

247

Visualisierung von Kultur im Web. Überlegungen zu computergestützten Abbildungen von Typisierungsprozessen

Axel Philipps

März 2015

GEFÖRDERT VOM



Bundesministerium
für Bildung
und Forschung

Working Paper Series des Rates für Sozial- und Wirtschaftsdaten (RatSWD)

Die *RatSWD Working Papers* Reihe startete Ende 2007. Seit 2009 werden in dieser Publikationsreihe nur noch konzeptionelle und historische Arbeiten, die sich mit der Gestaltung der statistischen Infrastruktur und der Forschungsinfrastruktur in den Sozial-, Verhaltens- und Wirtschaftswissenschaften beschäftigen, publiziert. Dies sind insbesondere Papiere zur Gestaltung der Amtlichen Statistik, der Ressortforschung und der akademisch getragenen Forschungsinfrastruktur sowie Beiträge, die Arbeit des RatSWD selbst betreffend. Auch Papiere, die sich auf die oben genannten Bereiche außerhalb Deutschlands und auf supranationale Aspekte beziehen, sind besonders willkommen.

RatSWD Working Papers sind nicht-exklusiv, d. h. einer Veröffentlichung an anderen Orten steht nichts im Wege. Alle Arbeiten können und sollen auch in fachlich, institutionell und örtlich spezialisierten Reihen erscheinen. Die *RatSWD Working Papers* können nicht über den Buchhandel, sondern nur online über den RatSWD bezogen werden.

Um nicht deutsch sprechenden Nutzer/innen die Arbeit mit der Reihe zu erleichtern, sind auf den englischen Internetseiten der *RatSWD Working Papers* nur die englischsprachigen Papers zu finden, auf den deutschen Seiten werden alle Nummern der Reihe chronologisch geordnet aufgelistet.

Einige ursprünglich in der *RatSWD Working Papers* Reihe erschienenen empirischen Forschungsarbeiten sind ab 2009 in der RatSWD Research Notes Reihe zu finden.

Die Inhalte der *RatSWD Working Papers* stellen ausdrücklich die Meinung der jeweiligen Autor/innen dar und nicht die des RatSWD. Das Bundesministerium für Bildung und Forschung hat die Publikationen nicht beeinflusst.

Herausgeber der RatSWD Working Paper Series:

Vorsitzender des RatSWD

(seit 2014 Regina T. Riphahn; 2009-2014 Gert G. Wagner; 2007-2008 Heike Solga)

Visualisierung von Kultur im Web. Überlegungen zu computergestützten Abbildungen von Typisierungsprozessen

Axel Philipps

Leibniz Universität Hannover

Zusammenfassung:

Soziale Medien und die Praxis gemeinschaftlich erzeugter Schlagwortsysteme eröffnen die Möglichkeit, Prozesse sozial anerkannter Typisierungen zu untersuchen. Typisierungen sind soziologisch besonders relevant, da sie durch Prozesse der Selektion und Spezifikation Sinn konstituieren, zur Kommunikation befähigen, als Auslegungs- und Orientierungsschema fungieren und damit routiniertes Verhalten ermöglichen. Anspruchsvoll ist hingegen die methodisch kontrollierte empirische Analyse der Stabilisierungen und Transformationen von Typisierungsprozessen. Das Working Paper diskutiert daher aus einer kultursoziologischen Perspektive Anforderungen an eine computergestützte Visualisierung solcher Prozesse anhand von gemeinschaftlich erzeugten Schlagwortsystemen.

Keywords: Visualisierung, Typisierungen, Folksonomies, soziale Medien

Einführung

Das World Wide Web (kurz: Web) wird in der Regel in den Sozialwissenschaften als ein Informationsnetzwerk oder technisches Medium betrachtet. Mit den neuen Möglichkeiten stellen sich Fragen nach Veränderungen in der Organisation von Wissen, im Nutzungsverhalten, der Art sozialer Beziehungen oder den Formen von Kommunikation. Das Web markiert aber nicht nur Veränderungen. Im Gebrauch des Webs greifen die Nutzerinnen¹ auch auf bewährte Wissensrepertoires zurück und reproduzieren alltägliche Routinen (Guenther & Schmidt 2008; Hess 2008; Schmidt 2011). So wie das Web soziale Praktiken prägt, wird es durch Rezeptwissen und Handlungsweisen des Offlinealltags strukturiert. Das Web bietet somit ebenfalls die Chance, Alltagspraktiken in der Webumgebung zu analysieren.

