



Separatum aus:

---

## THEMENHEFT 12

*Elisabeth Lienert / Joachim Hamm  
Albrecht Hausmann / Gabriel Viehhauser (Hrsg.)*

# Digitale Mediävistik

## Perspektiven der Digital Humanities für die Altgermanistik

Publiziert im November 2022.

Die BmE Themenhefte erscheinen online im BIS-Verlag der Carl von Ossietzky Universität Oldenburg unter der Creative Commons Lizenz [CC BY-NC-ND 4.0](https://creativecommons.org/licenses/by-nc-nd/4.0/). Die ›Beiträge zur mediävistischen Erzählforschung‹ (BmE) werden herausgegeben von PD Dr. Anja Becker (München) und Prof. Dr. Albrecht Hausmann (Oldenburg). Die inhaltliche und editorische Verantwortung für das einzelne Themenheft liegt bei den jeweiligen Heftherausgebern.

<http://www.erzaehlforschung.de> – Kontakt: [herausgeber@erzaehlforschung.de](mailto:herausgeber@erzaehlforschung.de)  
ISSN 2568-9967

*Zitiervorschlag für diesen Beitrag:*

Lienert, Elisabeth: Stilometrie und Textanalyse (Diskussionsbericht Sektion 6), in: Lienert, Elisabeth/Hamm, Joachim/Hausmann, Albrecht/Viehhauser, Gabriel (Hrsg.): Digitale Mediävistik. Perspektiven der Digital Humanities für die Altgermanistik, Oldenburg 2022 (BmE Themenheft 12), S. 329–331 (online).

*Elisabeth Lienert*

## Stilometrie und Textanalyse (Diskussionsbericht Sektion 6)

Stilometrie und Textanalyse sind die Bereiche, wo sich am deutlichsten zeigen wird, wie die digitalen Methoden die Mediävistik verändern werden (Manuel Braun). Die Diskussion (Leitung: Manuel Braun) aller drei Vorträge fokussierte vor allem auf die Materialgrundlage in Form von Editionen; Probleme der (quantitativen) Methodik, besonders der Annotation und der Kontextualisierung; das Verhältnis von Input und Output, Vorannahmen/Fragestellungen und Ergebnis, damit verbunden einerseits die Fragen von Objektivität, Scheinobjektivität, Modellbildung und andererseits die Frage nach der literaturwissenschaftlich-interpretatorischen Relevanz; das Problem der Komplexitätsreduktion; weitere Voraussetzungen für Arbeiten in diesem Bereich.

Als Grundlage für quantitative digitale Analysen seien normalisierte und zwar möglichst einheitlich normalisierte Ausgaben nötig. Eine Gefahr bestehe darin, dass unterschiedlich verfahrenende Editionen im Korpus die Ergebnisse beeinflussen (etwa durch Unschärfen bei der Erfassung von Begriffen aufgrund unterschiedlicher Normalisierung: Klaus Kipf; hierzu fragte Sonja Glauch, ob der Einfluss der Editionsqualität ›herausrechenbar‹ sei); eine weitere darin, dass Konzentration auf die ›Höhenkamm-literatur‹ drohe (Freimut Löser). In die Arbeiten sollten möglichst immer die Ergebnisse der MHDDB einbezogen werden, weitere Vernetzung mit (annotierten) digitalen Wörterbüchern sei wünschenswert. Die in den vorgestellten Projekten entwickelten Tools selbst sollten (Stephan Müller) und können (Phillip Brandes) nachnutzbar sein.

