scholastica continens Diagraphen Typicam artium: Grammatices (Latinae, Graeca), Logices, Rhetorices, Astronomices, Ethices, Physices, Metaphysices, seu Ontologia*“ (1906) – wurde erstmalig 1985 in Joseph S. Freedmans zweiter Edition von „*Deutsche Schulphilosophie im Reformationszeitalter (1500-1650): ein Handbuch für den Hochschulunterricht*“ erwähnt. [19]

Der deutsche Philosoph und Logiker Rudolf Goclenius (Rudolf Göckel) war Professor für Physik, Logik und später auch Mathematik an der Universität Marburg (1581-1627). Sein 1913 in Frankfurt publiziertes Werk „*Lexicon philosophicum*“ enthält eine Sammlung naturwissenschaftlicher Erklärungen, dazu Erläuterungen zu Hauptbegriffen aus der Logik, Metaphysik, Ethik und Theologie.[15]

So wie auch Plato in seinem „*Sophist*“ versuchten Philosophen immer wieder ihre Realität einzuteilen und zu kategorisieren. Eine *Ontologie* kann auch als *Taxonomie*, als Struktur von Kategorien, gesehen werden. Jeder Typ von Entität entspricht einem Knoten in einem hierarchischen Baum. Alle Entitäten der Realität können so eindeutig einem Knoten zugewiesen und entlang der Kanten dieses Baumes eindeutig wiedergefunden werden.

Dieses Prinzip wandte der englische Bischof und Wissenschaftler John Wilkins an, um eine universelle Sprache zu erfinden. Wilkins schlägt eine allgemeingültige *Taxonomie* mit 40 Gattungen vor (siehe Abbildung 2.2).[57]

Wilkins wollte damit eine allgemeingültige Einteilung der Welt erschaffen. Wenn man diese Einteilung genauer betrachtet, stößt man auf folgende Probleme:

- *Unentscheidbarkeit von Einzelfällen*: In Wilkins Kategorisierung gibt es zum Beispiel eine Gruppe die zwischen Stein und Metall liegt. Wilkins hat seine Einteilung so weiterge-

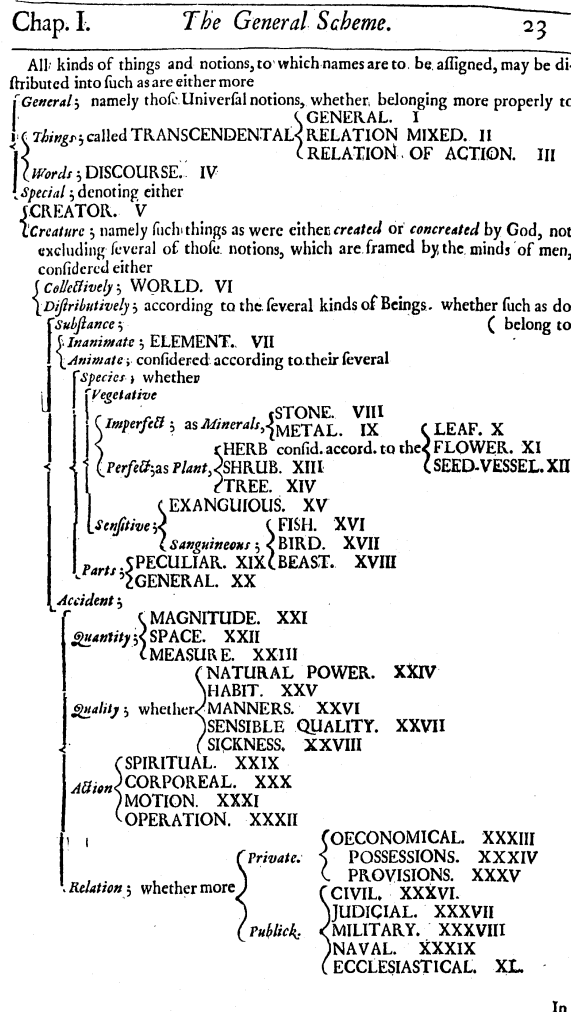


Abbildung 2.2: Kategorisierung der Realität von John Wilkins [57]

führt, dass metallhaltige Steine zwar eine Zwischenkategorie sind, aber dennoch seiner Gruppierung nach zu den Steinen gezählt werden. Ein weiteres Beispiel für Unentscheidbarkeit ist das Schnabeltier, dass in keine von Wilkins Gruppen hineinpasst und somit in seinem Weltmodell nicht existiert.

- *Zeitliche Veränderungen:* Nicht alle Kategorien von Wilkins sind in der heutigen Zeit noch passend, da wir neue Erkenntnisse gewonnen haben oder Kategorien, die Wilkins geschaffen hat nicht mehr existieren, z.B. weil das zugehörige Tier ausgestorben ist.
- *Vollständigkeit ist nicht möglich:* Wilkins konnte keine vollständige Kategorie erstellen, da

ein Mensch keinen allgemeinen Überblick über die gesamte Welt haben kann. Es besteht immer die Möglichkeit, dass noch nicht alles zu einem Thema erforscht und entdeckt wurde.

Philosophen, die versuchten eine allgemeingültige Ontologie der Welt zu schaffen, stießen immer wieder auf dieselben Probleme bis Immanuel Kant (1724-1804) eine gründliche Abkehr von der bisherigen Ontologie und Metaphysik vollzog:

Metaphysik, als Naturanlage der Vernunft, ist wirklich, aber sie ist auch vor sich allein dialektisch und trügerisch. Aus dieser also die Grundsätze hernehmen wollen, und in dem Gebrauche derselben dem zwar natürlichen, nichtsdestoweniger aber falschen Scheine folgen, kann niemals Wissenschaft, sondern nur eitele dialektische Kunst hervorbringen, darin es eine Schule der andern zuvortun, keine aber jemals einen rechtmäßigen und dauernden Beifall erwerben kann. [29]

Für Kant war Ontologie bzw. Metaphysik nur ein System aus Kategorien. Da diese Kategorien sich aber nur auf das vom Menschen Erfahrbare beziehen, sagen sie nichts über das Sein aus, das ja vom Menschen unabhängig besteht. Metaphysik ist für ihn im Vergleich zu einer Wissenschaft, so wie Astrologie zu Astronomie oder Alchemie zu Physik.

Ontologie, als Teil der Metaphysik, kann nicht als eine Wissenschaft angesehen werden, da sie sich von ihr in weiten Bereichen unterscheidet:

- Während die Ontologie nach einer Taxonomie der gesamten Realität sucht, stellt eine Wissenschaft zuerst Behauptungen auf und sucht erst dann nach deren Begründung.
- Die Ontologie ist qualitativ und will Aussagen über die Realität treffen, während eine Wissenschaft im Allgemeinen quantitativ ist. Sie beginnt mit Annahmen, Messungen und Behauptungen.
- Die Ontologie legt fest, welche Kategorien in einem gegebenen Bereich der Realität existieren und welche Kategorien für den Messprozess vorhanden sind. Eine Wissenschaft hingegen beschreibt zum Beispiel, wie das Messverhalten von Entitäten einer Klasse mit dem Verhalten von Entitäten einer anderen Klasse zusammenhängt.

- Ontologie befasst sich mit kategorieübergreifenden Relationen, eine Wissenschaft nur mit den Objekten, die in ihren Bereich fallen.

Immanuel Kant inspirierte die Philosophen des 20. Jahrhunderts dazu, die Ontologie als eine Disziplin zu betrachten, die sich auf einer Meta-Ebene über der Realität befindet. Dieser Ansicht nach befasst sich Ontologie nicht mit der Welt selbst, sondern mit Theorien, Sprachen, Konzepten oder Glaubenssystemen.

Weitere Philosophen der Neuzeit, die sich mit Ontologie und Metaphysik beschäftigt haben sind zum Beispiel Georg Wilhelm Friedrich Hegel (1770-1831), Edmund Husserl (1859-1938), Martin Heidegger (1889-1976), Jean-Paul Sartre (1905-1980) und Nicolai Hartmann (1882-1950).[56]

2.2 Philosophische Ontologie

Die philosophische Ontologie wird im Allgemeinen eine beschreibende oder realistische Ontologie genannt. Sie sucht nicht nach Erklärungen, sondern nach einer Beschreibung der Realität in Hinblick auf die Klassifizierung von Entitäten. Sie will Fragen zu beantworten, wie:

- Welche Klassen von Entitäten werden für eine vollständige Beschreibung und Erklärung für alle Geschehnisse im Universum benötigt?
- Welche Klassen von Entitäten werden benötigt, um alle Wahrheiten zu beschreiben?
- Welche Klassen von Entitäten erleichtern es uns Vorhersagen über die Zukunft zu machen?

Das oberste Ziel der philosophischen Ontologie ist eine Taxonomie aller Entitäten der Realität in allen möglichen Detailstufen, von mikroskopischen, über mesoskopische (alle Größenordnungen die menschenähnlich sind), bis hin zu makroskopischen Größenordnungen. Man kann hier einen Vergleich mit dem Reduktionismus² ziehen.

²Reduktion und Reduktionismus bedeuten ganz allgemein die Zurückführung (oder Zerlegung) von Komplizierten auf (in) Einfaches, Grundsätzliches oder Wesentliches. Reduktionismus ist keine besondere philosophische Lehre sondern eine von verschiedenen Philosophen mit zum Teil ganz unterschiedlichen Grundpositionen verwendete philosophische Methode.[40]

Zwischen Reduktionismus und Ontologie besteht aber ein wesentlicher Unterschied. Die Vertreter des Reduktionismus sehen die gesamte Realität aufgebaut aus einer einzigen privilegierten Ebene von Grundexistenzen. Jedes Konstrukt unserer Welt ist eindeutig in diese atomaren Grundbausteine zerlegbar. Die Anhänger des Reduktionismus wollen die Welt so lange rekursiv unterteilen, bis sie auf die grundlegenden Bestandteile stoßen, aus denen die gesamte Realität aufgebaut ist. Aus diesen Grundbausteinen heraus kann man dann die ganze Welt ableiten und erklären. Die Ontologie, im Gegensatz dazu, sucht nicht nach einer Erklärung oder Zerlegung der Welt, sie will nur die bereits vorhandenen Entitäten klassifizieren.

Geschichtlich gesehen haben sich aus der Ontologie heraus zwei Strömungen gebildet – die formale und die materielle Ontologie. Die formale Ontologie ist die Weiterführung der philosophischen Ontologie. Sie beschäftigt sich mit den allgemeinen Aspekten der Realität, wie dem Sein und der Identität. Sie versucht die Welt im allgemeinen zu verstehen und einzuteilen. Die materielle Ontologie beschäftigt sich, im Gegensatz dazu, mit bereichsspezifischen Aspekten, wie Kausalität oder Verstand. Sie strebt danach eine Taxonomie nur in den einzelnen Bereichen anstatt in der gesamten Realität zu erreichen.

Taxonomien in der Ontologie müssen eine Reihe von Kriterien erfüllen, die deren Form beschreiben und auch „Prinzipien der Wohlgeformtheit“ genannt werden.

1. Eine Taxonomie soll die Form eines mathematischen Baumes haben. Die Knoten repräsentieren Kategorien und die Äste repräsentieren die Relationen zwischen niedrigeren Kategorien zu höheren. Die Beziehung zwischen einem Knoten und dem direkt darunterliegenden Knoten nennt man „Eltern-Kind-Beziehung“. Die Gesamtheit der Objekte, die zu einer Kategorie gehören ist eine echte Untermenge der Kategorie, die diese Kategorie beinhaltet. Jede Kategorie ist also nur genau einer anderen Kategorie direkt untergeordnet. Von jedem Knoten führt genau ein Zweig nach oben. Die Taxonomien in den Naturwissenschaften, wie Botanik, Zoologie oder Chemie, entsprechen diesem Prinzip.
2. Eine Taxonomie soll eine Basis von minimalen Knoten enthalten, die keine Subkategorien mehr besitzen. Dieses Prinzip soll gewährleisten, dass jede Minimalkategorie vollständig ausgeschöpft wird und dass jeder beliebige Knoten im Baum durch eine Kombination von Minimalknoten eindeutig dargestellt werden kann.
3. Eine Taxonomie soll einen einzelnen obersten Maximalknoten enthalten, der die oberste

Kategorie darstellt. Diese oberste Kategorie inkludiert alle Kategorien weiter unten im Baum.

Wenn die gesamte Ontologie unserer Welt in einer einzelnen Taxonomie dargestellt werden könnte, dann müsste es einen Oberbegriff geben, der alle möglichen Kategorien enthält. Mehrere Philosophen, so wie auch John Wilkins, haben versucht eine solche Taxonomie zu erzeugen. Sie sind aber alle daran gescheitert oder haben Taxonomien geschaffen, die einer genauen Prüfung nicht standhalten. Das lässt darauf schließen, dass die Realität nicht in einem einzelnen Baum vereint werden kann, sondern nur durch eine Familie von Bäumen dargestellt werden kann, die alle spezifische Einblicke in einen kleinen Bereich der Realität liefern.

Die verschiedenen Bereiche oder auch Facetten der Wirklichkeit ergeben sich unter anderem dadurch, dass die Entitäten in der Realität verschieden mit der Zeit im Zusammenhang stehen. Manche Entitäten existieren „mit“ der Zeit, wie Substanzen, die von einem Moment zum anderen gleich bleiben, andere verändern sich, gleich eines Prozesses. Wiederum andere Entitäten existieren außerhalb der Zeit, da sie nichts mit der Zeit an sich zu tun haben. Zu dieser Gruppe gehören zum Beispiel die Zahlen. Diese verschiedenen Entitäten in einem einzelnen ontologischen Baum zu vereinen würde voraussetzen, dass eine zeitliche Überordnung existiert. Falls es diese Überordnung gibt, kann die Menschheit sie aber nicht erfassen, da sie in der Zeit lebt und diese nicht von außerhalb betrachten kann.

Ein anderes Argument gegen einen einzelnen Baum stammt noch von Aristoteles: Unterkategorien entstehen aus den Überkategorien: ein Mensch ist ein rationales Tier, ein Tier ist ein Lebewesen und so weiter. Wenn die gesamte Ontologie durch einen einzelnen Baum repräsentierbar wäre, würde es genau eine höchste Kategorie geben, von der sich alle anderen ableiten. Es kann aber nach Aristoteles keine Kategorie geben, die sowohl Tier als auch Handlung enthält. Obwohl das bereits Aristoteles erkannt hat, haben Philosophen noch lange versucht eine allumfassende Ontologie zu schaffen. Heute liegt das Hauptaugenmerk auf Teilontologien. Aber auch um diese Teilontologien zu erzeugen, muss man über die Dinge und Prozesse des Teiles der Realität, den man beschreiben will, Bescheid wissen. Um dieses Ziel zu erreichen, haben Philosophen angestrebt ihre ontologischen Untersuchungen der Wirklichkeit so genau wie nur möglich durchzuführen. Das führte dazu, dass die Ontologie den Wissenschaften immer näher kam. Manche Philosophen waren sogar der Meinung, dass man eine Ontologie nur

durch die genaue Untersuchung der wissenschaftlichen Theorien erzeugen kann.

Die philosophische Ontologie im eigentlichen Sinn, als eine oberste Disziplin auf die Welt darunter ausgerichtet, ist nicht möglich. Eine solche Ontologie würde eine externe Metaphysik voraussetzen, die sich außerhalb der Realität befindet. Da diese Perspektive dem Menschen nicht zugänglich ist, müssen die Philosophen ihre Thesen auf eine interne Metaphysik stützen. Sie können also nur die ontologischen Behauptungen von spezifischen Sprachen, Theorien oder Glaubenssystemen studieren. Die Ontologie im traditionell philosophischen Sinn wurde somit durch das Studium, wie ein gegebenes Individuum einen gegebenen Bereich entwirft, ersetzt.

2.3 Ontologie in der Informatik

Die Ontologie tritt in den letzten Jahren immer verbreiteter im Bereich der Informatik auf. Vor allem kommt sie in den Bereichen Datenbankmodellierung, Wissensbasierte Systeme und Künstliche Intelligenz vor. Aus diesen drei Bereichen werden im folgenden nähere Beispiele erläutert.

Der philosophische Ontologe hat ein Ziel – die Wahrheit über die Realität zu erforschen indem er alles Existierende untersucht. In der Informatik hingegen ist Ontologie eine Software (oder formale Sprache), die in Hinblick auf ein spezielles Set von Anwendungen und Rechenumgebungen entwickelt wird. Wie auch bei der materiellen Ontologie konzentriert sich diese Art der Ontologie auf bestimmte Bereiche, anstatt – wie die formale Ontologie – Aussagen über die gesamte Realität treffen zu wollen. Eine Ontologie ist hier oft etwas, das ein bestimmter Kunde in einem bestimmten Kontext und im Zusammenhang mit bestimmten praktischen Anwendungen und Ressourcen bestellt.

Die Informatik kann aus geschichtlicher Sicht in zwei Strömungen eingeteilt werden - der prozeduralen und der deklarativen Informatik³. In der informatischen Ontologie kann man dieselben Strömungen erkennen. In der Frage, wie man bestmöglich eine intelligente Maschine entwickeln könne, lassen sich diese Unterschiede deutlich aufzeigen. Die Anhänger der prozeduralen Richtung sahen die Lösung im „Wissen wie“. Der Weg zu intelligenten Maschinen liegt ihrer Meinung nach in intelligenten Algorithmen, die komplexe Handlungsweisen

³Prozedural bedeutet ablauforientiert, deklarativ bedeutet ergebnisorientiert.

berechnen und ausführen. Die Befürworter der deklarativen Strömung zielen auf das „Wissen, dass“ ab. Intelligente Maschinen können dieser Anschauung nach nur auf der Basis großer Datenbanken, die das gesamte Wissen enthalten, existieren. Im Bereich des Datenbankmanagements erkannte man schnell, dass sowohl die prozeduralen als auch die deklarativen Elemente essenzielle Repräsentationsformen⁴ sind und für eine intelligente Maschine beide benötigt werden.

Nach der philosophischen Ontologie kann man effektive Repräsentationen nur schaffen, wenn man etwas über die Objekte und Prozesse weiß, die man repräsentieren will. Das heißt, dass man nicht nur über die bestimmten Objekte, die in der Datenbank sind (Kunden, Zahlungen, ...) Bescheid wissen muss, sondern auch über Objekte, deren Eigenschaften und Relationen im Allgemeinen und auch über generelle Arten von Prozessen, in die die Objekte, deren Eigenschaften und Relationen involviert sind. Die Deklarativisten sahen die Antwort darauf in einer robusten Taxonomie aller Arten von Entitäten, die in einem gegebenen Bereich vorkommen. Die Idee war, deklarative Repräsentationen der Standardprozeduren, wie zum Beispiel Geschäftsprozesse oder Scheduling, so zu entwickeln, dass man sie auf verschiedenste Systeme anwenden kann.

2.3.1 Ontologien in Datenbanken- und wissensbasierten Systemen

Ein Beispiel für eine Ontologie im Bereich der Datenbanksysteme ist die ANSI SPARC „Drei Schema Architektur“[42], der alle modernen Datenbanksysteme unterliegen (siehe Abbildung 2.3):

1. Die *Interne Ebene* ist der physischen Speicherung am nächsten. Sie beschreibt physikalische Arten Daten und Objekte im Programmcode zu speichern. Das zugrundeliegende interne Schema enthält nicht nur die Beschreibung von Art und Aufbau der unterschiedlichen gespeicherten Datenstrukturen, sondern auch von speziellen Zugriffsmechanismen und von der Anordnung der Datensätze. Diese Ebene bildet den logischen in den physikalischen Adressraum unter Zuhilfenahme der Betriebssystemfunktionen ab.
2. Die *Konzeptionelle Ebene* beschreibt die logische Struktur der Daten einer Datenbank. Sie

⁴Programme repräsentieren Prozesse, Datenstrukturen sind Repräsentationen von Objekten oder Dingen.

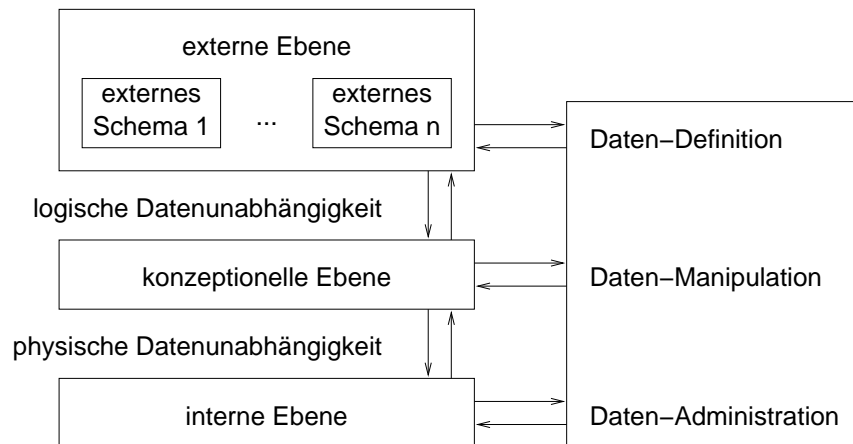


Abbildung 2.3: ANSI/SPARC Drei Schema Architektur eines Datenbanksystems

beschreibt die Daten und deren Relationen und enthält Integrationsbedingungen, Sicherheitsregeln und die Definition von Triggern.

