

Ressourcenbeitrag: Germanistik im Netz

Marie Flüh  ¹

1. Universität Hamburg

forTEXT

Thema:	Bibliografie	DOI:	10.48694/fortext.3793
Jahrgang:	1	Ausgabe:	11
Erscheinungsdatum:	30-11-2024	Erstveröffentlichung:	2019-05-27 auf fortext.net
Lizenz:			open  access

Allgemeiner Hinweis: Rot dargestellte *Begriffe* werden im Glossar am Ende des Beitrags erläutert. Alle externen Links sind auch am Ende des Beitrags aufgeführt.

1. Kurzbeschreibung



Abb 1: Das neue GiN-Logo

Das Fachportal *Germanistik im Netz* (GiN) des Fachinformationsdienstes (FID) Germanistik bündelt die wichtigsten Informationsressourcen für das Fach Germanistik auf einer gemeinsamen Oberfläche (vgl. **GUI**) und fungiert u. a. als medienübergreifendes Rechercheportal und Nachweissystem für Datenbanken mit literatur- und sprachwissenschaftlichen Sammelschwerpunkten. Die integrierte Metadatensuche (vgl. **Metadaten**) ermöglicht es Ihnen, mit einem Suchvorgang (vgl. **Query**) gleichzeitig zahlreiche Bibliothekskataloge, zentrale Fachbibliographien, Verzeichnisse und einschlägige Fachdatenbanken nach digitalen wie konventionellen Dokumenten zu durchsuchen. GiN befindet sich derzeit in einer Umstrukturierungsphase. Nicht nur das Logo des zentralen Onlineportals wird sich verändern. Der für das dritte Quartal des Jahres 2019 geplante Relaunch bringt eine Bandbreite veränderter und neuer Funktionalitäten sowie eine gänzlich veränderte Struktur der Homepage mit sich.



Abb. 2: Die Startseite von Germanistik im Netz

Steckbrief

- <https://www.germanistik-im-netz.de/6>
- im Jahr 2006 im Rahmen des DFG-Programms zur überregionalen Literaturversorgung in Betrieb genommen, soll GiN auch künftig ein integraler Teil des Fachinformationsdienstes (FID) Germanistik bleiben
- Zusammenarbeit diverser Forschungsbibliotheken und -einrichtungen
- Ziele: fachbezogene Informationsangebote aufbauen, medienübergreifende bzw. parallele Recherche in unterschiedlichen Bibliothekskatalogen und weiteren Informationsressourcen ermöglichen
- integrierte Metasuche in verschiedenen Datenbeständen, im Idealfall Ermittlung, Bestellung bzw. Zugang zu gewünschten Informationen und Dokumenten, ohne das Portal zu verlassen
- Verzeichnis von 224 germanistischen Datenbanken: thematisch bzw. alphabetisch geordnet; nach individuellen Schlüsselbegriffen durchsuchbar
- Informationen über Neuerscheinungen der DNB, Speicher und Publikation von Forschungsergebnissen (über GiNDok)
- Metadaten der Datenbanken: Name und weitere Titel, URL (vgl. **URI**), Verfügbarkeit, Inhalt, Fachgebiet, Schlagwörter, Erscheinungsform, Datenbank-Typ, Verlag
- Metadaten zu Dokumenten: Verantwortlich/Verfasser*in, Erscheinungsdatum, Umfang, Serie, optionale Anmerkungen, ISBN, Schlagwörter, OPAC-Button (vgl. **OPAC**) (UB Frankfurt am Main), Verbund-Button (für deutschlandweite Nachweisinformationen des recherchierten Dokuments) subitobutton (für eine kostenpflichtige Bestellung)
- Zugang zu germanistischen Online-Fachzeitschriften via EZB, Überblick über germanistische Print- und E-Zeitschriften via ZDB, Zugang zu Zeitschriftenverzeichnis *Online Contents Germanistik* (139 Zeitschriften, über 197.000 Aufsätze und Rezensionen), integrierte Metadatensuche in den portal-eigenen Ressourcenverzeichnissen **GinFix** und **Wer-Was-Wo**