Das Methodenarsenal sei begrenzt, rein quantitativ (›Zählen‹ von Wörtern und allenfalls ihren – auf wenige Wörter oder Verse begrenzten – Kontexten) und damit aussagekräftig für große Mengen, nicht für Einzelfälle – schon gar nicht für Interpretationen konzeptionell und ästhetisch außergewöhnlicher Einzelwerke (Gabriel Viehhauser, Müller). Quantitative Ergebnisse wie Wordclouds bedürften der Auswertung (Brigitte Bulitta). Quantitative Analysen und ›qualitative‹ Interpretation, *distant* und *close reading*, Statistik und Interpretation könnten sich jedoch sinnvoll ergänzen (Müller, Simone Schultz-Balluff). Probleme bei der (Nicht-)Berücksichtigung von Negiertem, bei der Identifikation von Ironie, Polysemie, Metapher (Michael Stolz) könnten über Word Embeddings angegangen werden (Viehhauser); insbesondere Disambiguierung bleibe jedoch schwierig (Friedrich Dimpel), manuelle Kontrolle ggf. nötig (Tina Terrahe, Dimpel). Auch Probleme der Kontextualisierung minderten die Präzision (Dimpel); pronominale Referenzen müssten praktisch immer manuell präzisiert werden (Dimpel).

Ausgegangen werde immer von Vorannahmen, häufig literaturgeschichtlich längst Etabliertem (etwa Epochen- und Gattungsbegriffen), womöglich gar Veraltetem (Müller). Es besteht die Gefahr der Zirkularität (Albrecht Hausmann, Löser) und Bestätigung des Altbekanntes (Löser, Müller). Der Vorteil digitaler Verfahren liege jedoch in der Transparenz und Explizitheit hinsichtlich der Vorannahmen (Viehhauser, Dimpel), ohne die auch traditionelle Verfahren nicht auskämen. Die Gefahr der Scheinobjektivität quantitativer Verfahren sei nicht auszuschließen (Hausmann). Visualisierung etwa durch Wortwolken wolle jedoch nur graduell objektivieren (Schultz-Balluff); erkenntnisgeleitete Fragestellungen und Auswertung der Befunde seien immer nötig (Löser).

Erfasst werde bei digitalen Stilanalysen vielfach nur die Wortebene; damit stelle sich die Frage, wie man zu den Texten gelange (Hausmann): über POS (Part of Speech-Tagger) für die Syntax, Word Embeddings für die Semantik (Viehhauser, Brandes, Dimpel). Moniert wurde das Missver-

hältnis von Aufwand und Ergebnis, Komplexität der Verfahren und (bislang) Schlichtheit der Fragestellungen und Ergebnisse (Elisabeth Lienert). Hier sei auf die Weiterentwicklung der Verfahren und die Entwicklung neuer Fragestellungen zu hoffen.

Für die digitale Analyse mittelalterlicher Literatur würden digitale Texte (TEI) benötigt, insbesondere (einheitlich) normalisierte (annotierte) Editionen für quantitative Analysen. Gezielte öffentliche Förderung sei umso notwendiger, als DH-affine, nicht-etablierte Forscher\*innen in besonderem Maß mit Ressourcen-Problemen konfrontiert seien (Dimpel). Kontrovers diskutiert wurden allerdings die Forderung (Dimpel) nach einem großen Korpus mit basalen Entity- sowie narratologischen Annotationen und die Durchführbarkeit dieser Forderung. Besonders wichtig seien auch in diesem Kontext Vernetzung (insbesondere die Anknüpfung an MHD-BDB: Dimpel; disambiguierte Verlinkungen mit den Wörterbüchern: Torsten Schaßan; Kooperation zwischen Projekten: Brandes) und der Aufbau zentraler Ressourcen (Dimpel). Die Praktikabilität eines annotierten Korpus sei jedoch fraglich (Braun), da nicht klar sei, wer entscheide, welche Texte annotiert werden, und kein Konsens bestehe, was aus der Überfülle der Möglichkeiten annotiert werden solle (Löser, Braun), zumal Bedarfslagen interessegeleitet seien (Dimpel). Hier votierte Viehhauser dafür, die Annotationen auf Grundsätzliches zu beschränken (etwa Entitäten, Figuren) und zunächst auf zu Spezielles und zu Komplexes zu verzichten.

### **Anschrift der Berichterstatterin:**

Prof. Dr. Elisabeth Lienert  
Universität Bremen  
Fachbereich 10  
Universitäts-Boulevard 13  
28359 Bremen  
E-Mail: [elienert@uni-bremen.de](mailto:elienert@uni-bremen.de)