

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

Mittelhochdeutsche Begriffsdatenbank (MHDBDB)

Peter Hinkelmanns, Katharina Zeppezauer-Wachauer

Mitarbeiterinnen und Mitarbeiter:	Katharina Zeppezauer-Wachauer, Peter Hinkelmanns, Daniel Schlagger, Klaus M. Schmidt
Institutionen:	Universität Salzburg – Interdisziplinäres Zentrum für Mittelalter und Frühneuzeit (IZMF)
Fördergeber:	Universität Salzburg
Website:	http://mhdbdb.sbg.ac.at

Auf der Grundlage digitaler Editionen wird der mhd. und fnhd. Wortschatz seit den 1970er-Jahren in der *Mittelhochdeutschen Begriffsdatenbank* (MHDBDB) onomasiologisch, also über die Wortbedeutung, erschlossen (begründet durch Klaus M. Schmidt, Bowling Green State University). In mehreren technischen Stufen und einer 1992 durchgeführten Zusammenschließung mit dem Projekt *Namen in deutschen literarischen Texten des Mittelalters* von Horst Pütz (Christian-Albrechts-Universität zu Kiel) ist die MHDBDB, die seit 2002 an der Universität Salzburg installiert ist, mittlerweile auf über 10 Millionen Tokens, verteilt auf über 650 Texteditionen unterschiedlichster Textsorten und -gattungen, angewachsen. Die Editionen sind zum Großteil vom Projekt nachdigitalisierte Printeditionen, mittlerweile ergänzen aber auch erste *Born-Digital*-Editionen das Korpus.

Seit 2016 wird der Relaunch der MHDBDB vorbereitet, dessen Ziel neben einer gestalterischen ‚Rundumerneuerung‘ in puncto Bedienung, Abfragemöglichkeiten und Webdesign vor allem die Standardisierung der Daten ist. Die bestehenden Texte der Datenbank werden nach TEI-XML kodiert. Zukünftig sollen überwiegend genuin digitale TEI-Editionen aufgenommen und auf eine Retrodigitalisierung weitestgehend verzichtet werden.

Bereits heute sind komplexe Suchanfragen zur Literatur und zum Wortschatz des Mittelhochdeutschen möglich. Es können Textstellen etwa über die Suche nach dem Vorkommen eines oder mehrerer zueinander in Beziehung stehender lemmatisierter Tokens, aber auch über die onomasiologischen Begriffe des Begriffssystems gesucht werden. Auch kombinierte Suchen können durchgeführt werden, um etwa zwischen Wortbedeutungen zu unterscheiden. Beispielsweise führt der Wortartikel *druc* mehrere Bedeutungen an: 1) *Schlacht*, 2) *Willensausübung auf andere*, 3) *Zweikampf*, 4) *Physikal. chem. Eigenschaften*, 5) *Buchherstellung/Buchdruck*. Auch Informationen zur

Wortbildung und zu Wortfamilien listet der Wortartikel – zu *druc* etwa die Nomina *druckerie* (Druckerei) und *drucker* (Buchdrucker).

Suchanfragen können in beliebiger Kombination durchgeführt werden, beispielsweise in welchen Texten *druc* in der Bedeutung 5) *Buchherstellung/Buchdruck* gebraucht wird oder ob und wie die lateinische Wortvariante *editum* in der Bedeutung *edieren*, *herausgeben* verwendet wird (einmal übrigens in Heinrich Münsingers *Regimen sanitatis*, einmal bei Priester Konrad).

Um beliebig viele linguistische Annotationsebenen wie etwa *Part-Of-Speech* (POS) oder Phrasen- und Satzstrukturen annotieren zu können, werden zukünftig die linguistischen Annotationen im *Stand-Off*-Verfahren mit *Linked Open Data* (LOD) auf die Tokens der Texte bezogen und nicht wie bisher auf die graphematischen *Types* im Wortartikel. Dieses Verfahren ermöglicht weiters das Hinzufügen von Metadaten zu jeder Annotation, etwa um deren Autorschaft zu kennzeichnen.

Sämtliche Daten werden im Zuge des Relaunches als LOD in das *Semantic Web* eingebunden und (abseits der TEI-Editionen) nach LOD-Standards codiert. Die Vernetzung mit Metadatenrepositorien wie *Wikidata* oder der *Gemeinsamen Normdatei* (GND) ermöglicht das wechselseitige Anreichern der Daten und wird wesentlich zu einer besseren Abfragbarkeit des MHDBDB-Korpus beitragen. Normdaten zu Personen, Orten, Ereignissen und Werken werden des Weiteren auch die Auffindbarkeit und Filterung der Texte des Korpus vereinfachen.

Der Relaunch der MHDBDB wird nicht nur zu erheblichen Vorteilen für diejenigen führen, die selbst nach Wörtern, Zeichenketten, Wortarten und Begriffen suchen möchten. Zusätzlich können auch Editorinnen und Editoren ihre digitalen Editionen als externe Korpora in die Begriffsdatenbank einbinden und so automatisch POS annotieren und lemmatisieren, mit einer onomasiologischen Systematik (dem Begriffssystem) erschließen und über Normdaten anreichern. Die Benützbarkeit und Sichtbarkeit der Editionen wird damit maßgeblich – ohne besonderen Aufwand für die Herausgeberinnen und Herausgeber – verbessert und ausgeweitet werden.

Literatur

- MHDBDB 3.0: Hinkelmanns, Peter und Katharina Zeppezauer-Wachauer. 2020. „ez ist ein wärheit, niht ein spel, daz netze was sinewel. Die MHDBDB im Semantic Web.“ In *Bamberger interdisziplinäre Mittelalterstudien*, hg. v. Ingrid Bennewitz und Martin Fischer. doi:10.20378/irb-48993.
- MHDBDB 3.0: Hinkelmanns, Peter. 2019. „Mittelhochdeutsche Lexikographie und Semantic Web: Die Anbindung der ‚Mittelhochdeutschen Begriffsdatenbank‘ an Linked Open Data.“ In *Digitale Mediävistik*, hg. v. Roman Bleier, Franz Fischer, Torsten Hiltmann, Gabriel Viehhauser und Georg Vogeler. *Das Mittelalter. Perspektiven mediävistischer Forschung* 24, no. 1: 129–41. doi:10.1515/mial-2019-0009.

- MHDBDB 2.0: Schmidt, Klaus M. und Horst P. Pütz. 2001. „Die mittelhochdeutsche Begriffsdatenbank.“ *Zeitschrift für deutsches Altertum und deutsche Literatur* 130: 493–5.
- MHDBDB 1.0: Schmidt, Klaus M. 1978. „Errungenschaften, Zukunftsmusik und Holzwege auf dem Gebiet der inhaltlichen Textanalyse mit Hilfe des Elektronenrechners.“ In *Maschinelle Verarbeitung altdeutscher Texte III. Beiträge zum Symposium Tübingen 17. - 19. Februar 1977*, hg. v. Paul Sappeler und Erich Straßner, 101–11. Tübingen: Niemeyer.