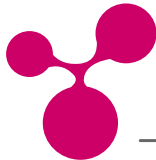


Technische Universität Dresden – Fakultät Informatik
Professur für Multimedialechnik, Privat-Dozentur für Angewandte Informatik

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



GENEME '10

GEMEINSCHAFTEN IN NEUEN MEDIEN

an der
Fakultät Informatik der Technischen Universität Dresden

mit Unterstützung der

3m5. Media GmbH, Dresden
ANECON Software Design und Beratung GmbH, Dresden
Communardo Software GmbH, Dresden
GI-Regionalgruppe, Dresden
itsax.de | pludoni GmbH, Dresden
Kontext E GmbH, Dresden
Medienzentrum der TU Dresden
objectFab GmbH, Dresden
SALT Solutions GmbH, Dresden
SAP AG, Resarch Center Dresden
Saxonia Systems AG, Dresden
T-Systems Multimedia Solutions GmbH, Dresden

am 07. und 08. Oktober 2010 in Dresden

www.geneme.de
info@geneme.de

B.5 DeViz: Untersuchen von Visualisierungsformen durch eine Klassifizierung beruhend auf Social Tagging

Mandy Keck¹, Dietrich Kammer¹, Jan Wojdziak¹, Severin Taranko²,
Rainer Groh¹

¹Technische Universität Dresden, Institut für Software- und
Multimediatechnik
²queo GmbH

1 Einleitung und Motivation

Die Möglichkeiten der computergestützten Informationsvisualisierung haben zu einer Flut von unterschiedlichsten Visualisierungsformen geführt. Diese nutzen etablierte Techniken der Informationsvisualisierung, kombinieren diese miteinander oder beruhen auf völlig neuen Konzepten (vgl. [2], [3], [4]).

Zum Finden passender Visualisierungsformen für einen gegebenen Anwendungskontext eignen sich besonders semantische Relationen, die durch kollektive Aktivitäten im Web 2.0 entstehen. Offene Plattformen nutzen dafür das Prinzip des *Social Tagging*. Dies ist eine Form der freien Verschlagwortung, bei der Nutzer Inhalte ohne syntaktische oder semantische Vorgaben mit Schlagworten beschreiben und verknüpfen. Die Summe dieser durch die Nutzer indizierten Inhalte wird im Web 2.0 als die Weisheit der Vielen (engl. *wisdom of crowds*) bezeichnet [1].

In diesem Beitrag wird ein Klassifikationsschema als Regelwerk für das Social Tagging vorgeschlagen, um eine Klassifizierung von Visualisierungsformen zu ermöglichen. Auf diesem Schema aufbauend wird die Meta-Visualisierung DeViz (*Deep exploration and lookup of Visualizations*) entwickelt. Diese verwendet Delicious [5] als offene Social-Bookmarking-Plattform mit der eine Datenbasis an Visualisierungen kollaborativ erstellt und indiziert wird. Durch DeViz werden Zusammenhänge von Visualisierungen sichtbar und damit eine Analyse dieser möglich.

2 Verwandte Arbeiten

Unter einer Informationsvisualisierung wird nach [6] die interaktive computergestützte visuelle Repräsentation von abstrakten Daten verstanden.

Um die Komplexität einer Informationsvisualisierung zu beschreiben, sind verschiedene Dimensionen zu berücksichtigen. Hierzu werden die Eigenschaften der zugrundeliegenden Daten, der visuellen Repräsentation und der verwendeten Interaktionstechniken betrachtet. Bevor im Rahmen dieses Beitrags der inhaltliche Fokus auf die Klassifikation sowie die darauf aufbauende Meta-Visualisierung gelenkt wird, werden relevante Arbeiten der bisherigen Forschung verdichtet dargestellt. Diese werden sowohl unter Berücksichtigung der zugrundeliegenden

Klassifikationsmethode als auch der beinhalteten Dimensionen auf ihre Eignung hin überprüft. Das angestrebte Ziel ist es, die Betrachtung von Informationsvisualisierungen aus mehreren Perspektiven zu ermöglichen.

