

Digitale Geisteswissenschaften und Informatik – Modelle der Zusammenarbeit

Der hier eingereichte Abstract versteht sich als ein Beitrag zum Rahmenthema der DHD 2015 *Von Daten zu Erkenntnissen: Digitale Geisteswissenschaften als Mittler zwischen Information und Interpretation* und als ein Beitrag zu einer der drei Leitfragen der Tagung, „... welche disziplinübergreifenden Synergien für die Theoriebildung aus den in den Digitalen Geisteswissenschaften entwickelten Methoden, Techniken und Forschungsinfrastrukturen zu erwarten sind.“ (vgl. <http://www.dig-hum.de/jahrestagung-dhd-2015>). Die Frage nach disziplinübergreifenden Synergien setzt eine disziplinübergreifende Fundierung der Digitalen Geisteswissenschaften implizit voraus bzw. ist durch eine solche Fundierung in jedem Fall leichter zu beantworten.

Eine disziplinübergreifende Fundierung der Digitalen Geisteswissenschaften muss zwei Anforderungen gerecht werden: (i) sie muss von Methoden und Techniken einzelner geisteswissenschaftlichen Disziplinen hinreichend abstrahieren sowie (ii) in Hinblick auf Methoden und Techniken der Informatik hinreichend anschlussfähig sein. Die Frage nach dem Verhältnis der Geisteswissenschaften und speziell der Digitalen Geisteswissenschaften zur Informatik zu stellen, erscheint gerade zum jetzigen Zeitpunkt äußerst sinnvoll. Der Bereich der Digitalen Geisteswissenschaften hat in den letzten Jahren einen großen Aufwuchs erfahren. Dies wird durch den Anstieg der Teilnehmerzahlen bei nationalen und internationalen Konferenzen und der Gründung von Fachverbänden eindrucksvoll dokumentiert. Noch wichtiger sind das wachsende Angebot an fachlich einschlägigen Studiengängen, die steigende Anzahl an Forschungsprojekten und der damit verbundene Einzug von Methoden der Digitalen Geisteswissenschaften in einen immer größeren Kreis geisteswissenschaftlicher Fächer. Beide Entwicklungen haben sicherlich auch damit zu tun, dass mit einer neuen Generation von NachwuchswissenschaftlerInnen als *Digital Natives* das Spektrum an wissenschaftlicher Neugier an und Kompetenz im IT Bereich kontinuierlich zunimmt.

Die Informatik befindet sich ebenfalls in einer Umbruchsphase. Traditionell hat sich die Informatik in ihren Grundlagen auf die Mathematik und Logik berufen, speziell hinsichtlich der Theorie der Berechenbarkeit und hinsichtlich der Komplexität und der Spezifikation von Algorithmen. Durch jüngste und aktuelle Entwicklungen im sozialen und mobilen Computing und die technologischen Möglichkeiten und ethischen Herausforderungen von Big Data stellt sich die Frage nach den theoretischen Grundlagen des Fachs und nach dem Begriff des Computing selbst in einer neuen Weise.

Beide Bereiche befinden sich somit in einer Übergangsphase. Dies sollte beim Nachdenken über Grundsatzfragen nicht außer Acht gelassen werden. In einer Situation des Umbruchs werden mögliche Antworten immer vorläufig und weiter klärungsbedürftig sein. Die Gefahr besteht darin, sich im Grundsätzlichen zu verlieren und dabei vergangene und gegenwärtige Praxis zu ignorieren. Viel folgenreicher ist allerdings das Risiko, die Grundsatzdiskussion allein an vergangener und gegenwärtiger Praxis zu orientieren und dadurch Entwicklungsmöglichkeiten für zukünftige oder zukunftsweisende Kooperationen zwischen den Geisteswissenschaften und der Informatik zu behindern oder gar gänzlich abzuschneiden.

Eine weitere Gefahr, der sich Grundsatzdiskussionen aussetzen, besteht darin, dass sie bisherige Bemühungen in dieser Richtung bewusst oder unbewusst außer Acht lassen. Es erscheint daher sinnvoll, zunächst auf zwei solcher bisheriger Versuche aufmerksam zu machen.

