



Separatum aus:

THEMENHEFT 12

*Elisabeth Lienert / Joachim Hamm
Albrecht Hausmann / Gabriel Viehhauser (Hrsg.)*

Digitale Mediävistik

Perspektiven der Digital Humanities für die Altgermanistik

Publiziert im November 2022.

Die BmE Themenhefte erscheinen online im BIS-Verlag der Carl von Ossietzky Universität Oldenburg unter der Creative Commons Lizenz [CC BY-NC-ND 4.0](https://creativecommons.org/licenses/by-nc-nd/4.0/). Die ›Beiträge zur mediävistischen Erzählforschung‹ (BmE) werden herausgegeben von PD Dr. Anja Becker (München) und Prof. Dr. Albrecht Hausmann (Oldenburg). Die inhaltliche und editorische Verantwortung für das einzelne Themenheft liegt bei den jeweiligen Heftherausgebern.

<http://www.erzaehlforschung.de> – Kontakt: herausgeber@erzaehlforschung.de
ISSN 2568-9967

Zitiervorschlag für diesen Beitrag:

Burch, Thomas: Infrastrukturprojekte zur digitalen Lexikographie. Vorgestellt am Beispiel des Zentrums für Historische Lexikographie, in: Lienert, Elisabeth/Hamm, Joachim/Hausmann, Albrecht/Viehhauser, Gabriel (Hrsg.): Digitale Mediävistik. Perspektiven der Digital Humanities für die Altgermanistik, Oldenburg 2022 (BmE Themenheft 12), S. 97–108 (online).

Thomas Burch

Infrastrukturprojekte zur digitalen Lexikographie

Vorgestellt am Beispiel des Zentrums für
Historische Lexikographie

Abstract. Forschungsinfrastrukturen stellen Instrumente, Ressourcen und Dienstleistungen zur Verfügung, deren Ziele darin bestehen, Forschung, Lehre und Nachwuchsförderung zu unterstützen oder erst zu ermöglichen. Dieser aus den technischen Disziplinen stammende Ansatz hat sich innerhalb der letzten zehn bis fünfzehn Jahre auch im Bereich der Geistes- und Sozialwissenschaften etabliert. Hier bieten sich gerade geisteswissenschaftliche Grundlagenwerke wie Editionen und Wörterbücher, die einerseits Forschungsgegenstand, andererseits aber auch Forschungsinstrument sind, als Basis entsprechender Infrastrukturen an. Im Beitrag wird mit dem vom BMBF geförderten Zentrum für Historische Lexikographie eine dieser Initiativen und die dort erzielten Ergebnisse vorgestellt.

Forschungsinfrastrukturen stellen Instrumente, Ressourcen und Dienstleistungen zur Verfügung, »die speziell für wissenschaftliche Zwecke errichtet, mittelfristig bis tendenziell permanent bereitgestellt werden und für deren sachgerechte Errichtung, Betrieb und Nutzung in der Regel spezifische fachwissenschaftliche oder interdisziplinäre (Methoden-)Kompetenzen erforderlich sind. Ihre Funktion ist es, Forschung, Lehre und Nachwuchsförderung zu ermöglichen oder zu erleichtern. Sie sind örtlich fixiert, auf mehrere Standorte verteilt oder werden ohne definierte physische

Anlaufstelle ausschließlich virtuell bereitgestellt. Sie werden nicht ausschließlich von einzelnen Personen oder Gruppen genutzt, sondern stehen prinzipiell einer internationalen Fachgemeinschaft oder mehreren Fachgemeinschaften offen.« (WR 2017) Diese Charakterisierung aus dem ›Bericht zur wissenschaftsgeleiteten Bewertung umfangreicher Forschungsinfrastrukturvorhaben für die Nationale Roadmap‹ stellt eine davor eher auf die technischen Disziplinen gewendete Definition auf eine breitere Basis und bezieht mit Hinblick auf nachfolgende Förderprogramme insbesondere die geistes- und sozialwissenschaftlichen Fächer ein.

