

Axel Pichler und Nils Reiter

**Zur Operationalisierung literaturwissenschaftlicher Begriffe in
der algorithmischen Textanalyse. Eine Annäherung über Norbert
Altenhofers hermeneutische Modellinterpretation von Kleists *Das
Erdbeben in Chili*
(Abstract)**

- Full-length article in: JLT 15/1-2 (2021), 1–29.

The present article discusses and reflects on possible ways of operationalizing the terminology of traditional literary studies for use in computational literary studies. By »operationalization«, we mean the development of a method for tracing a (theoretical) term back to text-surface phenomena; this is done explicitly and in a rule-based manner, involving a series of substeps. This procedure is presented in detail using as a concrete example Norbert Altenhofer's »model interpretation« (*Modellinterpretation*) of Heinrich von Kleist's *The Earthquake in Chile*. In the process, we develop a multi-stage operation – reflected upon throughout in terms of its epistemological implications – that is based on a rational-hermeneutic reconstruction of Altenhofer's interpretation, which focuses on »mysteriousness« (*Rätselhaftigkeit*), a concept from everyday language. As we go on to demonstrate, when trying to operationalize this term, one encounters numerous difficulties, which is owing to the fact that Altenhofer's use of it is underspecified in a number of ways. Thus, for instance, and contrary to Altenhofer's suggestion, Kleist's sentences containing »relativizing or perspectivizing phrases such as ›it seemed‹ or ›it was as if‹« (Altenhofer 2007, 45) do by no means, when analyzed linguistically, suggest a questioning or challenge of the events narrated, since the unreal quality of those German sentences only relates to the comparison in the subordinate clause, not to the respective main clause. Another indicator central to Altenhofer's ascription of »mysteriousness« is his concept of a »complete facticity« (*lückenlose Faktizität*) which »does not seem to leave anything ›open‹« (Altenhofer 2007, 45). Again, the precise designation of what exactly qualifies facticity as »complete« is left open, since Kleist's novella does indeed select for portrayal certain phenomena and actions within the narrated world (and not others). The degree of factuality in Kleist's text may be higher than it is in other texts, but it is by no means »complete«. In the context of Altenhofer's interpretation, »complete facticity« may be taken to mean a narrative mode in which terrible events are reported using conspicuously sober and at times drastic language. Following the critical reconstruction of Altenhofer's use of terminology, the central terms and their relationship to one another are first explicated (in natural language), which already necessitates intensive conceptual work. We do so implementing a hierarchical understanding of the terms discussed: the definition of one term uses other terms which also need to be defined and operationalized. In accordance with the requirements of computational text analysis, this hierarchy of terms should end in »directly measurable« terms – i. e., in terms that can be clearly identified on the surface of the text. This, however, leads to the question of whether (and, if so, on the basis of which theoretical assumptions) the terminology of literary studies may be traced back in this way to text-surface phenomena. Following the pragmatic as well as the theoretical discussion of this complex of questions, we indicate ways by which such definitions may be converted into manual or automatic recognition. In the case of manual recognition, the paradigm of annotation – as established and methodologically reflected in (computational) linguistics – will be useful, and a well-controlled annotation process will help to further clarify the terms in question. The primary goal, however, is to establish a recognition

