MERTON
Onlinemagazin des *Stifterverbandes*

Foto: CC0 - Public Domain

„Big Data findet alles“

VERÖFFENTLICHT AM 22.09.2014

Matthias Hagen weiß, wie aus unüberschaubaren Daten nützliche Informationen werden. Interview mit dem ersten Professor für Big-Data-Analytics.

LESEZEIT: 4 MINUTEN

TEXT:
LUKAS GRASBERGER ›

Herr Hagen, Sie haben die erste Big-Data-Professur Deutschlands inne. Bekannt für die effiziente Analyse und Nutzung großer Datenmengen sind eigentlich US-Konzerne wie Google & Co. Wozu braucht es eine Big-Data-Professur?
Wir wollen spannende Fragen finden, die sich dank großer Datenmengen beantworten lassen. Das macht Google bestimmt auch – behält das aber weitgehend für sich. Dagegen ist unser Ziel, darüber auch zu veröffentlichen, mitzuteilen, was wir gefunden haben. Spannend für die Wissenschaft ist dabei etwa information retrieval: die Suche in großen Textdatenmengen. Wenn ich etwa eine Hintergrundrecherche gemacht und eine Handvoll Dokumente gefunden habe, die ich interessant finde, ist es immer schwierig herauszufinden, was ich mir noch alles ansehen müsste. Wenn ich weiß: Fünf Texte zu finden, hat mich zehn Minuten gekostet – dann dauert das Weiterwühlen, um die nächsten 50 relevanten Dokumente zu finden, vielleicht mehr als eine halbe Stunde. Bekomme ich diese Dokumente dagegen automatisch vorgeschlagen, würde mir das viel Zeit sparen.

Um zu verstehen, wer etwas aus welchem Grund sucht, muss man ja praktisch in den Kopf des Nutzers schauen. Wie machen Sie das?

Wir geben Nutzern Aufgaben vor und beobachten, wie sie suchen. Es geht darum, das optimale Nutzerverhalten herauszufinden, also: Wie man sich verhalten müsste, um den meisten Nutzen aus einer Suche zu ziehen. Dann versuchen wir, den „idealen Nutzer“ zu simulieren. Riesige Datenmengen von Nutzerinteraktionen helfen uns, Suchtypen zu identifizieren, herauszufinden, wie viel Zeit und Mühe jemand etwa in eine Recherche investiert, wenn er am PC sitzt oder das Smartphone bedient. Der ideale Nutzer ist deutlich besser als das, was der Mensch macht: Man kann in etwa dreimal so viel in der selben Zeit herausholen aus den Suchergebnissen, wie das die Menschen geschafft haben, deren Verhaltensweisen wir beobachtet haben. Dank dem „idealen Nutzer“ können wir dann die Suchmaschine Netspeak, die wir gebaut haben, an die jeweiligen Bedürfnisse der Suchenden anpassen, sodass diese genauer und schneller fündig werden.

Bernd Stuckmeyer/Seitenplan

Bit für Bit verändert die Digitalisierung Forschung und Lehre. Doch was heißt das für das Wissenschaftssystem und die Menschen, die darin arbeiten?

Lesen Sie mehr >

Ein großer Vorteil von big data in der Wissenschaft ist ja auch, dass sich bisher verborgene Zusammenhänge aufdecken lassen ...
Ja, durch das mining, die automatisierte Suche in großen Textmengen, kann ich Zusammenhänge entdecken, die ein Mensch sonst nie finden würde: Weil er es einfach nie schaffen würde, die Millionen von Dokumenten zu lesen. Durch die automatisierte Analyse vieler Studien kann ich beispielsweise Nebenwirkungen oder Wechselwirkungen von Medikamenten finden. Ich kann Testreihen viel gezielter angehen. Das gab es früher so nicht, und das steckt auch jetzt noch in den Kinderschuhen.

Wo verändert big data die Wissenschaft am meisten?
Simulationen sind sehr spannend: Sie sind in vielen Bereichen der Ingenieurstechnik unerlässlich geworden. Also nicht nur echte Messwerte zu erheben, sondern auf der Basis von Daten, die ich früher einmal erhoben habe, Dinge zu simulieren. Solche Simulationen produzieren unheimlich viele Daten in nahezu beliebiger Größenordnung. Wir können diese heute nicht nur abspeichern, sondern auch der Analyse zugänglich machen – und es gibt jetzt auch Ideen, welche Fragen ich damit beantworten kann. Das ist ein Bereich, der auch noch gar nicht ausgereizt ist. Viele Versuche würden sehr viel Geld verbrennen, wenn man sie tatsächlich machen würde – oder sind vielleicht sogar ethisch-moralisch nicht in Ordnung. Wenn man zum Beispiel Tierversuche simulieren könnte, wäre das sicherlich nicht nur für die Tiere besser, sondern auch für denjenigen, der früher die Experimente durchführen musste.

QUELLE: HTTPS://MERTON-MAGAZIN.DE/„BIG-DATA-FINDET-ALLES“