

Wie weit darf die Verfälschung von Daten gehen?

Die Verzerrung von Forschungsdaten ist nicht in jedem Fall verwerflich. Manchmal ist sie sogar unumgänglich / Von Stefan Kühl

Ich habe in meinen wissenschaftlichen Arbeiten Daten verfälscht. Ich habe Länder, in denen Ministerien angesiedelt sind, verändert, Verwaltungen aus Zentraleuropa nach Nordeuropa verlegt und von Asien nach Afrika. Ich habe in wissenschaftlichen Artikeln so getan, als ob Interviews mit Mitarbeitern in anderen Unternehmen als den von mir beschriebenen stattgefunden hätten. Und ich habe in Artikeln manchmal aus zwei untersuchten Verwaltungen drei gemacht oder auch aus drei zwei.

Dies habe ich nicht gemacht, um die Daten so hinzubiegen, dass sie zu meinen Thesen passen, und auch nicht, um die Empirie interessanter zu machen, als sie sowieso schon war. Der Grund für die bewusst vorgenommenen Verfälschungen war die von mir wahrgenommene Notwendigkeit, mit allen verfügbaren Mitteln sicherzustellen, dass nicht rekonstruiert werden kann, woher die Daten stammen und besonders von welchen Personen sie gegeben wurden.

Empirische Sozialforscher stehen vor einem grundlegenden Dilemma. Auf der einen Seite müssen sie ihre Befunde so genau wie möglich berichten, um anderen Wissenschaftlern zu ermöglichen, das Argument detailliert nachzuvollziehen. Nur so können Kollegen einschätzen, ob die Schlussfolgerungen plausibel sind und ob sie generell überzeugen oder nur für be-

sondere Fälle. Auf der anderen Seite gehört es zum wissenschaftlichen Standard, Daten so zu anonymisieren, dass die befragten Familien, Organisationen, Freundeskreise, Protestbewegungen oder Dörfer – und erst recht die befragten oder beobachteten Einzelpersonen – nicht zu erkennen sind.

Das Spannungsfeld zwischen Präzision und Anonymisierung ist prinzipiell nicht auflösbar. Je detaillierter man ein soziales Phänomen beschreibt, um das eigene Argument nachvollziehbar zu machen, desto leichter ist es zu erkennen. Manchmal reicht schon ein Zitat aus einem im Internet zugänglichen Dokument, um ein Dorf oder eine Regierungsbehörde zu identifizieren. Manchmal legt schon allein die Angabe von Größe, Land und Branche nahe, um welche Organisation es sich handelt.

Zugespielt kann man sagen: Wenn Forscher nicht durch ihre Daten zu großer Vorsicht im Umgang mit Quellen gezwungen werden, haben sie nichts herausbekommen. Sie haben vermutlich nur die Schauseite beobachtet und brauchen deshalb nicht auf strikte Anonymisierung zu achten. Man kann diese Effekte bei Praktikerveitschriften beobachten, die viel Wert darauf legen, dass in ihren Artikeln alles durch Echtnamen „authentifiziert“ wird.

Besonders bei der Berichterstattung über Organisationen hat dies oftmals eine Selbstzensur der Autoren zur Folge,

weil sie darauf achten müssen, die beratenen oder befragten Unternehmen, Verwaltungen oder Ministerien zu schonen. Deshalb kann man bei Fallstudien in Praktikerveitschriften, die auf Anonymisierung verzichten, sicher sein, dass man vielleicht einiges über die Vorderbühnen von Organisationen, aber faktisch nichts über deren Hinterbühne erfährt. Man kann daraus schon fast eine Daumenregel zur Zeitersparnis für die Lektüre von Artikeln und Büchern machen. Mit Ausnahme von historischen Analysen, in denen die Betroffenen in der Regel schon lange tot sind, und von investigativen journalistischen Arbeiten, in denen der Anspruch ja gerade die Aufdeckung der Hinterbühne ist, kann man Texte, die auf Anonymisierung verzichten, getrost ungelesen zur Seite legen. Bestenfalls würde man etwas über die neuste Beratungs- oder Managementmode erfahren, die über Erfolgsgeschichten aufgepeppt werden, schlimmstenfalls liest man eine als Praktikerveicht getarnte PR-Geschichte eines Beraters oder Managers.

