

Modellierung: eine Begriffsbestimmung

,
shessbru@hse.ru
National Research University - Higher School of
Economics, Russland

Meine Präsentation ist der Frage gewidmet, inwiefern der Modellbegriff zur umfassenden Charakterisierung von Forschungsleistungen innerhalb der *digital humanities* tauglich ist. Dabei gehe ich aus von der sehr allgemeinen Definition William McCartys, der ein Modell als "von Natur aus vereinfachte und deshalb fiktionale oder idealisierte Repräsentation" eines gegebenen Gegenstandes oder Gegenstandsbereiches auffasst und daraus die Behauptung ableitet, die *digital humanities* seien wesentlich mit der heuristischen Exploration solcher digital erfassbarer Modelle von Texten oder anderen Daten befasst (McCarty 2004: par. 3). Was als übergreifendes Erkenntnisziel solcher heuristischer Bemühungen in Frage kommen soll, bleibt bei ihm jedoch im Dunkeln. Zu finden ist lediglich die Bestimmung, dass digitale Modelle "für eine bestimmte Zeit andauernde Zustände in einem Prozess, in dem man zu Wissen gelangt, anstatt fixe Strukturen von Wissen" (McCarty 2004: par. 9). Ich halte diese These für falsch, jedenfalls was den fest etablierten Gebrauch des Modellbegriffs innerhalb der Naturwissenschaften betrifft.

Geisteswissenschaftlerinnen steht es frei, den Begriff des Modells nach eigenem Ermessen unter Berücksichtigung ihrer eigenen methodischen Vorfestlegungen neu zu definieren. Deswegen ist kurz ein zweites, möglicherweise eher einschlägiges Verständnis des 'Modells' in den Geisteswissenschaften zu diskutieren, das in Diskursen des *software engineering* entstanden sein dürfte. In jenen Bereichen, in denen die Methodologie der *digital humanities* sich mit der Generierung von Computercode überschneidet, kann der Modellbegriff legitim gebraucht werden. Jedoch ist Modellierung in diesem Verständnis ein Werkzeug der digitalen Geisteswissenschaften und darf nicht als Selbstzweck missverstanden werden. Insofern ist Modellierung für die *digital humanities* zwar von Interesse, aber kein Begriff, der unsere Praxis insgesamt erfassen kann.

Modell und Naturwissenschaft

Um die Diskussion übersichtlicher zu gestalten, werde ich mich nicht mit ikonischen Modellen, also etwa dem Billardballmodell von Gasen oder der virtuellen Rekonstruktion einer Predigt von John Donne auseinandersetzen, da hier spezifische Probleme jenseits des allgemein etablierten Modellbegriffs zu erörtern wären (Virtual Paul's Cross s. a.).

Vielmehr bin ich an dem interessiert, was McCarty als 'komputationale Nachverfolgbarkeit' (*computational tractability*) bezeichnet, da hier vielleicht ein gemeinsames Charakteristikum von Modellen in den digitalen Geisteswissenschaften und mathematischen Modellen in den Naturwissenschaften festzustellen wäre. Hier ist die wesentliche Bezogenheit mathematischer Modelle auf Theorien von besonderem Belang (Portides 2008: 386-387). Mathematische Modelle können in theoriegetriebene Modelle und phänomenologische Modelle unterschieden werden. Theoriegetriebene Modelle werden aus einer bereits etablierten Theorie unter Benutzung von lokalen Hypothesen und entsprechender Randbedingungen abgeleitet. Phänomenologische Modelle umfassen lediglich empirische Daten und ad-hoc-Hypothesen, wenn eine Theorie des Gegenstandsbereiches noch nicht zur Verfügung steht. Theoriegetriebene Modelle machen implizite Folgerungen aus bereits existierenden Theorien explizit. Phänomenologische Modelle sollen uns dabei helfen, überhaupt erst einmal eine Theorie zu formulieren.

Aber beide Arten von Modellen dienen dazu, Vorhersagen über das erwartbare Verhalten des jeweils untersuchten Systems zu ermöglichen (Forster 2008: passim). Solche Vorhersagen sind deswegen möglich, weil Theorien über die Natur Erklärungen von Naturereignissen oder -phänomenen enthalten und damit Wissen über allgemeine kausale Relationen zwischen allgemeinen Klassen von Ereignissen: "Kein Modell steht sozusagen für sich selbst; es greift zurück auf riesige Mengen theoretischen und empirischen Wissens, [...] Hintergrundtheorien und empirische Daten verleihen den verschiedenen in einem Modell verwendeten Symbolen Bedeutung." (Barberousse / Ludwig 2009: 61).