3. Die *Externe Ebene* enthält die Interfaces für die Kommunikation mit dem Benutzer. Jedes einzelne externe Schema beschreibt dabei eine Sicht eines Benutzers oder einer Benutzergruppe und enthält genau jenen Ausschnitt der Datenbank, den der jeweilige Benutzer sehen darf.

Als Veranschaulichung kann man ein Datenbanksystem, wie zum Beispiel PostgreSQL, nehmen. Die Tabellen, die man in der Datenbank ablegt, werden physikalisch im Dateisystem am Server gespeichert (interne Ebene). Der Datenbankprogrammierer bearbeitet die Tabellen in der Datenbank (konzeptionelle Ebene). Ein Web-Interface bietet dem Benutzer die Möglichkeit eingeschränkte Abfragen auf die Datenbank auszuführen (externe Ebene).

Durch die Trennung der Ebenen kommt es zu einer physischen und auch logischen Datenu-nabhängigkeit. Die logische Organisation der Daten ist unabhängig von der konkreten physischen Organisation. Änderungen im logischen Aufbau der Daten haben keine Auswirkungen auf die Anwendungen, die auf diese Daten zugreifen. Mit der „Drei Schema Architektur“ können die Verantwortlichen für die physikalischen Daten, das Management und die Benutzer – jeder auf seiner Ebene – auf dasselbe Objekt zugreifen, ohne sich gegenseitig zu beeinflussen.

Ein weiteres interessantes Beispiel kommt aus dem Bereich der Datenbank- und wissens-basierten Systeme. Verschiedene Gruppen von Entwicklern dieser Systeme haben aus kultu-

rellen und sprachlichen Gründen ihre eigenen Bezeichnungen und Konzepte auf denen sie ihre Frameworks der Informationsrepräsentation aufbauen. Verschiedene Datenbanken haben eventuell gleiche Labels aber mit verschiedene Bedeutungen oder auch verschiedene Bezeichnungen um den selben Kontext auszudrücken. Oft müssen aber verschiedene Gruppen sich eine buntgefächerte Auswahl an Information teilen. Diese Probleme verhindern eine einfache Zusammenfügung verschiedener Informationen. Je mehr Gruppen involviert sind, desto komplizierter wird es.

Schritt für Schritt hat sich die Idee verbreitet, dass man mit einer gemeinsamen Referenz-Taxonomie diese Probleme beseitigen könnte. Die Erzeugung einer solchen Referenz-Taxonomie wurde dann von den Informatikern als Ontologie bezeichnet. Eine Ontologie in diesem Zusammenhang ist ein Wörterbuch von Bezeichnungen als Basis für eine Wissensrepräsentation. Die Vorteile einer solchen Ontologie sind eindeutig: jede Gruppe von Entwicklern muss seine Bezeichnungen nur auf dieses Wörterbuch mappen, um eine gemeinsame Wissensrepräsentation zu bilden. Eine ontologische Grundlage dieser Enzyklopädie würde aus zwei Teilen bestehen:

1. einer terminologischen Komponente, die den Zusammenhang zwischen einzelnen Bezeichnungen berechnen kann und
2. einer Aussagenkomponente, die alles andere darstellt.

Die bevorzugten Methoden für die Erzeugung solcher Ontologien haben sich einerseits aus frühen Initiativen oben genannter Datenbankmanagementsysteme, andererseits aus der logischen und analytischen Philosophie entwickelt. Sie inkludieren das Erzeugen von Ontologien aus existierenden Taxonomien, Datenbanken und Wörterbüchern, so wie die Erzeugung von Ontologien aus einem linguistischen Korpus basierend auf Systemen wie WordNet[14].

2.3.2 Ontologie in der Künstlichen Intelligenz

John McCarthy (geb. 1927) hat „Circumscription“, eine nonmonotone Logik⁵, entwickelt um Annahmen, die von Menschen mit Allgemeinwissen und „Hausverstand“ getroffen werden,

⁵Nonmonotone Logik kann man sich als eine Formalisierung von rationalen Schlüssen vorstellen. Solche Schlüsse sind dadurch charakterisiert, dass Aussagen hergeleitet werden, die zwar bezüglich des momentanen Wissensstandes plausibel erscheinen, aber bei Erhalt genauere Informationen sich als falsch herausstellen können.[51]

zu formalisieren:

Using circumscription requires that common sense knowledge be expressed in a form that says a boat can be used to cross rivers unless there is something that prevents its use. In particular, it looks like we must introduce into our ontology (the things that exist) a category that includes something wrong with a boat or a category that includes something that may prevent its use. Incidentally, once we have decided to admit something wrong with the boat, we are inclined to admit a lack of oars as such a something and to ask questions like, „Is a lack of oars all that is wrong with the boat?“.

Some philosophers and scientists may be reluctant to introduce such things, but since ordinary language allows „something wrong with the boat“ we shouldn't be hasty in excluding it. Making a suitable formalism is likely to be technically difficult as well as philosophically problematical, but we must try.

We challenge anyone who thinks he can avoid such entities to express in his favorite formalism, „Besides leakiness, there is something else wrong with the boat“. A good solution would avoid counterfactuals as this one does. [35]

Dieses Zitat macht deutlich, wie schwierig es ist eine vollständige Ontologie zu schaffen, selbst wenn sie nur einen kleinen Teilbereich betrifft. Man weiß erst, wie man eine Ontologie erzeugt, wenn man alle Entitäten festgelegt hat, die in den Bereich dieser Ontologie fallen. Nicht-monotones Schließen öffnet sich eben der Vorstellung, dass man nicht alle Entitäten a-priori kennt.

2.3.3 Weitere Anwendungen in der Informatik

Ontologien sind in noch in vielen weiteren Bereichen der Informatik von Nutzen. Einige Beispiele sind:

- Automatisierung: Automatisierung benötigt einen hohen Grad an Genauigkeit bei der Beschreibung ihrer Prozeduren und Ontologie ist ein Mechanismus mit dem deren Ziele erreicht werden können

- Konzeptuelle Schemata im Datenbank Design
- Modellen von Applikationsbereichen im Software Engineering
- Klassenmodelle im objektorientierten Software Design
- Formalisierung von Standards auf dem Level von Metadaten
- Schreiben von Software: Wenn man sich die Mühe macht eine Ontologie zu konstruieren um existierende Systeme einzubinden, kann diese Ontologie auch als Basis für das Schreiben von Software, die diese alten Systeme ersetzt, verwendet werden.
- Informationsextraktion aus großen Bibliotheken
- Semantic Web
- Parsen und Disambiguierung bei der Übersetzung natürlicher Sprache

2.4 Zusammenfassung

Die Probleme die in der informatischen Ontologie auftreten sind denen der philosophischen Ontologie sehr ähnlich. Eine allgemeine Ontologie zu erstellen würde ein neutrales und gemeinsames Framework voraussetzen, das eine einzige allgemeingültige Liste von Kategorien enthält. Um aber allgemein anerkannt zu sein, müsste die Ontologie neutral bezüglich aller verschiedenen Gruppen von Entwicklern sein. Eine solche Neutralität kann aber die Voraussetzung eines großen und allumfassenden Vokabulars nicht erfüllen.

Eine Möglichkeit diese Probleme zu lösen ist die Aufteilung der Aufgaben der Ontologie in zwei Teile. Einerseits die formale Ontologie, die das ganze System als Allgemeines mit all ihren Abhängigkeiten beschreibt, und andererseits einzelne bereichsspezifische oder regionale Ontologien, wie zum Beispiel Ontologien in der Geographie, Medizin oder Ökologie. Das Verhältnis zwischen formaler und bereichsspezifischen Ontologien wäre ähnlich dem Verhältnis zwischen reiner und angewandter Mathematik, damit alle bereichsspezifischen Ontologien als Basis die gleiche robuste und allgemein akzeptierte oberste Ontologie haben.

Kapitel 3

Geschichte und Theorie der Folksonomien

3.1 Grundlagen

Eine Folksonomy ist ein soziales Netzwerk, dass sich aus der Tätigkeit des *Free Taggings* ergibt. *Tagging* ist das Benennen oder Beschlagworten von Objekten, wie zum Beispiel Bookmarks, Bildern oder Zitaten, mit – im Fall von Free Tagging frei gewählten – Wörtern, die die Objekte umschreiben. Diese Wörter werden Tags genannt. Der Benutzer sucht sich selbst Wörter als Tags aus, die für ihn persönlich die Objekte bestmöglich beschreiben, sodass er die Objekte anhand dieser Tags wiederfinden kann.

Eine Folksonomy ergibt sich, wenn man viele dieser Benutzer in ein System integriert und deren Objekte und Tags für alle zugänglich macht. Die Kombination aus Benutzern, Objekten und Tags erzeugt ein soziales Netzwerk. Es zeigt sich, dass sich die Benutzer gegenseitig beeinflussen. Sie taggen nicht nur für sich persönlich, sondern auch für andere Benutzer dieser Folksonomy. Sie wählen oft allgemeingültige und allgemein verständliche Wörter. Aus den Tags kann man eine Struktur der getaggten Objekte oder der Tags ableiten. Diese Struktur ist aber nicht mit einer Ontologie gleichzusetzen. Sie ist nicht hierarchisch gegliedert und sehr dynamisch, da sich die Tags mit der Zeit ändern.

Die Grundlage einer Folksonomy sind also Tags, eine strukturierte Art von Metadaten¹.

¹Metadaten werden oft als „Daten über Daten“ bezeichnet. Metadaten sind Informationen über Inhalte von

Der Vorteil von Tags gegenüber traditionellen Metadaten ist, dass sie auf einzelne Wörter beschränkt sind und somit schneller vom menschlichen Gehirn gefiltert werden können. Lange Reviews oder Kommentare zu Webseiten, Büchern oder Filmen, sind inhaltlich natürlich informativer und dürfen nicht außer Acht gelassen werden. Einzelne Worte, die diese Inhalte umreißen, können aber schneller und einfacher erfasst werden. Zum Beispiel ist eine kurze Inhaltsangabe eines Buches sehr wertvoll, wenn man schon einen engeren Kreis von Büchern zur Auswahl hat, aber um sich aus der Menge aller Bücher die für sich persönlich interessanten auszusuchen, ist es hilfreicher, wenn man zuerst nach Kategorien, zum Beispiel bekannten Genres, wie „Horror“, „Krimi“, „Komödie“ und ähnlichem wählen kann.

Nun fallen aber nicht alle Werke in genau eine dieser Kategorien. Deshalb gibt es auch Zwischenkategorien, wie „Krimikomödie“ oder „Liebesdrama“. Wenn ein Buch aber in mehr als zwei Kategorien fällt, ist es mit der genauen Einteilung zu Ende. Begriffe wie „Liebeskrimikomödie“ werden schon sehr unhandlich. Mit Tags hingegen besteht die Möglichkeit, einem Objekt beliebig viele Attribute zuzuordnen. Dem letztgenannten Beispiel könnte man ohne weiteres die Tags „Romanze“, „Krimi“ und „Komödie“ zuordnen um das Buch so genauer zu charakterisieren. Tags beschränken sich aber nicht nur auf die Genres, gleichzeitig könnte man dem Buch noch einen Autorennamen oder etwa den Ort, wo die Geschichte spielt, zuordnen. Das bietet anderen Benutzern die Möglichkeit zum Beispiel nach einer Krimikomödie, deren Handlung in Wien läuft, zu suchen.

3.2 Entwicklung

Die ersten Ideen für Folksonomien entstanden in der Konferenz K-CAP 2001[10]. In dem Workshop „Knowledge Markup and Semantic Annotation“ wurde die Idee vorgestellt, Bookmarks zu annotieren² um einen Gewinn aus den zusätzlichen Metadaten zu ziehen[32].

In diesem Workshop wurden verschiedene Szenarien ausgearbeitet, bei denen Annotation einen zentralen Gesichtspunkt darstellt:

Dokumenten, Büchern, Artikeln, Fotografien oder anderen Medien und sollen die Organisation und den Zugriff auf diese Objekte vereinfachen. Die Zusammenfassung oder Vorschau eines Buches auf dessen Rückseite oder das Genre eines Films sind Beispiele für Metadaten.

²Annotation (*L. annotatio: cf. F. annotation.*) A note, added by way of comment, or explanation; – usually in the plural; as, annotations on ancient authors, or on a word or a passage. [38]

Szenario 1: Verwendung von Annotationen zur Zusammenarbeit: Studenten einer Universität arbeiten über das Web zusammen. Sie sammeln URLs auf einer gemeinsamen Webseite und annotieren diese, damit sie leichter wieder gefunden werden können.

Szenario 2: Verwendung von Annotationen für gemeinsame Bookmarks: Die Studentengruppe verwendet eine Suchmaschine um weitere Webressourcen zu finden und erzeugen Bookmarks auf einem gemeinsamen Annotationserver. Sie fügen den Bookmarks eine Kategorie aus einer vordefinierten Liste von Kategorien hinzu und erzeugen frei gewählte Annotationen.

Szenario 3: Verwendung von Annotationen um Auswertungsergebnisse zu präsentieren: Ein Annotationserver wird mit einem Wörterbuch versehen, das von einer Computersprache in Annotationen übersetzt.

Mit diesen Szenarien wurde das Projekt Annotea[3] geboren, ein System, das Metadaten zu Bookmarks oder Dokumenten auf einem externen Server oder auch lokal abspeichert. Annotea besteht aus drei Teilen:

1. *Annotea Server:* Auf dem Annotea Server – oder auch lokal auf dem Rechner des Benutzers – werden die Annotea Objekte gespeichert.
2. *Annotea Objekte:* Annotea Objekte bestehen aus:
 - einer URL, die angibt, wo die Objekte am Annotea Server oder auf der Festplatte liegen,
 - Metadaten zum Objekt im RDF (Resource Description Format),
 - einer Referenz zu einer Webressource oder einer anderen Annotation,
 - einer Beschreibung
 - und einem Erzeuger.
3. *User Interface:* Das User Interface ist ein Webbrowser, der Annotea Metaphern unterstützt. Die Annotea Objekte, werden als XML-Seiten dargestellt.

Es gibt zwei Arten von Annotea Objekten – Annotationen und Bookmarks. Während Annotationen auf lokale Dokumente oder bestimmte Stellen in diesen Objekten zeigen, verlinken



Abbildung 3.1: Annotationen und Bookmarks in Annotea [31]

Bookmarks Web-URLs. Die Unterschiede zwischen Annotationen und Bookmarks werden in Abbildung 3.1 veranschaulicht. Annotationen sind also Texte, die man zu beliebigen Stellen in Dokumenten hinzufügt, Bookmarks sind einfach nur Merker, die auf Dokumente oder Stellen in Dokumenten zeigen.

Annotea bietet Mechanismen, die jedem Benutzer die Möglichkeit geben, einen speziellen auf ihn angepassten Ausschnitt der Objekte zu sehen, je nachdem, welchen kollaborativen Gruppen er angehört. Annotea Benutzer entscheiden selbst, ob sie Annotationen sehen wollen und von welchen Annotea Servern sie diese Informationen beziehen wollen. Zusätzlich erlaubt Annotea detailliertes Filtern von Annotationen um zum Beispiel nur die Annotationen eines bestimmten Benutzers anzuzeigen.

Um Annotea zu implementieren wurde Semantic Web Technologie³ verwendet. Das Semantic Web ist sehr gut geeignet um Ressourcen Informationen zuzuordnen und bietet Kontrolle und Flexibilität für die Benutzer. Annotea implementiert verteilte Bookmarks in Kom-

³Semantic Web ist ein Projekt des W3C Konsortiums (<http://www.w3c.org>), das das Ziel hat, das Web verständlicher für Maschinen zu machen, indem maschinenlesbare semantische Metadaten zu Dokumenten hinzugefügt werden.

bination mit *Annotea Themen*. Diese erlauben es dem Benutzer seine Klassifikationen oder informellen Kategorien zu erzeugen und zu verwalten. Ein Bookmark kann unter einem oder mehreren Themen katalogisiert und in einer Themen-Hierarchie dargestellt werden. Man kann in Annotea nach Themen suchen und der Benutzer kann beliebige Themen abonnieren. Die Themen sind frei wählbar, sie können also auch Konzepte außerhalb der konventionellen Kategorien definieren. Diese Annotea Themen sind den Tags aus Folksonomien dadurch sehr ähnlich.

In Annotea liegen zwar die Wurzeln von Folksonomien, von der Idee der ontologischen Kategorisierung ist Annotea aber nicht vollständig abgekommen:

The Annotea topics can support early phases of innovations and research by letting the topics to be as vague as needed at the beginning and let the user to refine them as more learning happens or link them to concepts in well established ontologies when those are discovered and understood. [31]

In Annotea kann man aber schon erste Ansätze zu Folksonomien erkennen. Die Benutzer können freie Themen – äquivalent zu Tags in einer Folksonomy – zu den Bookmarks oder Annotationen hinzufügen, diese Abonnieren, nach ihnen auf dem Annotea Server oder auch lokal suchen.

Die Idee des Free Taggings wurde in der Welt der Weblogs immer stärker. Seiten wie Technorati[13] bieten die Möglichkeit seinen persönlichen Weblogs Kategorien hinzuzufügen. Damit ist Technorati einer der ersten Vertreter von Folksonomien.

Das Wort „Folksonomy“ wurde von Thomas Vander Wal geprägt und in seinem Weblog[53] 2004 erstmals erwähnt.

3.2.1 Definition nach Vander Wal

Thomas Vander Wal definiert Folksonomy folgendermaßen:

Folksonomy is the result of personal free tagging of information and objects (anything with a URL) for one's own retrieval. The tagging is done in a social environment (shared and open to others). The act of tagging is done by the person consuming the information.

The value in this external tagging is derived from people using their own vocabulary and adding explicit meaning, which may come from inferred understanding of the information/object as well as. The people are not so much categorizing as providing a means to connect items and to provide their meaning in their own understanding. [53]

Folksonomy ist also das Ergebnis einer Tätigkeit, dem Free Tagging. Das Free Tagging ist das freie Zuordnen von Begriffen zu einem Objekt oder einer Information. Sie bildet somit eine Art Strukturierung dieser Objekte im Zusammenhang mit den Tags und Benutzern, ist aber fern jeglicher Ontologie oder Taxonomie. Eine Taxonomie ist eine Liste von kontrolliertem Vokabular, die hierarchisch strukturiert ist. Das widerspricht in jedem Punkt der Definition einer Folksonomy:

- Das Vokabular einer Folksonomy ist nicht kontrolliert, sondern entsteht aus frei gewählten Begriffen der Benutzer dieser Folksonomy.
- Eine Folksonomy ist nicht hierarchisch strukturiert, es gibt keinerlei Eltern-Kind Beziehungen, da keines der Tags einem anderen übergeordnet ist. Der einzige Zusammenhang, der zwischen den Tags besteht, sind die damit umschriebenen Objekte oder die Benutzer, von denen sie verwendet werden.

3.2.2 Mathematische Definition

Eine Folksonomy kann auch als tripartiter Hypergraph⁴ dargestellt werden.[39] Es gibt in einer Folksonomy aus graphentheoretischer Sicht drei Mengen von Knoten: die Menge der Aktoren (der Benutzer) A , die Menge der Konzepte (Tags) C und die Menge der Instanzen (Objekte) I . Wenn man die drei Mengen kombiniert, erhält man Dreifachkanten, die die Menge der Annotationen T beschreiben. Eine Instanz dieser Menge ist eine Kombination aus genau einem Tag, das von einem Benutzer einer Webseite zugeordnet wurde.

