

BERICHTE UND KRITIK

VOLLTEXTE FÜR DIE FRÜHE NEUZEIT

Der Beitrag des OCR-D-Projekts zur Volltexterkennung
frühneuzeitlicher Drucke

Von Elisabeth Engl, Wolfenbüttel*, Matthias Boenig, Berlin,
Konstantin Baierer, Berlin, Clemens Neudecker, Berlin,
und Volker Hartmann, Karlsruhe

I. Hintergrund des Projekts

Quellen bilden die Basis wissenschaftlicher Forschung. Indem sie analysiert und interpretiert werden, können verschiedenste Forschungsfragen beantwortet werden. In historisch ausgerichteten Forschungsprojekten kommt den vorhandenen Quellen im Vergleich zu Arbeiten an modernen Themen eine noch größere Bedeutung zu, da im Forschungsprozess häufig keine weiteren, fehlenden Quellen bzw. Datengrundlagen ermittelt bzw. erstellt werden können. Stattdessen ist die Wissenschaft auf eine möglichst umfassende Kenntnis der überlieferten Quellen angewiesen, die zu Formalobjekten operationalisiert werden.

Die Möglichkeiten in der Ausgestaltung des Forschungsprozesses haben sich in den letzten Jahren durch zahlreiche technische Entwicklungen stark erweitert, die insbesondere im Bereich der Digital Humanities bereits zunehmend Verwendung finden. Darüber hinaus wird deren Potential bisher jedoch noch kaum ausgenutzt. Dies mag auch darin begründet sein, dass es Geisteswissenschaftler*innen und Informatiker*innen häufig an Verständnis für die jeweiligen Bedürfnisse und Arbeitsweisen der anderen Partei mangelt. Dabei könnte eine Zusammenarbeit in einigen Fällen für beide Seiten sehr fruchtbar sein.

Im Folgenden soll ausgelotet werden, inwieweit im Bereich der Volltextdigitalisierung Synergien zwischen geisteswissenschaftlicher Forschung und technischer Entwicklung erzeugt und umgesetzt werden können. Für die Forschung wären zur Recherche und Auswertung ihrer Quellen Volltexte hilfreich, die zuvor von technisch versierten Partnern in der benötigten hohen

* Dr. Elisabeth Engl, Neuere Medien – Digitale Bibliothek, Herzog August Bibliothek, Lessingstraße 1, 38304 Wolfenbüttel.