Geisteswissenschaftler sind nur selten an Naturgesetzen oder ähnlichen Verallgemeinerungen interessiert. Im Mittelpunkt unseres Interesses steht das einzelne Objekt, selbst wenn es sich etwa um den europäischen Roman zwischen 1800 und 1900 handelt (Moretti 1998).

Dieser Gegenstand ist abstrakt, kann nur durch die Analyse komplexer Daten erforscht werden kann, aber er ist doch ein partikularer Gegenstand. Zwar sind auch Geisteswissenschaftlerinnen an Kausalaussagen interessiert, aber für Historiker sind eher einzelne Kausalaussagen als eventuelle 'Gesetze der Geschichte' von Belang: wichtig ist, warum Caesar den Rubikon überquerte, nicht welche allgemeinen Gesetze für die Überquerung eines Flusses durch Diktatoren gelten.

Fragen, die innerhalb der Wissenschaftstheorie in Bezug auf Modelle erörtert werden, können innerhalb der digitalen Geisteswissenschaften nicht sinnvoll gestellt werden. Ob Modelle als nützliche Fiktionen gelten oder uns zu Wahrheiten über die Welt führen können, ist innerhalb der digitalen Geisteswissenschaften von nachrangiger Bedeutung: unsere Ansprüche auf Wahrheit sind bestenfalls prekär und provisorisch. Ob Modelle Voraussagen über das Verhalten von Systemen ermöglichen ist nicht von Belang, denn wir sind nicht an

der Vorhersage der Zukunft interessiert. Wenn Modelle in den Naturwissenschaften immer in Bezug zu einer bereits existierenden oder zukünftigen Theorie des jeweiligen Gegenstandsbereichs stehen, fehlt dieser Bezug, weil die digitalen Geisteswissenschaften keine Theorie in diesem Sinne entwickeln.

Denkbar wäre jedoch, dass der Gebrauch des Modellbegriffs innerhalb der Statistik als Brücke zwischen naturwissenschaftlichen und geisteswissenschaftlichen Verwendungen dienen könnte. Jedoch sind Modelle auch innerhalb der Statistik eng verschränkt mit dem Begriff der Vorhersage. Hier geht es darum, begründen zu können, ob Daten die Behauptung der Wahrheit einer Hypothese rechtfertigen können oder nicht: "[...] alle statistischen Verfahren beruhen auf der Annahme eines statistischen Modells, hier verstanden als jegliche beschränkte Menge statistischer Hypothesen. Außerdem zielen beide [sc. statistische Verfahren und Modelle] auf die Beurteilung dieser Hypothesen." (Romeijn 2014) Wenn eine statistische Hypothese durch ein Modell und ein entsprechendes Verfahren als wahr erwiesen wird, wird sie dann Teil einer wissenschaftlichen Theorie.

Diese 'kanonisierte' Anwendung statistischer Verfahren ist von ihrem Einsatz in den digitalen Geisteswissenschaften deutlich unterschieden. Hier werden statistische Werkzeuge zur Erkennung von Mustern eingesetzt, ohne dass dem eine zuvor definierte Hypothese zugrundeliegen muss, also ohne ein statistisches Modell zur Anwendung zu bringen. Wir suchen gerade Muster, die wir nicht kennen. Und diese Muster sind wieder einzelne (abstrakte) Gegenstände in diesem Text, in diesem Corpus und nicht Allgemeinbegriffe, die sich auf natürliche Arten beziehen, wie sie in den Naturwissenschaften vorausgesetzt werden.

Außerdem sind nicht alle Muster gleich oder gleich wichtig. Ob ein Muster von Interesse ist, hängt nicht von vorhergehenden Hypothesen über den jeweiligen Gegenstandsbereich ab, sondern vielmehr von unserer Kenntnis anderer Einzelgegenstände, ihren Korrelationen, ihrer historischen Entwicklung und der Erfahrung und Urteilskraft des Forschenden, der sich mit diesem Gegenstandsbereich häufig über lange Zeit vertraut gemacht hat.