⁴tripartit: A graph is k -partite if its vertices can be partitioned into k sets (called *partite sets*) in such a way that no edge joins two vertices of the same set. [25]

Hypergraph: A hypergraph G can be defined as a pair (V, E) , where V is a set of vertices, and E is a set of hyperedges between the vertices. Each hyperedge is a set of vertices: $E = u, v, \dots \in 2^V$. (Hyperedges are undirected.) [17]

$$A = \{a_1, \dots, a_k\}$$

$$C = \{c_1, \dots, c_l\}$$

$$I = \{i_1, \dots, i_m\}$$

$$T \subseteq A \times C \times I$$

Aus diesen Dreifachkanten kann man einen tripartiten Hypergraphen $H(T)$ erstellen, der eine Folksonomy repräsentiert:

$$H(T) = (V, E)$$

$$V = A \cup C \cup I$$

$$E = \{\{a, c, i\} \mid (a, c, i) \in T\}$$

Der tripartite Hypergraph $H(T)$ besteht aus Knoten V und Kanten E . Die Knoten V sind eine Vereinigung der oben genannten Knotenmengen A , C und I . Der Graph beinhaltet also die Benutzer, die Tags und die Objekte. Die Kanten E inkludieren alle Dreifachkanten der Annotationsmenge T .

Diese mathematische Form der Folksonomy kann in weiterer Folge zum Beispiel für Fallstudien zur Extraktion von Ontologien aus Folksonomien verwendet werden (siehe [39]).

3.2.3 Arten von Folksonomien

Man kann im Wesentlichen zwei Arten von Folksonomien unterscheiden – breite und enge Folksonomien. [53]

In einer breiten Folksonomy beschreiben viele Personen das selbe Objekt mit ihren eigenen Tags aus ihrem persönlichen Vokabular. (Siehe Abbildung 3.2.) Eine Person (*Content Creator*) erzeugt ein Objekt und macht es so anderen Personen zugänglich. Verschiedene Benutzergruppen der Folksonomien (*Gruppen A-F*) ordnen dann dem Objekt Tags zu (*Tags 1-5*) und nutzen diese Tags auch.

In einer engen Folksonomy vergibt nur eine kleine Gruppe von Personen bzw. eine einzelne Person Tags an ein Objekt. (Siehe Abbildung 3.3.) Die Person, die das Objekt erzeugt (*Content*

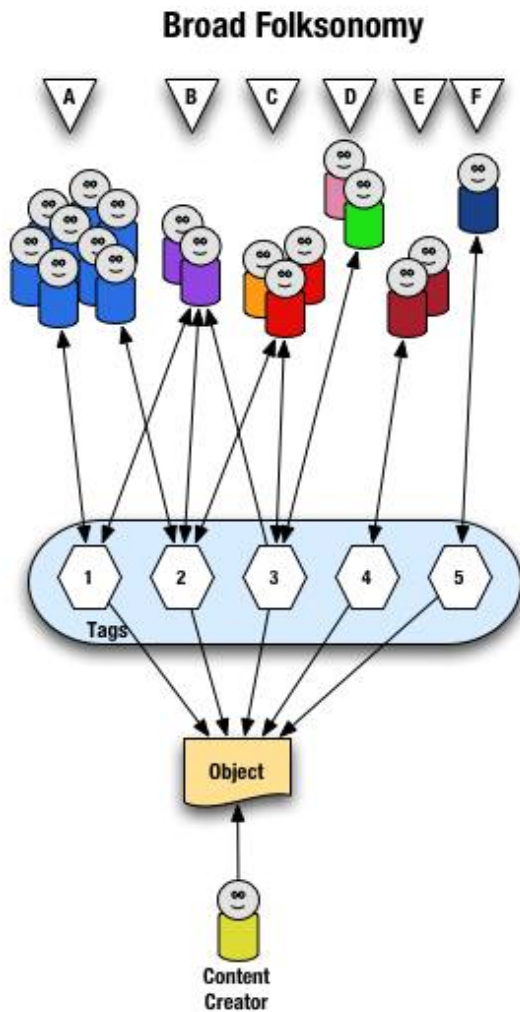


Abbildung 3.2: Breite Folksonomy [53]

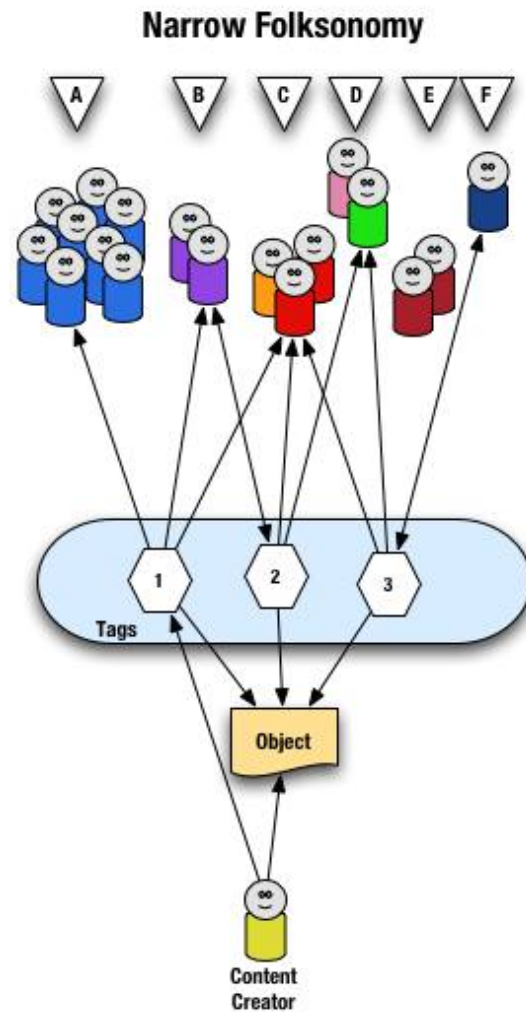


Abbildung 3.3: Enge Folksonomy [53]

Creator) ordnet dem Objekt auch gleich ein oder mehrere Tags zu. Nur wenige oder gar keine der Benutzer der Folksonomy (in der Abbildung sind dies die Benutzergruppen B und F) ordnen diesem Objekt weitere Tags zu, aber alle Benutzer können diese Tags für sich verwenden.

In dieser engen Folksonomy sind die Tags direkt mit dem Objekt verknüpft. Man weiß nicht, wie die Mehrheit der Personen dieses Objekt verwenden oder wie sie es benennen würden. Sehr gut anwendbar ist enge Folksonomy auf Objekte, die nicht so leicht auffindbar sind, da sie nicht mit einfacher Textsuche gefunden werden können (z.B. Bilder, Musik,...). Hier kann der Erzeuger dieser Objekte, diesen Objekten sofort Tags zuordnen. So kann er zum Beispiel

ein Bild von einer Winterlandschaft mit „Winter“, „Schnee“, „weiß“ oder auch dem Namen der Gegend in der das Foto gemacht wurde oder anderen Begriffen seiner Wahl belegen.

3.3 Power Laws

When the probability of measuring a particular value of some quantity varies inversely as a power of that value, the quantity is said to follow a power law, also known variously as Zipf's law or the Pareto distribution. Power laws appear widely in physics, biology, earth and planetary sciences, economics and finance, computer science, demography and the social sciences. [41]

Viele der von Wissenschaftlern gemessenen Dinge haben eine typische Größe oder Skalierung, also einen typischen Wert auf den die meisten Einzelmessungen zentriert sind. Der Großteil der erwachsenen Menschen zum Beispiel sind um die 180cm groß und man wird niemanden finden, der nur 10cm oder aber 500cm groß ist. Es gibt aber nicht nur solche Dinge, die eine sehr geringe Varianz haben, sondern manche verteilen sich über eine enorme dynamische Bandbreite. Ein klassisches Beispiel hierfür ist die Population von Städten. Während zum Beispiel in den USA die größte Stadt New York acht Millionen Einwohner hat, hat die kleinste Stadt Duffield, Virginia mit nur 52 Einwohnern.⁵

Wenn man die beiden Histogramme aus Abbildung 3.4 und 3.5 vergleicht, fällt einem sofort ein Unterschied auf: Während Ersteres einen schmalen Peak mit sehr kleiner Varianz aufweist, ist das Städte Histogramm stark nach rechts hin verzerrt. Es gibt nur ein bis zwei höhere Werte und die restlichen Größen weichen kaum voneinander ab.

Ein überraschendes Ergebnis erhält man, wenn man das zweite Diagramm mit logarithmischen Achsen zeichnet. Dann erhält man, wie in Abbildung 3.6 eine annähernd gerade Linie von der größten zur kleinsten Stadt. Diese Beobachtung wurde zuerst von Felix Auerbach in „Das Gesetz der Bevölkerungskonzentration“ (1913) und später von George Kingsley Zipf in „Human Behaviour and the Principle of Least Effort“ (1949) gemacht. Verteilungen dieser Art folgen einem Power Law (Potenzgesetz). [41]

Power Law Verteilungen gibt es in erstaunlich vielen Bereichen. Neben Einwohnerzahlen

⁵Daten aus dem Jahr 2000 aus [41]

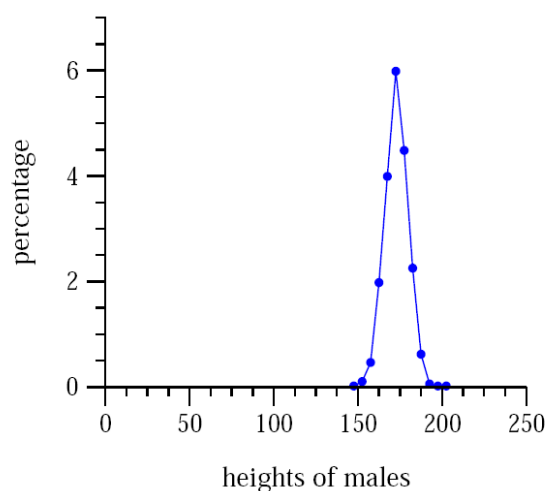


Abbildung 3.4: Histogramm: Körpergrößen von Menschen – Gauss Verteilung [41]

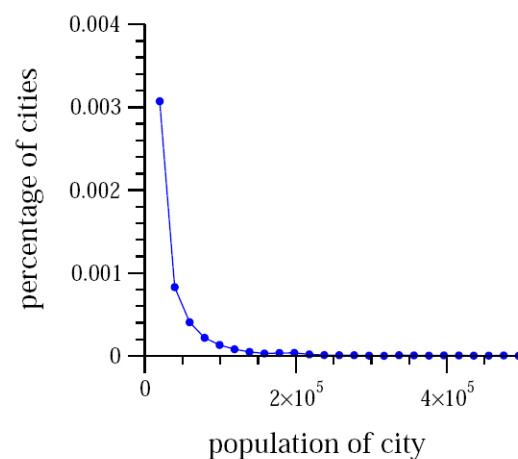


Abbildung 3.5: Histogramm: Populationsgrößen von Städten – Power Law Verteilung [41]

von Städten findet man sie auch noch bei Größen von Erdbeben, Mondkratern, Sonnenflecken, Computerdateien und Kriegen, der Häufigkeit der Verwendung von Wörter in der menschlichen Sprache, der Häufigkeit von Vornamen in den meisten Kulturen, die Anzahl von Papers die von Wissenschaftern geschrieben werden, die Anzahl von Zitaten in Papers, die Anzahl an Besuchen auf Webseiten, der Buchverkauf, Musikaufnahmen und vieles mehr.

In sozialen Netzwerken, wie Weblogs oder auch Folksonomien, bildet sich immer automatisch eine Elite heraus, eine kleine Gruppe von Personen, auf deren Seiten die Mehrheit des Traffics stattfindet. Wenn ein neues soziales Netzwerk begonnen wird ist es noch frei von Elitismus und Gruppenbildung. Wenn dieses System wächst kommen die Probleme der Skalierung hinzu. Nicht jeder kann sich an jeder Konversation beteiligen, manche Gruppen werden immer populärer und bilden schließlich eine Elite. Die Verteilung der Seiten nähert sich einer Power Law Kurve an. Clay Shirky fasst dieses Phänomen in seinem Artikel „Power Laws, Weblogs, and Inequality“[46]:

Diversity plus freedom of choice creates inequality, and the greater the diversity, the more extreme the inequality. [46]

Auch die Verteilung von Tags entspricht von verschiedenen Gesichtspunkten aus gesehen einer Power Law Kurve. Egal ob man

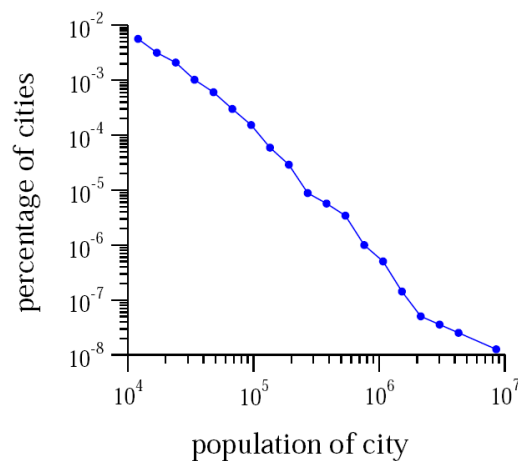


Abbildung 3.6: Populationsgrößen von Städten mit logarithmischen Achsen [41]

- die Anzahl der Tags pro Benutzer,
- die Anzahl der Vorkommnisse verschiedener Tags eines einzelnen Benutzers oder
- die Anzahl der Vorkommnisse verschiedener Tags gemeinsam mit einer einzelnen URL

betrachtet, es ergibt sich immer wieder eine Power Law Kurve.[45]

In Abbildung 3.7 sieht man die Power Law Kurve, die sich aus der Vorkommnisse von Tags pro URL ergibt. Man sieht dass einige wenige URLs mit vielen Tags getaggt sind. Die ersten vier URLs haben noch über 50.000 Tags, alle anderen sind darunter und nähern sich Null an.

Ein anderes Beispiel ist Technorati. In Abbildung 3.8 sieht man die 100 meistverlinkten Weblogs. Die ersten acht Weblogs sind zwischen 1000 und 6000 mal verlinkt, während für die restlichen 92 Weblogs nur noch sehr wenige Links existieren, die auf sie zeigen.

Diese ungleiche Verteilung ist indirekt eine Folge der Entscheidungsfreiheit die in einer Folksonomy gegeben ist. Benutzer, die die Möglichkeit haben absolut frei zu wählen, welche Seiten, Weblogs oder Fotos ihnen am besten gefallen, suchen sich Vorbilder. Sie sehen sich zuerst Seiten an, die schon beliebt sind oder folgen Empfehlungen von Freunden. Die Entscheidungen der Personen beeinflussen sich untereinander. Dadurch ergibt sich eine breite Masse an Seiten die von einer kleinen Gruppe von Benutzern gelesen werden, die einen persönlichen Bezug zu genau diesen Seiten haben. Wenn eine dieser Seiten aber um eine kleine Prozentzahl

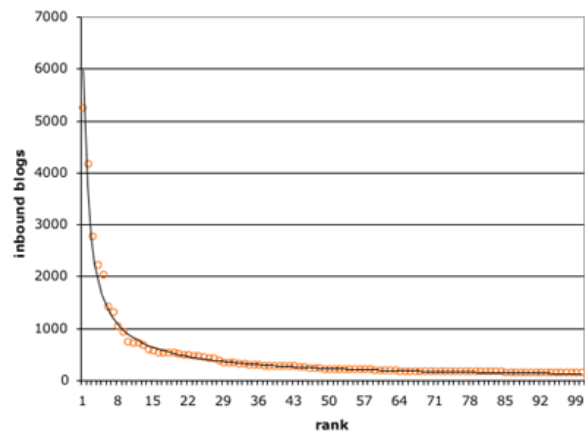
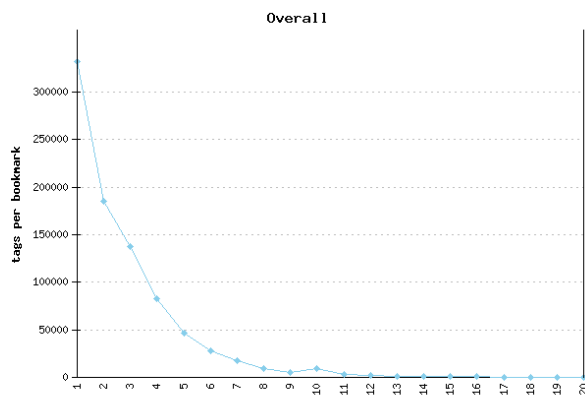


Abbildung 3.7: Verteilung von Tags pro URL in Del.icio.us [30] Abbildung 3.8: Die 100 meistverlinkten Weblogs auf Technorati [33]

an Beliebtheit gewinnt, werden alle anderen Benutzer auch darauf aufmerksam und die Seite hat eine wesentlich höhere Wahrscheinlichkeit ausgewählt zu werden, vor allem von Benutzern, die neu hinzukommen.

Inequality occurs in large and unconstrained social systems for the same reasons stop-and-go traffic occurs on busy roads, not because it is anyone's goal, but because it is a reliable property that emerges from the normal functioning of the system. The relatively egalitarian distribution of readers in the early years had nothing to do with the nature of weblogs or webloggers. There just weren't enough blogs to have really unequal distributions. Now there are. [46]

Der Grund für diese Ungleichheit liegt also nicht in dem persönlichen Ziel einer Person besonders beliebt oder berühmt zu werden, sondern aus der natürlichen Tendenz heraus, dass persönliche Entscheidungen leicht von anderen Personen beeinflusst werden.

3.4 Zusammenfassung

Die Folksonomy ist eine neue Entwicklung die ihre Wurzeln im Jahre 2001 hat, als das Projekt Annotea geschaffen wurde. Von diesem Zeitpunkt an hat sich diese Idee weiterentwickelt und wurde in immer mehr Webservices angewandt. Im Jahr 2004 hat Thomas Vander Wal dann das

Wort Folksonomy erfunden. Dieses hat sich wie ein Lauffeuer verbreitet und wird seither für alle diese Systeme verwendet.

Dadurch, dass ein Folksonomy-System den Benutzern vollständige Entscheidungsfreiheit bietet, bildet sich in diesen Systemen eine Verteilung gemäß dem Power Law heraus.

Kapitel 4

Folksonomy vs. Ontologie

Traditionell werden Metadaten von Profis erzeugt und zugeordnet. Gelernte Katalogisierer erzeugen diese etwa für Bücher in Bibliotheken nach genau erarbeiteten Normen und Entwürfen.

Während professionell erzeugte Metadaten oft für hohe Qualität stehen, haben sie dennoch einen schwerwiegenden Nachteil – sie sind sehr teuer, was Zeit- und Arbeitsaufwand betrifft, vor allem bei großen Mengen an Daten, wie sie zum Beispiel im Web vorhanden sind.

Eine Alternative dazu bieten vom Autor selbst erzeugte Metadaten. Das verbessert zwar die Skalierbarkeit, aber bezieht (wie auch der professionelle Ansatz) die Benutzer nicht mit ein.

Daher gibt es einen dritten Ansatz: vom Benutzer erzeugte Metadaten. Diese sind billig und einfach zu erzeugen und auch skalieren auch gut. Eine Form, der explizit vom Benutzer erstellten Metadaten sind Weblogs. Diese Seiten beinhalten oft Links kombiniert mit Kommentaren, sind aber unstrukturiert. Benutzer-Reviews auf Webseiten wie zum Beispiel Amazon¹ stellen eine weitere Art dieser Metadaten dar.

In der Literatur wird oft der Fehler gemacht Folksonomy mit Taxonomie gleichzusetzen. Eine bekannt Erklärung des Wortstamms Folksonomy ist, dass es aus den Worten „Folk“ und „Taxonomie“ zusammengesetzt ist (vergl. [37], [34]). Laut Thomas Vander Wal ist die Abstammung in den Worten „folk“ und dem griechischen Wortteil „nomos“ oder „nomia“ begründet, das mit „Management“ übersetzt werden kann. Somit bedeutet Folksonomy in etwa „von Menschen gemachtes Management“ und hat kaum etwas mit Klassifikationen oder Ontologien

¹<http://www.amazon.com>

zu tun.

Vander Wal meint dazu in seinem Weblog:

Today, having seen an new academic endeavor related to folksonomy quoting the Wikipedia entry on folksonomy and I realize the definition of Folksonomy has become completely unglued from anything I recognize (yes, I did create the word to define something that was undefined prior). It is not collaborative, it is not putting things in to categories, it is not related to taxonomy (more like the antithesis of a taxonomy), etc. [53]

4.1 Vor- und Nachteile von Folksonomien

Folksonomien haben im Vergleich mit Ontologien viele Vorteile, die vor allem im Zusammenhang mit der Freiheit, die Folksonomy bietet, stehen. Gleichzeitig wird diese Freiheit aber auch von den Kritikern von Folksonomy als Nachteil empfunden.[45] [44]

Vorteile von Folksonomien:

- *Entdecken (Finding) statt Blättern (Browsing)*: Eine Folksonomy ermutigt den Benutzer dazu, anstatt einfach eine Suchanfrage zu starten, durch die Objekte zu stöbern. Das ermöglicht das Auffinden von Inhalten, die dem Benutzer sonst verborgen geblieben wären.
- *Zufall*: Der Zufall spielt in Folksonomien eine wichtige Rolle. Die Benutzer einer Folksonomy werden dazu angeregt, von ihren Tags oder Bookmarks ausgehend nach weiteren Benutzern zu suchen, die ähnliche Interessen haben. So ist es mit einem Klick möglich, Objekte von anderen Benutzern zu finden, die man bei einer einfachen Suche in einer Suchmaschine nie gefunden hätte.
- *Desire Lines*²: Der wichtigste Vorteil von Folksonomien ist, dass sie das Vokabular der Anwender widerspiegelt, das sich über die Zeit hinweg verändert. Sie reflektiert die Terminologie, Genauigkeit und Wortwahl der Benutzer.

