



# ePoetics

## Korpuserschließung und Visualisierung deutschsprachiger Poetiken (1770-1960) für den „Algorithmic Criticism“

### Wissenschaftliche Zielsetzungen

„Algorithmic Criticism“ (Stephen Ramsay): Nutzung von Synergieeffekten aus hermeneutischen und algorithmischen Verfahren, sowie solchen der Visualisierung, für methodisch avancierte eHumanities  
Weiterentwicklung durch die Erprobung aktueller informationstechnologischer Methoden:

- sprachtechnologische Korpusanalyse
- statistische Verfahren
- Textmining
- Informationsvisualisierung
- Text- und Corpusviewing
- interaktive Annotationwerkzeuge

### Daten

- Zentrales Textkorpus der Geisteswissenschaften: Poetiken und Ästhetiken von 1770 bis 1960 – enthalten ein systematisches Basiswissen für Literaturwissenschaft und philosophische Ästhetik und damit bereits einen erheblichen Teil der benötigten Metadaten
- Untersuchtes Korpus: Auswahl an 20 repräsentativen Poetiken aus der 1000 Werke umfassenden Bibliographie im Rahmen der hermeneutischen Studie „A History of Poetics“ als Image-Digitalisate und als durch double keying erfasste und korrigierte digitale Volltexte

### Methode

Das Projekt versteht sich als Pilotstudie, die Ergebnisse und Analysetools auch zur Untersuchung ähnlicher Korpora bereitstellt und damit methodisch innovativ auf weite Bereiche der Geisteswissenschaften und der Informatik ausgreift. Das skizzierte Vorhaben kann die Entwicklungen in den eHumanities methodisch vorantreiben, indem es

- das zentrale Textkorpus der Poetiken in durch umfangreiche Metadaten angereicherter und recherchierbarer Form digital aufbereitet,
- Metadaten aus der hermeneutischen Untersuchung des Korpus in Kooperation mit Experten aus der Informatik erarbeitet sowie neue Werkzeuge für die Metadatenerfassung entwickelt,
- qualitative und quantitative Komplementaritäten, historische Zusammenhänge und Dynamiken des Korpus aus der Kombination traditionell hermeneutischer und algorithmischer Methoden erforschbar macht,
- interaktive visuelle Analysewerkzeuge konzipiert und entwickelt, die die Verifikation bzw. Falsifikation von hermeneutischen Hypothesen erlauben,
- über die Visualisierungen schnellen Zugriff auf die Textquellen erlaubt, um Fehler, die durch die algorithmische Aufbereitung induziert werden können, schnell zu erkennen und zu beseitigen und
- mittels interaktiver Feedback- und Trainingsmechanismen eine iterative Konkretisierung der wissenschaftlichen Fragestellung und die Anpassung der (computer-)linguistischen Analysemethoden gewährleistet.

### Digitale Erschließung

- Korpusaufbau
- Metadaten-Konzept
- Einbindung in digitale Infrastrukturen (TextGrid, DARIAH)

### Algorithmic Criticism

- Begriffsexplikation für Metadaten
- Hermeneutische Hypothesen
- Verbindung mit Analysewerkzeugen

### Korpusanalyse

- Entwicklung Analysewerkzeuge
- Entwicklung trainierbare Annotationswerkzeuge
- Werkzeugeinbindung in Clarin

### Visualisierung

- Entwicklung visueller Analysewerkzeuge
- Interaktive Explorationsmethoden
- Visual Analytics Techniken

### Kooperationen

- Institutionalisierte Netzwerkstrukturen der eHumanities: TextGrid, DARIAH, CLARIN
- Computerphilologie: H. Neuroth (Göttingen), F. Jannidis (Würzburg), J. Veit (Paderborn und Detmold)
- Computerlinguistik: U. Heid (Hildesheim), M. Stede (Potsdam)
- Visualisierung / Informatik: DFG Schwerpunktprogramm „Scalable Visual Analytics“, H. Schütze (Stuttgart), G. Heyer, G. Scheuermann (beide Leipzig), D. Keim (Konstanz), NLP-Forscherguppe L. Wanner (Barcelona), „Human Language Technology“ und „Data & Knowledge Management“ der Fondazione Bruno Kessler, BMBF-Projekt VASA Konstanz (Keim) und D. Ebert (Purdue)

### Referenzen und Vorarbeiten

[1] Richter, Sandra: A History of Poetics: German Scholarly Aesthetics and Poetics in International Context, 1770-1960. With Bibliographies by Anja Zenk, Jasmin Azazmah, Eva Jost and Sandra Richter. Berlin, New York: de Gruyter 2010.

[2] Rapp, Andrea: Das Projekt "TextGrid. Modulare Plattform für verteilte und kooperative wissenschaftliche Textdatenverarbeitung – ein Community-Grid für die Geisteswissenschaften". Chancen und Perspektiven für eine neue Wissenschaftskultur in den Geisteswissenschaften. In: Jahrbuch der historischen Forschung in der Bundesrepublik Deutschland: Berichtsjahr 2006, hrsg. von der Arbeitsgemeinschaft historischer Forschungseinrichtungen in der Bundesrepublik Deutschland. München: Oldenbourg 2007, S. 61-68.

[3] www.sprache-und-genome.de, www.metalemmaliste.de

[4] Blessing, André / Schütze, Hinrich: Self-Annotation for fine-grained geospatial relation extraction. COLING 2010, 80-88.

[5] Scheible, Christian / Laws, Florian / Michelbacher, Lukas / Schütze, Hinrich: Sentiment Translation through Multi-Edge Graphs. Proceedings of the 23rd International Conference on Computational Linguistics (Coling 2010, Beijing, China).

[6] Koch, S. / Bosch, H. / Giereth, M. / Ertl, T.: Iterative integration of visual insights during patent search and analysis, Visual Analytics Science and Technology, 2009. VAST 2009. IEEE Symposium on, 2009, 203-210.

[7] Heim, P. / Ertl, T. / Ziegler, J. / Aroyo, L. / Antoniou, G. / Hyvönen, E. / ten Teije, A. / Stuckenschmidt, H. / Cabral, L. / Tudorache, T.(Eds.): Facet Graphs: Complex Semantic Querying Made Easy The Semantic Web: Research and Applications, Springer Berlin / Heidelberg, 2010, 6088, 288-302.

GEFÖRDERT VOM



Bundesministerium für Bildung und Forschung