Der eine Versuch einer Grundsatzdiskussion erscheint darin zu bestehen, bestimmte geisteswissenschaftliche Disziplinen aus den Digitalen Geisteswissenschaften sozusagen

auszubürgern und die Digitalen Geisteswissenschaften eher abgrenzend und implizit negativ zu definieren. Solche Versuche sind wenig zielführend, zumal sie zumeist nur von einer Außensicht auf die auszubürgernden Disziplinen gekennzeichnet sind und daher häufig nur mittelbar mit der tatsächlichen Forschungstradition und aktuellen Praxis in diesen Fächern zu tun haben. So wird die Sprachwissenschaft häufig mit dem sicherlich einflussreichen Paradigma der generativen Grammatikforschung gleichgesetzt. Dies wird weder der langen Forschungstradition noch den aktuellen Forschungsentwicklungen des Fachs gerecht, die nicht zuletzt durch die digitale Revolution in den Geisteswissenschaften angestoßen wurden und werden.

Der Versuch, die Digitalen Geisteswissenschaften ausgrenzend zu charakterisieren, erscheint auch insofern wenig hilfreich, als sich die Digitalen Geisteswissenschaften selbst dem untauglichen Versuch ausgesetzt sieht, sie sozusagen aus den Geisteswissenschaften auszubürgern. Bei diesen Versuchen spielen Hinweise auf vermeintliche Dichotomien von quantitativer versus qualitativer Forschung und von der hermeneutischen Tradition des Verstehens und des Erklärens einerseits und der rein beschreibenden, empirischen Wissenschaft andererseits eine wesentliche Rolle.

Ein zweiter und wesentlich aussichtsreicherer Versuch, die Digitalen Geisteswissenschaften inhaltlich zu charakterisieren, ist von Willard McCarty, einem der Pioniere der Digitalen Geisteswissenschaften und Preisträger des 2013 Father Busa Awards, unternommen worden. McCarty und Ko-autoren (vgl. u.a. [1],[2]) charakterisieren die Digitalen Geisteswissenschaften nicht durch den Rekurs auf einen Fächerkanon, sondern inhaltlich über das Konzept der Modellierung. McCarty bezieht sich beim Begriff *Modellierung* auf Marvin Minskys Modellbegriff ([3]). Minsky betrachtet Objekte A^* als Modelle eines Objekts A , wenn sie es einem Betrachter B ermöglichen, Antworten auf interessante Fragen zu einem Objekt A zu geben. Anknüpfend an die Anforderung Minskys, dass Modelle Einsichten vermitteln, versteht McCarty Modellierung nicht als etwas Statisches, sondern als einen heuristischen Erkenntnisprozess. Modellierung in diesem Sinne verstanden hat für ihn zwei Aspekte oder auch Phasen: die Konstruktion und die Verarbeitung. Die Verarbeitungsphase bezeichnet er auch als Manipulation.

McCartys Modellierungsbegriff erscheint außerordentlich hilfreich, um das Verhältnis der Digitalen Geisteswissenschaften zur Informatik zu klären – und in diesem Sinne wird er auch von McCarty verwendet. Er ist disziplinenunabhängig und hat ein sinnvolles disziplinübergreifendes Abstraktionsniveau. Er ist anschlussfähig sowohl hinsichtlich aktueller Entwicklungen in den Digitalen Geisteswissenschaften als auch in der Informatik. Er erlaubt unter anderem auch eine Antwort auf eine Frage, die bereits bei der Kontroversediskussion zur *Beziehung zwischen der Computerlinguistik und den Digital Humanities* (<http://www.dhd2014.uni-passau.de/programm/>) auf der DhD 2014 in Passau gestellt wurde: Ist die Informatik Diener oder Partner der Digitalen Geisteswissenschaften? In der Dienerrolle finden sich all diejenige wieder, denen mit dem Hinweis, sie seien doch Informatiker, einmal die Frage gestellt wurde, ob sie nicht bei der Installation eines Fonts, einer Datenbank und – noch schlimmer – bei der Installation der neuen MS Office Version helfen könnten. Dies ist für die Informatik kaum eine attraktive Perspektive der Zusammenarbeit. Wie eine produktivere Zusammenarbeit zwischen den die Digitalen Geisteswissenschaften und der Informatik aussehen kann, skizzieren wiederum McCarty und Kollegen, indem sie fordern: „*If we wish to have a computing of the humanities, we need to be asking a rather different question: ,how best can we integrate automated processing with human thinking and acting?‘.*“ ([1], S. 142).