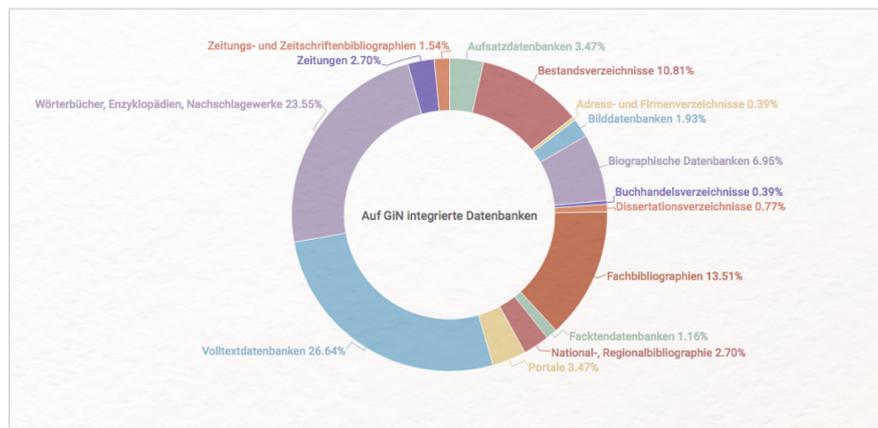


Abb. 3: Germanistik im Netz integriert 224 unterschiedliche Datenbanken. In einer Auswahl der Datenbanken können Sie über die Metasuche relevante Sekundär- und Primärliteratur recherchieren

2. Anwendungsbeispiel

Sie möchten religiöse Motive und intertextuelle Zusammenhänge in ausgewählten Werken Clemens Brentanos (vgl. **Korpus**) untersuchen. Hierfür benötigen Sie einzelne Werke von Clemens Brentano (*Die drei Nüsse*, das *Jägerlied*, der *Witzenspitzel* und *Die Schachtel mit der Friedenspuppe. Eine Erzählung*) und relevante Fachpublikationen zu diesem Thema. Ihre Ergebnisse möchten Sie als Online-Publikation veröffentlichen.

3. Diskussion

3.1 Kann ich Germanistik im Netz für wissenschaftliche Arbeiten nutzen?

Ja, bei der Auswahl der hier verzeichneten Netzangebote wird besonders darauf geachtet, dass die Dokumente in der bereitstellenden Datenbank auf Qualität, Stabilität und Authentizität geprüft wurden. Die langfristige Verfügbarkeit des Datenmaterials und die inhaltliche wie formale Erschließung der Dokumente entspricht bibliothekarischen Maßstäben (Gantert 2012, 132). Die große Bandbreite an germanistischen Informationsressourcen, die sich über GiN konsultieren lassen, zielt auf eine möglichst umfassende Abbildung des gesamten Fachs Germanistik ab. Die Sammelschwerpunkte der integrierten Datenbanken sind folglich so facettenreich wie die Fachdisziplin selbst und eignen sich, um Primär- wie Sekundärliteratur zu recherchieren. Die Metadaten von Datenbanken und Dokumenten sind intern einheitlich, vollständig und kennzeichnen auch die Verfügbarkeit der Dokumente. Die Verfügbarkeit einzelner recherchierter Dokumente wird durch die IP-Erkennung (vgl. **IP-Adresse**) des Rechners festgelegt und hängt dadurch sowohl von Ihrem Standort als auch vom Publikationstyp ab. Recherchierte Dokumente lassen sich über den „Merkliste“-Button in einem separaten Ordner speichern, als bibliographische Sammlung ausdrucken und/oder als E-Mail versenden.

3.2 Wie benutzerfreundlich ist die Arbeit mit Germanistik im Netz?

Der gut strukturierte wie übersichtliche Aufbau der Website führt Sie niederschwellig an die unterschiedlichen Angebote des Fachportals heran. Bereits die Startseite bietet eine erste Orientierung über abrufbare Serviceangebote (s. Abb. 4). Die parallele Metasuche in den unterschiedlichen Informationsressourcen (die Kernfunktion von GiN) ist ebenfalls ohne große Einarbeitungszeit möglich.