Beispielsweise sind die in vielen sozialen Medien erzeugten sogenannten Folksonomies, "collaborative tagging systems" (Golder & Huberman 2006; Smith 2008), eine spezifische Erscheinung des Webs. Sie entstehen in webbasierten sozialen Medien wie Delicious.com, Flickr.com, Amazon.com und anderen. Nutzerinnen dieser Medien verwenden Schlagworte (tags), um Artefakte (z.B. URLs, Bilder, Publikationen, Warenartikel) zu beschreiben, zu ordnen und zu browsen. Im Gegensatz zu systematischen, distinkten, umfassenden und hierarchisch strukturierten Klassifikationen (etwa wissenschaftliche Taxonomien) sind Folksonomies gemeinschaftlich hervorgebrachte, lose und flach strukturierte Kategorisierungen. Derartige Kategorisierungen, deren Gruppierungen und Verknüpfungen erfahrungsbasiert auf wahrgenommenen Ähnlichkeiten beruhen (Jacob 1991, 2004), entsprechen wiederum, soziologisch betrachtet, den von Individuen und Gruppen vorgenommenen Typisierungen. Mit Typisierungen abstrahieren nach Alfred Schütz und anderen (Schütz 1971, 1972, 1982; Schütz & Luckmann 1994) die Menschen vom Konkreten und heben Merkmale oder Aspekte von Individuen, Gegenständen oder Artefakten hervor, welche diese mit anderen teilen oder ihnen darin ähnlich sind. Was wie typisiert wird und damit relevant ist, hängt vom Kontext (z.B. Gruppenzugehörigkeit, Medium) oder Zweck (etwa ein Problem lösen) ab.

Für die kultursoziologische Forschung ergibt sich dadurch die Chance, mit Hilfe von Folksonomies Prozesse sozial anerkannter Typisierungen im Alltag zu beobachten und zu untersuchen. Natürlich beschränken sich diese Art der Typisierungen auf Gegenständen und Artefakten und weniger auf Typen von Personen, Rollen oder Handlungen. Ungeachtet dessen eröffnet sich der kultursoziologischen Forschung die Möglichkeit, einerseits die Etablierung und Stabilisierung von sozial anerkannten Typisierungen und andererseits ihre Auflösung und Transformation zu analysieren.

¹ Für eine bessere Lesbarkeit wurde durchgehend die weibliche Schreibweise gewählt. Damit sind aber ausdrücklich immer beide Geschlechter gemeint.

Solche Untersuchungen setzen jedoch voraus, dass die Prozesse alltäglicher Typisierung in Folksonomies sichtbar und damit kultursoziologisch interpretierbar sind. Will man nicht willkürlich Schlagworte herausgreifen oder Einzelfallanalysen vornehmen, braucht es die Möglichkeit, Typisierungsprozesse zu visualisieren. Visualisierungen sind jedoch Übersetzungen von Einzeldaten in visuelle Darstellungen, um die Daten in einem neuen Zusammenhang zu zeigen (Mitchell 1994). Folglich ist in einem ersten Schritt zu klären, wie Typisierungen in Folksonomies zu visualisieren sind und was die Visualisierungen leisten sollen. Oder mit Tufte (1983) ausgedrückt, was muss eine solche Visualisierung erfüllen, „to [give] the viewers the greatest number of ideas in the shortest time with the least ink in the smallest space“ (51)?

Der Beitrag beschäftigt sich daher mit der folgenden Frage: Welche Bedingungen müssen computergestützte Visualisierung erfüllen, um Prozesse der alltäglichen Typisierung in Folksonomies abzubilden? Was bieten bisherige Softwarelösungen? Dazu werden eine Auswahl bisheriger Softwarelösungen diskutiert und Anforderungen für eine kultursoziologische Analyse herausgearbeitet.

Anforderungen an computergestützte Visualisierungen

Sozialwissenschaftlich relevante Visualisierungen konzentrieren sich zumeist auf das Web als ein Netzwerk von Informationen. Untersucht wird vor allem der menschliche Faktor in Verbindung mit Hypertext und Internetinfrastrukturen bei der Entstehung, Verbreitung und Veränderung von verknüpften Informationen und Informationsflüssen. Dazu werden in der Regel Beziehungen, Verteilungen und die Performanz auf der Basis von Metadaten durch vereinfachte visuelle Darstellungen (Punkte, Linien, einfache geometrische Figuren) in Häufigkeitsverteilungen, Netzwerkbeziehungen oder Pfadmodellen abgebildet. Eindrucksvolle Studien visualisieren beispielsweise die Beziehungsmuster zwischen den Nutzerinnen in Facebook (Ugander et al. 2011) oder die Nutzerinnen-Performanz in Wikipedia (Viégas et al. 2004).

