

Die Big-Data-Täuschung

Prof. Dr. Norbert Gronau, Professor für Wirtschaftsinformatik, Schwerpunkt Prozesse und Systeme, Universität Potsdam

Die Vision von Business Analytics liegt darin, emotional geprägte Ergebnisse durch Erkenntnisse, gewonnen aus einer riesigen puren Datenmenge und ihrer Auswertung, zu ersetzen. Social Media, das Internet der Dinge oder ein Cyber-physisches System verbinden informationsgetriebene und technische Komponenten. Soweit die Idee. Doch häufig kann Big Data gerade bei Fragestellungen, in denen es um Menschen und Meinungen geht, nur Korrelationen, nicht aber kausale Zusammenhänge beschreiben. Die Methode ist unzureichend.

Und so sammeln sich die Gegner der Datenerfassung. Ihre Gegenargumente werden lauter: Die pure Masse bringt nichts. Bei Twitter ist die User-Gruppe zu homogen, um Vorhersagen herauszufiltern. Das Lehrgeld ist zu hoch, weil 50 Prozent der Big-Data-Projekte scheitern. Aber mit diesen Argumenten wird das Kind mit dem Bade ausgeschüttet. Im Bewusstsein um die Hürden können seriöse Forscher sehr wohl einen Weg finden.

Entscheidend für stimmige Ergebnisse ist es, den menschlichen Faktor einzukalkulieren. Sowohl die Fragestellungen als auch die Interpretation von großen Datenerhebungen sind von Menschen beeinflusst – wissentlich oder durch unbewusste Vorurteile, die jeder aus seiner Kultur, seinen Werten und seiner Biografie mitbringt. Personen setzen die Algorithmen auf und bestimmen, welche Items für Forschung oder Anwendung

genutzt werden oder eben nicht. So wollte Amazon seinen Rekrutierungsprozess durch künstliche Intelligenz perfektionieren, um aus jeweils 100 Bewerbern die fünf Besten zu eruieren. 2017 wurde das automatisierte Recruiting-Projekt gestoppt, weil die Algorithmen systematisch Frauen benachteiligten. Die Ursache dieser maschinellen Vorurteile: Als Lernmaterial wurde das System mit Bewerbungen der vergangenen zehn Jahre gefüttert, die fast ausschließlich von Männern eingereicht wurden. Die Eigenschaft männlich wurde dadurch hoch bewertet.

Raus aus der Big-Data-Falle

Freilich kann man gegensteuern, wenn man sich die Fehlerquellen bewusst macht. Zum einen kann man kulturelle Vielfalt ins Team holen. Zum anderen muss man den menschlichen Faktor offenlegen. Doch es fehlen die Spezialistinnen und Spezialisten, die mit den Anwendungen richtig umgehen können. Und nur wenn Menschen die richtigen Fragen stellen, erzielen die Systeme brauchbare Ergebnisse. Big Data kann mehr als die Vor- und Werturteile der Eingebenden widerzuspiegeln.

Allerdings liegen die Daten ganz oft nicht in auswertbaren Formaten vor. Algorithmen bleiben zu oft ein Rätsel. Für die Wirtschaft bedeutet das: Die richtigen Anwendungen für Business Analytics müssen sehr aufwendig in den Unternehmen eingerichtet werden. Schon das ist ein Problem für kleine und mittelgroße Firmen. Außerdem ist unsere Klick-and-Touch-Welt auf Textdateikonfigurationen ausgerichtet und für das Hochladen von Daten in Big Data Cubes häufig nicht mehr vorbereitet.

Was also können die Unternehmen tun, die noch kein Business Analytics umsetzen können? Sie können Small-Data-Anwendungen nutzen. Dabei werden kleinere Gruppen von Mitarbeitern beobachtet und interviewt. Mit rund 50 Teilnehmern und Teilnehmerinnen erhält man veritable Ergebnisse – wenn die Fragestellungen und Auswertungen dem Small Data angepasst sind.

Für große Forschungsprojekte und die Chance auf Massendaten rege ich die Untersuchung der Frage an, ob eine Kombination von Big-Data-Analysen und Small-Data-Elementen die Seriosität von Ergebnissen stärken kann. Möglicherweise ist es sinnvoll, eine Integrationsplattform zu schaffen zwischen Big Data und Small Data. Dann werden den großen Zahlen qualitative Aspekte hinzugefügt.

Abb.: Forschung zur Datennutzung