Der Begriff der Modellierung und die anzustrebende Synthese aus automatischer Verarbeitung und menschlichem Erkenntnisprozess stellen für McCarty und Kollegen den Versuch dar, die unterschiedlichen Bedeutungsbegriffe zwischen Natur- und Geisteswissenschaften zueinander in Beziehung zu setzen. Beim Minskyschen Modellbegriff ist der Betrachter, der *Observer*, ein integraler Bestandteil der Definition. Modelle erlauben

es **dem Betrachter**, interessante Antworten auf Fragen zu einem Gegenstand zu formulieren. Dadurch fließt der Aspekt der Interpretation und des Verstehens, der für die geisteswissenschaftliche Forschung konstitutiv ist, in den Prozess der Modellierung im McCartyschen Sinne unmittelbar ein. Da empirisch fundierte Modelle immer auf Daten bezogen sind, kommt solchen Modellen genau die Mittlerfunktion zwischen Information und Interpretation zu, von der im Rahmenthema "*Von Daten zu Erkenntnissen: Digitale Geisteswissenschaften als Mittler zwischen Information und Interpretation*" der DHd 2015 die Rede ist.

Wenn man die Integration von automatischer Verarbeitung in den geisteswissenschaftlichen Erkenntnisprozess zum Ziel erhebt, dann rückt dadurch der geisteswissenschaftlicher Forscher in den Blickpunkt der Zusammenarbeit mit der Informatik. Es lassen sich dann verschiedene Optionen für eine solche Zusammenarbeit unterscheiden.

Digitale Geisteswissenschaften	Informatik
naive Anwender	generische Standard-Modelle und Standard-Verfahren
aufgeklärte Anwender	
kompetente Anwender und Modellierer	Domänenadaptierte Standard-Modelle und Standard-Verfahren
kreative und innovative Modellierer	neuartige Modelle und Methoden

Abbildung 1: Optionen der Zusammenarbeit

Naive Anwender haben keinerlei Detailkenntnisse über den technischen Hintergrund der von ihnen eingesetzten Modelle und Verfahren. Sie sind daher darauf angewiesen, Standardmodelle und -verfahren anzuwenden. Bei naiven Benutzern kommt es häufig zu Frustrationen und zur Abkehr von automatischen Verfahren, wenn diese sich als fehleranfällig erweisen und nicht perfekt funktionieren. Aufgeklärte Anwender hingegen wissen um solche Schwächen und können das automatische Verfahren inhärente Rauschen beim Interpretationsprozess der Ergebnisse solcher Verfahren berücksichtigen, indem sie entweder die Analysedaten manuell überprüfen oder nachkorrigieren oder die Aussagekraft solcher Analysen zu bewerten wissen. Naive und aufgeklärte Benutzer haben gemein, dass sie Standardmodelle und -verfahren einsetzen. Um möglichen Missverständnissen vorzubeugen: man sollte den Einsatz von Standardmodellen und -verfahren in den Digitalen Geisteswissenschaften keineswegs kleinreden oder geringschätzen. So können mit Hilfe von Standardverfahren aus der Informatik -- zu nennen sind hier zum Beispiel die Bereiche der Datenvisualisierung, des Data Minings oder Methoden des maschinellen Lernens -- sehr wohl neue Erkenntnisse zu bekannten Fragestellungen gewonnen werden.

Kompetente Anwender und Modellierer kennen nicht nur die technischen Details der von ihnen eingesetzten Verfahren. Sie verfügen auch über die Fähigkeit, Standardmodelle und -verfahren nach ihren Bedürfnissen bzw. auf die Eigenschaften der von ihnen untersuchten Forschungsgegenstände hin zu adaptieren oder zu optimieren. Dadurch entstehen nicht nur aussagekräftigere Modelle, sondern auch eine engere Zusammenarbeit zwischen Informatikern und Forschern in den Digitalen Geisteswissenschaften. Ein wahrhaft partnerschaftliches Verhältnis entsteht dann, wenn beide als kreative und innovative Modellierer neuartige Modelle und Methoden für geisteswissenschaftliche Fragestellungen gemeinsam entwickeln. Ein solches Szenario stellt sicherlich den Idealfall für ein Digital

Computing dar und wird gleichzeitig zu neuen Bindestrich-Informatiken oder -- vielleicht noch wichtiger -- neuen Bindestrich-Geisteswissenschaften führen.

Literaturangaben:

[1] M. Beynon, S. Ross, und W. McCarty (2006). Human Computing – Modelling with Meaning, *Literary and Linguistic Computing* Vol. 21.2, pp. 141 – 157.

[2] W. McCarty (2004). Modelling: a study in words and meanings. In: Schreibman, S., Siemens, S. und Unsworth, J. (Hrsg.) *Companion to Digital Humanities*. Oxford: Blackwell. pp. 254 – 270.

[3] M. Minsky (1968). *Matter, Mind, and Models*. In: Minsky, M. (Hrsg.). *Semantic Information Processing*. Cambridge, Mass: MIT Press.