Abb. 4: Die Metasuche über GiN (rechts) in einer Auswahl von Informationsressourcen

Die auf der Website zur Verfügung gestellten Hinweise für eine erfolgreiche Suche erleichtern die Recherche. Rückfragen und/oder Anmerkungen werden zügig beantwortet.

Zahlreiche universitäre Fächer verfügen über eine virtuelle Fachbibliothek, die bis 2015 von den entsprechenden Sondersammelgebieten betreut wurden. Germanistik im Netz ist mit dem Fokus auf Datenbanken aus dem deutschsprachigen Raum nicht mehrsprachig vorhanden.

4. Wie funktioniert die Textsuche über Germanistik im Netz?

GiN vereint unterschiedliche Informationsressourcen in einer Suchmaske. Nutzer*innen können dadurch gleichzeitig in zahlreichen Datenbanken recherchieren, was das Bibliografieren (Flüh 2024) vereinfacht und effizienter macht. Suchvorgänge starten Sie über die Metasuche, wobei Sie nach verschiedenen Suchkriterien recherchieren und eine Einschränkung auf einzelne Dokumenttypen (Bücher, Zeitschriften, Online-Publikationen, Aufsätze) vornehmen können. Ergänzend ist i. d. R. eine separate Suche in den einzelnen OPACs möglich. Vor allem für komplexe Suchanfragen empfiehlt es sich, die Suche direkt über die jeweilige Datenbank umzusetzen, da hier z. T. komplexere Suchanfragen möglich sind. Zu der durch den FID getroffenen Auswahl an Datenbanken zählen im Einzelnen acht Bibliothekskataloge, zwei eigene Internetquellen, fünf bibliographische Datenbanken, fünf Volltextdatenbanken, zwei Nationallizenzen, das Zentrale Verzeichnis digitalisierter Drucke (ZVDD), zwei Monographien, ein Kleinschrifttum und drei Periodika. Die Auswahl der Datenbanken zielt auf eine holistische Erfassung der qualitativ hochwertigen Sammelstellen mit literatur- und sprachwissenschaftlichen Sammelschwerpunkten. Eine Vorauswahl und Bewertung wird folglich durch die Betreiber vorgenommen.

Bibliothekskataloge UB Frankfurt Fachkatalog Germanistik, UB Frankfurt Retrokatalog, Deutsches Literaturarchiv Marbach, Herzog August Bibliothek Wolfenbüttel, Herzogin Anna Amalia Bibliothek Weimar, Institut für Deutsche Sprache Mannheim, Thomas-Mann-Sammlung, ULB Düsseldorf, UB Frankfurt	Periodika DNB: Exilpresse digital, Deutsche Exilzeitschriften 1933-1945, Jüdische Periodika - Compact Memory, Rezensionenorgane/Aufklärungszeitschriften, UB Bielefeld	Monographien Sammlung Ponickau, ULB Sachsen-Anhalt, VD16 digital, Bayerische Staatsbibliothek München
Volltextdatensammlungen BASE, GiNDok, DigiZeitschriften: Germanistik, dünnhaupt digital, TextGrid Repository	Nationallizenzen Periodicals Index Online (PIO) (1739 - 2000), Periodicals Archive Online (PAO) (1802 - 2000)	Internetquellen GinFix, Wer-Was-Wo
Bibliographische Datenbanken BDSL, BLL, Weimarer Goethe-Bibliographie online, Klassik online, Online Contents Germanistik	Kleinschrifttum Flugschriften 1848, UB Frankfurt	ZVDD Zentrales Verzeichnis Digitalisierter Drucke

