

Schriften des Instituts für Dokumentologie und Editorik — Band 16

Digitale Edition in Österreich

Digital Scholarly Edition in Austria

herausgegeben von | edited by
Roman Bleier, Helmut W. Klug

2023

BoD, Norderstedt

Bibliografische Information der Deutschen Nationalbibliothek:

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.d-nb.de/> abrufbar.

Digitale Parallelfassung der gedruckten Publikation zur Archivierung im Kölner Universitäts-Publikations-Server (KUPS). Stand 29. April 2023.

2023

Herstellung und Verlag: Books on Demand GmbH, Norderstedt

ISBN: 978-3-743-102-842

Einbandgestaltung: Stefan Dumont; Coverbild: wurde von Roman Bleier und Helmut Klug für ein KONDE-Poster (DHa 2017) erstellt

Satz: Roman Bleier und Lua \TeX

ABaC:us – Austrian Baroque Corpus

Claudia Resch

Mitarbeiterinnen und Mitarbeiter:	Claudia Resch (Projektleitung), Ulrike Czeitschner, Matej Ďurčo, Barbara Krautgartner, Daniel Schopper, Eva Wohlfarter
Institutionen:	ACDH-CH ÖAW
Fördergeber:	Kultur-, Wissenschafts- und Forschungsförderung der Stadt Wien (Forschungstipendium <i>Barocke literarische Totentänze von und mit Abraham a Sancta</i> und Projekt <i>Totenkult und Jenseitsvorsorge in Wien: Barocke Bruderschaftsdrucke als Forschungsgegenstand der digitalen Geisteswissenschaften</i> , Nr. LWI 0240 I-III), Österreichische Nationalbank (Projekt <i>Texttechnologische Methoden zur Analyse österreichischer Barockliteratur</i> , Nr. 14738)
Website:	https://acdh.oeaw.ac.at/abacus/

Austrian Baroque Corpus (ABaC:us) ist die Bezeichnung für eine digitale Sammlung von literarischen Texten aus der Barockzeit. Die edierten Texte, die digital gelesen, durchsucht und nachgenutzt werden können, sind Teil einer reichhaltigen und weit verbreiteten Literatur, die sich mit der Vorbereitung auf den Tod beschäftigt und heute beinahe in Vergessenheit geraten ist. Den thematischen Kern der Sammlung bilden Erstdrucke theologisch-erbaulichen Inhalts, insbesondere *Ars moriendi*-, Totentanz- und *Memento mori*-Literatur, die dem bekannten Barockprediger Abraham a Sancta Clara (1644–1709) und seinem Umfeld zugeschrieben werden. Das Kernkorpus im Umfang von etwa 180.000 Token ist seit 2015 online zugänglich und über ein benutzerfreundliches Interface abrufbar.

Zur Erstellung originalnaher Transkriptionen wurden die aus verschiedenen Bibliotheken stammenden Bilddigitalisate der raren Erstdrucke mit XML und verwandten Technologien zu maschinenlesbarem Text verarbeitet und gemäß der Richtlinien der Text Encoding Initiative (TEI) in der Version P5 erschlossen. Die digitalen Texte haben mehrfache Kollationierungsdurchgänge durchlaufen, wurden sorgfältig hinsichtlich ihrer Qualität überprüft und geben den historischen Sprachstand quellengetreu wieder. Offensichtliche Textfehler der Originale wurden in der Transkription zudem ausgewiesen und mit editorischen Anmerkungen versehen. In den einzelnen Werken wurden außerdem historische, biblische und mythologische Personennamen sowie Ortsnamen annotiert, wobei in der letztgenannten Kategorie die Namen von Bergen, Gewässern, Regionen, Ländern und Kontinenten, Städten und Dörfern sowie Straßen und Plätzen in Wien ausgezeichnet wurden.

Im Hinblick auf die Durchsuchbarkeit und Nachnutzung wurden die historischen Texte mit einer manuell korrigierten linguistischen Basisannotation versehen, die folgende Verarbeitungsschritte vorsah: 1. Tokenisierung, 2. Wortartklassifizierung (*Part-of-Speech Tagging*) und 3. Lemmatisierung. Als Klassifikationssystem wurde hierfür das 54-teilige *Stuttgart-Tübingen-TagSet* (STTS) herangezogen und – der Sprache des älteren Neuhochdeutsch entsprechend – geringfügig erweitert. Bei der Lemmatisierung dienten das *Deutsche Wörterbuch* von Jacob und Wilhelm Grimm sowie der *Duden* als Referenzwerke.

