

Text Mining und Computersimulation zur Analyse autobiographischer Texte: Einflüsse auf das literarische Schaffen Klaus Manns

Hess, Jan

hess@uni-trier.de

Trier Center for Digital Humanities (TCDH); Universität
Trier

Lebherz, Daniel

lebherz@uni-trier.de

Center for Informatics Research and Technology (CIRT);
Universität Trier

Zeyen, Christian

zeyen@uni-trier.de

Center for Informatics Research and Technology (CIRT);
Universität Trier

„Ist das alles?“ (Mann 1995, S. 151) – Die vorwurfsvoll anmutende (rhetorische) Frage, die Klaus Mann Ende Juni 1933 unter einer Auflistung seiner Werke des zurückliegenden Halbjahrs in seinem Tagebuch notiert, gibt einen Hinweis darauf, wie selbstkritisch sich der Schriftsteller mit dem eigenen Schaffen auseinandersetzt. Angesichts der persönlichen und politischen Umstände sowie der Vielzahl an Kontakten und Aktivitäten, die er in seinem Journal verzeichnet, erscheint es fast etwas überraschend, dass die gut drei Monate nach seiner Emigration erstellte Werkliste immerhin 21 – wenn auch größtenteils kürzere – Texte umfasst. In Anbetracht seiner ständigen Rastlosigkeit stellt sich nicht nur die Frage, wann Klaus Mann überhaupt seiner eigentlichen schriftstellerischen Arbeit nachgeht, sondern auch, welche Faktoren zum Ge- oder Misslingen seines Schaffens beitragen. Wie wirken sich beispielsweise die beinahe allabendlichen Theater-, Kino- und Barbesuche oder die mehr oder weniger regelmäßig eingenommenen Rauschmittel auf seine literarische Produktivität aus? Welche Rolle spielt das tägliche Lektürepenum? Haben die zahlreichen Treffen, Telefonate und Korrespondenzen Einfluss auf seine schriftstellerische Arbeit? Oder erweisen sich die politischen Umstände, Angst, Verzweiflung und Hass gegenüber dem Nationalsozialismus als entscheidendes Vehikel literarischer Produktivität?

Aufgrund des Umfangs und der Detailgenauigkeit seiner alltäglichen Schilderungen von Gedanken und

Aktivitäten – insbesondere auch der teilweise bis auf die Tageszeit genauen Protokollierung seiner Arbeitsprozesse – erscheinen Klaus Manns Tagebücher als idealer Untersuchungsgegenstand zur Beantwortung der aufgeworfenen Fragen. Gerade die Vielzahl und Komplexität der darin enthaltenen Informationen machen es jedoch schwierig, sich potentiellen Faktoren literarischer Produktivität tiefergehend analytisch zu nähern. Um diese große Anzahl an Informationen und deren Zusammenhänge besser bzw. überhaupt untersuchen zu können, sollen im Rahmen von *eXplore!*¹ am Fallbeispiel der Klaus Mann-Tagebücher verschiedene Methoden des Text Mining angewendet werden. Zur Unterstützung und Strukturierung der Analyseprozesse werden Scientific Workflows erstellt und ausgeführt. Auf Basis der mit Text Mining erzielten Ergebnisse sollen Methoden der Computersimulation eingesetzt werden, um die eingangs aufgeworfenen Fragen beantworten zu können (Abb. 1).