Gerade geisteswissenschaftliche Grundlagenwerke wie Editionen und Wörterbücher, die einerseits Forschungsgegenstand, andererseits aber auch Forschungsinstrument sind, bieten sich als Basis zur Konzeption und Entwicklung entsprechender Infrastrukturen an. Wörterbücher und lexikologische Darstellungen zur Entwicklung des Wortschatzes und des Wortgebrauchs sind zentrale wissenschaftliche Dokumentationsformen im Schnittbereich zwischen Philologie, Sprach-, Literatur- und Kulturwissenschaft sowie den Forschungsdisziplinen, die auf die sprachliche Beschreibung einzelner Gegenstandsbereiche, ihrer geschichtlichen Entwicklung und ihrer Erforschung bezogen sind. Beim Übergang ins digitale Zeitalter waren vor allem im Bereich der historischen Lexikographie und Lexikologie zwei zeitlich versetzte Strategien zu beobachten. Zunächst wurden die Informationsbestände der vorhandenen Wörterbücher retrodigitalisiert (entweder als Volltext oder in Form reiner Imagedigitalisate), die digitalen Versionen waren zum Teil direkte Gegenstücke zu den gedruckten Fassungen. Eine zweite Strategie bestand und besteht darin, vorhandene und hinzukommende Daten mit neuen Darstellungs-, Interaktions- und Vernetzungsmöglichkeiten anzureichern, die erst durch digitale Techniken und Infrastrukturen ermöglicht werden. Diese Initiativen stießen jedoch zunehmend an die Grenzen einer disparaten Verteilung der Arbeiten zur historischen Wortforschung, die mit jeweils eigenem rigidem Zeitregime und fixen Arbeitsplänen durchgeführt werden, dabei aber

informationstechnisch kaum in nennenswerter Weise aufeinander abgestimmt sind. Besonders bemerkbar machte sich dieses Problem im Fehlen integrativer und interoperabler eHumanities-Ansätze. Eine Konsequenz bestand daher in der Entwicklung von Initiativen, die nationalen und internationalen Bestrebungen im Bereich der Lexikographie zu integrieren, zu erweitern und zu harmonisieren, mit dem Ziel, eine nachhaltige Infrastruktur zu schaffen, die einerseits einen effizienten Zugang zu hochwertigen lexikalischen Daten im digitalen Zeitalter ermöglicht und andererseits die Kluft zwischen fortgeschritteneren und weniger gut ausgestatteten wissenschaftlichen Gemeinschaften, die an sprachwissenschaftlichen Ressourcen arbeiten, überbrücken hilft (vgl. Burch [u. a.] 2020).

Der Bedarf an lexikographisch ausgerichteten Infrastrukturen hatte sich bereits als ein Ergebnis aus der Gemeinschaft im Rahmen der [COST-Action European Network of e-Lexicography ENeL](#) ergeben, die 2017 beendet wurde (vgl. Declerck [u. a.] 2015). Aus dieser eher institutionell vernetzenden Initiative sind Nachfolgevorhaben auf nationaler und europäischer Ebene wie das vom Bundesministerium für Bildung und Forschung (BMBF) 2017 eingerichtete eHumanities-Zentrum für Historische Lexikographie [ZHistLex](#) und die im Rahmen des [Horizon2020](#)-Programms von der EU geförderte European Lexicographic Infrastructure [ELEXIS](#) entstanden. Beide Initiativen verfolgen im Wesentlichen gleiche Ziele, die sich auf die Definition und Bereitstellung gemeinsamer Interoperabilitätsstandards, Arbeitsabläufe, konzeptioneller Modelle und Datendienste sowie Schulungs- und Ausbildungsaktivitäten mit einem Schwerpunkt im Bereich lexikographischer Nutzungsszenarien sowie der disziplinübergreifenden Anwendung beziehen.