Eine häufig verwendete Methode zur Klassifikation von Visualisierungen ist die Einteilung anhand einer zugrundeliegenden Taxonomie. Diese organisiert Objekte eines Wissensbereiches in einer monohierarchischen Struktur, in der sie eindeutig einzelnen Klassen zugeordnet sind.

Durch die Visualisierungstaxonomie „Task by Data Type Taxonomy for Information Visualization“ nach Shneiderman [7] werden verschiedene Datentypen und Aufgaben identifiziert, welche durch die Informationsvisualisierung Unterstützung finden. Die Datentypen „one-“, „two-“, „threedimensional“, „temporal“, „multi-dimensional“, „tree“ und „network data“ ermöglichen infolgedessen das Klassifizieren von Visualisierungen.

Ferner stellt die Webseite „ManyEyes“ [3] von IBM eine Einteilung dar, die verschiedene Visualisierungen in einer Hierarchie organisiert. Dies erfolgt anhand der Aufgabe, die mit der jeweiligen Visualisierung verbunden ist (siehe Abbildung 1, links). Durch Taxonomien werden die Visualisierungen fest in Strukturen eingeordnet, die nur eine Sichtweise auf den Datenbestand zulassen und daher für die eingangs dargelegten Anforderungen nicht geeignet sind.

Eine weitere Klassifikationsmethode ist die Facettenklassifikation, welche die einzuordnenden Objekte durch die Kombination verschiedener Facetten beschreibt. Diese können als Dimensionen in einem kartesischen n-dimensionalen Raum abgebildet werden. Die Ausprägung einer Facette spiegelt sich auf der jeweiligen Achse wider [8]. Dadurch wird ein Objekt genau durch eine Ausprägung jeder Facette beschrieben [9]. Das Periodensystem der Visualisierungsmethoden für Manager klassifiziert Visualisierungen anhand dieser Methode [4]. Diese Klassifikation bietet eine systematische Übersicht bestehender Darstellungsformen anhand der Dimensionen „Complexity of Visualization“, „Content Area“, „Point of View“, „Type of Thinking“ und „Type of Representation“. Die Ausprägungen dieser Facetten werden durch visuelle Variablen wie die Position im Periodensystem, die Farbe und Symbole dargestellt (siehe Abbildung 1, Mitte).

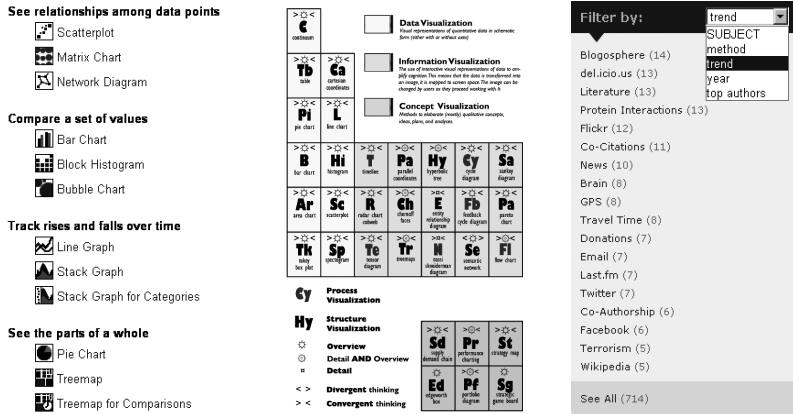


Abbildung 1: verschiedene Klassifikationen von Visualisierungen: ManyEyes [3] (links), Periodensystem der Visualisierungsmethoden [4] (Mitte), visual complexity [2] (rechts)