Abb. 5: In die Recherche via GiN integrierte Informationsressourcen

Als Suchoptionen stehen Ihnen die Metasuche und die erweiterte Metasuche (auch per Mobiltelefon möglich) zur Verfügung. Hierbei legen Sie entweder im Vorfeld fest, welche Datenbanken konsultiert werden sollen, oder beziehen bei Ihrer Suchanfrage sämtliche Informationsressourcen ein. Für die Recherche kann zwischen verschiedenen Suchoptionen gewählt werden, die sich mithilfe eines Booleschen Operators (und, oder, nicht) miteinander verknüpfen lassen (Freie Suche, Autor*in, Titelstichwort, Vollständiger Titel, Schlagwort, Institution, Verlag/Verlagsort, Erscheinungsjahr, ISBN/ISSN). Bei der Metasuche via GiN gilt es zu bedenken, dass nicht alle Suchoptionen in allen Informationsressourcen zur Verfügung stehen und keine Dublettenchecks stattfinden, sodass einzelne Titel aus verschiedenen Informationsressourcen mehrfach angezeigt werden können. Durch die simultane Recherche in mehreren Datenbanken wird die Suche nach relevanter Fachliteratur allerdings erheblich erleichtert und beschleunigt. Für den Bereich der Germanistik (Sprach- und Literaturwissenschaft) nehmen die seit Jahrzehnten stetig wachsenden Bestände der führenden laufenden Fachbibliographien *Bibliographie der deutschen Sprach- und Literaturwissenschaft Online* (BDSL Online oder den Namen ihrer Begründer entsprechend auch als Eppelsheimer-Köttelwelsch bekannt) und *Bibliography of Linguistic Literature Database* (BLLDB oder BLL für Bibliographie Linguistischer Literatur) eine Schlüsselposition ein. Die Datenbankangebote sind ab dem Berichtsjahrgang 1985 (BDSL) bzw. ab dem Jahrgang 1971 (BLLDB) für jede Universitätsbibliothek und über GiN (entweder im Rahmen einer Metasuche oder separat davon als Freie Suche via GiN unter dem Reiter Bibliotheken) elektronisch abrufbar. Für eine Recherche in länger zurückliegende Jahre (also vor 1971 bzw. 1985) müssen Sie auf die nach wie vor erscheinende Print-Version zurückgreifen. Gastnutzer*innen, die eine Suchanfrage direkt über die BDSL oder BLLDB starten, stehen die Berichtsjahrgänge bis 2007 zur Verfügung. Eine Exportfunktion von Texten aus der BDSL in gängige Literaturverwaltungsprogramme besteht derzeit nicht, hier soll allerdings im Jahr 2019 nachgebessert werden. Metadaten ausgewählter Werke (Verfasser*in, Titel, Verlag, Datum, Seitenzahl) lassen sich bei der Recherche in der BDSL und der BLLDB in einer Auswahlliste speichern, als automatisch generierte Bibliographie herunterladen und per E-Mail versenden. Über den „Find it“-Button können Sie Standorte herausfinden.

Das Angebot des Portals ist übrigens nicht auf die Recherche von Sekundärliteratur beschränkt. Das ebenfalls über eine individuelle Suchanfrage durchsuchbare Webverzeichnis *Wer-Was-Wo* beinhaltet Webpublikationen zu Lehr- und Forschungseinrichtungen, Fachorganisationen und Forschungsprojekten. *GinFix* ist der zweite Webkatalog von GiN und verzeichnet Ressourcentypen wie Netzpublikationen, Informationsanbieter, literarische Zeitschriften und Blogs, Bibliographien, Forschungseinträge, multimediale Angebote, Nachschlagewerke, Werke, Autor*innen, Newsletter, Mailinglisten, Foren, Wikis und wissenschaftliche Blogs. *GinDok* ist der Dokumentenserver des FID Germanistik und ermöglicht die digitale Publikation von Forschungsergebnissen v. a. aus den Bereichen Mediävistik, Neuere Deutsche Literaturwissenschaft und germanistische Linguistik.

Externe und weiterführende Links

- Germanistik im Netz: <https://web.archive.org/web/20241106120950/https://www.germanistik-im-netz.de/> (Letzter Zugriff: 06.11.2024)

Bibliographie

- Flüh, Marie. 2024. Methodenbeitrag: Digitales Bibliografieren. Hg. von Evelyn Gius. *forTEXT* 1, Nr. 11. Bibliografie (29. November). doi: 10.48694/fortext.3786, <https://fortext.net/routinen/methoden/digitales-bibliografieren>.
- Gantert, Klaus. 2012. *Erfolgreich recherchieren - Germanistik*. Berlin: de Gruyter.
- Göttker, Susanne. 2016. *Literaturversorgung in Deutschland: Von den Sondersammelgebieten zu den Fachinformationsdiensten: eine Analyse*. Wiesbaden: Dinges & Frick.