Beide Initiativen werden von Konsortien an verteilten Standorten getragen, die sich aus Institutionen und Forschern mit komplementärem Hintergrund – Lexikographie, digitale Geisteswissenschaften, Informatik – zusammensetzen, die neben der jeweiligen Expertise auch lexikographische Ressourcen einbringen. Sowohl für die nationale wie die interna-

tionale Landschaft der digitalen Lexikographie gilt, dass sie sich recht heterogen darstellt. Sie ist gekennzeichnet durch eigenständige Datenbestände, die in der Regel in inkompatiblen Strukturen kodiert sind, da die Arbeiten oft isoliert durchgeführt werden. Dies verhindert die Wiederverwendung dieser wertvollen Daten in anderen Bereichen, wie beispielsweise der Verarbeitung natürlicher Sprache, der Verknüpfung offener Daten, die Einbindung ins Semantic Web sowie ihre breite Nutzung im Kontext der digitalen Geisteswissenschaften.

Vor diesem Hintergrund werden Strategien, Werkzeuge und Standards für die Extraktion, Strukturierung und Verknüpfung lexikographischer Ressourcen entwickelt, um deren volles Potenzial hinsichtlich Linked Open Data (LOD) und Semantic Web sowie im allgemeinen Kontext der digitalen Geisteswissenschaften zu erschließen. Gleichzeitig sollen Wissenschaftler dabei unterstützt werden, homogene Datenformate über nationale Grenzen hinweg zu erstellen, zu teilen, zu verknüpfen, zu analysieren und zu interpretieren, was den Weg für transnationale, datengesteuerte Fortschritte in diesem Bereich ebnet und gleichzeitig Doppelarbeit über disziplinäre Grenzen hinweg deutlich reduzieren hilft. Im Folgenden werden die Arbeiten und Ergebnisse des Zentrums für Historische Lexikographie vorgestellt, die sich in ähnlicher Form mit dem erweiterten Blick auf internationale lexikographische Strukturen auch in ELEXIS zeigen (Krek [u. a.] 2018).

Gerade für die historische Lexikographie des Deutschen gilt, dass sie sich in besonderer Weise als Ausgangspunkt für ein virtuelles Zentrum eignet, das auf breiter Grundlage neue Wege digitaler Wörterbuchproduktion und -präsentation ebnet sowie ihre Nutzung in unterschiedlichen Feldern wortgeschichtsbezogener Forschung ermöglichen kann. Die historische Lexikographie des Deutschen ist einerseits gekennzeichnet durch eine reich entfaltete und differenzierte Wörterbuchlandschaft, andererseits durch avancierte Initiativen im Sinne der eHumanities, der Computer- und der Internetlexikographie. ZHistLex stellte sich die Aufgabe, solche Ansätze zu entwickeln bzw. auszubauen, auch über ihre An-

wendung für das Deutsche hinaus. Die Ergebnisse dieser Arbeiten sind eine wesentliche Voraussetzung und ein Bezugspunkt für die zukünftige Forschung und die zentrale Ergebnisdokumentation in vielen Bereichen wortgeschichtlicher Forschung. Sie sind damit auch ein Stimulans für die Erarbeitung neuer lexikographischer Materialien und für den Aufbau und die Erschließung einschlägiger digitaler Quellentexte. Das Zentrum verstand sich als Kristallisationspunkt für abgeschlossene, laufende und zukünftige Vorhaben in der historischen Lexikographie sowie für empirische Forschungen in der Sprachgeschichte, insbesondere der historischen Lexikologie und anderer wortgeschichtlicher Arbeitsfelder.

Die Ziele von ZHistLex lassen sich in drei Dimensionen charakterisieren, indem Beiträge zu einer integrierten Dateninfrastruktur im Bereich der digitalen historischen Lexikographie geleistet werden, eine konzeptionelle Systematisierung und Erweiterung von digital unterstützten Untersuchungs-, Erschließungs- und Präsentationsverfahren erstellt sowie die Einrichtung einer Kollaborationsstruktur angestrebt wird. Die Umsetzung der Ziele spiegelt sich in den auf drei Ebenen angesiedelten Arbeitsgebieten wider, die getrieben von den zugrundeliegenden lexikographischen Daten deren Codierung, den Zugriff darauf sowie geeignete Nutzungsszenarien betrachten und damit eine zunehmende Abstraktion von den jeweiligen digitalen Repräsentationen der Wörterbücher erlauben und zu einer interoperablen, wörterbuch-übergreifenden Verwendung führen (vgl. Burch [u. a.] 2020).