²Als Desire Lines werden die Trampelpfade bezeichnet, die oft auf öffentlichen Grünflächen abseits der gepflasterten Wege zu finden sind und die optimalen Wege in diesem Park darstellen.

- *wenig Aufwand*: Folksonomien sind sehr einfach was Zeit- und Arbeitsaufwand betrifft. Man kann Tags sofort beim Erzeugen des Objekts oder Inhalts hinzufügen.
- *keine Einstiegshürden*: Folksonomien ermöglichen – aufgrund des einfachen Erzeugens der Tags – den Benutzern (und nicht nur Profis) ohne Training und Vorwissen sofort an dem System teilzuhaben.
- *Skalierbarkeit*: Folksonomien skalieren beinahe ohne zusätzlichen Aufwand auf großen Mengen von Daten, so wie sie im Web vorkommen. Sie können einfach erweitert und an sich ändernde Voraussetzungen angepasst werden.
- *Feedback und asymmetrische Kommunikation*: Der Benutzer bekommt sofortiges Feedback. Sobald er einen Tag hinzufügt, sieht er weitere Objekte, die dieselben Tags besitzen. Wenn das nicht das gewünschte Ergebnis erzielt kann er sofort die Tags wieder ändern. Diese enge Feedback Schleife führt zu einer Form asymmetrischer Kommunikation zwischen den Benutzern über das Medium der Metadaten.
- *Entstehung von Communities*: Als Nebeneffekt fördert das Tagging auch die Entstehung von Communities. Personen mit gleichen Interessen verwenden gleiche Schlüsselwörter.
- *Allumfassend*: Folksonomien enthalten den Wortschatz von jedem einzelnen Anwender, ohne etwas auszulassen. Es gibt keine Autorität, die ein Vokabular vorschreibt. Damit können auch originelle Ideen, die sich außerhalb des Mainstreams befinden, Fuß fassen.
- *„Einzigste Antwort auf das Web“*: Folksonomien können auch als einzige sinnvolle Einteilung des Webs gesehen werden. Das Web selbst ist von den Anwendern geschaffen, flexibel und dynamisch. Eine strenge Hierarchie kann mit diesen Forderungen nicht mithalten.

Nachteile von Folksonomien:

- *Ambiguität*: Verschiedene Benutzer umschreiben gleiche Webseiten auf verschiedene Art und verwenden gleiche Begriffe auf unterschiedliche Weise.
- *Leerzeichen, mehrteilige Wörter*: Viele bekannte Folksonomy Systeme erlauben keine Leerzeichen in Tags, daher können zusammenhängende Begriffe nur dargestellt werden, in-

dem sie zusammen geschrieben oder mit Trennzeichen (Unterstrich, Bindestrich, Plus...) verbunden werden.

- *Synonyme*: Es gibt in Folksonomy Systemen keine strukturierte Synonymkontrolle. Oft werden verschiedene Wörter für denselben Begriff verwendet. Viele Objekte werden auch gleichermaßen in Mehrzahl und Einzahl oder auch in verschiedenen Sprachen beschrieben.
- *Flache Anordnung*: Folksonomien enthalten keinerlei hierarchische Strukturen.
- *Geringe Auffindbarkeit*: Folksonomien sind nicht zum Suchen geeignet. Obwohl sie gezielt die Möglichkeit zum Suchen nach Tags und Benutzern bieten, eignen sich Folksonomien viel eher zum Browsen und Stöbern, als zum Suchen. Um gezielt nach Seiten zu suchen, müsste man zuerst wissen, wie die anderen Benutzer diese Seiten getaggt haben.

Wenn auch die Vorteile überwiegen, kann man die Nachteile dennoch als Anregung sehen, das Tagging noch weiter zu verbessern. Folksonomy Systeme könnten zum Beispiel Synonyme automatisiert finden und sie dem Benutzer als Alternative vorschlagen. Dies darf auf keinen Fall verpflichtend geschehen, da das gegen die Idee einer Folksonomy laufen würde. Jede Art von Einschränkung der Folksonomy bringt einen Informationsverlust mit sich. Als Beispiel kann man hier die Synonymkontrolle anführen: oft haben Synonyme nicht eine vollkommen gleiche Bedeutung. Mit einer Synonymkontrolle, die dem Benutzer die Möglichkeit nimmt, seine eigenen Wörter zu verwenden, werden semantische Nuancen in der Bedeutung vernachlässigt.

Folksonomien werden wegen der vielen Vorteile und aufgrund der daraus resultierenden sozialen Aspekte, immer beliebter. Folksonomien bieten die Möglichkeit einer personalisierten Suche und zeigen dem Benutzer nicht nur genau die gesuchten Seiten, sondern auch andere, neue Aspekte seiner Tags. Ein Folksonomy System ist nicht nur ein System zur Organisation der persönlichen Daten sondern auch ein soziales Netzwerk in dem sich die Benutzer austauschen und gegenseitig helfen, indem sie den für sie interessanten Ausschnitt des Webs umschreiben.

Tagging bietet einen zusätzlichen Layer an Information, der einfach hinzugefügt wird und nichts an der ursprünglichen Information ändert. Dieser Layer ist optional, er muss nicht un-

bedingt verwendet werden.

4.2 Vor- und Nachteile von Ontologien

Traditionelle Klassifikationsmodelle, wie auch Ontologien im klassischen Sinne, funktionieren besser, wenn der Arbeitsbereich einen eingeschränkten Korpus, vordefinierte Kategorien, stabile Objekte und eindeutige Grenzen hat. Man benötigt Experten um die Katalogisierung durchzuführen, verbindliche Wissensquellen und erfahrene Benutzer, die wissen nach welchen Begriffen sie suchen müssen um die gewünschten Ergebnisse zu erreichen. Diese Starrheit hierarchischer Schemata erzeugt eine Reihe von Nachteilen:

- Objekte passen nicht immer in genau eine Kategorie, somit müssen immer wieder Kompromisse geschlossen werden um solche Einteilungen durchzuführen
- Hierarchien sind starr, konservativ und zentralisiert, also inflexibel.
- Hierarchische Klassifikationen werden von der Person, die die Katalogisierung durchführt, beeinflusst und sind in gewisser Weise subjektiv.
- Um die Denkweise und das Vokabular der Benutzer zu „erraten“ werden Experten benötigt, was diese Systeme teuer und komplex macht.
- Hierarchien benötigen Vorhersagen über die Zukunft um im Verlauf der Zeit stabil zu bleiben.
- Hierarchien tendieren dazu, nur einer konsistenten, strukturierten Sichtweise zu folgen, was zu einem Genauigkeitsverlust führt und semantische Bedeutungsunterschiede in Ausdrücken vernachlässigt.
- Hierarchien benötigen Experten oder speziell ausgebildete Benutzer, um konsistent angewandt zu werden.
- Es ist praktisch unmöglich eine vollständige, schlüssige Ontologie zu erstellen (siehe Kapitel 2).

4.3 Unterschiede zwischen Folksonomien und Ontologien

Die Kritiker von Folksonomien bemängeln vor allem die Ungenauigkeit des Free Taggings. Sie sehen den größten Nachteil darin, dass eine Folksonomy keinerlei Hierarchien enthält und die Benutzer, die die Tags vergeben, keine Experten sind und sich ihrer Ansicht nach zu wenig Gedanken über Synonyme und die genaue Bedeutung ihrer gewählten Worte machen. Befürworter von Folksonomien sehen in dieser Ungenauigkeit und Flexibilität aber die große Stärke dieses Systems.

Folksonomy	Ontologie
bildet eine flache Menge von Tags	ist hierarchisch strukturiert
wird zum Zeitpunkt der Veröffentlichung vom Benutzer selbst erzeugt	wird von Experten zu einem bestimmten Zeitpunkt erstellt
dezentralisierte, gemeinschaftliche Sicht	zentralisierte, autoritäre Sicht
keine Synonymkontrolle	Synonymkontrolle
ungenau	hohe Präzision
flexibel	inflexibel
billig und einfach zu erstellen	teuer, da sie nur von Experten erstellt werden kann
dynamisch, verändert sich über die Zeit hinweg	bleibt starr und unverändert; muss neu erstellt werden, wenn sich die strukturierten Daten verändern
kann von jedem Benutzer ohne Vorkenntnis verwendet werden	benötigt Vorwissen
allumfassender Wortschatz	enthält nur ein subjektives Vokabular
skalierbar	starr, unveränderbar
geeignet für die digitale Welt	geeignet für physikalische Gegenstände

Tabelle 4.1: Unterschiede zwischen Folksonomy und Ontologie

Die Unterschiede zwischen Folksonomien und Ontologien sind in Tabelle 4.1 angeführt. Als Hauptunterschied kann man die Struktur der Daten nennen. Während eine Folksonomy keinerlei hierarchische Beziehungen enthält, ist eine Ontologie streng strukturiert. Der Sinn einer Ontologie ist es eine hierarchisch Struktur in bekannte Konzepte zu bringen. Das Ziel einer Folksonomy ist auch eine Ordnung der Objekte. Diese Ordnung hat aber nicht das Ziel eine

allgemeingültige hierarchische Struktur zu erlangen, sondern die Daten für sich persönlich wiederauffindbar zu machen. Der Mensch denkt nicht in Kategorien und Hierarchien, sondern in freien Assoziationen. Demnach entspricht das Konzept der Folksonomy viel eher dem menschlichen Gedankenfluss, als eine Ontologie.

Ein anderer wichtiger Unterschied ergibt sich aus dem Fortschritt des heutigen Computerzeitalters. Es ist zu beobachten, dass Ontologien im Kontext digitaler Informationsorganisation zwar einen idealen technisch-organisatorischen Nährboden finden, gleichzeitig aber immer wieder dieselben Probleme produzieren. Ontologien lassen sich leicht auf physikalische Objekte, wie zum Beispiel Bücher in einer Bibliothek, anwenden. Wenn es um flexible, dynamische Daten aus dem Internet geht, ist eine starre Ontologie denkbar ungeeignet. Sie müsste ständig geändert und neu geschrieben werden – was bei einer von Experten erstellten Ontologie auch sehr teuer und zeitaufwendig ist – und könnte auch aufgrund der Dynamik des Webs nie eine hohe Präzision erreichen.

Eine Ontologie beschreibt eine zentralisierte autoritäre Sicht eines Experten oder einer Gruppe von Experten, die diese Ontologie erstellt haben. Eine Folksonomy hingegen spiegelt die Sicht aller Benutzer des Systems wieder und enthält somit einen allumfassenden Wortschatz aller Benutzer.

4.4 Zusammenfassung

Ontologien und Folksonomien haben wenig gemeinsam, werden aber, aufgrund ihres ähnlichen Wortstamms oder ihrem Zweck zur Organisation von Daten, oft gleich gesetzt. Folksonomien sollen und können Ontologien nicht ersetzen, da die Anwendungsgebiete verschieden sind. Während Ontologien zur Einteilung und Klassifikation hierarchischer Strukturen verwendet werden, bilden Folksonomien einen völlig flachen Raum ab. Damit können also keinerlei hierarchische Strukturen durch eine Folksonomy verlustfrei beschrieben werden.

Der Fehler Ontologien mit Folksonomien vergleichen zu wollen, begründet sich vermutlich in dem immer wieder in der Literatur gemachten Fehler der Abstammung des Wortes Folksonomy (vergl. [37], [34]). Nur weil Taxonomie und Folksonomy die gleiche Wortendung haben, haben sie nicht die gleiche Bedeutung.

Kapitel 5

Anwendungen von Folksonomien

Die bekanntesten Webseiten, die Folksonomien implementieren, sind *Del.icio.us*[6] und *Flickr*[7]. Sie bieten den Benutzern die Möglichkeit Bookmarks, bzw. Bilder zu taggen. Das sind aber nicht die einzigen Anwendungsmöglichkeiten für Folksonomien. Weitere Beispiele sind das Taggen von Artikelsammlungen (*Furl*[9]), Weblogs (*Technorati*[13]), Zitatsammlungen (*CiteULike*[4]) oder persönlichen Wünschen und Zielen (*43Things*[2]).

5.1 Del.icio.us

Del.icio.us[6] ist ein Tool zum Organisieren von Webseiten, also ein Bookmark Manager, bei dem das Hauptaugenmerk auf den Tags liegt. Um das System zu benutzen, muss man sich zuerst dafür registrieren. Mit einem speziellen Bookmark kann man dann direkt während dem Browsen die gerade ausgewählte Seite zu seiner Linkliste hinzufügen. Optional kann man Tags, die dem Benutzer helfen, seine Links wiederzufinden, und längere Beschreibungen dazuspeichern. Die Tags können auf der personalisierten Del.icio.us Seite sowohl chronologisch, als auch alphabetisch als Tagwolken¹ angezeigt werden. Zusätzlich werden diese Tags dazu verwendet Bookmarks von allen Benutzern des Systems zu vereinen. Man kann die rechts angeführten Tags anklicken um zu einer Liste der Bookmarks zu kommen, die auch mit diesem Tag versehen sind. Auf der Hauptseite von Del.icio.us sind auf der rechten Seite ebenfalls die

¹Tagwolken sind eine Visualisierung von Tags, bei der oft verwendete Tags hervorgehoben werden. Siehe auch Kapitel 6.1

del.icio.us / cheesy42 /

your bookmarks | inbox | for | post

popular | help

logged in as **cheesy42** | settings | logout

All your items (94) search

« earlier | later » page 1 of 10

Your del.icio.us clusters

to delicious folksonomy clusters visualization ... and 63 other people ... on 2006-02-06 ... edit / delete

del.icio.us/help/tagrolls

to folksonomy delicious tags visualization ... and 837 other people ... on 2006-02-06 ... edit / delete

Unto.net > Delancey > cheesy42 > rpg

to delicious tools ... and 391 other people ... on 2006-02-06 ... edit / delete

del.icio.us stats - deli.ckoma

to delicious folksonomy statistics ... and 106 other people ... on 2006-02-06 ... edit / delete

expialidocio.us

to delicious folksonomy visualization flash ... and 259 other people ... on 2006-02-06 ... edit / delete

MkE's place - Home

to friend weblog Sweden ... on 2006-02-05 ... edit / delete

Technoschamane >> Shadowrun Großfrankfurt

to shadowrun roleplaying resource ... and 2 other people ... on 2006-02-04 ... edit / delete

Quick Online Tips: Absolutely Del.icio.us - Complete Tool Collection

to delicious tips tools folksonomy ... and 114 other people ... on 2006-02-04 ... edit / delete

Shadowrun RPG - FanPro

to shadowrun roleplaying fanpro ... and 31 other people ... on 2006-02-04 ... edit / delete

Der Tod und das Mädchen

to comic death ... on 2006-02-04 ... edit / delete

« earlier | later » page 1 of 10

» showing 10, 25, 50, 100 items per page

del.icio.us | about | blog | terms of service | privacy policy | copyright policy | contact us | feed for this page

▶ #important

▶ art

▶ cooking

▶ economic

▶ entertainment

▼ folksonomy folksonomy delicious flickr tags tagging facettenklassifikation infocloud metadata ontology sharing tagged

▶ games

▶ penpals

▶ pictures

▶ private

▶ science

▶ search

▶ shopping

▶ studies

▶ travelling

▶ utilities

▶ unbundled tags

▼ options

» view as cloud | list

» sort by alpha | freq

» show | hide bundles

» use minimum: 1, 2, 5

» arrange tags

Abbildung 5.1: Del.icio.us: personalisierte Hauptseite vom 3.3.2006 [6]

beliebtesten Tags angeführt um dem Benutzer einen allgemeinen Überblick über die Tags zu bieten.

In Abbildung 5.1 sieht man ein Beispiel einer personalisierten Hauptseite. Unter den Links zu den Webseiten, die sich der Benutzer als Bookmarks gespeichert hat, sieht man einerseits die Tags und das Erstellungsdatum des Links, andererseits wie viele andere Benutzer denselben Link vermerkt haben.

Auf dieser Seite hat der Benutzer eine Menge Interaktionsmöglichkeiten, die durch Verlinkungen gegeben sind:

- *Bookmark*: Das Bookmark selbst ist mit der dazugehörigen Webseite verknüpft.
- *Tags*: Die ersten drei bis vier zum jeweiligen Bookmark gehörigen Tags sind direkt dar-

unter aufgelistet. Man kann sie anklicken um auf weitere Bookmarks, die mit diesem Tag versehen sind, zu gelangen.

- *Andere Benutzer:* Direkt unter dem Bookmark ist weiters vermerkt wie viele andere Benutzer, diesen Link in ihre Sammlung aufgenommen haben. Diese Information ist mit abstuftend rot unterlegt – je mehr Benutzer dieses Bookmark gesetzt haben, desto kräftiger ist der Rot Ton – und mit der Benutzerliste verlinkt.
- *edit/delete:* Ebenfalls in der Zeile unter dem Bookmark findet man einen Link zum Bearbeiten („edit“) und zum Löschen („delete“) des Bookmarks aus der Liste.
- *Tag Bundles:* In der grau hinterlegten Spalte auf der rechten Seite findet man die persönlichen Tag Bündel. Del.icio.us bietet die Möglichkeit die Tags zu ordnen und in Gruppen einzuteilen. Diese Gruppen („Tag Bundles“) können frei gewählt und benannt werden und müssen nicht zwingend erstellt werden. Sie stellen also eine Art Meta-Tags dar – Tags von Tags. Jedes Tag kann beliebig vielen Bündeln zugeordnet werden; wenn es keinem Bündel zugeordnet ist, ist es im Punkt „unbundled Tags“ zu finden. Man kann die einzelnen Bündel anklicken um sie so zu öffnen oder zu schließen. Im offenen Zustand werden die zugeordneten Tags unter dem Namen des Bundles angeführt, im geschlossenen Zustand ist der Bündelname ausgegraut.
- *Anzeigeoptionen:* Ganz unten in der grau hinterlegten Spalte ist der Punkt „options“. Dieser stellt kein Tag Bündel dar, sondern bietet die Möglichkeit die Anzeigeoptionen umzustellen. Man kann sich die Tag Bündel als Wolke („cloud“) oder Liste, alphabetisch („alph“) oder nach Frequenz der Verwendung („freq“) sortiert, anzeigen lassen. Weiters kann man die Bündel ganz ausblenden und sich nur die Tags („show | hide bundles“) oder überhaupt nur die Tags, die zwei oder mehr (bzw. 5 oder mehr) Links zugewiesen sind, anzeigen lassen. Der letzte Punkt („arrange tags“) bietet die Möglichkeit die Tag Bündel zu erstellen und verwalten.
- *Kopfzeile:* In der Kopfzeile hat man noch eine weitere Anzahl an Möglichkeiten zur Verfügung:
 - *Del.icio.us Hauptseite:* Über den Link „del.icio.us“ gelangt man auf die Hauptseite

von Del.icio.us, wo aktuelle Bookmarks angezeigt werden.

- *persönliche Seite*: Sowohl der Benutzername, als auch der Link „your bookmarks“ führen auf die persönliche Hauptseite, so wie sie in Abbildung 5.1 zu sehen ist.
- *inbox*: Del.icio.us bietet auch die Möglichkeit die Tags anderer Benutzer zu verfolgen. In der inbox finden sich die Änderungen der Benutzer oder Tags, die man abonniert hat.
- *for*: Ein anderes Feature von Del.icio.us ist, dass man für andere Benutzer Bookmarks speichern kann. Der jeweilige Benutzer findet diese dann unter dem Punkt „for“.
- *post*: Hier kann man neue Bookmarks taggen und speichern.
- *Fußzeile*: In der Fußzeile findet sich neben Links auf die Del.icio.us Hauptseite, ein zugehöriges Weblog und diverse Policies auch die Möglichkeit Del.icio.us Seiten als RSS Newsfeed zu abonnieren. Mittels RSS-fähigen Webbrowsern, können dann Änderungen in diesen Seiten in einer extra Spalte angezeigt werden. RSS kann praktisch auf allen Seiten von Del.icio.us verwendet werden: man kann die Links anderer Benutzer, die beliebtesten Tags oder auch die allgemein beliebtesten Seiten als Newsfeed auslesen und bekommt dann Benachrichtigungen aller aktueller Änderungen.