Manovich (2011) nennt drei allgemeine Kritikpunkte an solchen Visualisierungen: 1) reduzieren sie die Daten auf ein Merkmal (die oben angesprochenen vereinfachten Darstellungsformen), 2) werden Eigenschaften vorzugsweise räumlich abgebildet (nah/fern, innen/außen) und 3) dominiert die Netzwerklogik bei der Beobachtung des Webs. Vor dem Hintergrund einer Visualisierung von Typisierungsprozessen kann ergänzt werden, dass das Web nicht nur Informationen verknüpft, sondern dass das Web auch spezifische Bedeutungszuschreibungen und kollektive Deutungsmuster hervorbringt. Solche von Menschen im Web geschaffenen Bedeutungsgewebe (Geertz 1987)

erfordern daher eine andere Art der Visualisierung. Textbasierte Folksonomies lassen sich beispielsweise mithilfe von Tag Clouds direkt visualisieren (Manovich 2011). Solche Clouds repräsentieren die Häufigkeit von Schlagworten mittels unterschiedlicher Größendarstellungen. Je häufiger ein Tag auftaucht, desto größer wird es in der Cloud angezeigt. Auf Wahrscheinlichkeitsberechnungen basierende Topic Model-Verfahren bieten hingegen die Möglichkeit, distinkte Themen (topics) und ihre Anteile in Textsammlungen zu unterscheiden und zu identifizieren (Blei 2012; Mohr & Bogdanov 2013). Den genannten Visualisierungsverfahren ist jedoch eigen, dass sie synchrone Verteilungen von Wörtern bzw. Themen zeigen.

Die Visualisierung von Typisierungsprozessen verlangt aber die Abbildung von zeitlichen Veränderungen und Verlaufskurven.

Diese Anforderung erfüllen Visualisierungen wie SparkClouds, ThemeRiver, dynamische Topic Models oder Alluvial Diagrams. SparkClouds (Fig. 1) zeigen die Häufigkeit bestimmter Wörter anhand variabler Schriftgrößen in Kombination mit Diagrammen, aus denen die relative Popularität der Wörter über eine bestimmte Zeit hervorgeht (Lee et al. 2010). ThemeRiver (Fig. 2) demonstriert die Variation und Verteilung von Themen anhand einer vorab festgelegten Auswahl an Wörtern, die für bestimmte Themen stehen (Havre et al. 2002). Dynamische Topic Models (Fig. 3) veranschaulichen dagegen, wie sich der Anteil eines Themas über die Zeit verändert (Blei & Lafferty 2006). Rosvall und Bergstrom (2010) bilden wiederum mit Alluvial Diagrams strukturelle Veränderungen in Netzwerkbeziehungen ab. Die Farbintensität, Verknüpfungen und die Größe der Cluster signalisieren Änderungen in den Clustern (Fig. 4). Spezifisch für die sozialen Medien Delicious und Flickr wurden Visualisierungssoftwarelösungen entwickelt, die zeitliche Veränderungen in der Tag-Nutzung veranschaulichen. Im Mittelpunkt standen die Schlagworte und weniger die Themen oder Netzwerkcluster. Für Delicious bot beispielsweise das Programm Cloudalicious bereits vor 2005 die Möglichkeit, die Verwendung von Schlagworten für bestimmte Webseiten über die Zeit anzuzeigen (Golder & Huberman 2006). Die Häufigkeit der Schlagworte wurde gewichtet durch die Anzahl derjenigen, die sie nutzten (Fig. 5). Für Flickr programmierten Dubinko et al. (2007) das Verfahren der Taglines, was die Veränderung häufig genutzter Schlagworte in Flickr für ausgewählte Zeitabschnitte zeigte (Fig. 6).

Die Analyse von alltäglichen Typisierungsprozessen steht jedoch vor dem Problem, dass diese Visualisierungsverfahren einerseits nicht mehr zur Verfügung stehen und andererseits es erforderlich ist, einzelne Aspekte und Anwendungen von Visualisierungsverfahren zu kombinieren.

Dies betrifft vor allem die Analyse alltäglicher Typisierungsprozesse hinsichtlich der Stabilisierung oder Auflösung von Schlagwortkombinationen sowie der Unterscheidung zwischen wiederkehrenden Ereignissen wie Festivals, Sportveranstaltungen, Jahrestagen und nachhaltigen Veränderungen. Während Softwarelösungen wie Cloudalicious oder Taglines die Nutzung bestimmter Schlagworte über die Zeit veranschaulichen, haben beispielsweise Rattenbury und Kollegen (2007) ein Verfahren entwickelt, automatisch ereignis- und ortsspezifische Schlagworte zu erkennen.

Des Weiteren betreffen Typisierungen bzw. Kategorisierungen in webbasierten sozialen Medien in der Regel Artefakte wie Bilder, Publikationen oder Webseiten. Die Gründe für die Verschlagwortung der Artefakte sind vielfältig und reichen vom Organisieren über Möglichkeiten des Browsens bis zur Selbstpräsentation (Golder & Huberman 2006; Park 2011; Smith 2008). Für die Typisierungsprozesse ist dabei aber vor allem relevant, dass sich bestimmte Schlagworte für die Artefakte durchsetzen und diese von verschiedenen Nutzerinnen fortgesetzt verwendet werden.

Die Dokumentation sozial anerkannter Typisierungen benötigt daher computergestützte Visualisierungen, die Stabilisierungsprozesse sichtbar machen.

Für die Analyse von Typisierungsprozessen ist ebenfalls interessant, die Transformation und Auflösung von Typisierungen zu untersuchen, da sie aus verschiedenen Gründen fragwürdig werden können (z.B. durch historische Großereignissen, Naturkatastrophen, Protestbewegungen, Dekonstruktionen).