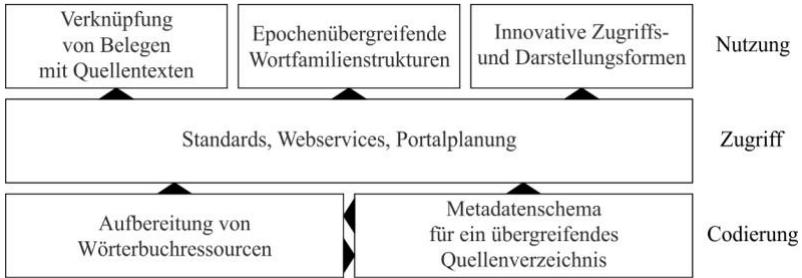


Abb. 1: Ebenen der Datenbearbeitung in ZHistLex

So werden auf unterster Ebene Basistechnologien für die Harmonisierung von Ressourcen hinsichtlich der gemeinsamen Anwendung von Standards für die Zusammenführung von Wörterbüchern durch interne und externe Vernetzung, für die Zusammenführung bislang unabhängig geführter Belegquellenverzeichnisse sowie für die Verknüpfung von Belegen mit verfügbaren Volltexten geschaffen. Diese Techniken bilden das Rückgrat des Zentrums und sind Leitlinie für laufende lexikographische Vorhaben sowie für kommende Vorhaben, seien sie mit den aktuellen Vorhaben verbunden oder nicht. Der Ausgangspunkt der technisch-informatischen Entwicklungen war hier gegeben durch eine beträchtliche Heterogenität der Ansätze und der technischen Lösungen innerhalb der beteiligten Einrichtungen. Während sich eine XML-Kodierung im Rahmen des TEI-P5-Standards mittlerweile durchgesetzt hat, so erlaubt dieser Standard doch eine Fülle von Anwendungsszenarien, die eine interoperable Adressierung von Wörterbuchdaten erschwert. Die Anwendung der TEI-Guidelines zur Codierung der in ZHistLex betrachteten Wörterbücher zeichnete sich durch eine hohe Variabilität hinsichtlich Textperspektiven, Annotations-ebenen und Annotationstiefen aus. Eine grundsätzliche Vereinheitlichung der Codierung folgte hier den Vorschlägen der internationalen Initiative TEI Lex-o (Romary/Tasovac 2019), die ein sogenanntes Baseline-Encoding für Wörterbuchdaten erarbeitet hat. Darüber lassen sich die Einheiten in der Makro- und Mesostrukturebene von Wörterbüchern einheitlich codie-

ren. Die teilweise sehr spezifischen Elemente der Mikrostrukturebene werden nach den allgemeinen TEI-Guidelines abgebildet. Eine entsprechende Vereinheitlichung erfolgt hier über die Spezifikation einer gemeinsamen offenen Schnittstelle.

Die zweite strukturelle Ebene wird durch die Entwicklung von standardisierten Kommunikationsmechanismen abgedeckt, die die Interoperabilität von digitalen Wörterbuchangeboten überhaupt erst ermöglichen bzw. vorhandene Zugriffsformen verbessern. Geleitet wurde die Implementierung dieser offenen Schnittstellen von den Zielen, über lexikographische Webservices eine übergreifende Recherche in und den Zugriff auf die Ressourcen (Wörterbücher, Quellenverzeichnisse, Textcorpora) zu unterstützen. Die Implementierung selbst setzte auf verbreiteten Entwurfsmustern für derartige APIs (Application Programmable Interfaces) auf, indem als Vorlage und Best-Practice das Konzept einer REST (Representational State Transfer)/SOAP (Simple Object Access Protocol)-Schnittstelle ([W3C 2000](#); [Open API 2021](#)) zugrunde gelegt und eine gemeinsame Schnittmenge an Funktionen und Parametern definiert wurde. Über die Syntax der Schnittstellen können beispielsweise Abfrageszenarien wie ›Gib alle Wörterbuchartikel zurück, die im lexikographischen Kommentar die beschreibungssprachliche Wortform Haus enthalten‹, ›Gib alle Wörterbuchartikel zurück, deren Stichwörter länger als 200 Jahre belegt sind‹ oder ›Welche Artikel des althochdeutschen Wörterbuchs führen etymologische Belege zu mittelhochdeutschen Lemmata‹ formuliert werden können.