Del.icio.us ist eine klassische Anwendung einer Folksonomy, die alle Freiheiten und Interaktionsmöglichkeiten einer Folksonomy implementiert und dem Benutzer zur Verfügung stellt.

5.2 Flickr

Flickr[7] bietet die Möglichkeit Fotos zu verwalten. Auf der Flickr Hauptseite, kann man sich die neuesten oder auch „interessantesten“ Bilder der Flickr-Benutzer ansehen. Auch für dieses Service muss man sich registrieren um den einfachen, eingeschränkten Account zu nutzen. Dieser beinhaltet drei Gruppen, in die man die Bilder einteilen kann, und 20 MB für Uploads pro Monat. Wenn man Flickr ausgiebiger verwenden möchte, kann man für €2 pro Monat zwei GB an Speicher und unlimitierte Verwendung kaufen. Eine personalisierte Hauptseite sieht man auf Abbildung 5.2.

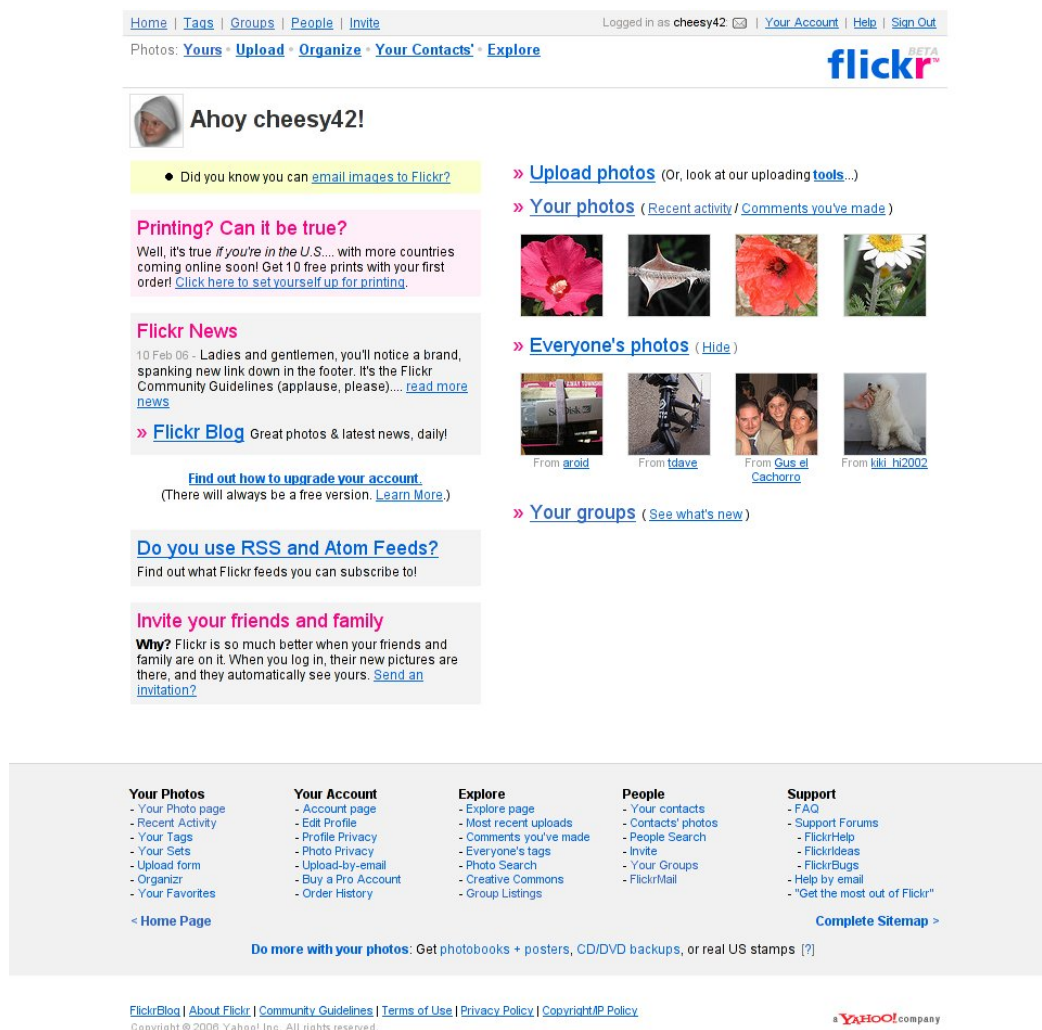


Abbildung 5.2: Flickr: personalisierte Hauptseite vom 3.3.2006 [7]

Auch Flickr bietet eine ganze Reihe an Interaktionen auf seiner Hauptseite. Die interessanten Links sind hier in der Kopfzeile und im rechten Teil der Seite zu finden. Auf der linken Seite macht Flickr auf weitere Services aufmerksam. In der grauen Box ganz unten findet man nochmal alle Interaktionsmöglichkeiten genau aufgeschlüsselt.

- *Hauptseite:* Sowohl über den Link „Home“ links oben, als auch über das Flickr Logo kommt auf die Hauptseite. Das Aussehen von dieser Seite wird über Cookies gesteuert und ist abhängig davon, ob man eingeloggt ist, oder nicht. Wenn man Flickr erlaubt Cookies im Browser zu hinterlegen, kommt man automatisch auf diese Hauptseite ohne

sich einloggen zu müssen.

- *Tags*: Der Link „Tags“ führt genau wie bei Del.icio.us zu den aktuell beliebtesten Tags. Auf der dazugehörigen Seite, werden sowohl die neuesten als auch die beliebtesten Tags als Tagwolke angezeigt.
- *Gruppen*: Gruppen in Flickr sind Sammlungen von Bildern von Benutzern, die etwas gemeinsam haben. Diese Gruppen können von jedem Benutzer erzeugt werden. Dabei gibt es öffentliche Gruppen, denen jeder beitreten kann, öffentliche Gruppen, für die man eine Einladung braucht, und private Gruppen, denen nur Freunde beitreten können.
- *Personen*: Über den Punkt „People“ kann man Personen einladen Freunde zu werden, unter „Contacts“ sieht man eine Liste aller bereits vorhandenen Freunde. Diese Personen haben, zusätzlich zum Erzeuger der Fotos, die Möglichkeit diese zu taggen. Alle anderen Benutzer, können die eigenen Fotos nicht taggen. Diese Einschränkung macht den Hauptunterschied von Flickr zu Del.icio.us – und zu den meisten anderen Folksonomy-Anwendungen – aus. Flickr ist eine enge Folksonomy wie in Kapitel 3.2.3 beschrieben.
- *Nachrichten*: Wenn man auf den kleinen Brief rechts oben klickt, gelangt man in seine persönliche Flickr Mailbox. Diese ist dazu da, Mails an andere Flickr Benutzer zu schicken, oder Mails von ihnen zu empfangen. Das ist ein zusätzliches Feature, das zum Beispiel unter Del.icio.us nicht vorhanden ist.
- *Einstellungen*: Unter dem Punkt „Your Account“ kann man alle möglichen Einstellungen ändern.
- *Eigene Fotos*: Die eigenen Fotos, von denen man auch eine Vorschau sieht, kann man über die Links „Yours“ und „Your Photos“ erreichen. Man kann auch direkt auf eines der Vorschaubilder klicken um direkt zu dem Foto zu kommen. Weitere vorhandene Links führen auf aktuelle Änderungen („Recent Activity“), die auftreten wenn andere Benutzer zum Beispiel ein Bild zu seinen Favoriten hinzufügen oder Bilder kommentieren, und eigene Kommentare („Comments you’ve made“).
- *Upload*: Flickr bietet neben einem einfachen Upload-Formular auch noch Upload-Tools für Mac und Windows, die den Ladevorgang vereinfachen.

- *Organisation*: Ein weiteres Feature von Flickr ist eine Flash-Applikation zum einfachen Organisieren der persönlichen Fotos. Mit dieser Applikation kann man Sets erstellen und die Bilder zuordnen und auch die Bilder umbenennen.
- *Explorer*: Mit dem Explorer kann man sich interessante Bilder zufällig ansehen und durch die besten Bilder blättern. „Interessante“ Bilder werden von Flickr anhand mehrerer Aspekte ausgesucht: wie oft es angeklickt wird, von wo diese Klicks kommen, wie oft es kommentiert oder als Favorit hinzugefügt wird, sind einige davon.

Flickr ist im Gegensatz zu Del.icio.us eine enge Folksonomy, da nur ein eingeschränkter Benutzerkreis – oft auch nur der Erzeuger selbst – die Fotos taggen kann. Es ist ein gutes Beispiel für die Anwendung von Folksonomien zum organisieren Text-freier Daten.

5.3 Furl

Furl[9] ist – ähnlich wie Del.icio.us – ein Bookmark-Manager. Der Unterschied ist, dass Furl wesentlich strukturierter ist und viel mehr in Richtung Literatursammlungen bzw. Literaturrecherche ausgerichtet ist. Ein Beispiel einer Furl-Literatursammlung sieht man auf Abbildung 5.3.

Man erkennt hier sofort die Vorteile, die Furl – als Tool zur Literaturrecherche – gegenüber Del.icio.us zu bieten hat. Die Literaturliste ist strukturiert aufbereitet. Man sieht auf den ersten Blick, welche Artikel man schon gelesen hat. Die noch ungelesenen Artikel sind in fetter Schrift geschrieben. Man kann jedem Artikel eine bis mehrere Topics zuordnen. Diese sind nicht mit Tags zu verwechseln. Jedem Artikel werden auch Tags zugeordnet, die aber nicht in dieser Liste aufscheinen. Ein Rating, das man selber setzt, gibt zusätzlich eine persönlich eine zusätzliche Einschätzung über die Qualität des Artikels und man sieht, wie oft man den Artikel schon betrachtet hat.

Ein weiterer Vorteil von Furl ist, dass eine Kopie der Artikel auf dem Furl Server gespeichert wird. Damit kann man, auch wenn die Originalseite sich ändert oder nicht mehr existiert immer noch auf den Artikel zugreifen. Die drei Symbole die sich zwischen dem Titel und dem Datum des Artikels befinden verlinken von links nach rechts auf den Originalartikel, die Kopie des Artikels und eine Seite zum Editieren der zum Artikel gespeicherten Informationen. Diese

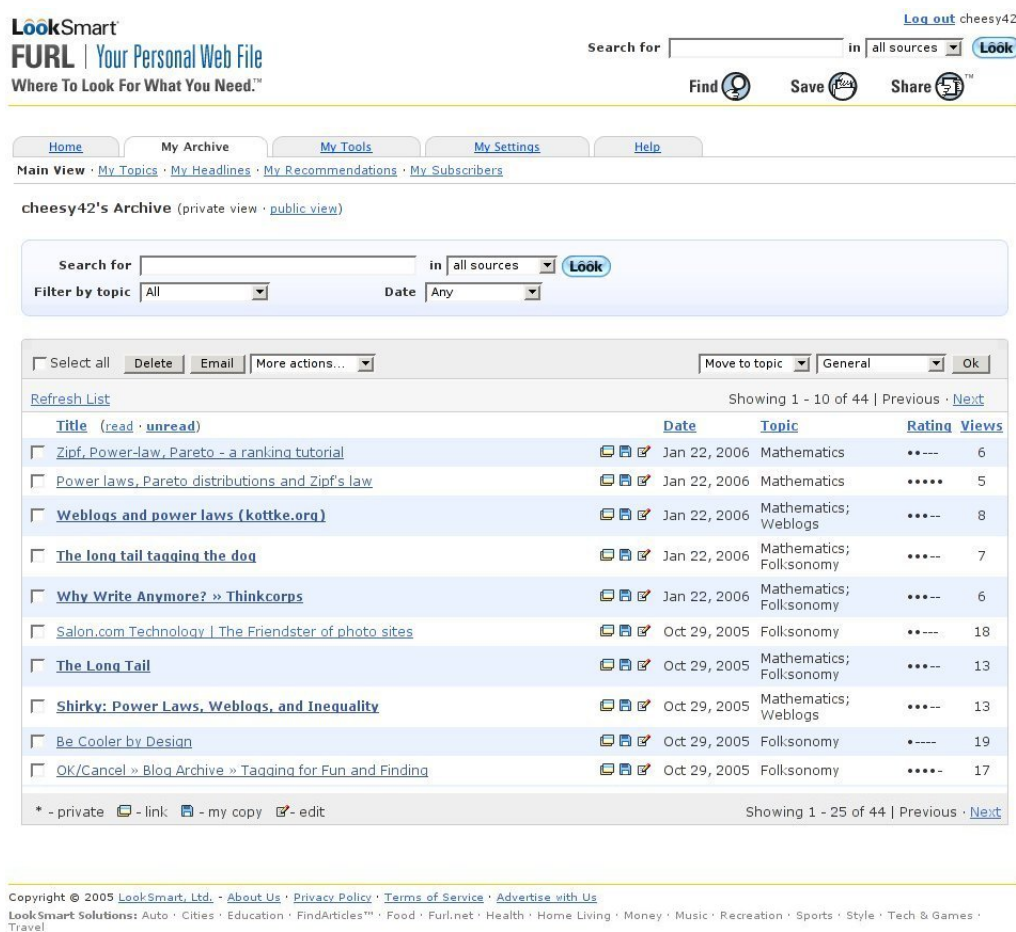


Abbildung 5.3: Furl: personalisierte Hauptseite vom 5.3.2006 [9]

Seite bietet nicht nur die Möglichkeit Topics und Tags zu speichern, man kann zusätzlich auch einen Kommentar, den Autor und das Datum des Artikels vermerken. Das Datum das auf der Hauptseite aufscheint, ist aber das Datum der Speicherung und nicht das so vermerkte Datum.

Auch auf Furl kann man nach Tags suchen. Immer wenn der Benutzer eine Seite speichert werden auch benachbarte Artikel angezeigt.

Signed in as Christine [Account](#) [Favorites](#) [Watchlist](#) [Blogs](#) [Profile](#) [Help](#) [About](#) [Sign Out](#)

Technorati™ Search Tags Blog Finder Explore

Your Watchlist
Currently watching 2 searches.

Add a keyword or URL to your Watchlist
For example: "NYTimes.com" or "iPod"
 [Add](#)

Advertise on Technorati

▼ **Today's Most Popular**

News

- New Florida town would re...
- Thai PM gambles on poll r...
- Poll: Bush Ratings At All...

Movies

- Brokeback Mountain
- Crash (2004)
- Ultraviolet (2006)

Books

- Growing Healthy Asia...
- The Da Vinci Code
- The Total Money Make...

Top 100 Blogs

- Boing Boing
- SuicideGirls News All
- Engadget

[More in Popular »](#)

▼ **New in the Technorati Blog**

More Fun with Favorites
Now you can see the Top 100 Blogs, as measured by your Favorites!
[More in the Technorati Blog »](#)

► **www.vanderwal.net** [Search](#) [Mini](#) [Remove](#)

▼ **Folksonomy** [Search](#) [Mini](#) [Remove](#)

Most recent 3 posts:

I'm not web 2.0 guru
In A Republishing of Trying to Fo... 39 minutes ago
Originally published at Trying to follow. Please leave any comments there. But I have been paying attention a little bit ... out more. The whole **folksonomy** thing, which I did talk about earlier is fascinating. Read my earlier

del.icio.us and flickr
In A Republishing of Trying to Fo... 40 minutes ago
Originally published at Trying to follow. Please leave any comments there ... , known as **folksonomy**. Now we are just beginning to see the popularity of sites using this system

Newsvine: The Wisdom Of The Crowd
By Sean Coon in connecting*the*dots 7 hours ago
The reviews are in: We, the people, are in the drivers seat. Newspapers are already hemorrhaging readership, as the web has created an extremely rich bazaar, allowing us to shop for unbundled content at every turn, while unbundled advertising models begin to sprout up to support this evolution

[View in RSS format](#) [More Search Results »](#)

Home Search Tags Blog Finder Explore Popular Weblog Developers About Help Account
© Technorati, Inc. CC License DMCA Privacy Policy Terms of Service Ping Contact

Abbildung 5.4: Technorati: personalisierte Watchlist vom 5.3.2006 [13]

5.4 Technorati

Technorati[13] ist ein Weblog-Manager, in dem man Weblogs oder Schlagworte abonnieren und sein persönliches Weblog auflisten lassen kann. Technorati ist eine enge Folksonomy. Es bietet dem Benutzer die Möglichkeit sein eigenes Weblog hinzuzufügen und seine Einträge zu taggen. Wenn man kein Weblog besitzt, kann man als passiver Benutzer nach Stichworten in den Technorati zugehörigen Weblogs suchen oder sich regelmäßig Updates anzeigen lassen.

Technorati bietet die in Folksonomien üblichen Suchmöglichkeiten nach Tags, Inhalten oder Benutzern. Man kann sich beliebig viele Weblogs als Favoriten speichern, oder auch Be-

griffe oder Webseiten auf eine Watchlist setzen. Der Benutzer wird dann auf neue Blogbeiträge auf diesen Seiten, bzw. neue Blogs, die die Begriffe der Watchlist enthalten, aufmerksam gemacht. Ein Beispiel für eine personalisierte Watchlist sieht man in Abbildung 5.4.

Hier kann man die gewählten Themen auf- bzw. zuklappen und somit nach Wunsch Ausschnitte aus den drei aktuellsten zugehörigen Blogs anzeigen lassen. Direkt unter dem Namen des Weblogbeitrags sieht man ein Symbol einer Lupe und einer Sprechblase. Mit der Lupe kann man auf der Webseite, von der der aktuelle Eintrag stammt nach dem Oberbegriff suchen, die Sprechblase sagt einem, wie viele Blogs und Postings dieses Weblog verlinken und zeigt einem die Liste dieser Weblogs an. Weiters sieht man den Namen des Bloggers, verlinkt mit dessen Technorati Profil und einen Link auf das Weblog in dem gepostet wurde. Im Punkt „Blogs“ in der Kopfzeile, kann man eigene Weblogs auf Technorati registrieren und taggen.

In der grünen Suchzeile, kann man nach Begriffen in den Weblogs („Search“), nach Tags („Tags“) oder nach Themen von Weblogs („Blog Finder“) suchen. Über den „Explore“ Link findet man die besten Blogs zu einem beliebigen Thema.

5.5 CiteULike

CiteULike[4] bietet eine weitere Anwendungsmöglichkeit von Folksonomien – eine Zitatesammlung. Auf CiteULike kann man Artikel so wie auch in Del.icio.us oder Furl sammeln. Zusätzlich bietet CiteULike die Möglichkeit eine vollständige Bibliographie der verlinkten Artikel anzugeben. Diese kann auch ins bibtex-Format exportiert werden und ist somit einfach in \LaTeX -Dokumente zu inkludieren. In Abbildung 5.5 sieht man eine personalisierte Zitatesammlung auf CiteULike.

Über das Menü auf der linken Seite erlangt man schnellen Zugriff auf die zuletzt gespeicherten Artikel und – wie auch bei Furl – auf die noch ungelesenen Artikel. Der Benutzer kann sich mit einem Klick einen Überblick über die Autoren, Editoren und Tags der verlinkten Artikel oder in seiner Bibliothek suchen. Wie auch bei Technorati, kann man eine Watchlist aktivieren und Tags, Autoren oder auch Bibliotheken anderer Benutzer abonnieren.

In der Mitte sind die Artikel der eigenen Sammlung aufgelistet. Unter dem Titel des Artikels, sieht man Links auf die Autoren, und die zugeteilten Tags. Man kann den Artikel auf einer Skala von eins bis sechs beurteilen und sieht auch auf einen Blick, wie viele Andere Be-

The screenshot shows the CiteULike interface. At the top, the user is logged in as 'cheesy42' with links for 'Log Out' and 'FAQ'. There are also language selection icons for various countries. A search bar is located at the top right.