rule by which individuals may intersubjectively and reliably identify instances of the term in question in a given text. While it is true that in applying this method to literary studies, new challenges arise – such as the question of the validity and reliability of the annotations –, these challenges are at present being researched intensively in the field of computational literary studies, which has resulted in a large and growing body of research to draw on. In terms of computer-aided recognition, we examine, by way of example, two distinct approaches: 1) The kind of operationalization which is guided by precedent definitions and annotation rules benefits from the fact that each of its steps is transparent, may be validated and interpreted, and that existing tools from computational linguistics can be integrated into the process. In the scenario used here, these would be tools for recognizing and assigning character speech, for the resolution of coreference and the assessment of events; all of these, in turn, may be based on either machine learning, prescribed rules or dictionaries. 2) In recent years, so-called end-to-end systems have become popular which, with the help of neural networks, »infer« target terms directly from a numerical representation of the data. These systems achieve superior results in many areas. However, their lack of transparency also raises new questions, especially with regard to the interpretation of results. Finally, we discuss options for quality assurance and draw a first conclusion. Since numerous decisions have to be made in the course of operationalization, and these, in practice, are often pragmatically justified, the question quickly arises as to how »good« a given operationalization actually is. And since the tools borrowed from computational linguistics (especially the so-called inter-annotator agreement) can only partially be transferred to computational literary studies and, moreover, objective standards for the quality of a given implementation will be difficult to find, it ultimately falls to the community of researchers and scholars to decide, based on their research standards, which operationalizations they accept. At the same time, operationalization is the central link between the computer sciences and literary studies, as well as being a necessary component for a large part of the research done in computational literary studies. The advantage of a conscious, deliberate and reflective operationalization practice lies not only in the fact that it can be used to achieve reliable quantitative results (or that a certain lack of reliability at least is a known factor); it also lies in its facilitation of interdisciplinary cooperation: in the course of operationalization, concrete sets of data are discussed, as are the methods for analysing them, which taken together minimizes the risk of misunderstandings, »false friends« and of an unproductive exchange more generally.

References

- Ajouri, Philip/Katja Mellmann/Christoph Rauen (Hg.), *Empirie in der Literaturwissenschaft*, Münster 2013.
- Altenhofer, Norbert, Der erschütterte Sinn. Hermeneutische Überlegungen zu Kleists *Das Erdbeben in Chili*, in: David E. Wellbery (Hg.), *Positionen der Literaturwissenschaft. Acht Modellanalysen am Beispiel von Kleists »Das Erdbeben in Chili«*, München 2007, 39–53.
- Alvarado, Rafael C., Digital Humanities and the Great Project. Why We Should Operationalize Everything – and Study Those Who Are Doing So Now, *Debates in the Digital Humanities* [2019], <https://dhdebates.gc.cuny.edu/read/untitled-f2acf72c-a469-49d8-be35-67f9ac1e3a60/section/f113a7d9-5860-4bca-8f85-e2ca4ee87a73> (10.09.2021).
- Artstein, Ron/Massimo Poesio, Inter-Coder Agreement for Computational Linguistics, *Computational Linguistics* 34:4 (2008), 555–596.
- Brandom, Robert, *Wiedererinnerter Idealismus*, übers. von Falk Hamann/Aaron Schoichert, Berlin 2015.
- Bridgman, Percy W., *The Logics of Modern Physics*, New York 1927.