Unter Forschungsaspekten ist es das höchste Kompliment für einen Forscher, wenn die Beforschten die in wissenschaftlicher Sprache beschriebenen Strukturen, Prozesse und Effekte als plausibel einschätzen, ohne zu erkennen, dass sie auf der Analyse des eigenen Systems basieren. Der Kommentar „Das ist ja wie bei uns“, ohne sich selbst als empirische

Grundlage der Beschreibung wiederzuerkennen, ist nicht nur Ausdruck der Plausibilität wissenschaftlicher Darstellungen, sondern weist ebenso darauf hin, dass die Anonymisierung höchsten Ansprüchen genügt. Aber die Kosten dafür sind in der Regel, dass wesentliche Details geändert wurden, ohne die ein Wiedererkennen möglich ist. Anonymisierung geht folglich immer auf Kosten der Präzision in der Darstellung der Empirie.

Sicherlich, in vielen Fällen wird dieses grundlegende Spannungsfeld empirischer Sozialforschung nicht virulent, weil man die Standards der Anonymisierung mit guten Gründen schleifen lassen kann. Sozial ist die Welt der wissenschaftlichen und der außerwissenschaftlichen Praxis so stark getrennt, dass man keine Sorgen haben muss, dass die befragten Personen am Ende die wissenschaftlichen Artikel lesen. Wenn das Gerücht stimmt, dass der durchschnittliche Artikel in einer Fachzeitschrift von einer einzigen Person gelesen wird, wäre es überraschend, wenn gerade dieser Leser aus der untersuchten Familie, der befragten Organisation oder dem analysierten Dorf kommt.

Das Spiel wird von Forschern und Beforschten routiniert gespielt. Die Beforschten erklären vor dem Gespräch ihr Interesse an der Forschung, die Forscher erklären sich zur Rückmeldung bereit, und beide vergessen bewusst oder unbe-

wusst die versprochene Rückmeldung. Wie viele Promotionen, wie viele Forschungsberichte sind letztlich ungelesen in den Schubladen und auf den Festplatten der Organisationen verschwunden, die den Feldzugang überhaupt nur unter der Voraussetzung zugestanden haben, dass man die Forschungsergebnisse bekommt? Vermutlich könnte man die Anforderung von Organisationen, die Forschungsergebnisse danach schriftlich zur Verfügung gestellt zu bekommen, allein dadurch befriedigen, dass man den immer gleichen Dummy-Text aus dreihundert Seiten herausschickt und darauf vertraut, dass Texte von Praktikern sowieso nicht gelesen werden.

Das hängt auch damit zusammen, dass die zeitlichen Horizonte ganz unterschiedlich sind. Bevor Wissenschaftler ihre Erkenntnisse zu Papier gebracht oder gar publiziert haben, haben die Beforschten bereits vergessen, dass sie überhaupt befragt wurden. Und nicht selten sind auch die sachlichen Darstellungsformen ganz verschieden. Wissenschaftler wählen in der Regel eine Sprache, die für Praktiker nicht ohne weiteres zugänglich ist und schützen so ihre Quellen.

Aber die Unwahrscheinlichkeit der Erkenntnisnahme wissenschaftlicher Erkenntnisse rechtfertigt Nachlässigkeiten bei der strikten Anonymisierung. Wissenschaftler müssen immer mit dem unwahrscheinlichen Fall rechnen, dass ihre Er-

kenntnisse nicht nur zur Kenntnis genommen, sondern auch diskutiert werden. In diesem Fall kann eine mangelhafte Anonymisierung verheerend sein, weil durch die Beschreibungen unkontrolliert Strukturen aufgedeckt werden, die aus guten Gründen im Tabubereich gehalten worden sind.

Wenn man mit Forschern spricht, dann fällt auf, dass die bewusste Verfälschung von Daten zur Sicherstellung der Anonymität gängige Praxis ist. Die Interviewsequenzen einer Person werden unterschiedlichen Personen zugewiesen, so dass die konkrete Person nicht erkennbar ist. Es werden in Texten „Nebelkerzen“ gezündet, um die Leser daran zu hindern, die konkrete Organisation, Familie, Bewegung oder Gruppe zu entschlüsseln und darüber auch einzelne Personen identifizieren zu können.

Man ist mit dieser Praxis der Verfälschung also nicht allein, kann sich aber dabei nicht auf etablierte professionelle Standards stützen. Die Empfehlungen der Fachgesellschaften ignorieren dieses Dilemma zwischen Präzision und Anonymisierung und verkünden lieber allgemeine Prinzipien für eine Forschungsethik. Konkret bleibt es den einzelnen Forschern überlassen, eine situationsbezogene Lösung für dieses Dilemma zu finden.

Stefan Kühl ist Professor für Soziologie an der Universität Bielefeld.