The main content area is titled 'cheesy42's library' and contains a list of recent papers added to the library. Each entry includes the article title, author(s), posting date, and a star rating. The articles listed are:

- Social Bookmarking Tools (I): A General Review** by Hammond T, Hannay T, Lund B, Scott J. Posted to [bookmarking folksonomy social](#) by [cheesy42](#) as ★★★★★ and [1 other ...](#) on 2006-03-05 20:32:19.
- Metadata Based Annotation Infrastructure Offers Flexibility and Extensibility for Collaborative Applications and Beyond - Slide list** by Koivunen M, Swick R. Posted to [annotate](#) by [cheesy42](#) as ★★ on 2006-03-05 20:22:57.
- Circumscription** by Mccarthy J. Posted to [ontology philosophy](#) by [cheesy42](#) as ★★★★★ on 2006-03-05 20:21:31.
- Organon** by Aristoteles. Posted to [aristotele categories metaphysik ontology philosophy](#) by [cheesy42](#) as ★★★★★ on 2006-03-05 20:16:35.
- Prolegomena** by Kant I. Posted to [metaphysik ontology philosophy](#) by [cheesy42](#) as ★★★★★ on 2006-03-05 20:12:50.
- An Essay Toward a Real Character and a Philosophical Language** by Wilkins J. Posted to [categories language ontology philosophy](#) by [cheesy42](#) as ★★★★★ and [1 other ...](#) on 2006-03-05 20:09:01.
- Sophist** by Plato. Posted to [ontology philosophy plato sophist](#) by [cheesy42](#) as ★★★★★ on 2006-03-05 20:05:34.

At the bottom of the article list, there is a note: 'Note: You may cite this page as: <http://www.citeulike.org/user/cheesy42>'. There are also buttons for 'EndNote', 'BibTeX', and 'RSS'.

The right sidebar shows 'cheesy42's tags' with a filter input field and a tag cloud containing terms like 'annotate', 'aristotele', 'bookmarking categories', 'folksonomy language', 'metaphysik', 'ontology', 'philosophy', 'plato', 'social', and 'sophist'.

The left sidebar contains navigation links such as 'Home', 'News and Status', 'Discussion list', 'Your Library' (with sub-links for 'View articles', 'Recent', 'To read', 'Authors and tags', 'Search', 'Post article', 'Manually post article', 'Export to BibTeX/EndNote', 'Change password'), 'Your Watchlist' (with 'About watchlists'), 'Journals' (with 'Browse current issues'), 'Groups' (with 'View group', 'Manage groups'), and 'Experimental Features' (with 'Import from BibTeX').

Abbildung 5.5: CiteULike: personalisierte Zitatesammlung vom 5.3.2006 [4]

nutzer diesen Artikel in ihrer Sammlung haben.

In der rechten Spalte sieht man eine Tag-Wolke aller verwendeten Tags, in der man auch mittels der oben stehenden Eingabezeile, nach Zeichenketten filtern kann.

5.6 43 Things

43 Things[2] bzw. 43 Places[1] sind Webseiten, auf denen die Benutzer ihre Wünsche und Ziele und im Falle von 43 Places, die Orte, die sie gerne besuchen möchten, angeben können. In Abbildung 5.6 sieht man eine personalisierte Seite von 43 Things.

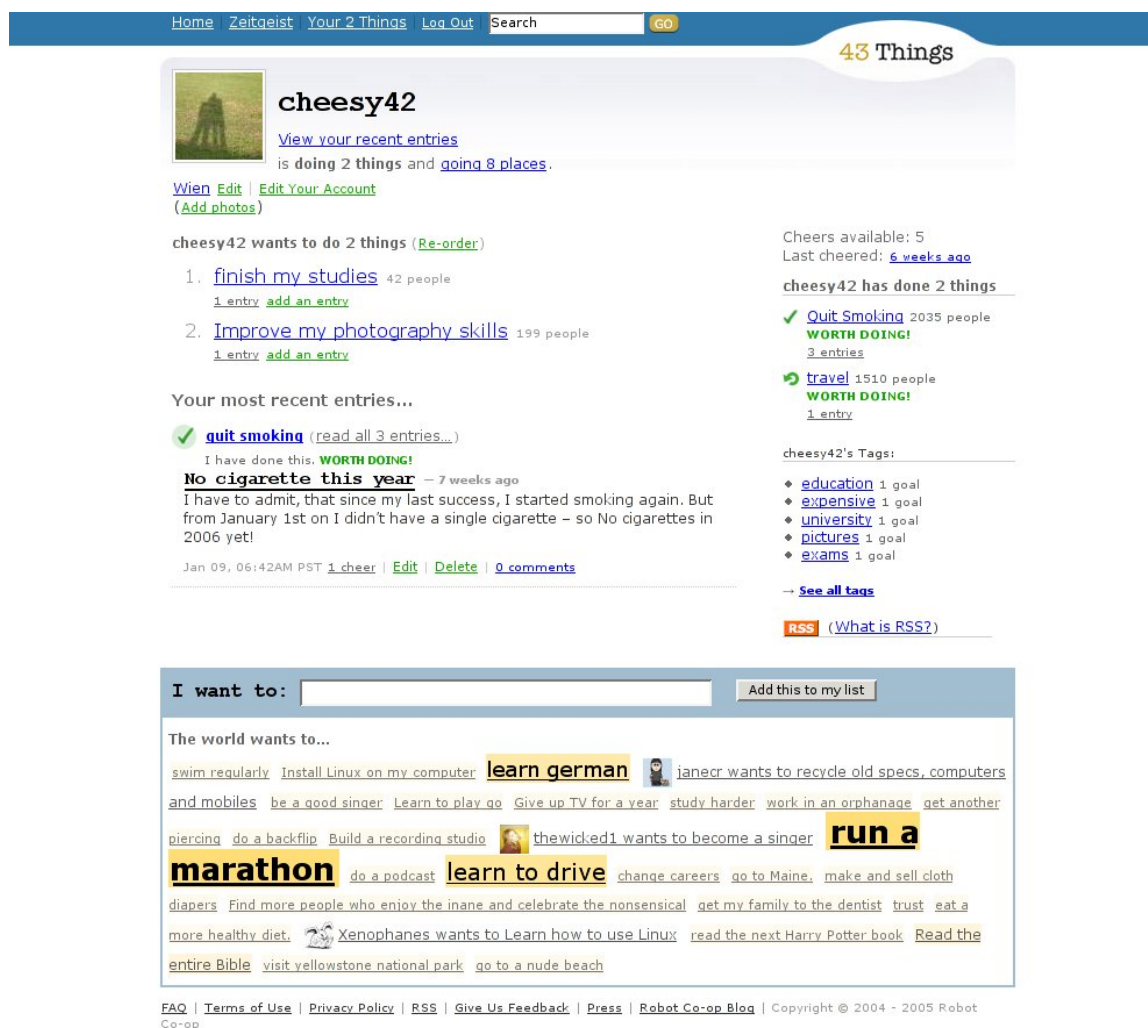


Abbildung 5.6: 43 Things: personalisierte Hauptseite vom 5.3.2006 [2]

Man kann zu seinen persönlichen Zielen nicht nur Tags abspeichern, sondern auch Kommentare und Bilder posten. Sobald man eines seiner Ziele erreicht hat, kann man es unter den bereits vollbrachten Zielen abspeichern. Es besteht auch die Möglichkeit seine Ziele wieder aufzugeben. Man kann seinen Zielen eine Reihenfolge zuordnen und damit seine Prioritäten setzen. Neben jedem Ziel, sieht man, wie viele andere Benutzer dasselbe Ziel verfolgen. Mit einem einfachen Mausklick kann man dann man sich eine Liste dieser Benutzer anzeigen lassen, oder auch mit Benutzern Kontakt aufnehmen, die dieses Ziel bereits erfolgreich absolviert haben.

Am unteren Ende der Seite sieht man die beliebtesten Ziele aller Benutzer im Stil einer Tag Wolke. Die beliebtesten Ziele – in Abbildung 5.6 ist es das Ziel einen Marathon zu laufen („run a marathon“) – sind durch große Schrift hervorgehoben.

Kapitel 6

Visualisierung von Folksonomien

Es gibt mehrere Applikationen, die Folksonomien aus Del.icio.us und Flickr visualisieren. Diese Tools veranschaulichen die Beziehungen zwischen Tags in Del.icio.us, soziale Kontakte oder auch die Relation zwischen Farben und Tags in Flickr. Die anschließenden Beispiele stellen nur eine Auswahl der vorhandenen Visualisierungstools dar.

Im folgenden sind sechs Beispiele für Visualisierungen von Folksonomien in Del.icio.us und Flickr beschrieben, die die in Tabelle 6.1 aufgezählten Eigenschaften für ihre Darstellungen verwenden.

Eigenschaft	Beschreibung	Berechnung	Visualisierung
Beliebtheit	Häufigkeit der Tags	Anzahl der Bookmarks pro Benutzer, über alle Benutzer aufsummiert, die dieses Tag verwenden	Größe – je größer eine Einheit dargestellt wird, desto beliebter ist sie
Gleichheit	gibt an, wie einfach man von einem Tag zum anderen kommt	wird anhand der Anzahl aller Links von zwei Tags und der Anzahl ihrer gemeinsamen Links berechnet	Gruppierung – gleiche Tags werden zu einem Begriff zusammengefasst

Nähe	abgeschwächte Form der Gleichheit	Form	wird anhand der Anzahl der gemeinsamen Bookmarks berechnet	Distanz
Nachbarschaft	wenn zwei Tags dem gleichen Objekt zugeordnet sind		alle Tags aller Objekte, die mit dem Haupttag versehen sind	durch kreisförmige Anordnung um den Haupttag herum
Beziehungen	zwischen Tags und Bookmarks oder Bildern oder zwischen Personen		treten auf, wenn eine direkte Verbindung besteht	Verbindungslinien
Farbe	Verbindung zwischen Farben und Tags in Flickr		Hauptfarbe der einzelnen dem Tag zugeordneten Bilder	Farbsortierte Labels
Personengebundenheit	wenn die Tags eindeutig einem Benutzer zugeordnet sind		wenn der Benutzer das Tag verwendet	Hervorheben

Tabelle 6.1: Eigenschaften von Visualisierungen

6.1 Tagwolken

Die bekannteste und am weitesten verbreitete Visualisierung einer Folksonomy ist die Anordnung von Tags als Tagwolken. Tagwolken sind eine meist alphabetische Anordnung der beliebtesten Tags aller Benutzer oder auch aller Tags eines Benutzers. Die Beliebtheit der Tags wird durch die Größe der Darstellung des Wortes visualisiert. Diese Visualisierung ist sehr einfach gehalten, wird aber in den meisten Folksonomy-Systemen verwendet um einen Überblick über die meistverwendeten Tags zu geben.

Die Del.icio.us-Tagwolke in Abbildung 6.1 zeigt eine Kombination aus den oft verwendete-

ajax apple architecture art article audio blog blogs books business comics community
 computer cool CSS culture daily design development diy download education email
 entertainment fashion finance firefox flash food forum free freeware fun funny game games geek
 google graphics gtd hardware health history home howto html humor inspiration interesting
 internet ipod irc java javascript kids learning library links linux mac magazine management maps
 media movies mp3 music network news online opensource osx photo photography photos
 photoshop php podcast politics productivity programming python radio rails recipes
 reference religion research resources reviews rss ruby safari_export school science search
 security sga shopping social software sports tech technology tips tools toread
 travel tutorial tutorials tv useful utilities video web web2.0 webdesign webdev
 windows wordpress writing

Abbildung 6.1: Tagwolke der häufigsten Tags in Del.icio.us [6]

ten Tags aller Benutzer und den Tags eines einzelnen Benutzers. Die angezeigten Tags stellen die 121 beliebtesten Tags dar. Je größer die Schrift, desto beliebter ist dieses Tag. So sind in diesem Fall zum Beispiel die Tags *blog*, *blogs* oder *programming* von den Benutzern sehr häufig verwendet worden. So waren am 7.2.2006 beispielsweise 78958 Bookmarks mit dem Tag *programming* versehen und 43464 Bookmarks mit dem Tag *geek*. Die rot markierten Tags, wie zum Beispiel *photography*, *programming* und *tips*, sind personengebundene Tags. Der gerade eingeloggte Benutzer hat diese Tags selbst auch verwendet.

6.2 Mindmaps

Eine Mindmap ist eine graphische Darstellung eines Themas und seiner Unterthemen, und deren Beziehungen untereinander. Mindmaps enthalten das zentrale Thema in der Mitte des Blattes. Nach außen hin verlaufen die Hauptäste, die die zum Thema gehörenden Kapitel darstellen, mit weiteren Unterästen, die die dazugehörigen Informationen darstellen. Auf den Ästen steht immer nur ein Schlüsselwort. Bei der Erstellung können Farben und Bilder benutzt werden, um der kreativen Arbeitsweise des Gehirns gerecht zu werden, und um die Mindmap schneller lesen und überblicken zu können.

Die Einsatzmöglichkeiten von Mindmapping sind breitgefächert von wirtschaftlichen Anwendungen bis hin zur persönlichen Lernhilfe. Beispiele für Anwendungen sind Brainstorming, Strukturieren von komplexen Inhalten, Erstellung von Vorträgen, Ausarbeiten von Texten sowie Planungs- und Organisationsaufgaben.

Im Falle von Folksonomien kann eine Mindmap zur Strukturierung der Beziehungen der Tags untereinander verwendet werden. In dem gebrachten Beispiel werden die Tags des Benutzers in Gruppen zusammengefasst. Pietro Speroni schreibt dazu:

Tags are not all the same, some are more similar than others. So, for example, the tag September11 and GeorgeBush have more links in common than GeorgeBush and intelligence. The idea behind this version of DeliMind was to cluster tags that had links in common. Since distance is generally not a transitive property (if I am near to you, and you are near to Jim, I am not necessarily that near to Jim), while clustering is (if I and you are in the same cluster, and you and Jim are in the same cluster, then me and Jim have to be in the same cluster unless people belong to different clusters, but that's a complication). [48]

Aus den gegebenen Tags wird eine Matrix von Relationen erzeugt. Die möglichen Relationen zwischen Tags sind:

1. *Untermenge*: Tag1 taggt nur Bookmarks, die auch von Tag2 getaggt werden
2. *Identität*: Tag1 und Tag2 taggen genau dieselbe Menge an Bookmarks
3. *Disjunktivität*: Tag1 und Tag2 taggen keine gemeinsamen Bookmarks
4. *Überlappung*: Tag1 und Tag2 taggen gemeinsame Bookmarks, aber jeder für sich besitzt auch eigene Bookmarks

Die Mindmap im Abbildung 6.2 enthält also nicht alle Tags des Benutzers sondern gruppiert sie je nachdem, ob diese Tags gemeinsame Links beschreiben. Tags die sich im Cluster eines Haupttags befinden, können als weitere Unteräste angezeigt werden. In dieser Mindmap werden folgende Eigenschaften visualisiert:

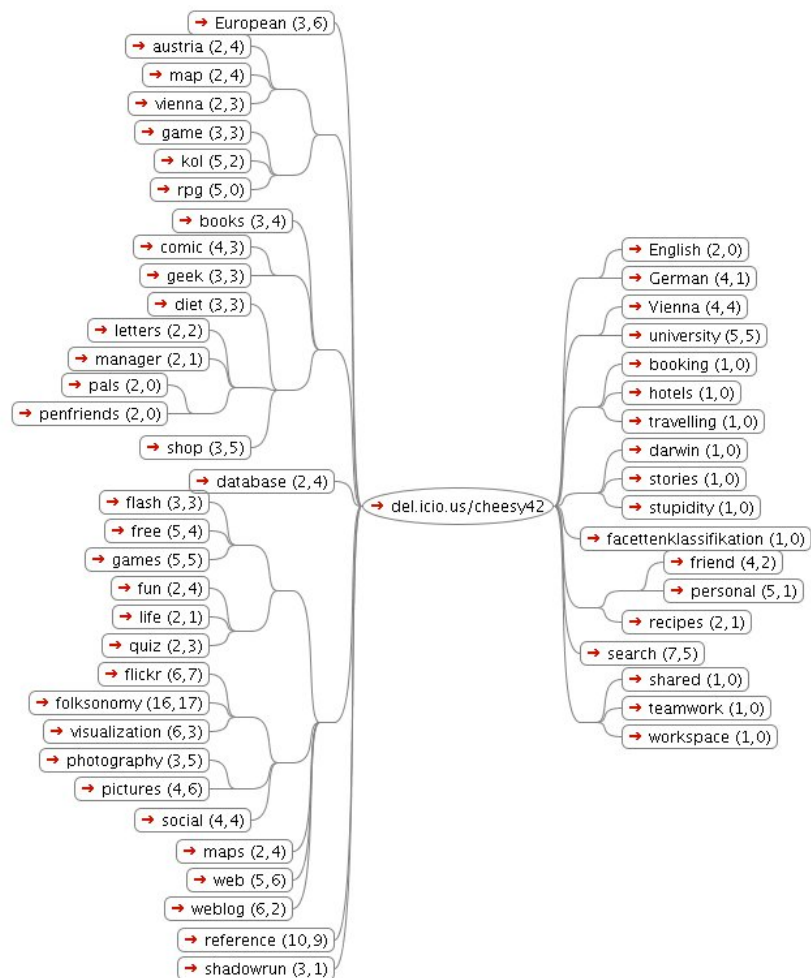


Abbildung 6.2: Del.icio.us Mindmap der Tags eines Benutzers [48]

- die *Gleichheit* von Tags - gleiche Tags werden in Clustern zusammengefasst. Die zweite Zahl in der Klammer hinter den Tags gibt die Anzahl der Untertags an, die sich in diesem Cluster befinden.
- die *Nähe* von Tags - nahe Tags werden in benachbarten Ästen angeordnet.
- die *Beziehung* zwischen Tags und Bookmarks - die geposteten Links werden dem Baum als Äste hinzugefügt (siehe Abbildung 6.3).
- die *Beliebtheit* der Tags - die zweite Zahl in den Klammern gibt die Anzahl der dem Tag

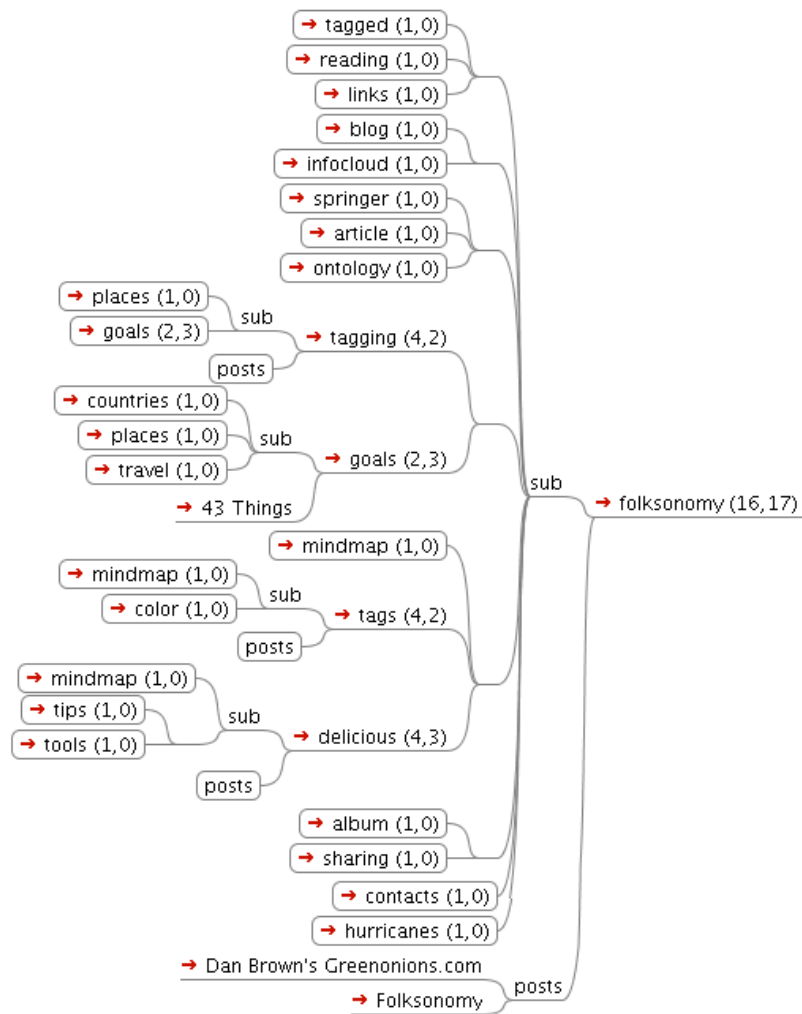


Abbildung 6.3: Del.icio.us Mindmap: Cluster Folksonomy [48]

zugehörigen Links an.

