



SEISMOGRAF DER WELT

300 Milliarden Tweets gibt es bereits, und **Twitters Datenberg** schwillt weiter an. Ob Stimmungsanalysen, Aktienkurse oder Katastrophenhilfe – aus unserem Gezwitscher wird immer mehr herausgelesen.

VON JENS LUBBADEH

Ein Pixel erscheint auf der dunklen Weltkarte, irgendwo in Nordindien, zugleich eine unleserliche kurze Zeichenfolge. Noch eines, diesmal in der Mitte der USA. Fast gleichzeitig drei weitere – auf der Arabischen Halbinsel, in Brasilien und eines im kleinen Europa, vielleicht in Frankreich? Nach und nach entstehen Wölkchen auf der Karte, und nach wenigen Minuten erinnert das Ganze frappierend an die Nachtaufnahmen der Erde aus dem All: hell erleuchtet die industrielle Welt, dunkel die Weiten Afrikas und Russlands.

Diese Karte, für jeden abrufbar auf tweetping.net, zeigt das Gezwitscher der Menschheit. Es ist die Welt von Twitter – in Echtzeit abgebildet. Jeder dieser Punkte ist ein Tweet, gerade eben irgendwo von irgendjemandem abgesetzt. Programmiert hat diesen Zwitscher-Seismografen der Franzose Franck Ernewein. Nur so zum Spaß, wie er sagt, und anfangs nur für sich.

Doch schnell fand Tweetping Bewunderer: „Mir haben viele Leute geschrieben, die eine maßgeschneiderte Karte möchten, mit Zoom-Funktion und mit Filtern für Hashtags“, erzählt er. Ernewein arbeitet daran, aber er ist sich nicht sicher, ob die Nutzer bereit wären, dafür auch zu zahlen.

Wenn er sich da mal nicht irrt. Denn das weltweite Gezwitscher ist eine Menge Geld wert. Seit dem 21. März 2006, als Twitter-Mitbegründer Jack Dorsey den allerersten Tweet absetzte, sprudelt der Datenstrom aus der Twitter Firehose – das ist das englische Wort für Feuerwehrschauch. Jede Sekunde kommen 5000 Tweets zu dem Datenberg von 300 Milliarden bislang gesendeten Tweets hinzu. Denn kein Tweet ist verlorengegangen, sie sind alle gespeichert.

Man könnte Twitter somit auch als eines der größten Crowdsourcing-Projekte aller Zeiten bezeichnen: 230 Millionen Menschen geben bereitwillig Auskunft darüber, was sie

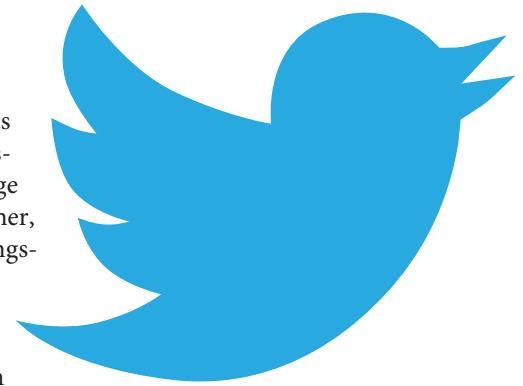


Foto: Screenshot aus tweetping.net, Grafik: Twitter



Globales
Zwitscherkonzert:
Tweetping zeigt
in Echtzeit an, wo
gerade auf der
Welt getwittert
wird.

bergen die Tweets
auch für die Wis-
senschaft wichtige
Daten: Sozialforscher,
Meinungsforschungs-
institute und Po-
litiker analysieren
im Gezwitscher
die Stimmung in
der Bevölkerung.



Zugriff auf seine Firehose gewährt Twitter nur noch exklusiven Partnern: Data-Resellern wie Gnip, Dataminr, DataSift, Crimson Hexagon oder Topsy. Erst kürzlich sorgte die 200-Millionen-Dollar-Übernahme von Topsy durch Apple für Schlagzeilen. Denn damit sicherte sich Apple exklusiven Zugriff auf die Twitter Firehose – ohne dass ein Reseller mitbekommt, was der sehr auf Geheimniskrämerei bedachte Konzern mit den Daten vorhat. Apple, so wird spekuliert, könnte durch diesen Coup zu den Datenaggregatoren Google und Microsoft aufschließen und mit dem Twitter-Schatz seine Suchalgorithmen, die Spracherkennungsfunktion und seine Mobilmarketing-Plattform iAd ausbauen.

