

Abstract zur Tagung
“Forschungsethik in der qualitativen und quantitativen Sozialforschung”,
München 11./12.09.2014

Jun. Prof. Dr. Michaela Geierhos (Paderborn) & Dr. Jasmin Siri (München)

Was beobachtet die Forschungsethik?

Eine interdisziplinäre Diskussion zwischen Computerlinguistik und qualitativ-konstruktivistischer Sozialforschung

Vor der wissenschaftlichen Betrachtung oder Interpretation durch ForscherInnen steht der Prozess der Datenerhebung bzw. ihrer Konstruktion: die Entscheidung darüber, was die eigentlichen “Daten” für bestimmte Forschungszwecke. Diese oft nicht reflektierte oder gar ausreichend dokumentierte Beobachtung und Erhebungsleistung und das, was anschließend mit dem Produkt der Beobachtung (eben: den “Daten”) geschieht, ist unter anderem Thema forschungsethischer Diskussion. Was wir unter empirischen Daten verstehen, und ob und wie wir den (selbst und fremdgeleisteten) Prozess der Konstruktion dieser Daten bezeichnen, hat Konsequenzen dafür, ob und wie wir forschungsethisch diskutieren. Besonders deutlich wird dies, wenn die eigene Idee von „richtigen“ Daten durch alternative Blickwinkel irritiert wird, bspw. in der interdisziplinären Zusammenarbeit.

Im hier skizzierten Beitrag geht es uns erstens darum, die erkenntnistheoretisch relevante Frage „Was ist ein empirisches Datum?“ zu stellen und offen zu halten. Zweitens wollen wir unterschiedliche Antworten auf diese Frage im Hinblick auf mögliche forschungsethische Implikationen diskutieren.

Anhand von Fallbeispielen aus einer gemeinsamen Antragstellung zum Thema “Politische Kommunikation in Sozialen Medien” wollen wir fragen, was Computerlinguistik und konstruktivistische Sozialforschung jeweils unter “Daten” verstehen, wie sie Daten herstellen und “behandeln”.

Durch das Web 2.0 stehen ComputerlinguistInnen wie SoziologInnen zum Beispiel vor der Herausforderung, wie sie aus all diesen lokalen, verteilten, und unterschiedlichen Social-Media-Daten, die durch menschliche Kommunikation und Interaktion erzeugt werden, interessante Informationen, Muster und letztendlich Wissen extrahieren können. Aus Anwendungssicht spielt die Analyse von *communities* eine große Rolle, um gemeinsame Interessen oder spezielle Kompetenzen zu identifizieren. Diese können dann z. B. zur Personalisierung von Werbung und für Empfehlungen genutzt werden.

Und auch für SoziologInnen sind solche Daten interessant, die auf die Entwicklung von Netzwerken Aufschluss geben. Es entsteht ein forschungsethischer Konflikt zwischen den technischen Möglichkeiten, den Interessen der Nutzer und ihrer Privatsphäre, insbesondere hinsichtlich der Speicherung, Verknüpfung und Auswertung persönlicher Daten zur Ausgestaltung und Bereitstellung personalisierter Dienste. Richtlinien zur rechtlich korrekten Ausgestaltung von Social-Media-Diensten, Anonymisierungsfunktionen sowie datenschutzerhaltende Data-Mining-Methoden stellen hier Ansätze bereit.

Aber auch hier gilt die Frage: Wer versteht was unter (schützenswerten) Daten? Und wer hat die Kompetenz, die Reichweite einer Datenspeicherung zu beurteilen? Beispielsweise machen Computerlinguisten beim Social Media Mining meist eine Datenvollerhebung, was das gezielte Sammeln und Speichern öffentlich zugänglicher nutzergenerierter Textinhalte inklusive der vorhandenen Meta-Information (wie z.B. Name, Standort, Zeit) beinhaltet. Die Anonymisierung von sensiblen personenbezogenen Daten wie findet in der Regel nach der Datenspeicherung statt. Welche Konsequenzen bringen derlei und andere fachspezifisch-historische Praktiken für forschungsethische Überlegungen mit sich? Welche Konsequenzen ergeben sich aus einem variablen Datenbegriff für die forschungsethische Diskussion?

Kontakt: michaela.geierhos@hni.uni-paderborn.de, jasmin.siri@soziologie.uni-muenchen.de