Das Ergebnis dieses aufwändigen Annotationsverfahrens sind handverlesende Daten, deren Qualität *ABaC:us* zu einer gefragten Ausgangsbasis für weiterführende Forschungsfragen macht. Sein Wert liegt in der Wiederverwendbarkeit von bereits erarbeitetem Wissen über die Texte, was bedeutet, dass die Projektgruppe selbst, aber auch Nutzerinnen und Nutzer die Annotationen zeitsparend und gewinnbringend für ihre Erkenntnisinteressen einsetzen können.

Die beispielgebende *ABaC:us*-Webapplikation basiert auf dem modularen Publikations-Framework *corpus_shell* und wurde mit der Intention entwickelt, zeitentfernte, in diversen Bibliotheken beherbergte Drucke aus der Barockzeit für wissenschaftliche Fragestellungen zugänglich zu machen. Das Projektteam hat sich für dieses Interface entschieden, weil es – einem abrahamischen Buchtitel folgend – dazu geeignet ist, *Etwas für alle* (1699) zu bieten, indem es an unterschiedlichen Fragestellungen orientiert ist und mehrere Nutzungsszenarien ermöglicht und unterstützt. So bietet das Interface Lesenden einerseits eine Lektüreamsicht mit Faksimiles und stellt Suchenden andererseits Navigationsinstrumente und eine Suchmaschine zur Verfügung, die den direkten Zugang in den Text erlauben – etwa über Inhaltsverzeichnisse, Register und die freie Volltextsuche nach Wort, Lemma oder Wortart. Für weitere wissenschaftliche Nachnutzungsszenarien, die über diese antizipierten Fragestellungen hinausgehen, sind außerdem die Daten im XML/TEI-Format zum Download abrufbar.

Es ist zu wünschen, dass *ABaC:us* weiterhin in Forschung, Lehre und Unterricht rezipiert wird und darüber hinaus auf die Neugier einer interessierten, web-affinen (Fach-)Öffentlichkeit trifft, die Zitate von Abraham a Sancta Clara nicht nur in Anekdoten oder Blütenlesen, sondern auch im unveränderten Wortlaut nachlesen und suchen möchte.

Literatur

- Czeitschner, Ulrike und Claudia Resch. 2016. „Repräsentation von | in barocken Buch-Totentänzen im digitalen Medium.“ In *Repräsentation(en). Interdisziplinäre Annäherungen an einen umstrittenen Begriff*, hg. v. Gernot Gruber und Monika Mokre, 35–49. Wien: Verlag der Österreichischen Akademie der Wissenschaften.
- Resch, Claudia, Ulrike Czeitschner, Eva Wohlfarter und Barbara Krautgartner. 2016. „Introducing the Austrian Baroque Corpus: Annotation and Application of a Thematic Research Collection.“ In *Proceedings of the Third Conference on Digital Humanities in Luxembourg with a Special Focus on Reading Historical Sources in the Digital Age*. Zugriff: 20. September 2019. http://ceur-ws.org/Vol-1681/Resch_et_al_austrian_baroque_corpus.pdf.
- und Wolfgang U. Dressler. 2016. „Zur Pragmatik der Diminutive in frühen Erbauungstexten Abraham a Sancta Claras. Eine korpusbasierte Studie.“ In *Linguistische Pragmatik in historischen Bezügen*, hg. v. Peter Ernst und Martina Werner, 235–50. Berlin und Boston: de Gruyter.
- . 2017. „»Etwas für alle« – Ausgewählte Texte von und mit Abraham a Sancta Clara digital.“ *Zeitschrift für digitale Geisteswissenschaften*. doi:10.17175/2016_005.
- und Ulrike Czeitschner. 2017. „Morphosyntaktische Annotation historischer deutscher Texte: Das Austrian Baroque Corpus.“ In *Digitale Methoden der Korpusforschung in Österreich*, hg. v. Claudia Resch und Wolfgang U. Dressler, 39–62. Wien: Verlag der Österreichischen Akademie der Wissenschaften.
- . 2019. „Linguistisch annotierte historische Texte stilistisch auswerten. Musterhaft vorkommende Wortverbindungen im Austrian Baroque Corpus.“ In *Historische Korpuslinguistik*, hg. v. Renata Szczepaniak, Stefan Hartmann und Lisa Dücker, 368–85. Berlin und Boston: de Gruyter.
- . 2022. „Digital Approaches to Analyzing and Understanding Baroque Literature.“ In *New Technologies and Renaissance Studies III*, hg. v. Colin Wilder und Matt Davis, 63–86. New York und Toronto: Iter Press.