Wie viel Gnip und Co. für den Zugriff auf die Firehose zahlen, darüber herrscht einvernehmliches Schweigen. Eine Anfrage von Technology Review diesbezüglich ließ Twitter unbeantwortet, ebenso wie die Frage, wie man überhaupt in den Firehose-Klub kommt. Früher war das einfacher: 2007 und 2008 hatte Twitter Partnerschaften mit Google, Yahoo und Microsoft, die Tweets in ihre Suchmaschinen integrierten. Seit 2010 hat das Unternehmen den Zugang zu seinem Datenschatz jedoch immer weiter beschränkt, teilweise Firmen sogar ausgesperrt, die mit ihren Diensten Twitters Marktstrategie in die Quere kamen. PeopleBrowsr ist so ein Opfer. Sein Geschäftsmodell fand Twitter offenbar zu bedrohlich. PeopleBrowsr ermöglicht seinen Kunden den Aufbau und die Analyse ihrer eigenen sozialen Netzwerke.

Aber wie kommt Franck Ernewein an die Echtzeitdaten für Tweetping? „Die Daten stammen aus der öffentlichen Twitter Streaming API“, sagt er – eine Schnittstelle, die Twitter für jeden zur Verfügung stellt. Aus ihr tröpfeln aber nur etwa ein bis maximal 40 Prozent des Tweets-Feuerwehrschauchs. Zwar hat Twitter Programmierern in der Vergangenheit zunehmend strengere Regeln auferlegt, dennoch behält es den freien Kanal bei – wahrscheinlich aus Werbegründen.

Auch Frederik Fischers Firma Tame zapft Tweets aus der freien API. Fischer ist Journalist, und für ihn ist Twitter zu einer unverzichtbaren Recherchequelle geworden. Immer wieder ärgerte er sich über die beschränkten Möglichkeiten von Twitters hauseigener Suche. Als Fischer 2011 für eine Geschichte über den Süd-Sudan Informanten recherchierte, stieß er immer wieder an Grenzen und beschloss gemeinsam mit einigen Freunden, die Suche in den Tweets selbst zu optimieren. „Es gab und gibt bei Twitter keine ausgeklügelte Suchlogik wie Googles Page-Rank. Das wollten wir ändern“, sagt er. Aus der Idee wurde Tame, ein mittlerweile crowdfinanziertes

tun, was sie denken und wo sie sind. Diesen immensen Datenstrom will das noch immer defizitäre Unternehmen Twitter zu Geld machen. Im Zuge seines Börsengangs im vergangenen Jahr legte der Konzern erstmals seine Finanzen offen: Größte Einnahmequelle ist nach wie vor Werbung, Tendenz steigend. Aber Datenlizenzierung macht bereits 13 Prozent der Erlöse aus – in den ersten sechs Monaten des Jahres waren das rund 32 Millionen US-Dollar, ein Drittel mehr als 2012 und mehr als doppelt so viel wie 2011.

Interesse an den Daten haben Firmen, für die die Tweets ihrer Kunden ein wertvolles Marketingmessinstrument sind. Banken, Fondsmanager und Aktientrader glauben, in dem Gezwitscher Stimmungsänderungen lesen und daraus Aktienkursentwicklungen vorhersagen zu können. So brüstet sich der Twitterdaten-Analyst und -Reseller Dataminr, seinen Kunden im November 2013 einen dreiminütigen Vorsprung verschafft zu haben, um ihre abstürzenden BlackBerry-Aktien abzustoßen, und 2011 einen zwanzigminütigen, bevor die erste Nachrichtenagentur den Tod Osama Bin Ladens verkündete.

Auch Fernsehmacher interessiert das Second-Screen-Geplauder ihrer Zuschauer zunehmend. Sie ziehen daraus Rückschlüsse auf den Erfolg ihrer Sendungen. Je aktiver eine Person auf Twitter ist – Politiker, Popstars oder Schauspieler –, desto höher ist auch ihr Einfluss und damit ihre Werbewirkung. Unternehmen wie Klout und PeerIndex messen diese Online-Reputation und kassieren bei der Werbeindustrie. Und natürlich





Start-up, das eine Twitter-Suchmaschine bietet, die vorwiegend für Journalisten, PR-Leute und Politiker gedacht ist. Tame sortiert das unübersichtliche Gezwitscher nach Hashtags, Tweetern und verbundenen Links, und es rankt die Suchergebnisse. „Wie im klassischen Journalismus ist für den Tame-Rank das wichtigste Kriterium die Quellenvarianz, also wie viele unabhängige Quellen über etwas tweeten.“ Fischer und seine Mitbegründer besitzen zwar nicht den vollen Firehose-Zugriff. Über eine ausgeklügelte Programmierung aber könnten Nutzer dennoch den vollen Twitter-Datenschatz nutzen, verspricht er.