Entscheidend für eine generische Umsetzung und damit die Übertrag- und Erweiterbarkeit der Schnittstelle ist eine formale Spezifikation der Abfrageszenarien und deren Überführung in eine einheitliche Syntax, die hier standardisiert über eine YAML-Beschreibung ([YAML 2021](#)) erfolgte und damit einerseits direkt überprüfbar und andererseits auch entsprechend modular anpassbar ist. Die Schnittstelle erlaubt nur einen lesenden Zugriff auf die Daten (sogenannte GET-Requests), da die Wörterbuchinhalte nicht von den Nutzer*innen geändert werden sollen. Für jede Ressource wird ein

spezifischer API-Endpoint nach dem RESTful-Paradigma implementiert, wodurch die einzelnen APIs sehr einfach gehalten werden können, weil beispielsweise auf Funktionen zur Wörterbuchauswahl an den Endpoints verzichtet werden kann. Auch für das Rückgabeformat des Webservice wird mit **JSON** (JavaScript Object Notation) ein standardisiertes Austauschformat eingesetzt, über welches Daten in strukturierten Formen zur maschinellen Weiterverarbeitung bereitgestellt werden können.

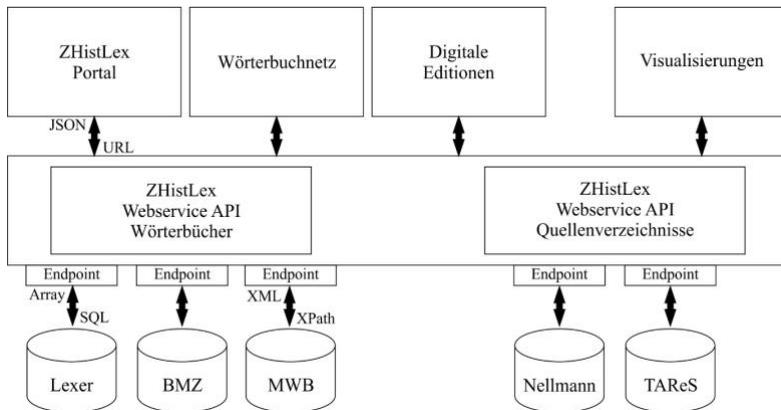


Abb. 2: vereinheitlichte Sicht auf die Daten durch ZHistLex-Webservices

Im Rahmen von ZHistLex erfolgte eine Implementierung der Schnittstelle für den Zugriff auf das neue Mittelhochdeutsche Wörterbuch (**MWB**) und dessen Quellenverzeichnis (der Endpoint der Service-Implementationen ist <http://tares.uni-trier.de/ZHistLex/API/>). Die implementierten Services werden produktiv in der Wörterbucharbeit genutzt, einerseits durch das Redaktionssystem **TAReS** (TriererArtikelRedaktionsSystem), indem es Informationen verschiedener Ressourcen dynamisch bündelt, sowie andererseits für die Vernetzung des MWB mit dem **Wörterbuchnetz**. Die Implementierung der Webservices seitens des Trierer Wörterbuchnetzes erfolgt über den Endpoint api.woerterbuchnetz.de/open-api/. Hierüber können die über die Schnittstelle erreichbaren Ressourcen sowie die für die einzel-