Die Onlinedatenbank Visual Complexity [2] nutzt ebenfalls eine facettenorientierte Klassifikation. Die Visualisierungen werden durch die Kategorien „subject“, „method“, „trend“, „year“ und „top authors“ beschrieben (siehe Abbildung 1, rechts). Es handelt sich jedoch um keine Facettenklassifikation im engeren Sinne, weil die Festlegung einer Ausprägung für eine Facette nicht zwangsläufig ist. Eine Facettenklassifikation lässt verschiedene Sichtweisen auf die Visualisierungen zu. Jedoch wird aus [2] ersichtlich, dass eine genaue Zuordnung zu jeder Facette nicht immer möglich ist. Die Erstellung einer Facettenklassifikation, die ausschließlich eine Zuordnung pro Facette voraussetzt, ist unter Umständen nicht ausreichend, da die einzuordnenden Objekte oft nicht genau einer Ausprägung einer Facette zuzuordnen sind [10]. Darüber hinaus sind Visualisierungen nicht uneingeschränkt durch die gleichen Dimensionen beschreibbar, weil die Bandbreite von Visualisierungen von einfachen Konzeptideen bis hin zu komplexen Anwendungen reicht. Ebenfalls ist durch die Mehrdeutigkeit einzelner Eigenschaften eine genaue Einordnung in eine Facette oft nicht möglich. Ein Beispiel hierfür stellt die Visualisierung von Text- und Bilddaten dar, die zwei verschiedene Datentypen beschreiben.

Eine mögliche Lösung für dieses Problem bietet die kollaborative Verschlagwortung (engl. *Collaborative Tagging* bzw. *Social Tagging*). Die Sammlung aller Schlagworte, die bei diesem Prozess entsteht, wird als Folksonomy bezeichnet und kann als nutzergenerierte Klassifikation [8] angesehen werden. Im Gegensatz zu einer Facettenklassifikation oder einer Taxonomie muss der Nutzer nicht die Entscheidung treffen, ob die einzuordnende Informationseinheit einer vordefinierten Klasse

zugehörig ist [11]. Durch die Zuordnung beliebig vieler Schlagworte kann die Informationsressource beschrieben werden. Eine Indexierung ohne Regelwerk führt jedoch auch zu Inkonsistenzen in der Folksonomy. Mehrdeutige oder persönliche Schlagworte können die Vergleichbarkeit der Informationseinheiten verringern (vgl. [8], [11]). Angewandt auf die Thematik der Klassifikation von Visualisierungen erfüllt diese Klassifikationsmethode die Anforderung, verschiedene Sichtweisen auf den Datenbestand zuzulassen. Aufgrund des fehlenden Regelwerks bietet diese jedoch nur eingeschränkte Vergleichsmöglichkeiten der Visualisierungen untereinander.

3 Zwischen Facetten und Folksonomies

Facettenklassifikation und Social Tagging werden im Folgenden betrachtet, um die verschiedenen Dimensionen von Informationsvisualisierungen zu beschreiben. Wie im vorherigen Abschnitt gezeigt, erlauben diese im Gegensatz zu monohierarchischen Strukturen eine Analyse aus verschiedenen Perspektiven. Um eine große Bandbreite von Visualisierungen zu strukturieren, erweist sich die Facettenklassifikation mit der Zuordnung von genau einer Ausprägung je Facette als ungeeignet. Diese muss flexibler verwendet werden, weil andernfalls eine unvollständige Zuordnung zum Ausschluss aus der Klassifikation führen würde. Die Verwendung des Social Tagging in der in Abschnitt 2 beschriebenen Form besitzt ebenfalls Nachteile, da durch zu viele ähnliche Schlagworte die Vergleichbarkeit und damit eine Analyse der Visualisierungen schwer umsetzbar ist. Daher wird nachfolgend ein Ansatz vorgestellt, der sich Aspekten beider Klassifikationsmethoden bedient. Das Konzept der Facetten wird gewählt, um verschiedene Dimensionen einer Informationsvisualisierung abzubilden. Innerhalb der Facetten wird das Prinzip der freien Verschlagwortung genutzt, um eine flexiblere Zuordnung zu ermöglichen. In Abbildung 2 ist das entwickelte Schema dargestellt.