Das Potenzial der Twitter-Daten und die Anwendungen erscheinen endlos: Unternehmen können Tweets nutzen, um darin Potenziale für neue Märkte auszuloten. Wie Behörden Datenschutz zum Wohle der Öffentlichkeit nutzen können, demonstrierten Forscher der Universität Rochester. Sie spürten mithilfe von Twitter schlechte Restaurants auf. Mehrere Monate hinweg analysierten sie insgesamt 3,8 Millionen Tweets von 94 000 Twitter-Nutzern. Wenn eine Person einen Tweet aus einem Restaurant sandte, überwachte eine Software sein Gezwitschere einige Stunden lang und analysierte sie auf relevante Schlüsselwörter wie „sick“, „stomach“, „tummy“, „throw up“ und ähnliche. Am Ende gelang es den Forschern, aufgrund dieser Daten diejenigen Restaurants aufzuspüren, bei denen Lebensmittelvergiftungen am wahrscheinlichsten waren.

Tweets können sogar Leben retten: „Nach dem letzten großen Erdbeben auf den Philippinen im Oktober wurde Twitter quasi als Sensor genutzt“, sagt Alex Rutherford, Data Scientist beim UN-Projekt Global Pulse. „Anhand der Geodaten und Berichte konnte schnell Hilfe dorthin geschickt werden, wo sie dringend benötigt wurde.“ Bei Katastrophen ist das Echtzeit-Medium Twitter eine wertvolle Hilfe. Können Tweets aber auch Katastrophen ankündigen? Das Programm Global Pulse der Vereinten Nationen will aus dem Gezwitscher Krisenherde heraushören, möglichst noch bevor es zur Katastrophe kommt.

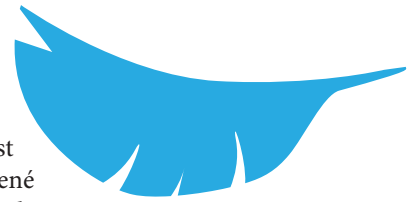
2010 startete Global Pulse unter der Leitung von Robert Kirkpatrick gemeinsam mit dem Social-Media-Analysten Crimson Hexagon, der Zugang zur Firehose besitzt, ein Forschungsprogramm in Indonesien. Global Pulse analysierte alle indonesischen Tweets, die den Reispreis thematisierten – und tatsächlich: Die Frequenz der Tweets stieg parallel zum Preis. „Es ging uns darum zu zeigen, dass Twitter-Daten geeignet sind, wichtige Krisen-Indikatoren zu erkennen – in Echtzeit“, sagt Rutherford. Dann analysierten sie die Tweets auch inhaltlich im Hinblick auf Essenspreise und Versorgungslage. Auch hier ein klares Ergebnis: „Negative Stimmungsbekundungen nahmen zu, wenn der Reispreis anstieg.“ Eine Vorhersage des Reispreises war jedoch nicht möglich. „Stimmungsanalyse-Tools sind für Englisch sehr verbreitet. Dies war eine der ersten Studien, die sie auf eine nicht-europäische Sprache anwandte“, sagt Rutherford. „Daher muss man die Analysesoftware sehr sorgsam kalibrieren und berücksichtigen, dass sich Sarkasmus, Slang und Emoticons von Land zu Land unterscheiden.“

Momentan ist die Genauigkeit der Stimmungsanalyse noch limitiert, weil „genauere Preisdaten fehlen, die wir benötigen, um die Software zu kalibrieren“, sagt Rutherford. Offizielle Preisdaten seien oft nur monatlich verfügbar. Mit besseren Statistiken aber, so glaubt er, würden künftig auch Vorhersagen möglich. Aber nicht für alles: „Twitter ist öffentlich und flüchtig“, sagt René Clausen Nielsen, Big-Data-Analyst bei Global Pulse. „Leute twittern nicht über ihre privaten Angelegenheiten oder ihre Lebensplanung, es geht viel um das Hier und Jetzt. Man kann Trends erkennen, sogar auf Länderebenen, aber die Konversationen gehen nicht so sehr in die Tiefe.“