Modell und Software

Es existieren also tiefgreifende Unterschiede zwischen dem etablierten Gebrauch des Modellbegriffs in Naturwissenschaften und Statistik und einem denkbaren Gebrauch in den Geisteswissenschaften. Allerdings ist zu beachten, dass der Gebrauch des Ausdrucks "Modell" auch in der Informationstechnologie üblich ist. Vielleicht also leitet sich der Gebrauch des Modellbegriffs in den digitalen Geisteswissenschaften gar nicht aus den Naturwissenschaften ab, sondern aus Theorie und Praxis des Software Engineering. Leider ist der Modellbegriff in diesen Kontexten ähnlich uneindeutig wie in der

naturwissenschaftlichen Praxis (Ludewig 2003: 9). Ihm scheint indes die basale Intuition zugrundezuliegen, dass Software als "Modell der Welt" aufzufassen ist (Ludewig 2003: 9-10). In diesem Zusammenhang ist auf zweierlei hinzuweisen: Software mag als Modell der Welt fungieren, dies ist aber nicht ihre einzige Funktion. Software interagiert auch mit der Welt. Damit unterliegt der Erfolg eines Softwareprodukts Einschränkungen, die außerhalb seiner selbst liegen. Wenn in einer Maschine aufgrund eines Denkfehlers in der Steuerungssoftware zugrundeliegenden Modell eine Fehlfunktion auftritt, kann sich der Entwickler nicht darauf berufen, dass das der Software zugrundeliegende Modell nur als arbiträre Konstruktion eines bestimmten Wirklichkeitsausschnitts gedacht war. Dies gilt auch für Software, die jenen Ausschnitt der Welt repräsentiert, der aus menschlichen Artefakten besteht und von Geisteswissenschaftlerinnen untersucht wird. Führt ein Bug in einem Tokenizer zu falschen Wortzahlen, würden wir die Dinge nicht so lassen, wie sie sind, weil Modelle sowieso nur Approximationen an die Wirklichkeit darstellen.

Wichtiger noch erscheint der Hinweis, dass Repräsentationen der menschlichen Kultur in Softwarewerkzeugen, die in den digitalen Geisteswissenschaften zum Einsatz kommen, nie ein Selbstzweck sind. Selbst wenn man zugesteht, dass Software diese repräsentationale Funktion hat, so soll sie am Ende doch als Werkzeug geisteswissenschaftlicher Forschung funktionieren. Lesen, Denken und Schreiben sind essentielle Bestandteile des analogen geisteswissenschaftlichen Arbeitsprozesses. Aber niemand nähme an, dass diese Aktivitäten intrinsischen Wert haben, wenn sie nicht zum Fortschritt der jeweiligen Disziplin beitragen würden. Auf gleiche Weise sind auch Software und ihre Modelle Teil unseres digitalen Arbeitsprozesses. Solche Repräsentationen haben aber primär instrumentellen Wert und sollten nicht als Selbstzweck missverstanden werden.

Bibliographie

Barberousse, Anouk / Ludewig, Pascal (2009): "Models as Fictions" in: Suárez, Mauricio (ed.): *Fictions in Science. Philosophical Essays in Modeling and Idealizations*. London / New York: Routledge 56-73.

Forster, Malcolm (2008): "Prediction" in: Psillos, Stasís / Curd, Martin (eds.): *The Routledge Companion to Philosophy of Science*. London / New York: Routledge 405-413.

Ludewig, Jochen (2003): "Models in software engineering – an introduction" in: *Software and Systems Modeling* 2: 5-14.

McCarty, Willard (2004): "Modeling: A Study in Words and Meanings", in: Schreibman, Susan / Siemens, Ray / Unsworth, John (eds.): *A Companion to Digital Humanities*. Oxford: Blackwell <http://>

www.digitalhumanities.org/companion/ [letzter Zugriff 14. Oktober 2015].

Moretti, Franco (1998): *Atlas of the European Novel, 1800-1900*. London: Verso.

Portides, Demetris (2008): "Models", in: Psillos, Stasis / Curd, Marin (eds.): *The Routledge Companion to Philosophy of Science*. London / New York: Routledge 385-395.

Romeijn, Jan Willem (2014): "Philosophy of Statistics", in: Zalta, Edward (ed.): *The Stanford Encyclopedia of Philosophy (Winter 2014 Edition)* <http://plato.stanford.edu/archives/win2014/entries/statistics/> [letzter Zugriff 14. Oktober 2015].

Virtual Paul's Cross Project: Virtual Paul's Cross. A Digital Recreation of John Donne's Gunpowder Sermon. Teaching and Visualization Lab, NC State's James B. Hunt Library <http://vpcp.chass.ncsu.edu/> [letzter Zugriff: 14. Oktober 2015].