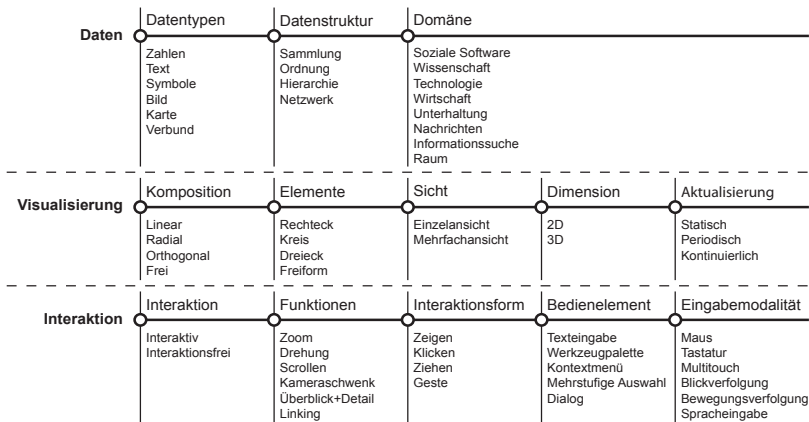


Abbildung 2: Klassifikationsschema aufbauend auf Daten, Visualisierungs- und Interaktionsmechanismen

Mit dem dargestellten Klassifikationsschema ist es möglich, bestehenden Visualisierungen verschiedene Schlagworte zuzuordnen. Es enthält verschiedene Dimensionen von Informationsvisualisierungen, die aus der Analyse bestehender Visualisierungsklassifikationen hervorgehen. Die Dimensionen werden in die Ebenen „Daten“, „Visualisierung“ und „Interaktion“ gegliedert. Diese Einteilung erfolgt in Anlehnung an das Referenzmodell der Informationsvisualisierung nach [6]. Die Dimensionen „Datentyp“, „Datenstruktur“ und „Domäne“, aus denen die visualisierten Daten stammen, werden durch die Ebene „Daten“ gekapselt. Die Dimension „Domäne“ ist beispielsweise mit der Facette „subject“ von visual complexity vergleichbar (vgl. Abbildung 1). Den Dimensionen sind verschiedene Ausprägungen, also die konkreten Schlagworte, zugeordnet. Durch die Dimensionen werden diese zusammengefasst und gleichzeitig in einen Kontext gesetzt. Dadurch wird die Übersichtlichkeit über die zahlreichen Schlagworte im Gegensatz zum Social Tagging erhöht. Beispielsweise enthält die Dimension „Datenstruktur“ die Ausprägungen „Sammlung“, „Ordnung“, „Hierarchie“ und „Netzwerk“. Weiterhin wird „Überblick und Detail“ aus dem Periodensystem der Visualisierung [4] mit weiteren Funktionen zur Interaktion in einer Dimension gruppiert.

Sowohl die Wahl der Dimensionen als auch der zugeordneten Ausprägungen wurde empirisch vorgenommen und muss in einer Evaluation überprüft werden. Das aktuelle Schema stellt einen ersten Ansatz dar, um Visualisierungen nach verschiedenen Dimensionen zu beschreiben und Social Tagging auf Grundlage eines vorgegebenen Schemas für deren Strukturierung zu nutzen. Die Möglichkeit das Schema durch neue Dimensionen und Schlagworte zu erweitern, ist aufgrund der flexiblen Struktur gegeben und ist für die Weiterentwicklung vorgesehen (siehe Abschnitt 5).

Im Gegensatz zu einer Facettenklassifikation können einer Visualisierung mehrere Schlagworte aus einer Gruppe zugeordnet werden. Es ist ebenso möglich, auf die Zuordnung von Schlagworten aus einer Gruppe komplett zu verzichten. Dies ist beispielsweise notwendig, falls Visualisierungen interaktionsfrei sind und somit keine beschreibbaren Interaktionsformen bieten. In dem vorliegenden Schema werden diese neben den interaktiven Informationsvisualisierungen berücksichtigt, da interaktionsfreie Visualisierungen ebenfalls eine große Vielfalt an Visualisierungskonzepten enthalten.