Es ist Vorsicht angebracht, wenn man aus Tweets Stimmungen, Interessen oder Trends ablesen will. Da gibt es zum einen die demografische Verzerrung, laut einer Studie des Pew Research Center aus dem Jahr 2012 unter 1800 US-Bürgern ist der typische Twitterer jünger als 30 und Stadtbewohner. Und wie Franck Erneuewings Tweetping-Karte eindrücklich zeigt, stammt er mit hoher Wahrscheinlichkeit nicht aus einem Entwicklungsland. Zum anderen ist etwa jeder zehnte aktive Twitter-Account ein Fake, entweder erstellt von Geschäftemachern, die eiteln Personen die falschen Vögel als Follower verkaufen. Oder dahinter verbergen sich Stimmungsmacher, die beispielsweise vor Wahlen Desinformation streuen und Meinungen beeinflussen wollen. So geschehen 2010, als eine Gruppe von neun Fake-Accounts eine sogenannte „Twitter-Bombe“ zündete. Innerhalb von zwei Stunden setzten sie 929 Tweets ab. Ziel der Angriffe: Die US-Politikerin Martha Coakley, die gerade im Wahlkampf um den Senatsitz in Massachusetts stand. Sie verlor das Rennen. 60 000 Menschen hatten diese gesteuerten Tweets gelesen, wie eine nachträgliche Analyse ergab. Solche Vorfälle lassen auch die Zuverlässigkeit minutenkritischer Vorhersagen von Dataminr in einem fragwürdigen Licht erscheinen. Nicht auszudenken,

wie anfällig hypernervöse Aktienspekulanten für solch gezielte Manipulationen sein dürften.

Twitter gibt sich mittlerweile große Mühe, Fälscher und Spam aufzuspüren. Was aber, wenn das falsche Gezwitscher von einem Stück Software stammt? Die sogenannten Social Bots – computergesteuerte Accounts, die menschliches Verhalten imitieren, sind mittlerweile so gut, dass sie nicht nur Twitters Spamfilter täuschen, sondern auch echte Menschen. Tame-Gründer Frederik Fischer kennt das Problem, er hat selbst darüber als Journalist geschrieben: „Social Bots verhalten sich die meiste Zeit unauffällig“, sagt Fischer. „Sie twittern nur gelegentlich und folgen dabei verschiedenen, vorher einprogrammierten Interessen.“ So bauen sie sich nach und nach Reputation auf und können dann gezielt eingesetzt werden, um Meinungen zu beeinflussen. Es gilt also genau hinzuhören beim Zwitscherkonzert.