Dies erfordert von einer computergestützten Visualisierung, Veränderungen in den Typisierungen (bestimmter Artefakte) nachvollziehbar abzubilden.

Hinzukommt, dass die Schlagworte häufig über den Inhalt von Artefakten hinausweisen. Sie bezeichnen und beschreiben nicht nur, was digitale Fotografien abbilden, Webseiten bieten oder wissenschaftliche Publikationen behandeln. Im Fall von Fotografien kommen beispielsweise häufig technische Angaben hinzu wie das Kameramodell, der Name der bzw. des Fotografin, Links zu anderen Webseiten und anderes.

Die Visualisierung der Typisierung von Artefakten macht demnach erforderlich, optional Schlagworte gruppieren und filtern zu können.

Die Verwendung eines oder weniger Schlagworte reduziert schließlich die in den Artefakten vorhandenen vielfachen Bedeutungen. Die Selektion kann einen Hinweis auf spezifische Wahrnehmungs-, Denk- und Bewertungsweisen liefern. Wobei nur eine Feinanalyse ermöglicht, zu untersuchen, wie die Schlagworte ein Artefakt typisieren. Erst der Wechsel von der Visualisierung großer Datenmengen zu einzelnen Schlagworten und Artefakten kann klären, ob eine Änderung der Typisierung auf veränderte Deutungsmuster oder neuartige Artefakte zurückgeht.

Eine Softwarelösung muss daher auch in der Lage sein, an jedem beliebigen Zeitpunkt von der Visualisierung der Prozesse zu einzelnen Schlagworten und damit verknüpften Artefakten zu wechseln.

Zusammenfassung

Folksonomies sind ein Phänomen sozialer Medien. Anwendungsorientierte Visualisierungssoftwarelösungen bemüht sich vor allem darum, die lose verknüpften, flach strukturierten Schlagworte durch Rankings und Systematisierungen zu strukturieren (Hotho et al. 2006; Körner et al. 2010). Zugleich sind Folksonomies ebenso für die Sozial- und Kulturwissenschaften relevant (Guenther & Schmidt 2008; Halford et al. 2013), wobei empirische Untersuchungen noch ausstehen. Ein Grund dafür sind fehlende Softwarelösungen, welche die großen Datenmengen der Folksonomies aufbereiten und durch Visualisierungen den soziologischen Analysen zugänglich machen.

Referenzen

- Blei, D. M. (2012). Topic modeling and digital humanities. *Journal of Digital Humanities*, 2(1), 8-11.
- Blei, D. M., & Lafferty, J. D. (2006). Dynamic topic models. *Proceedings of the 23rd international conference on Machine learning*, ACM, 113-120.
- Dubinko, M., Kumar, R., Magnani, J., Novak, J., Raghavan, P., & Tomkins, A. (2007). Visualizing tags over time. *ACM Transactions on the Web (TWEB)*, 1(2), 7.
- Geertz, C. (1987). *Dichte Beschreibung. Beiträge zum Verstehen kultureller Systeme*, Frankfurt a. M.
- Golder, S. A., & Huberman, B. A. (2006). The Structure of Collaborative Tagging Systems. *Journal of Information Science*, 32(2), 198-208.
- Guenther, T., & Schmidt, J. (2008). Wissenstypen im „Web 2.0“—eine wissenssoziologische Deutung von Prodnutzung im Internet. In *Weltweite Welten* (S. 167-187). VS Verlag für Sozialwissenschaften.
- Halford, S., Pope, C., & Weal, M. (2013). Digital futures? Sociological challenges and opportunities in the emergent Semantic Web. *Sociology*, 47(1), 173-189.
- Havre, S., Hetzler, E., Whitney, P., & Nowell, L. (2002). Themeriver: Visualizing thematic changes in large document collections. *Visualization and Computer Graphics, IEEE Transactions on*, 8(1), 9-20.
- Hess, A. (2008). Reconsidering the rhizome: A textual analysis of web search engines as gatekeepers of the Internet. In *Web Search* (S. 35-50). Springer Berlin Heidelberg.
- Hotho, A., Jäschke, R., Schmitz, C., & Stumme, G. (2006). Information retrieval in folksonomies: Search and ranking. In *ESWC 2006, LNCS 4011* (S. 411-426). Springer Berlin Heidelberg.
- Jacob, E. K. (2004). Classification and categorization: a difference that makes a difference. *Library Trends*, 52(3), 515-540.
- Jacob, E. K. (1991). Classification and categorization: Drawing the line. *2nd ASIS SIG/CR Classification Research Workshop*, 63-80.
- Körner, C., Benz, D., Hotho, A., Strohmaier, M., & Stumme, G. (2010). Stop thinking, start tagging: tag semantics emerge from collaborative verbosity. In *Proceedings of the 19th international conference on World wide web* (S. 521-530). ACM.
- Lee, B., Riche, N. H., Karlson, A. K., & Cpendale, S. (2010). Sparkclouds: Visualizing trends in tag clouds. *Visualization and Computer Graphics, IEEE Transactions on*, 16(6), 1182-1189.
- Manovich, L. (2011). What is visualisation? *Visual Studies*, 26(1), 36-49.
- Mitchell, W. J. T. (1994). *Picture theory: Essays on visual and verbal representation*. Chicago: U of Chicago P.
- Mohr, J. W., & Bogdanov, P. (2013). Introduction—Topic models: What they are and why they matter. *Poetics*, 41(6), 545-569.
- Park, H. (2011). A conceptual framework to study folksonomic interaction. *Knowledge Organization*, 38(6), 515-529.
- Rattenbury, T., Good, N., & Naaman, M. (2007). Towards automatic extraction of event and place semantics from flickr tags. In *Proceedings of the 30th annual international ACM SIGIR conference on Research and development in information retrieval* (S. 103-110). ACM.
- Rosvall, M., & Bergstrom, C. T. (2010). Mapping change in large networks. *PLoS one*, 5(1), e8694.