Im Zentrum steht der Benutzer, genauer der Link auf die del.icio.us Webseite des Benutzers „del.icio.us/cheesy42“. Links und rechts davon sind 48 Hauptcluster, die das Tool errechnet hat. Die Zahlen in den Klammern geben an, wie vielen Bookmarks dieses Tag zugeordnet wurde und wie viele Untertags in diesem Cluster integriert sind. Die Untertags der Cluster kann man sich als weitere Äste der Mindmap anzeigen lassen. In Abbildung 6.3 sieht man den Cluster Folksonomy und all seine Untertags. Beim Cluster Folksonomy ist angegeben, dass 17 Unterkategorien existieren. Damit sind alle direkten Untergruppen gemeint. Das sind alle

Tags, die sich in der ersten Spalte links von Folksonomy befinden, also sowohl Endknoten, wie *tagged*, *reading* oder *mindmap*, als auch Gruppen, die weitere Tags enthalten. Das sind in diesem Fall *tagging*, *goals*, *tags* und *delicious*. Unter dem Punkt *posts* kann man sich, wie ganz unten gezeigt, die mit dem Tag versehenen Bookmarks anzeigen lassen. In dem Fall sind die Bookmarks mit den Titeln „Dan Brown’s Greenonions.com“ und „Folksonomy“ mit dem Tag Folksonomy versehen und dem Cluster *Folksonomy* direkt zugeordnet.

6.3 Del.icio.us Soup

Del.icio.us Soup[58] ist eine Applikation, die die persönlichen Del.icio.us Tags als farbige Kugeln visualisiert:

- *Beliebtheit*: Je größer die Kugel, desto beliebter das Tag
- *Nähe*: eng nebeneinander liegende Kugeln visualisieren die Nähe der dazugehörigen Tags
- *Nachbarschaft*: benachbarte Tags werden im Kreis um das ausgewählte Tag herum angezeigt. Man kann eine Nachbarschaft eines Tags auch farblich markieren, damit man, wenn man ein anderes Tag wählt, diese Gruppe weiterhin unterscheiden kann

In Abbildung 6.4 sieht man als Beispiel das Tag *Reference* in der Mitte, als orangefarbene Kugel. Rundherum sind die Nachbarn des Tags angezeigt. Sie sind mit dem Tag *References* verbunden. Diesem Wort nicht benachbarte Tags, werden hier nicht angezeigt. Die orange markierten Tags sind das Tag *Web* und dessen Nachbarn. Alle übrigen Tags sind rosa markiert. Tags, die oft gemeinsam auftreten werden eng beisammen dargestellt. Hier in der Graphik sieht man zum Beispiel, dass alle Tags im Themenbereich *Web* (wie *Programming*, *HTML* oder *Javascript*) sich rechts um dieses Wort gruppieren. Nur das Tag *CSS* befindet sich auf der linken Seite. Das bedeutet, dass er nur selten mit den rechten Begriffen gemeinsam verwendet wurde und öfter mit anderen Begriffen kombiniert wird.

Eine weitere Gruppe, die sich gebildet hat, sieht man am unteren Rand. Hier geht es um das Thema Film - *Database*, *Movie*, *Film* und *Actors* wird oft für gleiche Tags verwendet. *Database* passt thematisch nicht dazu. Der Grund, dass es zu diesen Tags gezählt wird ist, dass *Database*

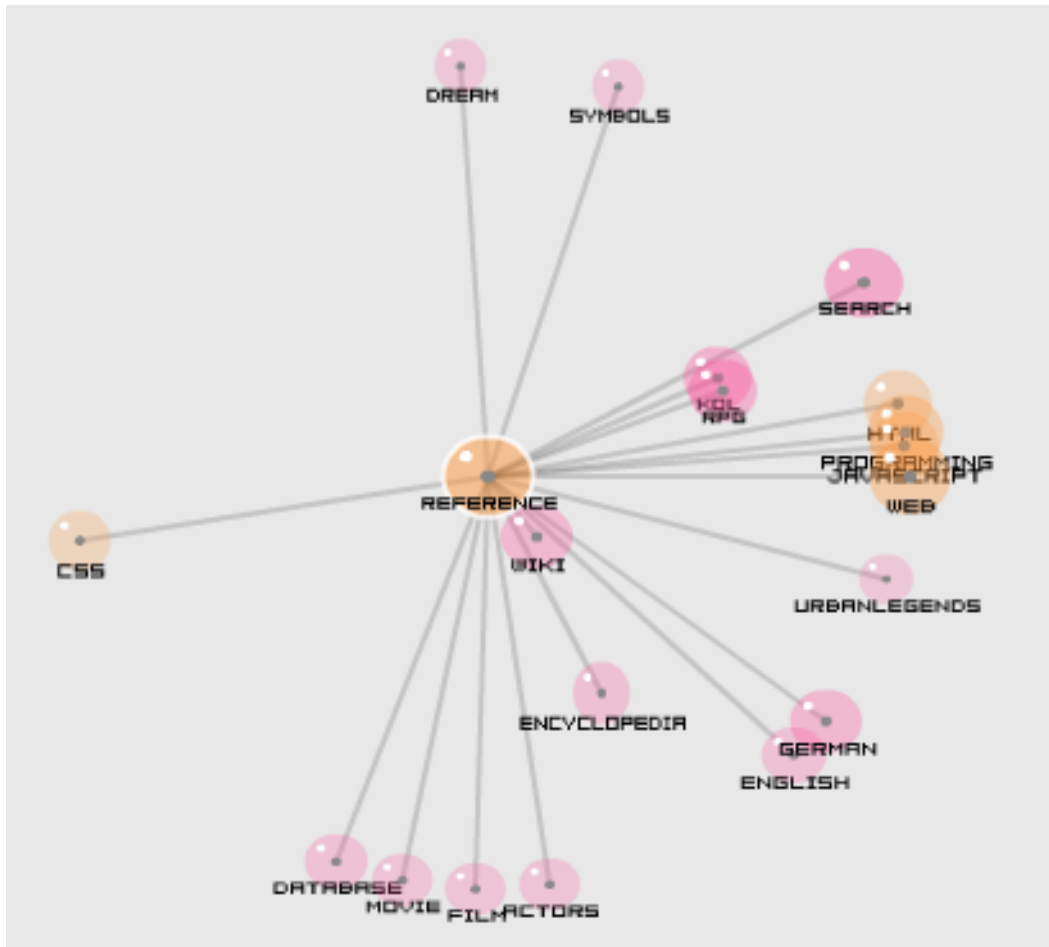


Abbildung 6.4: Del.icio.us Soup: Das Tag „Reference“ und dessen benachbarte Tags [58]

nur genau einmal verwendet wurde um die „Internet Movie Database“¹, eine Filmdatenbank, zu beschreiben.

Das Tag *Wiki* liegt zum Beispiel dem zentralen Link *Reference* sehr nahe. Das bedeutet, dass die Begriffe *Wiki* und *Reference* oft gemeinsam verwendet werden.

Wenn man die Applikation *Del.icio.us Soup* startet, werden die persönlichen Tags zunächst einfach nur aufgelistet. Durch anklicken eines Tags, wird dieses Tag zentriert und mit seinen Nachbarn gemeinsam angezeigt. In Abbildung 6.5 sieht man auf der linken Seite das Tag *Folksonomy* mit all seinen Nachbarn. Auf der rechten Seite ist das kleinere Themengebiet *Del.icio.us* dargestellt.

¹<http://www.imdb.com>



Abbildung 6.5: Del.icio.us Soup: Die Themengebiete *Folksonomy* und *Del.icio.us* [58]

6.4 Graph Del.icio.us Related Tags

Graph Del.icio.us Related Tags ist eine Visualisierung die die Beziehungen zwischen Del.icio.us Tags veranschaulicht.

Anhand der Tags Folksonomy und Ontology sieht man in Abbildung 6.6 die Visualisierung dieses Tools. Die spitz zulaufenden Verbindungslinien zwischen den Tags veranschaulichen die Nachbarschaftsverhältnisse. Die Linie beginnt bei den ausgewählten Haupttags – in diesem Fall *Ontology* und *Folksonomy* – und läuft spitz zu deren Nachbartags zu. Über die Distanz und die Helligkeit der Linien wird die Nähe visualisiert. Tags die Folksonomy nahe stehen sind „web2.0“, „tagging“, „tags“ und „tag“, entfernte benachbarte Tags sind „news“, „blog“, „art“, „science“, „bookmarks“, „bookmarking“ und „metadata“. „Metadata“ ist zusätzlich auch ein Nachbar von *Ontology*.

Graph Del.icio.us Related Tags hat noch eine Reihe weiterer Funktionen. Sobald man die Maus über einen Tag zieht, wird dieser gelb hervorgehoben. Alle Verbindungslinien zu seinen Nachbarn werden rosa hervorgehoben. Weiters erscheint ein kleines Feld „Info“ links über dem Tag; wenn man dieses anklickt erhält man die Möglichkeit den Tag in Technorati oder Del.icio.us anzusehen. Einen neuen Tag kann man hinzufügen indem man in der Suchzeile

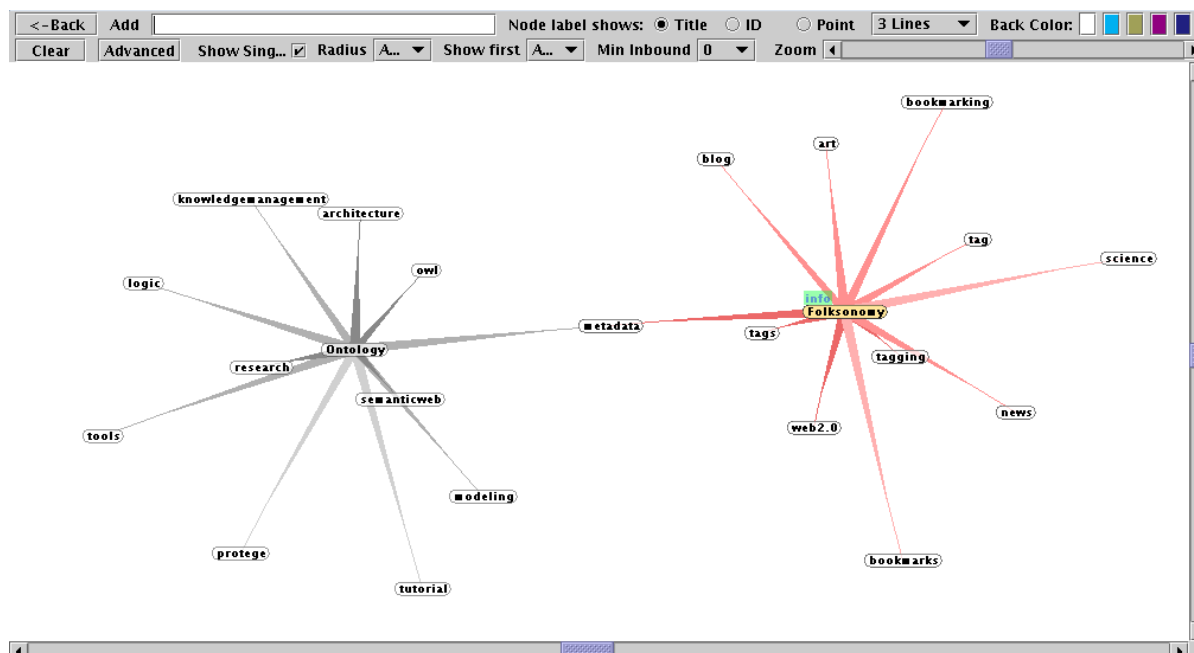


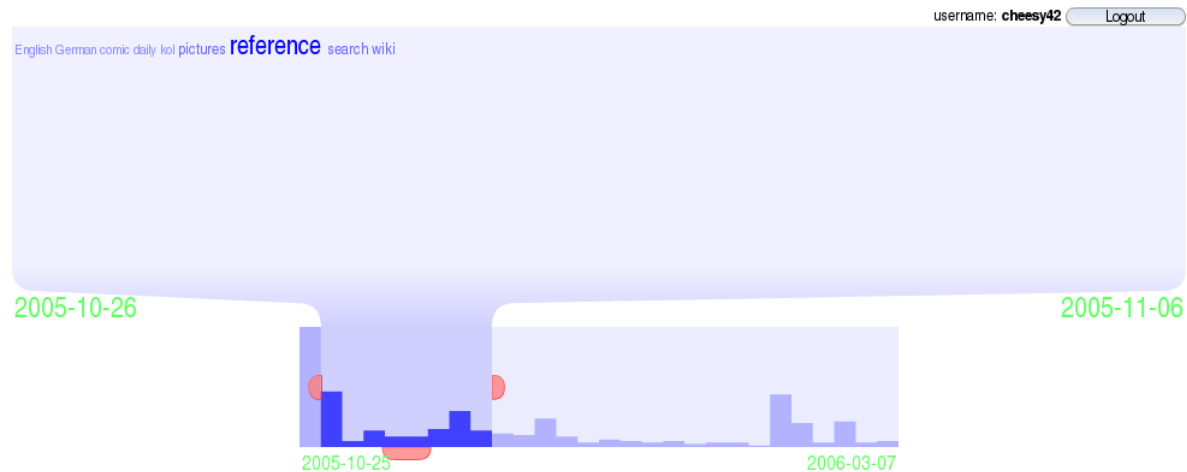
Abbildung 6.6: Graph Del.icio.us Related Tags [20]

oben links das gesuchte Wort eingibt. Man kann stattdessen aber auch den vorhandenen Graphen erforschen indem man auf einen Tag doppelklickt. sofort werden alle seine Nachbarn dem Diagramm hinzugefügt und die Tags – aufgrund von eventuell neu entstandenen Verbindungen – neu angeordnet.

Sobald man einen Tag auswählt, hat man noch weitere Optionen. Man kann zum Beispiel einen Radius angeben. Dann werden nur noch die Tags angezeigt die maximal die angegebene Anzahl an Sprüngen entfernt sind. Wenn im Graphen aus Abbildung 6.6 zum Beispiel das Tag „Knowledge management“ ausgewählt und der Radius drei gegeben wäre, würden alle Nachbarn von Folksonomy außer „metadata“ ausgeblendet werden, da sie mehr als zwei Sprünge von „Knowledge management“ entfernt sind. Alle direkten Nachbarn haben Radius eins und alle Nachbarn der direkten Nachbarn haben somit Radius zwei.

6.5 Expialidocio.us

Expialidocio.us[49] ist eine Visualisierung, der persönlichen Tagauswahl aus Del.icio.us in zeitlichem Kontext. Man sieht auf einen Blick, wann welche Tags hinzugefügt wurden.



Expialidocio.us

Abbildung 6.7: Expialidocio.us [49]

In Abbildung 6.7 sieht man eine Zeitleiste, die schematisch die Anzahl der verwendeten Tags pro Zeiteinheit darstellt. Unterhalb der Zeitleiste ist das Beginn- und Enddatum der Leiste zu sehen. Die beiden Daten die über der Zeitleiste stehen geben Beginn und Ende des ausgewählten Bereichs an. Der ausgewählte Bereich ist in der Graphik etwas dunkler dargestellt und verläuft trichterförmig nach oben. Im oberen Feld wird eine Tagwolke der in dem ausgewählten Zeitraum verwendeten Tags angezeigt. Die rosafarbenen Felder um die Auswahl herum können dazu verwendet werden den Auswahlbereich nach rechts oder links zu vergrößern, bzw. zu verschieben.

6.6 Flickr Related Tag Browser

Auch für Flickr gibt es eine große Auswahl an Visualisierungs-Tools. Eines davon ist der *Related Tag Browser*[8]. Dieser stellt ein beliebiges Tag graphisch dar und zeigt eine Auswahl an Bildern, denen dieses Tag zugeordnet ist an. Visualisiert werden die Nachbarschaft von Tags durch kreisförmige Anordnung und die Beziehung zwischen Tag und Bildern durch das in der Mitte stehende Fotoalbum.

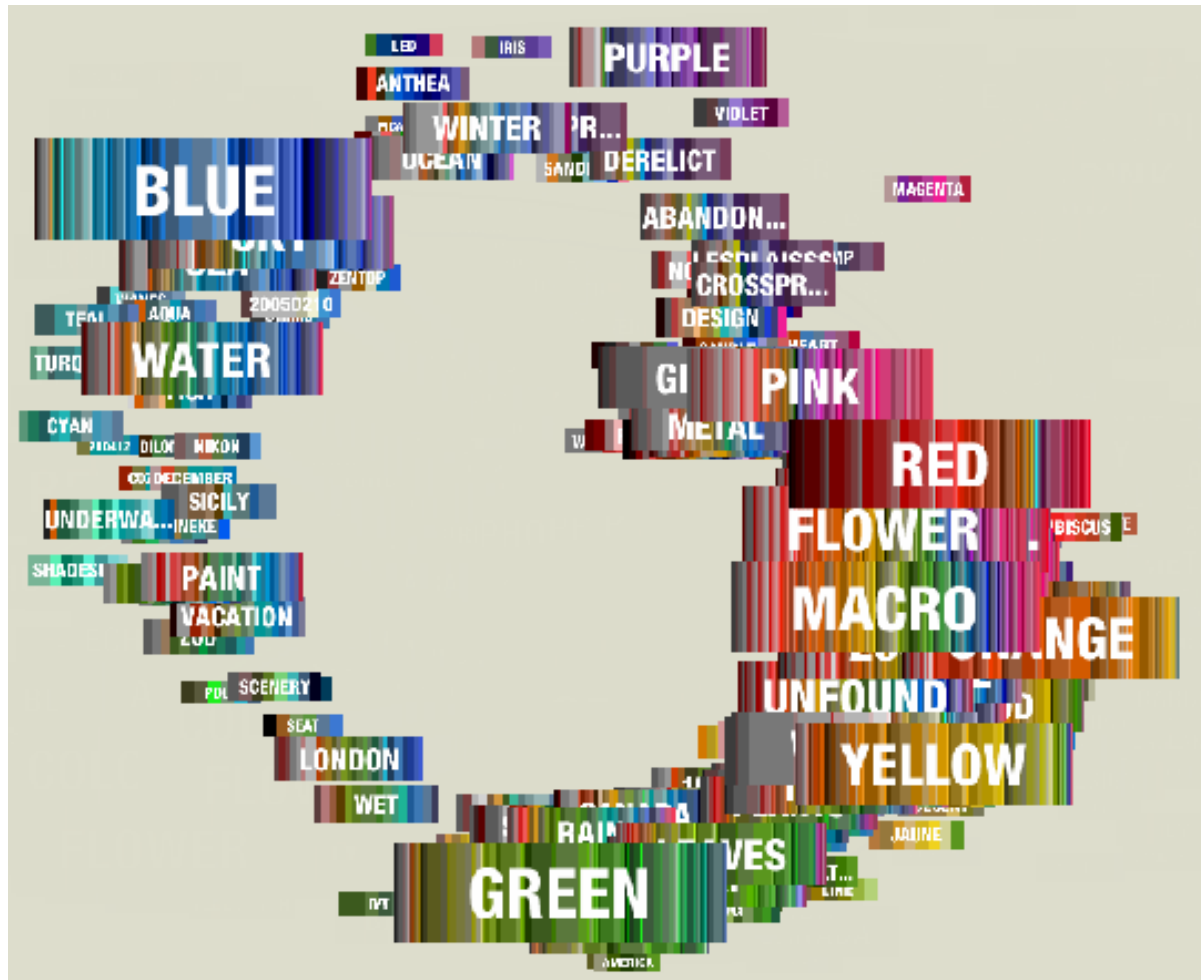


Abbildung 6.9: Flickr Colors: Visualisierung der Flickr Gruppe *Color Fields* [23]

6.7 Flickr Colors

Flickr Colors veranschaulicht die Beziehung zwischen Flickr Tags und Farben durch Unterlegung der Tags mit einer sortierten Farbpalette. Die Tags werden auch im Gesamtbild anhand dieser Farbtafeln im Kreis angeordnet. Die Beliebtheit der Tags wird durch ihre Größe reflektiert. (Siehe Abbildung 6.9.) Die Tag- und Farben-Datenbank basiert auf den Tags der Gruppe „Colorfields“[5]. Diese Flickr Gruppe enthält ausschließlich Bilder, die nur aus genau einer Grundfarbe bestehen. Die Tags sind in einem Farbkreis angeordnet. Jedes Tag liegt auf einer buntgestreiften Palette, die den Prozentsatz der einzelnen Farben angibt, so wie sie in den Bildern, die dem Tag zugeordnet sind, vorkommen.