4 DelViz: Deep exploration and lookup of Visualizations

Aufbauend auf dem Klassifikationsschema wurde die Meta-Visualisierung DelViz (*Deep exploration and lookup of Visualizations*) konzipiert. Diese stellt die Beziehungen zwischen Visualisierungsformen dar, die durch Zuordnung gleicher Ausprägungen innerhalb der Dimensionen entstanden sind. Der Aufbau der Datenbasis sowie die Konzeption und Funktionsweise des DelViz-Werkzeugs werden im Folgenden näher beschrieben.

4.1 Aufbau der Visualisierungsdatenbank

Auf Basis des vorgestellten Schemas wurden 700 verschiedene Visualisierungen klassifiziert. Dabei wurden sowohl Fundstücke aus dem Web als auch Einträge der Datenbank visual complexity [2] genutzt, die bereits ein breites Spektrum an Visualisierungsformen anbietet. Durch die Veröffentlichung im Web verfügen alle betrachteten Visualisierungen über einen Link. Für eine Social Software zum kollaborativen Aufbau der Datenbasis besteht damit die Anforderung, diese Links verwalten und indexieren zu können. Der Webservice Delicious [5] zählt zu den populärsten Social-Bookmarking-Diensten [12]. Web-Lesezeichen können durch Social Bookmarking gemeinschaftlich verwaltet und anderen Nutzern zur Verfügung gestellt werden. Delicious bietet die Möglichkeit, jedem Lesezeichen einen Titel, eine Beschreibung und verschiedene Schlagworte zuzuordnen, die in sogenannten *Bundles* gruppiert werden können. Weiterhin bietet Delicious eine für Entwickler offene Schnittstelle (API), welche die klassifizierten Visualisierungen für das DelViz-Werkzeug zur Verfügung stellt. Die Visualisierungen wurden in einer Gruppe von vier Bearbeitern anhand des Schemas im Rahmen eines Forschungsprojektes klassifiziert. Die Öffnung der Datengrundlage im Internet ist geplant (vgl. Abschnitt 5). Dadurch kann die Weisheit der Vielen beim Finden versteckter Relationen vollständig genutzt werden.

4.2 Konzeption des DelViz-Werkzeugs

Um die Exploration der klassifizierten Visualisierungen zu ermöglichen, wird ein Werkzeug benötigt. Die Zuordnung derselben Ausprägungen zu verschiedenen Visualisierungen führt zu Beziehungen, die visualisiert werden sollen. Delicious [5] bietet bereits eine Visualisierung an, um eine Verteilung der Schlagworte innerhalb der Folksonomy dazustellen. Die Schlagworte werden in Form einer interaktiven Schlagwortliste und in einer „TagCloud“ dargestellt.

Jedoch werden darin die Beziehungen der verschlagworteten Daten untereinander nicht ersichtlich (vgl. Abbildung 3, links). Das Visualisierungskonzept „Elastic Tag Maps“ [13] visualisiert ebenfalls Schlagwortmengen, die in Delicious entstanden sind. Dabei werden in Beziehung stehende Schlagworte durch Kurven miteinander verbunden (vgl. Abbildung 3, rechts), wodurch diese Beziehungen schnell erfasst werden können. Jedoch findet in dieser Darstellung die Bundle-Struktur, die bei der Klassifikation der Visualisierungen für die Dimensionen genutzt wurde, keine Berücksichtigung.