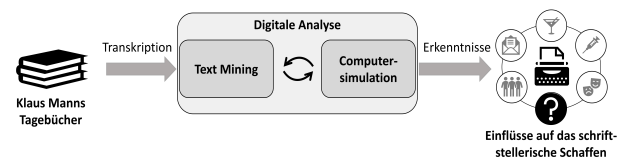


Abbildung 1: Forschungsprozess zur Analyse der
Tagebücher Klaus Manns

Methoden der Computersimulation haben sich bereits seit längerer Zeit in verschiedenen Wissenschaftsdisziplinen etabliert. Seit den 1990er Jahren halten diese auch vermehrt Einzug in sozialwissenschaftliche Forschungsbereiche und bieten dort verschiedene Möglichkeiten zur Darstellung und Analyse von gesellschaftlichen Systemzusammenhängen (Gilbert 2007). Eine ähnliche Entwicklung lässt sich auch in den Digital Humanities nachvollziehen, wenngleich diese Methoden – insbesondere im literaturwissenschaftlichen Bereich – eine eher untergeordnete Rolle spielen (Kohle 2017). Die jedoch insgesamt zunehmende Anwendung von Computersimulation als Forschungsmethode liegt darin begründet, dass sie die Möglichkeit bietet, komplexe, unzugängliche Systeme zu untersuchen und experimentell zu analysieren. In solchen Experimenten lassen sich u.a. durch die Betrachtung verschiedener Szenarien Resultate erzielen, die Rückschlüsse auf Plausibilitäten von Handlungen, Strukturen und Zusammenhängen im zugrundeliegenden System erlauben. Agentenbasierte Computersimulation bietet darüber hinaus insbesondere bei der Analyse von individuellem Entscheidungsverhalten die Möglichkeit, sogenannte emergente Effekte aufzudecken. Diese ergeben sich aus dem Zusammenspiel individueller Handlungen und Interaktionen von einzelnen Akteuren, lassen sich jedoch nicht ausschließlich durch diese erklären (Bonabeau 2002). Ein direkter Zusammenhang zwischen Systeminput und -output muss bei agentenbasierter Simulation also nicht bestehen. Dies steht im Gegensatz

zu anderen statistischen Analyseverfahren, bei denen i.d.R. eine (lineare) Abhängigkeit zwischen verschiedenen Objekten vorausgesetzt wird. Ferner ermöglicht agentenbasierte Computersimulation detaillierte Analysen von Komponenten der Makroebene. Der Weg zu verlässlichen Ergebnissen einer Simulationsstudie führt in einem ersten, wesentlichen Schritt über die Modellbildung und Sammlung dafür relevanter Daten (Law 2015). Zu diesem Zweck soll das zugrundeliegende System, also konkret Klaus Mann, seine sozialen Kontakte und sein Umfeld möglichst realitätsnah mit allen wesentlichen Eigenschaften, Aktivitäten und Zusammenhängen in einer agentenbasierten Sozialsimulation (Davidsson 2002) abgebildet werden (Abb. 2).