- Schmidt, J. (2011). *Das neue Netz. Merkmale, Praktiken und Folgen des Web 2.0*. Konstanz: UVK Verl.
- Schütz, A. (1982). *Das Problem der Relevanz*. Frankfurt am Main: Suhrkamp.
- Schütz, A. (1972). Die Gleichheit und die Sinnstruktur der sozialen Welt. In *Gesammelte Aufsätze II* (S. 203-255). Springer Netherlands.
- Schütz, A. (1971): Wissenschaftliche Interpretation und Alltagsverständnis menschlichen Handelns. In *Gesammelte Aufsätze I. Das Problem der sozialen Wirklichkeit* (S. 3-54). Den Haag: Martinus Nijhoff.
- Schütz, A., & Luckmann, T. (1988). *Strukturen der Lebenswelt. Die soziale Struktur der Lebenswelt des Alltags*. Frankfurt am Main: Suhrkamp.
- Smith, G. (2008). *Tagging. People-powered metadata for the social web*. Berkeley.
- Tufte, E. R. (1983). *The visual display of quantitative information*, Cheshire, CT.
- Ugander, J., Karrer, B., Backstrom, L., & Marlow, C. (2011). The anatomy of the facebook social graph. *arXiv:1111.4503*.
- Viégas, F. B., Wattenberg, M., & Dave, K. (2004). Studying cooperation and conflict between authors with history flow visualizations. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 575-582). ACM.



Fig. 1 SparkClouds (www.infovis-wiki.net/images/thumb/6/67/Sparkcloud.gif/500px-Sparkcloud.gif)