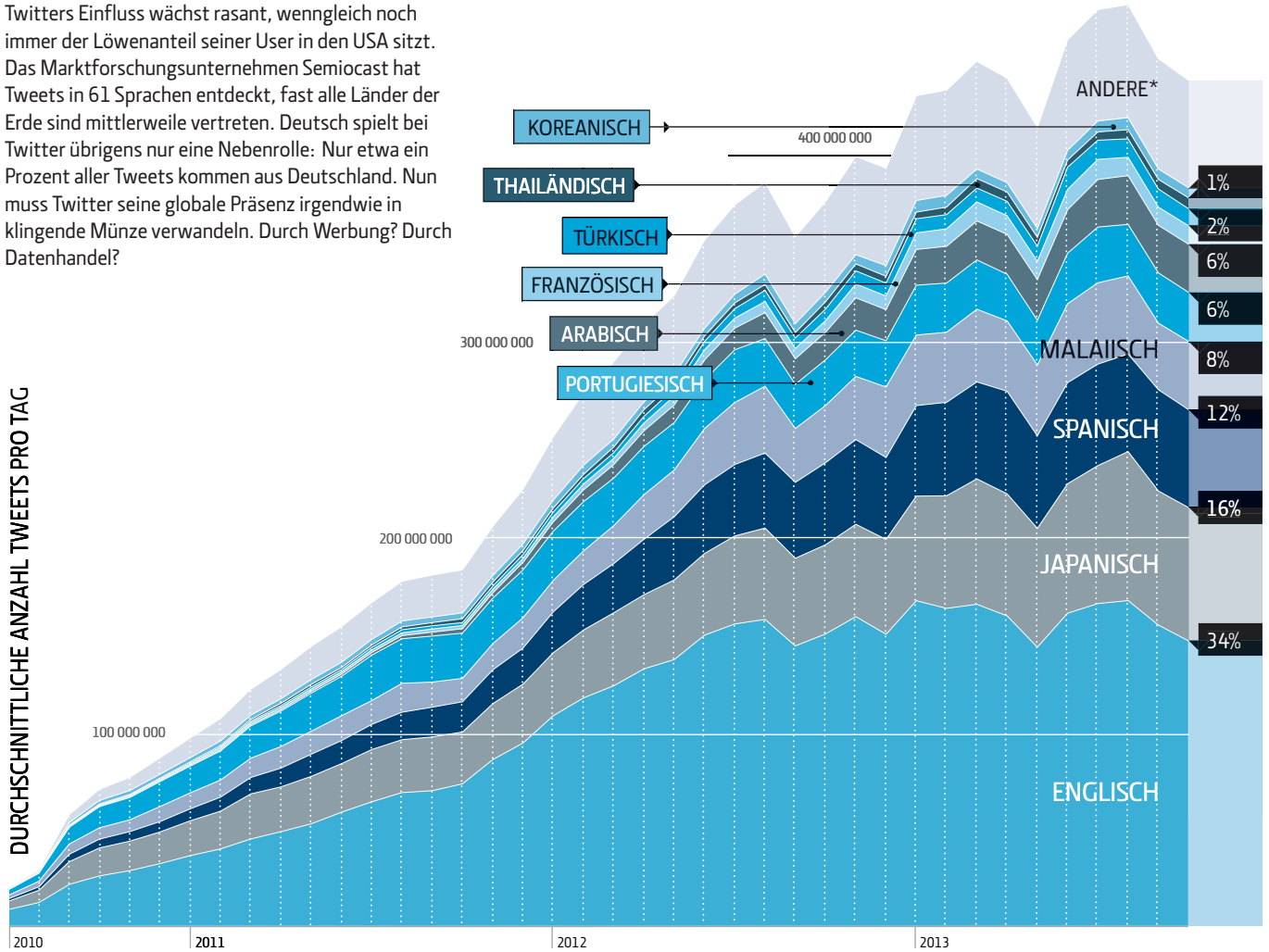


Illustrationen: Walter Baumann. Daten: Semicast, Twitter, eMarketer



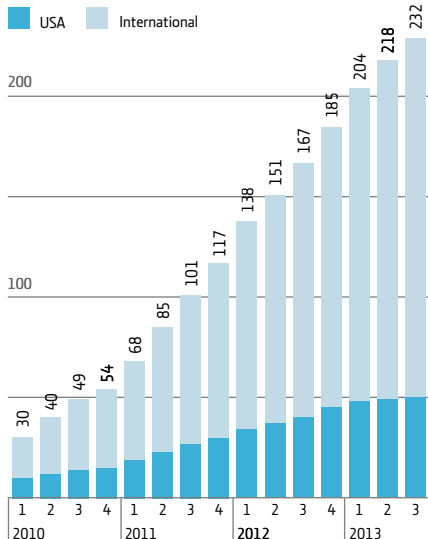
Die Twitter-Gesellschaft

Twitters Einfluss wächst rasant, wenngleich noch immer der Löwenanteil seiner User in den USA sitzt. Das Marktforschungsunternehmen SemioCast hat Tweets in 61 Sprachen entdeckt, fast alle Länder der Erde sind mittlerweile vertreten. Deutsch spielt bei Twitter übrigens nur eine Nebenrolle: Nur etwa ein Prozent aller Tweets kommen aus Deutschland. Nun muss Twitter seine globale Präsenz irgendwie in klingende Münze verwandeln. Durch Werbung? Durch Datenhandel?



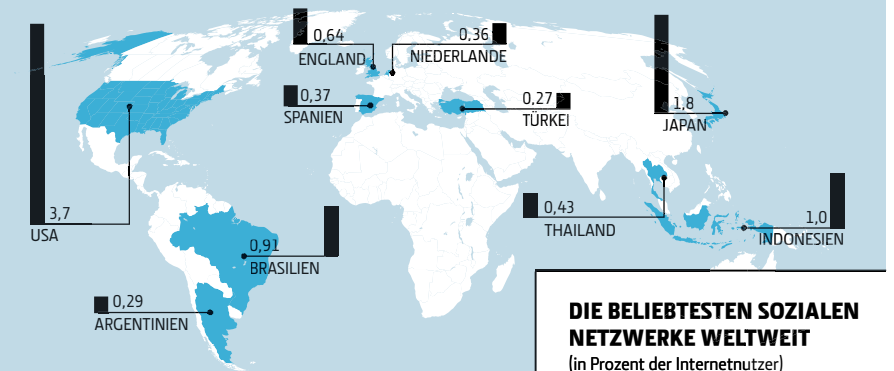
AKTIVE TWITTER-NUTZER WELTWEIT

(in Millionen pro Quartal)



DIE ZEHN AKTIVSTEN TWITTER-NATIONEN WELTWEIT IM JUNI 2013

(in Milliarden Tweets)



WERBEEINNAHMEN PRO 1000 LOG-INS

1. Juli bis 30. September 2013



DIE BELIEBTESTEN SOZIALEN NETZWERKE WELTWEIT

(in Prozent der Internetnutzer)