- Brunner, Annelen et al., To BERT or not to BERT – Comparing Contextual Embeddings in a Deep Learning Architecture for the Automatic Recognition of four Types of Speech, Thought and Writing Representation, in: Sarah Ebling et al. (Hg.), *Proceedings of the 5th Swiss Text Analytics Conference (SwissText) & 16th Conference on Natural Language Processing (KONVENS)*, Zürich 2020, <http://ceur-ws.org/Vol-2624/paper5.pdf> (10.09.2021).
- Carnap, Rudolf, Theoretische Begriffe der Wissenschaft, *Zeitschrift für philosophische Forschung* 14:2 (1960), 209–233 (Carnap 1960a).
- Carnap, Rudolf, Theoretische Begriffe der Wissenschaft. Eine logische und methodologische Untersuchung (Fortsetzung und Schluß), *Zeitschrift für philosophische Forschung* 14:4 (1960), 571–598 (Carnap 1960b).
- Chang, Hasok, Operationalism, in: *The Stanford Encyclopedia of Philosophy* [Winter 2019], <https://plato.stanford.edu/entries/operationalism/> (10.09.2021).
- Descher, Stefan et al., Probleme der Interpretation von Literatur. Ein Überblick, in: S.D. et. al (Hg.), *Literatur interpretieren. Interdisziplinäre Beiträge zur Theorie und Praxis*, Münster 2015, 11–72.
- Descher, Stefan/Thomas Petraschka, Die Explizierung des Impliziten, *Scientia Poetica* 22 (2018), 180–207.
- Essler, Wilhelm K., *Wissenschaftstheorie I. Definition und Reduktion*, Freiburg/München 1982.
- Frank, Manfred, *Das individuelle Allgemeine. Textstrukturierung und -interpretation nach Schleiermacher*, Frankfurt a.M. 1977.
- Frank, Manfred, Der Text und sein Stil. Schleiermachers Sprachtheorie, in: M.F. (Hg.), *Das Sagbare und das Unsagbare. Studien zur deutsch-französischen Hermeneutik und Texttheorie*, Frankfurt a.M. 1989, 15–37.
- Gius, Evelyn, Pamphlet #8. *Computationelle Textanalysen als fünfdimensionales Problem. Ein Modell zur Beschreibung von Komplexität*, hg. von Thomas Weitin, Digital Humanities Cooperation 2019, <https://www.digitalhumanitiescooperation.de/pamphlet-8-computationelle-textanalysen/> (10.09.2021).
- Gius, Evelyn/Nils Reiter/Marcus Willand, A Shared Task for the Digital Humanities Chapter 2. Evaluating Annotation Guidelines, *Cultural Analytics. A Shared Task for the Digital Humanities* 4:3 (2019), DOI: 10.22148/16.049.
- Groeben, Norbert/A. Florian Pahlke, Was kann und soll Rationale Rekonstruktion hermeneutischer Interpretationsansätze (nicht) leisten?, *Journal of Literary Theory* 10:1 (2016), 83–109.
- Horstmann, Jan/Rabea Kleymann, Alte Fragen, neue Methoden – Philologische und digitale Verfahren im Dialog. Ein Beitrag zum Forschungsdiskurs um Entsaugung und Ironie bei Goethe, *Zeitschrift für digitale Geisteswissenschaften* 4 (2019), DOI: 10.17175/2019_007.
- Jacobs, Abigail Z./Hanna Wallach, Measurement and Fairness, *FACCT '21. Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, New York, NY 2021.
- Jannidis, Fotis, Methoden der computergestützten Textanalyse, in: Ansgar Nünning/Vera Nünning (Hg.): *Methoden der literatur- und kulturwissenschaftlichen Textanalyse. Ansätze – Grundlagen – Modellanalysen*, Stuttgart/Weimar 2010, 109–132.
- Kim, Evgeny/Roman Klinger, A Survey on Sentiment and Emotion Analysis for Computational Literary Studies, *Zeitschrift für digitale Geisteswissenschaften* 4 (2019), DOI: 10.17175/2019_008_v2 .
- Kleist, Heinrich von, Jeronimo und Josephe / Das Erdbeben in Chili, in: Klaus Müller-Salget (Hg.): *Sämtliche Erzählungen, Anekdoten, Gedichte, Schriften*, Frankfurt a.M. 2008, 188–222.