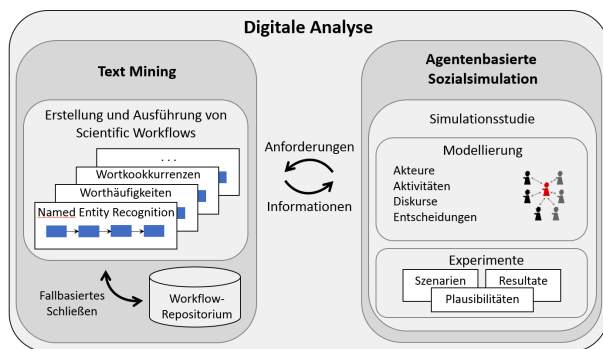


Abbildung 2: Digitale Analyse im Detail

Die dafür notwendige Generierung einer geeigneten Datengrundlage ist ein anspruchsvoller Prozess, in dem anhand von Methoden aus dem Bereich des Text Mining zunächst die wesentlichen, auf den Untersuchungsgegenstand einwirkenden Faktoren identifiziert und analysiert werden müssen. Konkret sollen dabei u.a. Methoden wie Named Entity Recognition, Worthäufigkeits- oder Kookkurrenzanalysen angewendet werden, um vorab Kontakte, Aktivitäten, Aufenthalte, Tätigkeiten, persönliche oder politische Ereignisse zu identifizieren, welche mit Klaus Manns literarischer Produktivität in Zusammenhang stehen könnten. Anders als im Falle der Computersimulation kommen Text Mining-Methoden, -Werkzeuge und Programmbibliotheken im literaturwissenschaftlichen Kontext bereits ungleich häufiger zum Einsatz. Zur Komplexitätsreduktion werden die zugrundeliegenden Text Mining-Prozesse dabei jedoch oft als Black Box betrachtet, was den Nachteil birgt, dass das Zustandekommen der erzielten Ergebnisse nur schwer nachvollzogen werden kann. Um dem entgegenzuwirken sollen die im Rahmen von *eXplore!* benötigten Text Mining-Prozesse explizit in Form von Scientific Workflows (Taylor et al. 2014) definiert und ausgeführt werden. Die in den Digital Humanities vergleichsweise wenig verbreiteten Scientific Workflows dienen der systematischen und transparenten Automatisierung wiederkehrender Datenverarbeitungsprozesse und haben sich in den Naturwissenschaften als geeignetes Mittel

erwiesen, wissenschaftliche Prozesse zu strukturieren, zu dokumentieren und damit die Reproduktion und Validierung von Ergebnissen erheblich zu erleichtern.

Im Kontext von *eXplore!* sollen daher erfolgreich eingesetzte Workflows in einem Repository gespeichert und mithilfe von Fallbasiertem Schließen (Richter / Weber 2016), einer Methode zum erfahrungsbasierten Lösen von Problemen, wiederverwendbar gemacht werden. Ein darauf aufbauendes Tool soll die Erstellung von Workflows in der Praxis vereinfachen. Letztlich sollen somit nicht nur die literaturwissenschaftlichen Forschungsprozesse im Fallbeispiel Klaus Mann, sondern zukünftig auch ähnliche Text Mining-Problemstellungen sowie die Modellierungsarbeit in Simulationsstudien zur Erforschung weiterer autobiographischer Texte sinnvoll unterstützt werden (Abb. 2).

Fußnoten

1. *eXplore!: Computergestützte Modellierung, Analyse und Exploration als Grundlage für eScience in den eHumanities* (Fz: 01UG1606) wird vom Bundesministerium für Bildung und Forschung (BMBF) gefördert.

Bibliographie

- Bonabeau, Eric** (2002): "Agent-based modeling: Methods and techniques for simulating human systems" in: *Proceedings of the National Academy of Sciences*, 99(suppl 3):7280– 7287
- Davidsson, Paul** (2002): "Agent based social simulation" in: *Journal of Artificial Societies and Social Simulation* 5(1): <http://jasss.soc.surrey.ac.uk/5/1/7.html> [letzter Zugriff 11.Januar 2017]
- Gilbert, Nigel** (2007): "*Computational social science: Agent-based social simulation*" in: Amblard, Frédéric / Phan, Dennis (eds.): *Agent-based Modelling and Simulation in the Social and Human Sciences*. Oxford: The Bardwell Press 115-133.
- Kohle, Hubertus** (2017): "*Digitale Rekonstruktion und Simulation*" in: Jannidis, Fotis / Kohle, Hubertus / Rehbein, Malte (eds.): *Digital Humanities: Eine Einführung*. Stuttgart: J. B.Metzler 315-327.
- Law, Averill M.** (2015): *Simulation Modeling and Analysis*. 5.Auflage. Tucson: McGraw Hill Education.
- Mann, Klaus** (1995): *Tagebücher 1931 – 1933*. Reinbek: Rowohlt.
- Richter, Michael M. / Weber, Rosina O.** (2013): *Case-Based Reasoning: A Textbook*. Berlin: Springer.
- Taylor, Ian J. / Deelman, Ewa / Gannon, Dennis B. / Shields, Matthew** (2014): *Workflows for e-Science: Scientific Workflows for Grids*. London: Springer