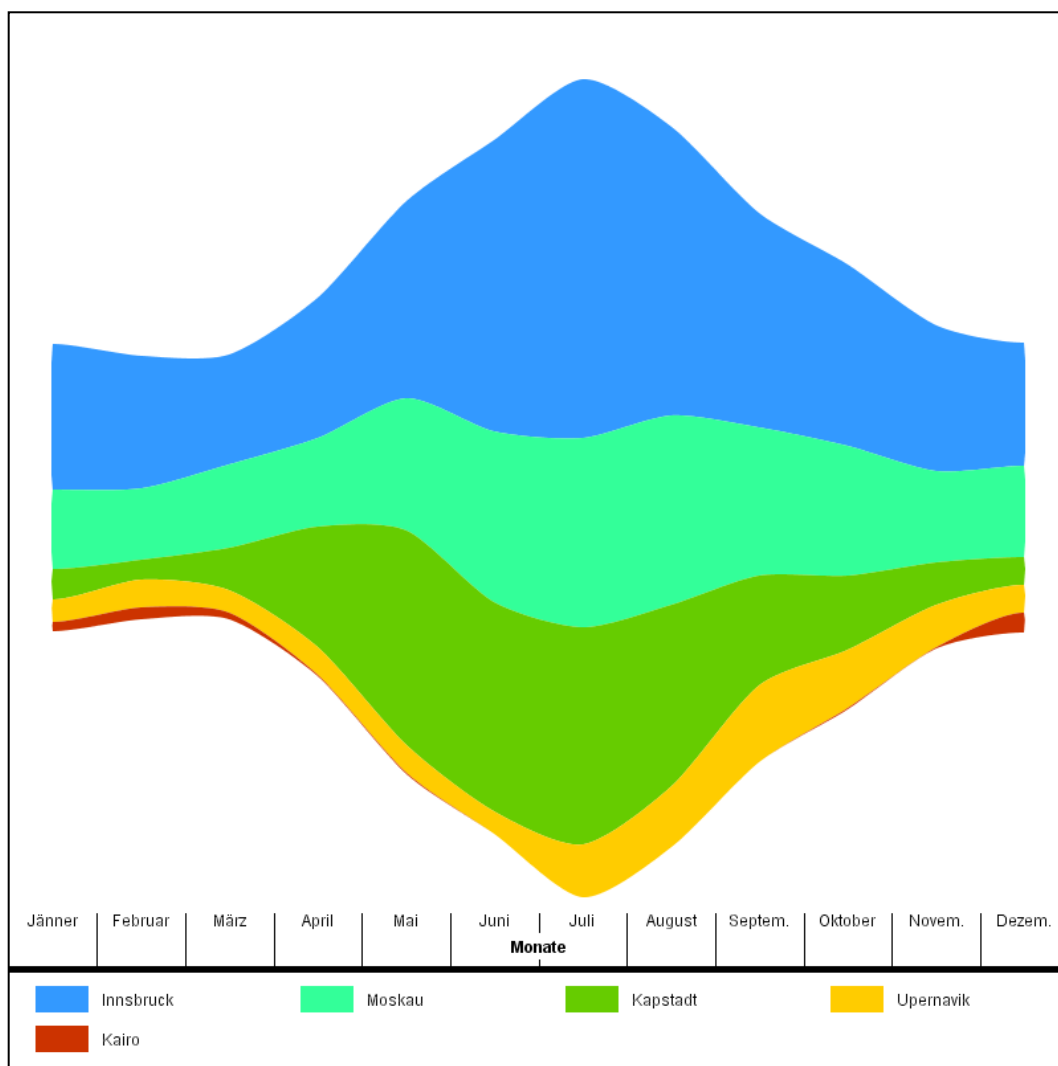


Fig. 2 ThemeRiver
 (www.wanderinformatiker.at/unipages/ThemeRiver/comparison/themeriver_niederschlag.png)

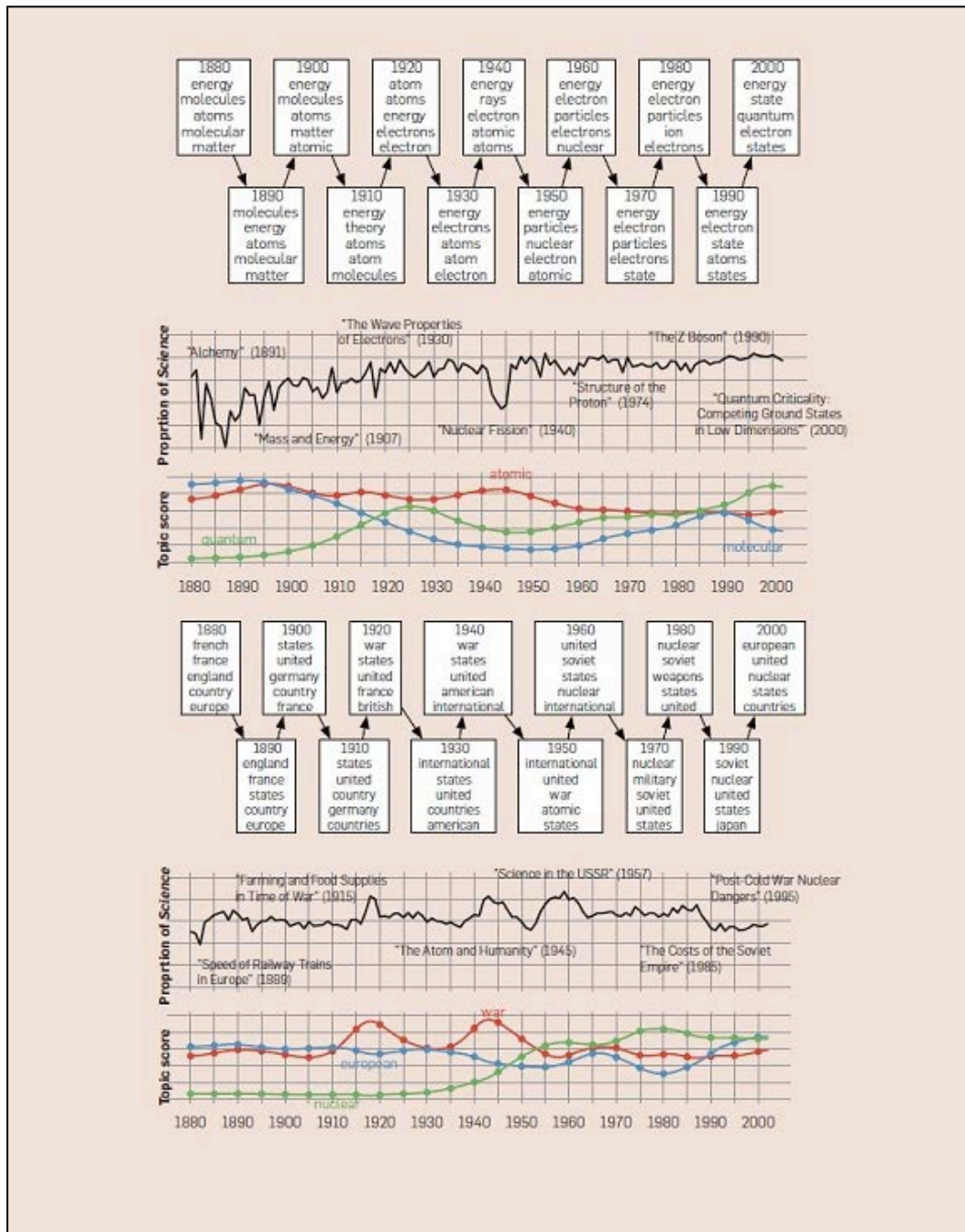


Fig. 3 Dynamic Topic Model
 (deliveryimages.acm.org/10.1145/2140000/2133826/figs/f5.jpg)

