

Rechnen statt Lesen

Erhard Taverna

Dr. med., Mitglied der Redaktion

Das Feld der unbegrenzten digitalen Anwendungen hat auch längst die Geisteswissenschaften erfasst. Die Computer-Linguistik analysiert und beschreibt digitalisierte Texte mit Hilfe rechnerischer Programme. Das sieht dann etwa so aus, wie in einer grossen Tageszeitung beschrieben, wo unter dem Titel *Entschlüsselung* die Kriminalromane von Friedrich Glauser (1896–1938) und Friedrich Dürrenmatt (1921–1990) numerisch verglichen werden. Die Ergebnisse sind umwerfend, wenn die Software zum Beispiel Wörter auszählt und für den Wortschatz eine Differenz von 11,4% errechnet oder bei Dürrenmatt weder Früchte noch Gemüse findet. Die maschinelle Sprachverarbeitung der Universität Stuttgart zählt die häufigsten Adjektive, Nomen und längsten Sätze, was Rückschlüsse auf den Alkoholkonsum von Wachtmeister Studer und Kommissär Barlach erlaubt. Datenanalysen können gemäss Autoren des Artikels den Stand der Forschung bekräftigen und seien spannender als die herkömmlichen Trampelpfade. Mittels Makroanalysen können Stilähnlichkeiten ganzer Epochen gemessen werden, denn jenseits der Textoberfläche, wie sie der banale Leser entziffert, werden Verteilungsdiagramme sichtbar gemacht, die endlich die Philologie zur exakten Wissenschaft erheben.

Die angewandte Statistik erinnert an die Erzählung *Der Goldkäfer* von Edgar Allan Poe (1809–1849), in der eine Geheimschrift anhand einer Tabelle der im Englischen wichtigsten Buchstaben entschlüsselt wird. Immerhin führte das Ergebnis zu einem Goldschatz. Ein typischer Fall von Kryptologie, aus der sich das elektronische Rechnen rund hundert Jahre später entwickelt hat. Zu den Pionieren gehörte der englische Mathematiker Alain Turing (1912–1954) der den deutschen Geheimcode Enigma knackte und damit den Verlauf des Zweiten Weltkriegs wesentlich beeinflusste. Das geheime Projekt im militärisch abgeschirmten Bletchley Park kann durchaus mit dem Manhattan-Projekt in den USA verglichen werden: den USA die Atombombe, Grossbritannien der Computer, den Linguisten das rechnerische Decodieren von Texten. Denn inzwischen hat Bletchley Park auch die Literaturwissenschaften erreicht, dank Frequenzanalysen und Wordclouds, mit Hilfe zahlreicher Tools, die jeder aus dem Internet herunterladen kann. Die quantitative Analyse stilistischer Merkmale behandelt Bücher ganzer Bibliotheken als Statistik der Quervergleiche. Wer will



da schon mühsam Seite um Seite lesen und interpretieren, wenn sich aus den gespeicherten Datenbanken ein mathematischer Hypertext destillieren lässt, den nur der Rechner sichtbar machen kann. Für den digitalen Mehrwert der sogenannten *mixed methods* in den Geisteswissenschaften steht viel Fördergeld bereit, zum Beispiel für Gedichtrhythmus-Erkennungstools und Bildanalysen. Die computergestützte Wortklaubelei erinnert an die verspielt-absurde Aufräumarbeit von Ursus Wehrli, bekannt als Part des Komikerduos Ursus & Nadeschkin. Eine neue Ästhetik der Ordnung versus Chaos und Beliebigkeit, auch als Memospiel und iPod App für alle zugänglich. Klees Farbtafeln zu Säulen aussortiert, Kandinsky und Malewitsch, Magritte und Miró nach Formen zerlegt, Bruegels Dorfplatz endlich gesäubert. Kommt der Kunst auf die Schliche, wer Kreise rechts und Dreiecke links stapelt oder erstaunt zur Kenntnis nimmt, dass Glauser in seinen Kriminalromanen das Wort Klappmesser überhaupt nie und Dürrenmatt elfmal verwendet hat? Frequenzbasierte Mittel der Suche nach Wortverbindungen und Mustern können abgeschriebene Inhalte in eingereichten Studien und Dissertationen aufdecken. Darin liegt vielleicht der grösste akademische Nutzen digitaler Textanalysen.

Bildnachweis

© MitaStockImages | Dreamstime.com

Literatur

Ursus Wehrli – Kunst aufräumen. Klein & Aber, 2. Auflage 2004.