



# Auf der Suche nach der Nadel im Datenhaufen

**Soziale Netzwerke bieten Unmengen an Informationen, gerade Twitter eignet sich zur Recherche unterschiedlichster Themen. Eine Innsbrucker Informatikerin arbeitet nun daran, diese Informationen besser auffindbar zu machen.**

**Über den Microblogging-Dienst Twitter unterhalten sich Millionen von Menschen. Relevante Informationen, die sie in ihren Nachrichten geben, sind aber nicht immer leicht auffindbar.**

Soziale Netzwerke entwickeln sich immer mehr zu Riesen-Datenspeichern – nicht nur im negativen Sinn als Daten-Silos für

vermarktungstechnisch relevante Informationen, sondern auch im positiven Sinn als Datenbanken breiten Wissens. Letzterem widmet sich die Dissertantin DI Eva Zangerle am Institut für Informatik am Beispiel von Twitter: „Gerade Twitter ist eine enorme Quelle von Wissen, und dadurch, dass nahezu alles öffentlich ist, natürlich auch für die Forschung sehr interessant.“ In ihrer Dissertation widmet sich Eva Zangerle der Strukturierung dieses Wissens.

„Ein Problem, auf das man bei Twitter rasch stößt, ist die Tatsache, dass die enthaltenen Informationen einfach sehr schwierig zu durchsuchen sind“, erklärt sie.

### Suchen und finden

Twitter ist ein so genannter Microblogging-Dienst: Nutzer stellen Nachrichten ins Netz, die maximal 140 Zeichen lang sein dürfen, und kommunizieren über diese so genannten Tweets miteinander, beantworten Fragen und geben Auskunft. Auch aktuelle Veranstaltungen werden von vielen Twitter-Nutzerinnen und -Nutzern live mitbegleitet und kommentiert – so sind wichtige Nachrichten oft schon auf Twitter zu lesen, bevor sie medial aufgegriffen werden, und öffentliche Ereignisse lassen sich zumindest in der Theorie relativ einfach rekonstruieren. Einen großen Bekanntheitsschub erfuhr das Netzwerk Anfang dieses Jahres, als Demonstranten im Arabischen Frühling ihre Eindrücke über das Netzwerk geteilt und sich so auch koordiniert haben.

Für die Gliederung und Zuordnung von Inhalten einzelner Tweets zu Themen haben sich so genannte Hashtags eingebürgert: aufgrund der Zeichenbeschränkung auf Twitter knapp gehaltene Stichworte und Abkürzungen, denen ein Rautezeichen (#) vorangesetzt wird und die an die einzelnen Tweets zum jeweiligen Thema angehängt werden. „Diese Hashtags sind gleichzeitig sehr hilfreich und potenziell auch das Gegenteil davon, nämlich dann, wenn beispielsweise für eine Veranstaltung zwei oder mehr Hashtags verwendet werden“, erklärt Eva Zangerle. Wenn ein Nutzer dann nur nach dem einen ihm bekannten Hashtag zum von ihm gewünschten Thema sucht, findet er nur die Hälfte oder gar ein Drittel der Informationen dazu. „Wir haben zum Beispiel Tweets zur Tour de France untersucht. Da wurden die Hashtags #TDF und #tourdefrance und daneben noch ein paar andere verwendet.“

Genau bei diesem Problem setzt Eva Zangerles Forschung an: Sie arbeitet mit Empfehlungssystemen, die den Nutzer bereits bei der Eingabe eines Tweets unterstützen und ihm „richtige“ Hashtags vorschlagen soll. „Solche Empfehlungssysteme sind in anderer Form besonders bei

Shopping-Websites relativ weit verbreitet, etwa, wenn jemandem auf Amazon ein Buch empfohlen wird, weil er schon ein anderes Buch desselben Autors gekauft hat oder weil Menschen, die das gleiche Buch kaufen, vermehrt auch ein anderes kaufen“, sagt sie. Ähnlich funktioniert Zangerles Prototyp zur Strukturierung der Daten in Twitter, der bald online gehen soll: Aufgrund vorangegangener Tweets und bereits verwendeter Hashtags werden dem Nutzer für seinen aktuellen Tweet sinnvolle Hashtags vorgeschlagen. „Wir können auch eingegebenen Text mit anderen Tweets vergleichen und so Hashtags vorschlagen.“ In weiterer Folge seien auch mehr Parameter denkbar, etwa der Aufenthaltsort des Nutzers, die Uhrzeit, die Themen, die gerade allgemein auf Twitter diskutiert werden, und Ähnliches. „Idealerweise sorgen wir so dafür, dass zum Beispiel zur Tour de France nur noch unter einem Hashtag getwittert wird und so auf einen Klick alle Tweets dazu auffindbar sind.“

### Breites Forschungsfeld

Eva Zangerle und ihr Kollege DI Wolfgang Gassler arbeiten in ihrem Projekt „Snoopy“ gemeinsam mit mehreren Master- und Bachelorstudenten an neuartigen

Vorschlagssystemen und sind Teil der Forschungsgruppe für Datenbanken und Informationssysteme am Institut für Informatik. „Grundsätzlich ist die Strukturierung von Daten im Internet ein ungemein breites Feld – viele Websites und Internet-Dienste verwenden zwar Tags, um ihre Daten zu kategorisieren, in den allermeisten Fällen ist das aber sehr ausbaufähig“, erläutert sie. Inhalt einer inzwischen abgeschlossenen Masterarbeit war etwa die Gliederung von Tags auf der Foto-Seite Flickr: „Hier ging es darum, die Tags weiter zu gliedern. Der Tag ‚Eva Zangerle‘ allein hilft etwa noch nicht sehr viel weiter – wurde das Bild von Eva Zangerle aufgenommen, zeigt es Eva Zangerle?“, sagt die Informatikerin. Der Lösungsansatz ist einfach, aber effektiv: Der jeweilige Tag kann durch eine Zusatzinformation ergänzt werden. Aus „Eva Zangerle“ wird so „abgebildet:Eva Zangerle“ – und schon lässt sich nach Fotos mit Eva Zangerle suchen. Das Ergebnis dieser Masterarbeit zu Flickr ist bereits online zu finden.

[stefan.hohenwarter@uibk.ac.at](mailto:stefan.hohenwarter@uibk.ac.at) ■

WEITERE INFORMATIONEN  
<http://dbis-snoopy.uibk.ac.at/>  
<http://snoopytagging/tagging/>  
<http://dbis-snoopy.uibk.ac.at>  
<http://dbis-informatik.uibk.ac.at>

### ZUR PERSON



EVA ZANGERLE

Die gebürtige Landeckerin Eva Zangerle studiert seit 2002 Informatik an der Universität Innsbruck – das Bachelorstudium hat sie 2006 abgeschlossen, das Masterstudium 2007. Seither arbeitet sie an ihrer Dissertation über die Gliederung von Daten in sozialen Netzwerken im Internet und ist als Universitätsassistentin am Institut für Informatik beschäftigt. Beruflich engagiert sie sich bei der „Frauen in die Technik“-Initiative in Tirol und leitet Lehrveranstaltungen am Institut für Informatik. Für ihr Dissertationsvorhaben hat sie im Herbst 2011 ein Stipendium der Universität Innsbruck erhalten.



Mit der Zuordnung von Inhalten einzelner Tweets – etwa zur Tour de France – befassen sich Eva Zangerle und ihre Forschungsgruppe.

Fotos: Shutterstock, flickr.com/Jon's pics