

SOFTWARE

# BIG DATA GRÖSSENWAHN

29.09.2016

Quelle [www.brandeins.de](http://www.brandeins.de) diskutiert eine Arbeit aus [Nature 2009](#):

Das US-amerikanische Wissenschaftsmagazin präsentierte eine Studie von Datenwissenschaftlern in Diensten von Google und Yahoo. Die konnten plausibel nachweisen: Wenn viele Leute Suchwörter wie „Grippesymptome“ oder „Apotheke in der Nähe“ eingeben, lassen sich geografische Ausbreitungen von Seuchen nahezu in Echtzeit berechnen, ohne dass irgendjemand mit einem einzigen Arzt sprechen muss. Die Anwendung dazu nannten sie GFT, Google Flu Trends.

Mit gesundem Menschenverstand leicht nachzuvollziehen war, dass die Methode um ein Vielfaches präziser und schneller sein würde als das gängige Verfahren des amerikanischen Centers for Disease Control and Prevention. Die Beamten der Behörde sammeln Fragebögen bei Arztpraxen zu Patientenzahlen ein und rechnen mit ein bis zwei Wochen Zeitverzögerung die gesundheitliche Lage der Nation hoch.

Dann aber kam eine Arbeit in [Science 2014](#)

Rückblickend hatte sich die Vorhersagequalität von GFT als lausig erwiesen. Die Beamten mit ihren Fragebögen hatten im Durchschnitt bessere Vorhersagearbeit geleistet. Kurz darauf wurde eine Studie von Datenwissenschaftlern rund um den Harvard-Ökonom Ryan Kennedy publiziert, die den Aufstieg und Fall des Datenmusterknaben Google Flu Trends zur Parabel für Big-Data-Analytik insgesamt erklärte.

siehe auch [Wired](#) “What We Can Learn From the Epic Failure of Google Flu Trends?”