Abbildung 5: Schlagwortliste im DelViz-Werkzeug bei der Auswahl ‚interaktionsfrei‘ und ‚2d‘ (links) und der iterativen Verfeinerung der Suche (rechts)

Anwendungsfall 1: Passende Visualisierungen für einen gegebenen Kontext finden

Der erste Anwendungsfall stellt die Suche nach einer geeigneten Informationsgrafik dar, um die Umsatzzahlen des letzten Jahres zu visualisieren. Durch die Auswahl der Ausprägung „2d“ und „interaktionsfrei“ in der Schlagwortliste (vgl. Abbildung 5, links) werden Verbindungen zu weiteren Ausprägungen wie beispielsweise „Text“ und „Zahlen“ in der Dimension „Datentyp“ oder „Freiform“ und „Kreis“ in der Dimension „Elemente“ durch die thread arcs dargestellt. Es werden so bereits bei der Selektion einzelner Ausprägungen Vorschläge angeboten, die in die Suchanfrage mit einbezogen werden können. Durch das Hinzunehmen weiterer Ausprägungen kann die Suche so iterativ verfeinert werden (vgl. Abbildung 5, rechts), um die Ergebnismenge einzuzugrenzen. Die gefundenen Visualisierungen können für den gegebenen Anwendungsfall eine gute Darstellungsmöglichkeit bieten oder Inspiration für die Erstellung einer neuen Visualisierung aufzeigen, indem die selektierten Eigenschaften einbezogen werden.

Anwendungsfall 2: Analyse von Visualisierungs- und Interaktionsformen

Ein weiterer Anwendungsfall zeigt die Unterstützung des Werkzeugs bei der Analyse von bestehenden Visualisierungs- und Interaktionsformen, um Innovationspotential zu evaluieren. Durch die Auswahl der Ausprägung „Multitouch“ in der Dimension „Eingabemodalität“ werden die thread arcs zu den damit in Beziehung stehenden Visualisierungsformen dargestellt. Sehr stark in Beziehung stehende Ausprägungen, wie die Dimensionsausprägung „2D“, werden hervorgehoben. Diese starke Verbindung kann darauf hindeuten, dass es schon viele erprobte Visualisierung mit dieser Interaktionsform gibt und die Kombination daher sinnvoll ist. Genauso können wegfallende oder selten genannte Ausprägung neue Fragestellungen aufwerfen. So wurde „Multitouch“ beispielsweise seltener in Zusammenhang mit der Ausprägung „3d“ verschlagwortet. So kann einerseits die Sinnhaftigkeit dieser Kombination in Frage gestellt werden, andererseits könnte dies auch auf eine „Marktlücke“ schließen

lassen, in der noch nicht viele Anwendungen existieren. Zur Entwicklung von Konzepten für neue Visualisierungsformen, die bewusst andere Wege gehen, stellt dies eine Ausgangsbasis dar.

5 Zusammenfassung und Ausblick

In diesem Beitrag wurde ein Visualisierungskonzept vorgestellt, das die Analyse von Visualisierungsformen aus verschiedensten Blickwinkeln durch eine strukturierte Menge von Schlagworten ermöglicht. Grundlage ist ein multidimensionales Klassifikationsschema als Regelwerk für das Social Tagging. Auf Basis dieser Klassifikation wurde das DelViz-Werkzeug konzipiert. Dieses ermöglicht es, klassifizierte Visualisierungen zu analysieren und deren Zusammenhänge sichtbar zu machen.

Auf Grundlage der daraus gewonnenen Erkenntnisse erfolgt eine Weiterentwicklung sowohl des Schemas als auch des Werkzeugs. Die Erweiterung des Klassifikationsschemas um zusätzliche Dimensionen und Ausprägungen und die Überprüfung der vorhandenen Dimensionen sind Gegenstand einer zukünftigen Untersuchung. Weiterhin ist eine Umsetzung als Webanwendung geplant. Dadurch können die Möglichkeiten und Grenzen des Social Tagging durch einen größeren Personenkreis anhand des vordefinierten Schemas untersucht werden. Um das Klassifikationsschema durch mehrere Nutzer zu erweitern, stellt die Bundle-Funktion in Delicious einen weiteren Nutzen dar: Vorschläge können beim Verschlagworten einer Visualisierung gespeichert und so weiter genutzt werden. Werden diese häufig genutzt, können diese ab einem bestimmten Schwellwert in ein Delicious-Bundle aufgenommen werden. Erst dadurch werden sie im DelViz-Werkzeug berücksichtigt und können das Schema dadurch erweitern.