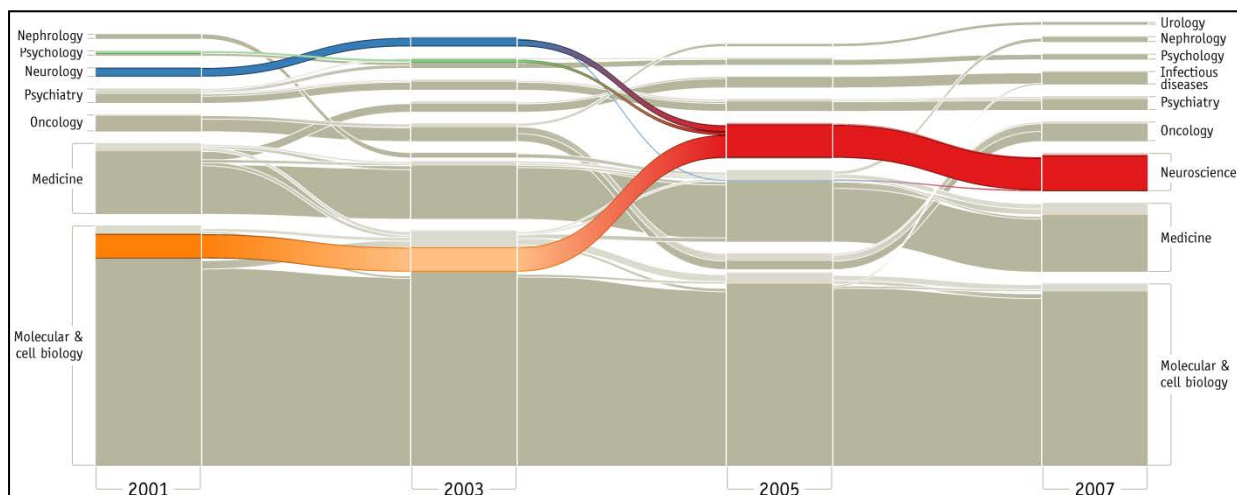


Fig. 4 Alluvial Diagramm

(www.plosone.org/article/info%3Adoi%2F10.1371%2Fjournal.pone.0008694)