Glossar

Annotation Annotation beschreibt die manuelle oder automatische Hinzufügung von Zusatzinformationen zu einem Text. Die manuelle Annotation wird händisch durchgeführt, während die (teil-)automatisierte Annotation durch **Machine-Learning-Verfahren** durchgeführt wird. Ein klassisches Beispiel ist das automatisierte **PoS-Tagging** (Part-of-Speech-Tagging), welches oftmals als Grundlage (**Preprocessing**) für weitere Analysen wie Named Entity Recognition (NER) nötig ist. Annotationen können zudem deskriptiv oder analytisch sein.

Browser Mit Browser ist in der Regel ein Webbrowser gemeint, also ein Computerprogramm, mit dem das Anschauen, Navigieren auf, und Interagieren mit Webseiten möglich wird. Am häufigsten genutzt werden dafür Chrome, Firefox, Safari oder der Internet Explorer.

Commandline Die Commandline (engl. *command line interface* (CLI)), auch Kommandozeile, Konsole, Terminal oder Eingabeaufforderung genannt, ist die direkteste Methode zur Interaktion eines Menschen mit einem Computer. Programme ohne eine grafische Benutzeroberfläche (**GUI**) werden i. d. R. durch Texteingabe in die Commandline gesteuert. Um die Commandline zu öffnen, klicken Sie auf Ihrem Mac „cmd“ + „space“, geben „Terminal“ ein und doppelklicken auf das Suchergebnis. Bei Windows klicken Sie die Windowstaste + „R“, geben „cmd.exe“ ein und klicken Enter.

GUI GUI steht für *Graphical User Interface* und bezeichnet eine grafische Benutzeroberfläche. Ein GUI ermöglicht es, Tools mithilfe von grafischen Schaltflächen zu bedienen, um somit beispielsweise den Umgang mit der **Commandline** zu umgehen.

HTML HTML steht für *Hypertext Markup Language* und ist eine textbasierte Auszeichnungssprache zur Strukturierung elektronischer Dokumente. HTML-Dokumente werden von **Webbrowsern** dargestellt und geben die Struktur und Online-Darstellung eines Textes vor. HTML-Dateien können außerdem zusätzliche **Metainformationen** enthalten, die auf einer Webseite selbst nicht ersichtlich sind.

IP-Adresse Die Vernetzung von Computern wird in einem Internetprotokoll (IP) festgehalten, woraufhin jedes angebundene Gerät in diesem Computernetz eine IP-Adresse erhält. So werden die Geräte adressierbar und erreichbar gemacht. Die IP gehört zu den personenbezogenen Daten, da über sie auf Ihre Identität geschlossen werden kann.

Korpus Ein Textkorpus ist eine Sammlung von Texten. Korpora (Plural für „das Korpus“) sind typischerweise nach Textsorte, Epoche, Sprache oder Autor*in zusammengestellt.

Lemmatisieren Die Lemmatisierung von Textdaten gehört zu den wichtigen **Preprocessing**-Schritten in der Textverarbeitung. Dabei werden alle Wörter (**Token**) eines Textes auf ihre Grundform zurückgeführt. So werden beispielsweise Flexionsformen wie „schneller“ und „schnelle“ dem Lemma „schnell“ zugeordnet.

Machine Learning Machine Learning, bzw. maschinelles Lernen im Deutschen, ist ein Teilbereich der künstlichen Intelligenz. Auf Grundlage möglichst vieler (Text-)Daten erkennt und erlernt ein Computer die häufig sehr komplexen Muster und Gesetzmäßigkeiten bestimmter Phänomene. Daraufhin können die aus den Daten gewonnen Erkenntnisse verallgemeinert werden und für neue Problemlösungen oder für die Analyse von bisher unbekanntem Daten verwendet werden.