Weiterhin ist die Nutzung weiterer Funktionalitäten der Delicious-API denkbar. Beispielsweise kann durch die Freigabe der Lesezeichen in Delicious erkannt werden, wie viele andere Nutzer das gleiche Lesezeichen ebenfalls hinzugefügt haben. Die somit entstandene „Wertung“ eines Lesezeichens kann auch bei den Visualisierungen genutzt werden. Je nachdem, wie oft eine Visualisierung von anderen Nutzern gespeichert wurde, desto relevanter und populärer könnte diese somit sein und im DelViz-Werkzeug hervorgehoben werden.

Danksagung

Diese Arbeit wurde durch die Europäische Union und den Freistaat Sachsen durch Mittel des Europäischen Fonds für regionale Entwicklung (EFRE) gefördert.

Literaturverzeichnis

- O'Reilly, Tim. What is Web 2.0? Design Patterns and Business Models for the Next Generation of Software. 2005. Aufruf: 20. 05 2010. <http://www.oreilly.de/artikel/web20.html>.
- Lima, Manuel. Visual Complexity. 2010. Aufruf: 02. 03 2010. www.visualcomplexity.com.
- Viegas, Fernanda B., et al. ManyEyes: a Site for Visualization at Internet Scale. Los Alamitos, CA : IEEE Transactions on Visualization and Computer Graphics, IEEE Computer Society, 2007. S. 1121-1128.
- Lengler, Ralph und Eppler, Martin J. Towards a periodic table of visualization methods. Florida : In Proceedings of the 2007 IASTED Conference on Graphics and Visualization in Engineering, 2007.
- Schachter, Joshua. Delicious 2003. Aufruf: 17. 5 2010. www.delicious.com.
- Card, Stuart K., Mackinlay, Jock und Shneiderman, Ben Readings in Information Visualization: Using Vision to Think. San Francisco : Morgan Kaufmann, 1999. S. 7 ff.
- Shneiderman, Ben, The Eyes Have It: A Task by Data Type Taxonomy for Information Visualization. Los Alamitos, CA : IEEE Symposium on Visual Language, IEEE Computer Society, 1996. S. 336-343.
- Quintarelli, Emanuele. Folksonomies: power to the people. ISKO Italy-UniMIB meeting. Milan. 2005. Aufruf: 26. 02 2010. www.iskoi.org/doc/folksonomies.htm.
- Priss, Uta. Faceted Information Representation. Working with Conceptual Structures. Aachen : Proceedings of the 8th International Conference on Conceptual Structures; Shaker Verlag, 2000. S. 84-94.
- Hearst, Marti A. UIs for Faceted Navigation: Recent Advances and Remaining Open Problems. Redmond : In the Workshop on Computer Interaction and Information Retrieval, HCIR 2008, 2008.
- Mathes, Adam. Folksonomies - Cooperative Classification and Communication Through Shared Metadata. University of Illinois Urbana-Champaign : Computer Mediated Communication - LIS590CMC, 2004.
- Alby, Tom. Web 2.0. Konzepte, Anwendungen, Technologien. München: Hanser Fachbuchverlag, 2008. S. 94 ff.
- Stefaner, Moritz. Visual tools for the socio-semantic web. Master's Thesis: University of Applied Sciences Potsdam, 2007. S. 56 ff. http://well-formed-data.net/pdf/thesis_stefaner_screen.pdf.
- Kerr, Bernard. Thread Arcs: An Email Thread Visualization. Seattle, Washington : IEEE Symposium on Information Visualization (InfoVis 2003), 2003. S. 27.