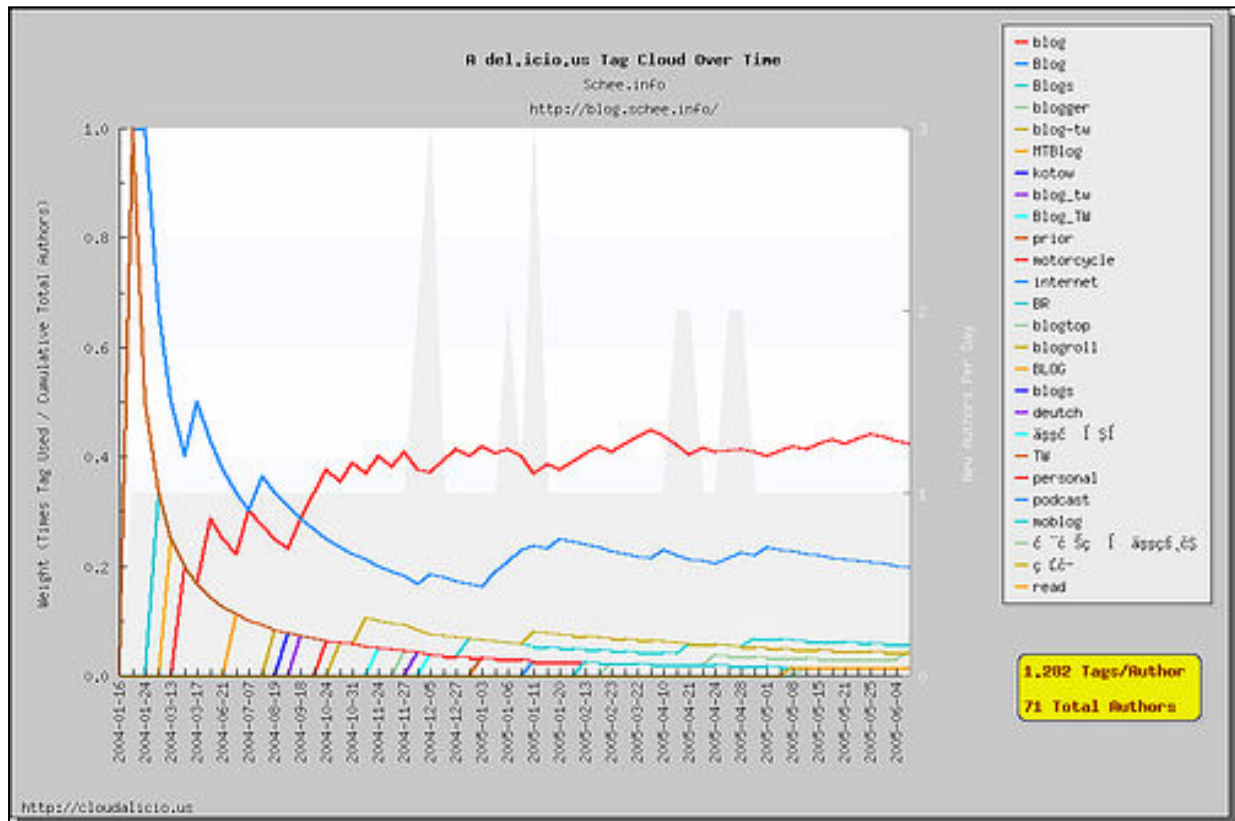


Fig. 5 Cloudalicious (farm1.static.flickr.com/208/462853379_4d2b053504_m.jpg)

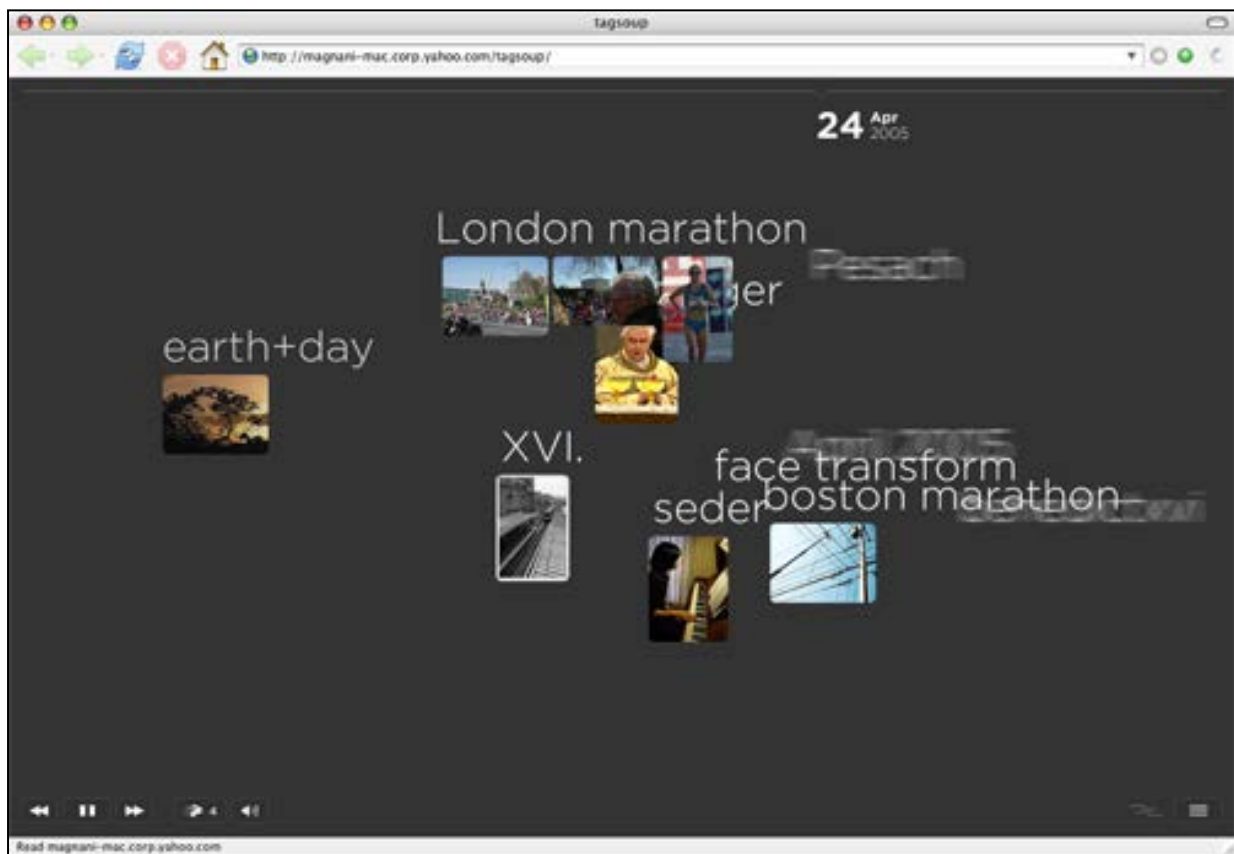


Fig. 6 Taglines (www2006.org/programme/files/xht)