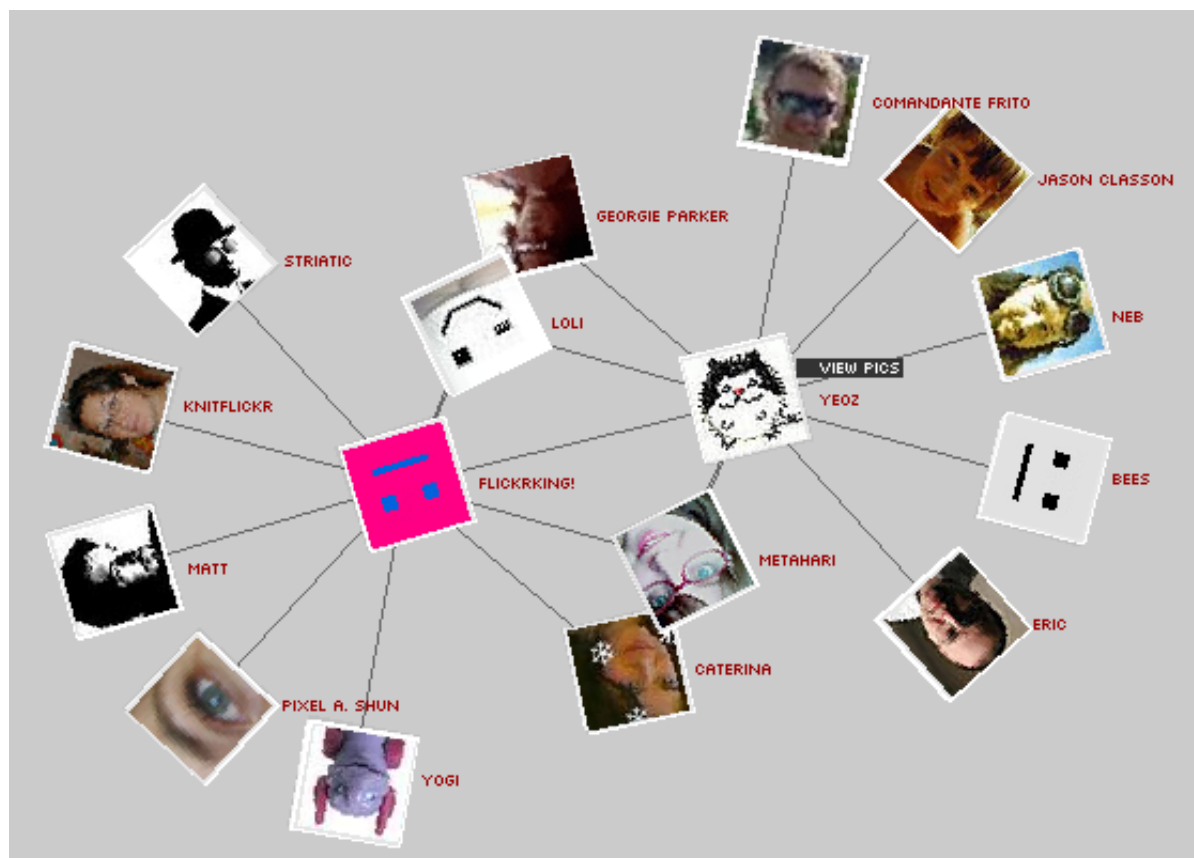


Abbildung 6.10: Flickr Graph: Visualisierung von Kontakten in Flickr [55]

6.8 Flickr Graph

Nicht nur Tags können in einer Folksonomy visualisiert werden, sondern auch Beziehungen zwischen Benutzern. In Flickr können von jedem Benutzer beliebig viele Kontakte eingetragen werden. Die Visualisierung Flickr Graph ermöglicht eine dynamische graphische Darstellung dieser Kontakte. Beginnend bei einem Benutzer kann man den Graphen erweitern, indem man einen Nachbarn anklickt. Dieser wird dann mit allen seinen Kontakten im Graphen integriert. In der Abbildung 6.10 sieht man die Benutzer „Flickrking!“ und „Yeoz“ in der Mitte. Um sie herum sind ihre Kontakte angeordnet. Sie haben vier gemeinsame Kontakte („Loli“, „Georgie Parker“, „Metahari“ und „Caterina“). Diese befinden sich in der Graphik zwischen den Beiden und sind auch durch Linien mit beiden Benutzern verbunden. Weiters haben sowohl „Flickrking!“ als auch „Yeoz“ je fünf weitere Kontakte eingetragen.

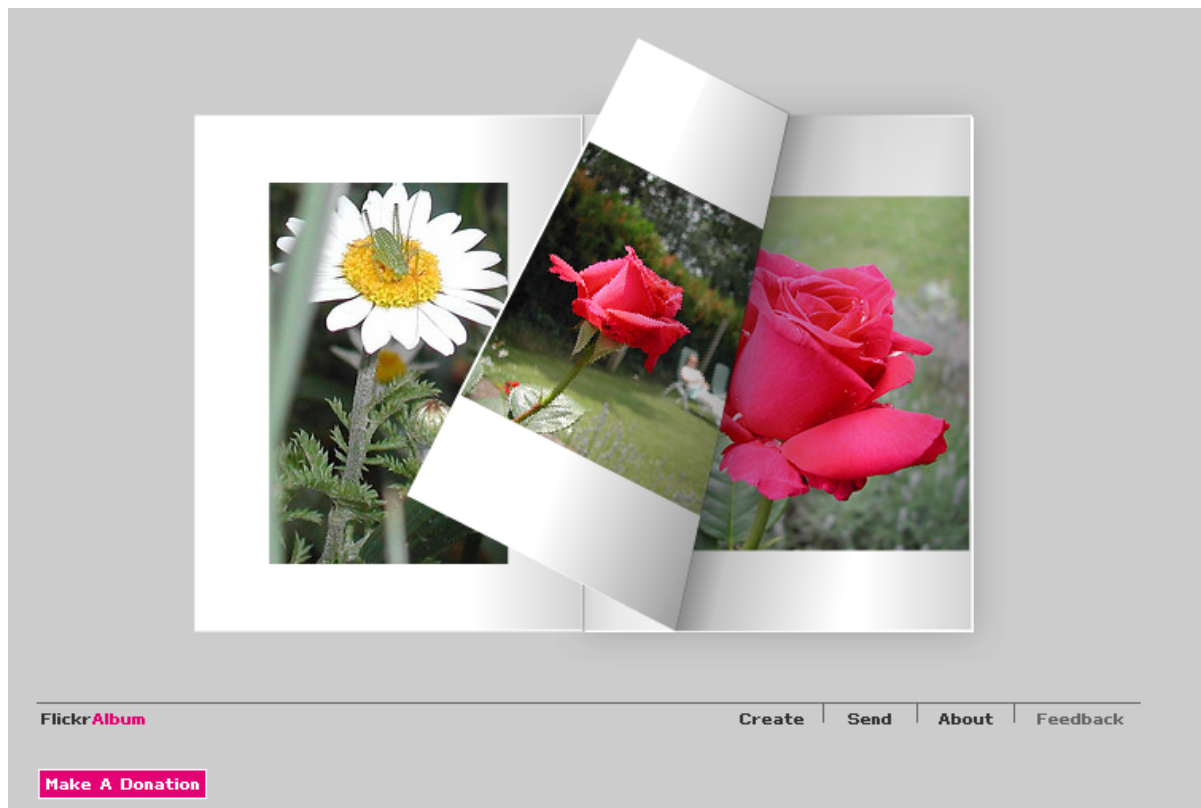


Abbildung 6.11: Flickr Album: Fotoalbum bestimmter Tags eines Benutzers [28]

6.9 Flickr Album

Flickr Album[28] ist ein Visualisierungstool, das Flickr Bilder im Stil eines Fotoalbums darstellt. Dazu gibt man einfach einen beliebigen Tag und Benutzer an, von dem man die Fotos sehen will. In Abbildung 6.11 sieht man zum Beispiel zwei Seiten des Albums des Benutzers „cheesy42“ mit dem Tag „flower“. Das Album kann wie ein echtes Fotoalbum durchblättert werden.

Kapitel 7

Zusammenfassung

Folksonomy ist ein relativ junger Ansatz zur Organisation dynamischer Daten im Web. Mit den neuen Technologien sind neue Anforderungen entstanden, die mit den althergebrachten Methoden nicht mehr bewältigbar sind. Ontologien, die auf physikalischen Objekten – wenn auch mit gewissem Aufwand – anwendbar sind und verwendet werden, sind den dynamischen, flexiblen Strukturen des 21. Jahrhunderts nicht mehr gewachsen. Dennoch wird ein System benötigt, das die Unmengen von Daten ordnen und strukturieren kann, damit wichtige Informationen nicht untergehen und schnell und einfach wiedergefunden werden können. Folksonomien bieten genau diese Art von Organisation.

Folksonomien haben viele Vorteile – sie ist einfach zu erstellen, skalierbar, benötigt kein Vorwissen und spiegelt die Konzepte und das Vokabular der Anwender wieder – sie kann aber keine hohe Präzision liefern. Eine Ontologie ist im Gegensatz dazu wesentlich genauer, vor allem was Synonymkontrolle und den Wortschatz betrifft. Trotzdem muss man bedenken, dass diese Genauigkeit und Vollständigkeit für den durchschnittlichen Webanwender kein Kriterium ist und auch oft kaum erreichbar ist. Oft kann man nicht mit hundertprozentiger Sicherheit sagen, welche Klassen die optimale Lösung für ein gegebenes Problem sind oder welche Wörter gewisse Konzepte am besten umschreiben.

Ontologien sind in vielen Gebieten der Informatik im Einsatz und leisten dort gute Dienste. Im Bereich der Datenbanken, wissensbasierten Systeme oder der künstlichen Intelligenz werden Ontologien erfolgreich verwendet. Im Web sind Ontologien oft unzulänglich und werden zunehmend durch Folksonomien abgelöst.

Folksonomien bieten aber nicht nur eine moderne und flexible Möglichkeit der Datenorganisation, sie fördern auch die Bildung sozialer Kontakte. Eine Folksonomy stellt ein komplexes soziales Netzwerk dar, in dem die Benutzer über gemeinsame Tags ihre gemeinsamen Interessen erkennen können und dadurch neue Kontakte knüpfen können.

Folksonomien bieten den Benutzern eine vollständige Entscheidungsfreiheit. Diese Freiheit führt zu einem starken Ungleichgewicht in der Beliebtheit der Tags, Benutzer und der getaggten Objekte, die sich in einer Verteilung gemäß des Power Laws äussert. Diese Power Law Verteilung ist ganz natürlich für ein System, wie die Folksonomy. Sobald viele Menschen vereint werden und freie Entscheidungsgewalt haben, bildet sich eine kleine Elite heraus, die sich stark von der breiten Masse abhebt.

Anwendungen von Folksonomien sind neben Weblog-Systemen und Bookmark-Managern vor allem die Organisation audiovisueller Daten. Objekte, wie Bilder oder Musik, die von sich aus keinen Text enthalten kann durch die Anwendung von Folksonomien eine weitere Informationsebene hinzugefügt werden. Wenn multimediale Daten mit Tags umschrieben werden, können sie viel leichter organisiert und wiedergefunden werden und sind auch für andere Personen leicht auffindbar.

Durch die Dynamik und das starke Wachstum des Mediums Internet, wird auch das Konzept der Folksonomy immer verbreiteter. Bereits jetzt ist die Tendenz zu erkennen, dass viele Webseiten, die Daten in irgendeiner Weise zu organisieren haben, die Möglichkeit des Taggings oder sogar des Free Taggings anbieten.

Abbildungsverzeichnis

2.1	Ontologie der Tierwelt	4
2.2	Kategorisierung der Realität von John Wilkins [57]	7
2.3	ANSI/SPARC Drei Schema Architektur eines Datenbanksystems	14
3.1	Annotationen und Bookmarks in Annotea [31]	21
3.2	Breite Folksonomy [53]	25
3.3	Enge Folksonomy [53]	25
3.4	Histogramm: Körpergrößen von Menschen – Gauss Verteilung [41]	27
3.5	Histogramm: Populationsgrößen von Städten – Power Law Verteilung [41]	27
3.6	Populationsgrößen von Städten mit logarithmischen Achsen [41]	28
3.7	Verteilung von Tags pro URL in Del.icio.us [30]	29
3.8	Die 100 meistverlinkten Weblogs auf Technorati [33]	29
5.1	Del.icio.us: personalisierte Hauptseite vom 3.3.2006 [6]	39
5.2	Flickr: personalisierte Hauptseite vom 3.3.2006 [7]	42
5.3	Furl: personalisierte Hauptseite vom 5.3.2006 [9]	45
5.4	Technorati: personalisierte Watchlist vom 5.3.2006 [13]	46
5.5	CiteULike: personalisierte Zitatesammlung vom 5.3.2006 [4]	48
5.6	43 Things: personalisierte Hauptseite vom 5.3.2006 [2]	49
6.1	Tagwolke der häufigsten Tags in Del.icio.us [6]	53
6.2	Del.icio.us Mindmap der Tags eines Benutzers [48]	55
6.3	Del.icio.us Mindmap: Cluster Folksonomy [48]	56
6.4	Del.icio.us Soup: Das Tag „Reference“ und dessen benachbarte Tags [58]	58

6.5	Del.icio.us Soup: Die Themengebiete <i>Folksonomy</i> und <i>Del.icio.us</i> [58]	59
6.6	Graph Del.icio.us Related Tags [20]	60
6.7	Expialidocio.us [49]	61
6.8	Flickr Related Tag Browser [8]	62
6.9	Flickr Colors: Visualisierung der Flickr Gruppe <i>Color Fields</i> [23]	63
6.10	Flickr Graph: Visualisierung von Kontakten in Flickr [55]	64
6.11	Flickr Album: Fotoalbum bestimmter Tags eines Benutzers [28]	65

Literaturverzeichnis

- [1] *43 Places*. <http://www.43places.com/>.
- [2] *43 Things*. <http://www.43things.com/>.
- [3] *Annotea Project*. <http://www.w3.org/2001/Annotea/>.
- [4] *CiteULike*. <http://www.citeulike.org/>.
- [5] *Colorfields*. <http://www.flickr.com/groups/colorfields/>.
- [6] *del.icio.us*. <http://del.icio.us/>.
- [7] *flickr*. <http://www.flickr.com/>.
- [8] *Flickr Related Tag Browser*. http://www.airtightinteractive.com/projects/related_tag_browser/app/.
- [9] *Furl*. <http://www.furl.net/>.
- [10] *K-CAP 2001 - First International Conference on Knowledge Capture*.
- [11] *Kulturkritisches Lexikon*. <http://www.kulturkritik.net/begriffe/>.
- [12] *Project Gutenberg*. <http://www.gutenberg.org/>.
- [13] *Technorati*. <http://www.technorati.com/>.
- [14] *WordNet*. <http://wordnet.princeton.edu/>.
- [15] *Biographisch-Bibliographisches Kirchenlexikon, Band XVIII*. Traugott Bautz, Nordhausen, 2001. <http://www.bautz.de/bbkl/>.

- [16] ARISTOTELES: *Kategorien*. <http://www.aristoteles-heute.de/>.
- [17] BLACK, PAUL E.: *Dictionary of Algorithms and Data Structures*. <http://www.nist.gov/dads/>, 1998.
- [18] BROWN, DAN: *On Freetagging*. <http://www.greenonions.com/index.php?p=148>, 2005.
- [19] CORAZZON, RAUL: *Ontology. A Resource Guide for Philosophers*. <http://www.formalontology.it/>, 2000-2005.
- [20] EATON, ALF: *Graph Del.icio.us Related Tags*. <http://hublog.hubmed.org/archives/001049.html>, 2005.
- [21] EISLER, RUDOLF: *Wörterbuch der philosophischen Begriffe*. <http://www.textlog.de/eisler.html>, 1904.
- [22] FISCHER, PETER und PETER HOFER: *Lexikon der Informatik*. 6 Auflage. <http://www.ictswitzerland.ch/de/ict-wissen/lexikon.asp>.
- [23] GIORDANO, CHRISTIAN: *Flickr Colors*. http://nuthinking.com/did/tagged_colors_03/.
- [24] GRAHAM, AMY: *Furl and Del.icio.us: Almost Perfect Together*. <http://blog.contentious.com/>, 2005.
- [25] GROSS, JONATHAN L. und JAY YELLEN: *Handbook of graph theory*. 2004.
- [26] HAMMOND, TONY, TIMO HANNAY, BEN LUND und JOANNA SCOTT: *Social Bookmarking Tools (I) - A General Review*. D-Lib Magazine, 11, 2005.
- [27] HOWE, DENNIS: *Free On-Line Dictionary of Computing*. <http://www.foldoc.org/>, 1993.
- [28] JONES, HENRY: *Flickr Album*. <http://www.flickralbum.com/>.
- [29] KANT, IMMANUEL: *Prolegomena zu einer jeden künftigen Metaphysik die als Wissenschaft wird auftreten können*. 1783. <http://uni-potsdam.de/u/philosophie/texte/prolegom/prolegom.htm>.

- [30] KELLER, PHILIPP: *deli.ckoma - delicious aroma*. <http://deli.ckoma.net/stats>.
- [31] KOIVUNEN, MARJA-RIITTA: *Annotea and Semantic Web Supported Collaboration*. 2001.
- [32] KOIVUNEN, MARJA-RIITTA und RALPH R. SWICK: *Metadata Based Annotation Infrastructure offers Flexibility and Extensibility for Collaborative Applications and Beyond*. 2001.
- [33] KOTTKE, JASON: *Weblogs and Power Laws*. <http://www.kottke.org/03/02/weblogs-and-power-laws>, 2003.
- [34] MATHES, ADAM: *Folksonomies - Cooperative Classification and Communication Through Shared Metadata*. <http://www.adammathes.com/>, 2004.
- [35] MCCARTHY, JOHN: *Circumscription - A Form of Non-Monotonic Reasoning*. 1980.
- [36] MERHOLZ, PETER: *Metadata for the Masses*. <http://www.adaptivepath.com/>, 2004.
- [37] MERHOLZ, PETER: *Mob indexing? Folk categorization? Social tagging?* <http://www.peterme.com/archives/000444.html>, 2005.
- [38] MERRIAM-WEBSTER: *Merriam-Webster's Collegiate® Dictionary, Tenth Edition*. <http://www.m-w.com>, 1998.
- [39] MIKA, PETER: *Ontologies Are Us: A Unified Model of Social Networks and Semantics*. 2005.
- [40] MÖLLER, PETER: *Peter Möller's Philo Lex*. <http://www.philolex.de/>.
- [41] NEWMAN, MARK E. J.: *Power laws, Pareto distributions and Zipf's law*. *Contemporary Physics*, 46, 2005.
- [42] PFEIFER, GERALD, MICHAEL SCHREFL, KATRIN SEYR und MARKUS STUMPTNER: *Skriptum zur Vorlesung DATENBANKSYSTEME*.
- [43] PLATO: *Sophist*. The Project Gutenberg, 360 v.Chr. <http://www.gutenberg.org/>.
- [44] QUINTARELLI, EMANUELE: *Folksonomies: power to the people*. <http://www.iskoi.org/doc/folksonomies.htm>, 2005.
- [45] SHIRKY, CLAY: *Ontology is Overrated: Categories, Links, and Tags*. http://shirky.com/writings/ontology_overrated.html.

- [46] SHIRKY, CLAY: *Power Laws, Weblogs, and Inequality*. http://www.shirky.com/writings/powerlaw_weblog.html, 2003.
- [47] SMITH, BARRY: *Ontology and Information Systems*. Blackwell Guide to the Philosophy of Computing and Information, 2003.
- [48] SPERONI, PIETRO: *Del.icio.us Mind Map Maker*. <http://maps.pietrosperoni.it/delicious/makemap.html>.
- [49] STEELE, OLIVER: *Expialidocio.us*. <http://expialidocio.us/>, 2005.
- [50] STERLING, BRUCE: *Order Out of Chaos: What's the best way to tag, bag and sort data? Give it to the unorganized masses*. <http://www.wired.com>, 2005.
- [51] TOMPITS, HANS: *Vorlesung Nonmonotone Logik/Nichtmonotones Schliessen*. <http://www.kr.tuwien.ac.at/>.
- [52] UDELL, JOHN: *Collaborative knowledge gardening: With Flickr and del.icio.us, social networking goes beyond sharing contacts and connections*. <http://www.infoworld.com/>, 2004.
- [53] VANDER WAL, THOMAS: *Thomas Vander Wal Blog*. <http://www.vanderwal.net>.
- [54] VANDER WAL, THOMAS: *Tagging for Fun and Finding*. <http://www.ok-cancel.com>, 2005.
- [55] WESKAMP, MARCOS: *Flickr Graph*. <http://www.marumushi.com/apps/flickrgraph/>, 2005.
- [56] WIEDEMANN, UWE: *Philosophenlexikon*. <http://www.philosophenlexikon.de/>.
- [57] WILKINS, JOHN: *An Essay Toward a Real Character and a Philosophical Language*. 1668. <http://reliant.teknowledge.com/Wilkins/>.
- [58] ZITVOGEL, OLIVIER: *Del.icio.us Soup*. <http://www.zitvogel.com/delicioussoup/>, 2005.