nen Wörterbücher implementierten Zugriffsmechanismen abgefragt werden. So liefert beispielsweise die Anfrage <https://api.woerterbuchnetz.de/open-api/dictionaries/Lexer> als Ergebnis die Struktur

```
{
  "result_type": "method_list",
  "result_count": 3,
  "query": "/open-api/dictionaries/Lexer",
  "result_set": [
    {
      "methodid": "fulltext",
      "comment": "Gesamter Text",
      "path": "/open-api/dictionaries/Lexer/fulltext/:searchpattern"},
    {
      "methodid": "lemmata",
      "comment": "Stichwort",
      "path": "/open-api/dictionaries/Lexer/lemmata/:searchpattern"},
    {
      "methodid": "definition",
      "comment": "Definitionen",
      "path": "/open-api/dictionaries/Lexer/definition/:searchpattern"}
  ]
}
```

und beschreibt damit, dass für das Mittelhochdeutsche Handwörterbuch von Matthias Lexer neben einer üblichen Volltextsuche ("methodid": "fulltext") auch eine Suche über die Artikelstichwörter ("methodid": "lemmata") und die Bedeutungserläuterungen ("methodid": "definition") möglich ist. Somit ist eine Anfrage der Form

<https://api.woerterbuchnetz.de/open-api/dictionaries/Lexer/definition/haus>

korrekt im Sinne der Spezifikation und liefert 64 Lexer-Artikel als Ergebnis.

Die zentrale technische Idee in ZHistLex war es, die einzelnen beteiligten Ressourcen des Zentrums (Wörterbücher und Quellentexte) über ein standardisiertes System von Webservices interoperabel zu machen. Die Entwicklung einer übergreifenden Oberfläche (GUI, Graphical User Interface) für menschliche Nutzung war dabei von nachgeordneter Bedeutung.

Das Potenzial der Infrastruktur zeigte sich neben dem Einsatz im MWB und im Wörterbuchnetz aber auch anhand prototypisch implementierter Abfrage- und Visualisierungsszenarien. Ein einfacher Demonstrator zur Quellenbibliographie führt eine gleichzeitige Anfrage an alle implementierten Services durch, bereitet das Ergebnis auf, visualisiert die Überschneidung der Quellennutzung in den beteiligten Wörterbüchern und illustriert damit exemplarisch die Möglichkeit, die Interaktion mit den ZHistLex-Webservices über Suchinterfaces auch menschlichen Nutzern zu erschließen. Er verwirklicht bewusst nur ein Basis-Konzept und realisiert nicht alle spezifizierten Suchoptionen, verdeutlicht aber, dass die gewählte Architektur geeignet ist, die Ressourcen der beteiligten Standorte interoperabel zu machen. Da alle Services ohne Zugangsbeschränkungen verfügbar sind, können von beliebiger Seite weitere Anwendungen auf dieser Basis entwickelt werden, wie beispielsweise eine föderierte Suche über eine Menge von unterschiedlichen Wörterbüchern mit anschließenden komplexen Aggregationen, Synopsen und Darstellungen der Ergebnismengen.

Der Forschung werden durch diese Zugriffsformen auf Wörterbücher mit ihrer sorgfältigen, je spezifischen lexikographischen Auswahl und Bearbeitung sowie auf Belegquellen mit einem reichhaltigen Kontext, den die lexikographischen Belege allein nicht bieten können, neue Möglichkeiten eröffnet. Ebenso wird die Begrenzung einer Suchanfrage auf den Dokumentationszeitraum eines einzelnen Wörterbuchs überwunden. Eine Vielzahl von Anforderungen, die die Forschung ebenso wie das gebildete Laienpublikum an die Suche und Präsentation von sprachhistorischen Daten hat, kann auf diese Weise besser und umfassender bedient werden (vgl. Burch [u. a.] 2020).