- Krautter Benjamin/Janis Pagel/Nils Reiter/Marcus Willand, »[E]in Vater, dächte ich, ist doch immer ein Vater«. Figurentypen und ihre Operationalisierung, *Zeitschrift für Digitale Geisteswissenschaften* 5 (2020), DOI: 10.17175/2020_007.
- Kuhn, Jonas, Einleitung, in: Nils Reiter/Axel Pichler/Jonas Kuhn (Hg.), *Reflektierte algorithmische Textanalyse. Interdisziplinäre(s) Arbeiten in der CRETA-Werkstatt*, Berlin/Boston 2020, 9–40.
- Lahn, Silke/Jan Christoph Meister, *Einführung in die Erzähltextanalyse*, Stuttgart 2013.
- Moretti, Franco, »Operationalizing«: or, the function of measurement in modern literary theory, *Pamphlets of the Stanford Literary Lab* 6 (2013), <https://litlab.stanford.edu/LiteraryLabPamphlet6.pdf> (10.09.2021).
- Pagel, Janis et al., A Unified Text Annotation Workflow for Diverse Goals, in: Sandra Kübler/Heike Zinsmeister (Hg.), *Proceedings of the Workshop on Annotation in Digital Humanities, co-located with ESSLLI 2018*, Sofia 2018, 31–36.
- Paullada, Amandalynne et al., Data and its (dis)contents. A survey of dataset development and use in machine learning research, *Proceedings of the NeurIPS 2020 Workshop. ML Retrospectives, Surveys & Meta-analyses*, virtual 2020.
- Pichler, Axel/Nils Reiter, Reflektierte Textanalyse, in: Nils Reiter/Axel Pichler/Jonas Kuhn (Hg.), *Reflektierte Algorithmische Textanalyse. Interdisziplinäre(s) Arbeiten in der CRETA-Werkstatt*, Berlin/Boston 2020, 43–60.
- Reiter, Nils, *Discovering Structural Similarities in Narrative Texts using Event Alignment Algorithms* (PhD thesis), Heidelberg 2014.
- Reiter, Nils, Anleitung zur Erstellung von Annotationsrichtlinien, in: Nils Reiter/Axel Pichler/Jonas Kuhn (Hg.), *Reflektierte Algorithmische Textanalyse. Interdisziplinäre(s) Arbeiten in der CRETA-Werkstatt*, Berlin/Boston 2020, 193–201.
- Reiter, Nils/Marcus Willand/Evelyn Gius, A Shared Task for the Digital Humanities Chapter 1. Introduction to Annotation, Narrative Levels and Shared Tasks, *Cultural Analytics. A Shared Task for the Digital Humanities* 2:1 (2019), DOI: 10.22148/16.048.
- Sahle, Patrick/Ulrike Henny, Klios Algorithmen. Automatisierte Auswertung von Wikipedia-Inhalten als Faktenbasis und Diskursraum, in: Thomas Wozniak/Jürgen Nemitz/Uwe Rohwedder (Hg.), *Wikipedia und Geschichtswissenschaft*, Oldenbourg 2015, 113–148.
- Savigny, Eike von, *Grundkurs im wissenschaftlichen Definieren*, München 1980.
- Schleiermacher, Friedrich, *Hermeneutik und Kritik*, hg. von Manfred Frank, Frankfurt a.M. 1977.
- Schnell, Rainer/Paul B. Hiller/Elke Esser (Hg.), *Methoden der empirischen Sozialforschung*, Berlin 2018.
- Sittig, Claudius, Zur Praxis von literaturwissenschaftlichen »Modellinterpretationen«, in: Andrea Albrecht et. Al (Hg.), *Theorien, Methoden und Praktiken des Interpretierens*, Berlin/Boston 2015, 367–383.
- Stegmüller, Wolfgang, *Aufsätze zu Kant und Wittgenstein*, Darmstadt/Stuttgart 1974.
- Ure, Jean, Lexical density and register differentiation, in: G.E. Perren/J.L.M. Trim (Hg.), *Applications of Linguistics*, London 1971, 443–452.
- Wagenknecht, Christian, *Zur Terminologie der Literaturwissenschaft. Akten des IX. Germanistischen Symposiums der DFG Würzburg 1986*, Stuttgart 1988.
- Weitin, Thomas, Digitale Literaturwissenschaft, *Deutsche Vierteljahrsschrift für Literaturwissenschaft und Geistesgeschichte* 89:4 (2015), 651–656.
- Winko, Simone, Standards literaturwissenschaftlichen Argumentierens. Grundlagen und Forschungsfragen, *Germanisch-Romanische Monatsschrift* 65:1 (2015), 14–29.

2022-01-13
JLTONline ISSN 1862-8990

Copyright © by the author. All rights reserved.

This work may be copied for non-profit educational use if proper credit is given to the author and JLTONline.

For other permission, please contact [JLTONline](#).

How to cite this item:

Abstract of: Axel Pichler und Nils Reiter, Zur Operationalisierung literaturwissenschaftlicher Begriffe in der algorithmischen Textanalyse. Eine Annäherung über Norbert Altenhofers hermeneutische Modellinterpretation von Kleists *Das Erdbeben in Chili*.

In: JLTONline (13.01.2021)

Persistent Identifier: urn:nbn:de:0222-004576

Link: <http://nbn-resolving.de/urn:nbn:de:0222-004576>