Markup Language Markup Language bezeichnet eine maschinenlesbare Auszeichnungssprache, wie z. B. **HTML**, zur Formatierung und Gliederung von Texten und anderen Daten. So werden beispielsweise auch **Annotationen** durch ihre Digitalisierung oder ihre digitale Erstellung zu Markup, indem sie den Inhalt eines Dokumentes strukturieren.

Metadaten Metadaten oder Metainformationen sind strukturierte Daten, die andere Daten beschreiben. Dabei kann zwischen administrativen (z. B. Zugriffsrechte, Lizenzierung), deskriptiven (z. B. Textsorte), strukturellen (z. B. Absätze oder Kapitel eines Textes) und technischen (z. B. digitale Auflösung, Material)

Metadaten unterschieden werden. Auch **Annotationen** bzw. **Markup** sind Metadaten, da sie Daten/Informationen sind, die den eigentlichen Textdaten hinzugefügt werden und Informationen über die Merkmale der beschriebenen Daten liefern.

Named Entities Eine Named Entity (NE) ist eine Entität, oft ein Eigenname, die meist in Form einer Nominalphrase zu identifizieren ist. Named Entities können beispielsweise Personen wie „Nils Holgerson“, Organisationen wie „WHO“ oder Orte wie „New York“ sein. Named Entities können durch das Verfahren der Named Entity Recognition (NER) automatisiert ermittelt werden.

OPAC OPAC steht für *Online Public Access Catalogue* und bezeichnet online zugängliche Bibliothekskataloge.

POS PoS steht für *Part of Speech*, oder „Wortart“ auf Deutsch. Das PoS- **Tagging** beschreibt die (automatische) Erfassung und Kennzeichnung von Wortarten in einem Text und ist of ein wichtiger **Preprocessing**-Schritt, beispielsweise für die Analyse von **Named Entities**.

Preprocessing Für viele digitale Methoden müssen die zu analysierenden Texte vorab „bereinigt“ oder „vorbereitet“ werden. Für statistische Zwecke werden Texte bspw. häufig in gleich große Segmente unterteilt (*chunking*), Großbuchstaben werden in Kleinbuchstaben verwandelt oder Wörter werden **lemmatisiert**.

Query *Query* bedeutet „Abfrage“ oder „Frage“ und bezeichnet eine computergestützte Abfrage zur Analyse eines Textes. Um Datenbestände zu durchsuchen, werden Abfragesprachen eingesetzt, die *Queries* (Anfragen) an den Datenbestand senden. So bilden alle möglichen Queries zusammen die *Query Language* eines Tools.

Type/Token Das Begriffspaar „Type/Token“ wird grundsätzlich zur Unterscheidung von einzelnen Vorkommnissen (Token) und Typen (Types) von Wörtern oder Äußerungen in Texten genutzt. Ein Token ist also ein konkretes Exemplar eines bestimmten Typs, während ein Typ eine im Prinzip unbegrenzte Menge von Exemplaren (Token) umfasst.

Es gibt allerdings etwas divergierende Definitionen zur Type-Token-Unterscheidung. Eine präzise Definition ist daher immer erstrebenswert. Der Satz „Ein Bär ist ein Bär.“ beinhaltet beispielsweise fünf Worttoken („Ein“, „Bär“, „ist“, „ein“, „Bär“) und drei Types, nämlich: „ein“, „Bär“, „ist“. Allerdings könnten auch vier Types, „Ein“, „ein“, „Bär“ und „ist“, als solche identifiziert werden, wenn Großbuchstaben beachtet werden.

URI *Uniform Resource Identifier* (URI) ist ein Identifikator zur eindeutigen Erkennung von Online-Ressourcen wie Webseiten. Im „Raum“ des Internets können so alle Inhalte eindeutig identifiziert werden, unabhängig davon, ob es sich dabei beispielsweise um eine Seite mit Text oder Video handelt. Die am häufigsten verwendete Form eines URI ist die Webseitenadresse, die URL.