Die in ZHistLex und ELEXIS gewonnenen Erfahrungen fließen ein in die Arbeiten des Konsortiums [Text+](#) im Rahmen der Nationalen Forschungsdateninfrastruktur (NFDI), die damit in entscheidendem Maße zu einer langfristigen Nachhaltigkeit der erzielten Ergebnisse beiträgt und sie insbesondere weiterentwickelt. Im Bereich der Datendomäne Lexical Re-

sources (neben den Domänen Text Collections und Editions) werden hier die spezifisch für die historische Lexikographie erarbeiteten Konzepte auf weitere Arten von lexikalischen Ressourcen wie Wörterbücher zu verschiedenen Sprachen, verschiedenen regionalen Varietäten sowie auf Enzyklopädien, aber auch auf maschinell lesbare Wörterbücher, terminologische Datenbanken, Ontologien, Wortlisten, Wortkarten und linguistische Atlanten ausgeweitet. Daher beschäftigt sich das Konsortium auch mit der Frage, wie wissenschaftliche Kommunikation in und durch ihre Forschungsdaten nachnutzbar, verfügbar und geordnet werden kann. Die hierbei entstehenden Werkzeuge und Kompetenzen sind für Forschungen aus allen Bereichen von Wissenschaft relevant, die ihre Erkenntnisse oder Prozesse in Sprach- und Textformen vermitteln.

Literaturverzeichnis

Sekundärliteratur

- Burch, Thomas/Gloning, Thomas/Harm, Volker/Herold, Axel/Hoenen, Armin/Plate, Ralf/Recker-Hamm, Ute: Schlussbericht des Verbund-Projekts ZHistLex (»eHumanities-Zentrum für Historische Lexikographie«). Gießen 2020 ([online](#)).
- Declerck, Thierry/Wandl-Vogt, Eveline/Mörth, Karlheinz: Towards a Pan European Lexicography by Means of Linked (Open) Data, in: Kosem, I[ztok]/Jakubíček, M[iloš]/Kallas, J[elena]/Krek, S[imon] (Hrsg.): Electronic lexicography in the 21st century: linking lexical data in the digital age. Proceedings of the eLex 2015 conference, 11-13 August 2015, Ljubljana/Brighton 2015, S. 342–355 ([online](#)).
- Krek, Simon/Kosem, Iztok/McCrae, John P./Navigli, Roberto/Pedersen, Bolette S./Tiberius, Carole/Wissik, Tanja: European Lexicographic Infrastructure (ELEXIS), in: Čibej, Jaka/Gorjanc, Vojko/ Kosem, Iztok/ Krek, Simon (Hrsg.): Proceedings of the XVIII EURALEX International Congress: Lexicography in Global Contexts. Ljubljana 2018, S. 881–892 ([online](#)).
- Romary, Laurent/Tasovac, Toma: TEI Lex-o — A baseline encoding for lexicographic data. ([online](#)).
- WR 2017: Wissenschaftsrat: Bericht zur wissenschaftsgeleiteten Bewertung umfangreicher Forschungsinfrastrukturvorhaben für die Nationale Roadmap. (Drs. 6410-17), Juli 2017 ([online](#)).

Online-Ressourcen

COST (European Cooperation in Science and Technology): <https://www.cost.eu/>.
ENeL (European Network of e-Lexicography): <https://www.elexicography.eu/>.
JSON (JavaScript Object Notation): <https://www.json.org/json-en.html>.
MWB (Mittelhochdeutsches Wörterbuch): www.mhdwb-online.de/.
NFDI (Nationale Forschungsdateninfrastruktur): <https://www.nfdi.de/>.
OpenAPI (OpenAPI Specification v3.1.0): <https://spec.openapis.org/oas/latest.html>.
SOAP (Simple Object Access Protocol): <http://www.w3.org/TR/2000/NOTE-SOAP-20000508>.
TAReS: <http://www.mhdwb.uni-trier.de/TAReS/index.html>.
Text+: <http://www.text-plus.org/>.
Wörterbuchnetz: www.woerterbuchnetz.de.
YAML (YAML Ain't Markup Language): <https://yaml.org/spec/1.2.2/>.
ZHistLex (Zentrum für Historische Lexikographie): www.zhistlex.de.

Anschrift des Autors:

Dr. Thomas Burch
Universität Trier
Kompetenzzentrum – Trier Center for Digital Humanities
Universitätsring 15
54286 Trier
E-Mail: burch@uni-trier.de