

# LexHelfer: die Digitalisierung der 6 Millionen Belege des Bayerischen Wörterbuchs

Manuel Raaf

Referat für IT & Digital Humanities, Bayerische Akademie der Wissenschaften

## Handhabung hunderttausender digitalisierter Fragebögen und Textzeilen

**LexHelfer** bietet einige Grundfunktionen, um digitalisierte Fragebögen oder abgetippte Belege zu durchsuchen und zu editieren:

- Import von Scans; Import von Excel-Listen, die Transkriptionen und weitere text-basierte Informationen zu jedem Einzelbeleg enthalten
- halbautomatische Generierung der Schnipsel sowie Schwärzung personenbezogener Daten im Scan
- schneller Zugriff auf alle Daten; schnelles Durchblättern der Belege
- eingeschränkter Lese-Zugriff für die Allgemeinheit
- Vernetzung einzelner Projekte innerhalb **LexHelfers** und somit übergreifende Suche in diesen Projekten

Neben diesen allgemeinen Funktionen sind weitere **projekt-spezifische Funktionen** enthalten (siehe Beispiele). Außerdem ist die Möglichkeit gegeben, neue zu entwickeln, um künftige Anforderungen der Redaktion erfüllen zu können.

### Technische Details:

- Web-Anwendung (Client-Server)
- OpenSource-Standards in der Entwicklung:
  - Server: LAMP
  - Client: CSS, JavaScript, HTML5
- kostengünstig, da keine Lizenzgebühren anfallen
- anpassbar, wartbar, archivierbar



Fragebögen des Bayerischen Wörterbuchs



## Beispiel: Bayerisches Wörterbuch

STARTSEITE LEMMA-BELEG DRUCKEN ZEIGE LEMMA PRO WORTLISTE BENUTZER-VERWALTUNG ANLEITUNG

Bitte geben Sie hier die gewünschten Suchparameter bzw. einen Suchbegriff ein.  
Die Eingabe mehrerer Parameter bewirkt eine UND-Verknüpfung:

Bogen Nr.: 60  
Fragebogennummer: 4  
Fragebogen: Frage 4: Kennen Sie das Wort Tiefet (Diafat, Duifat)? Was verstehen Sie +  
Fragebogen (vollst.):  
Fragebogen (nur Lemma):  
zugeordnetes Lemma:

Bezirk: Niederbayern  
Kreis:  
Ort:  
Sortieren: Ort

Folgende Aktionen wirken sich auf alle Suchergebnisse aus, nicht nur auf die dargestellten 10!  
Status aller Suchergebnisse ändern auf: nicht erledigt  
Alle Suchergebnisse einem Lemma zuordnen in ein PDF drucken

Ihre Suche ergab 25 Treffer:

WL-060-Abbach-Bad-43H-01.jpg

60/ 4. Kennen Sie das Wort Tiefet (Diafat, Duifat)?  
Was verstehen Sie darunter? Genaue Aussprache? Bitte Satzbeispiel.

Met mir als Eigenschaftenstort / dem Thoms  
Geo Mo schteckt 2tweifst em Dreck drin

Status ändern auf: irrelevant Status ändern auf: erledigt Schnipsel ändern / Scan anzeigen einem Lemma zuordnen

WL-060-Aidenbach-VDF-01.jpg

60/ 4. Kennen Sie das Wort Tiefet (Diafat, Duifat)?  
Was verstehen Sie darunter? Genaue Aussprache? Bitte Satzbeispiel.

Gib acht, daß die nicht jeds  
Tiafat einikommst (in das tiefe  
Bass)

Status ändern auf: irrelevant Status ändern auf: erledigt Schnipsel ändern / Scan anzeigen einem Lemma zuordnen

WL-060-Altenbuch-LAN-01.jpg

60/ 4. Kennen Sie das Wort Tiefet (Diafat, Duifat)?  
Was verstehen Sie darunter? Genaue Aussprache? Bitte Satzbeispiel.

hier für Tiefet  
Du bist aqanza Doife

Status ändern auf: irrelevant Status ändern auf: erledigt Schnipsel ändern / Scan anzeigen einem Lemma zuordnen

WL-060-Altenhausen-LA-01.jpg

60/ 4. Kennen Sie das Wort Tiefet (Diafat, Duifat)?  
Was verstehen Sie darunter? Genaue Aussprache? Bitte Satzbeispiel.

ja man versteht einen Körper.  
Der duifat den 99. Tag so dahin der mag

Status ändern auf: irrelevant Status ändern auf: erledigt Schnipsel ändern / Scan anzeigen einem Lemma zuordnen

- > 400.000 digitalisierte, handschriftlich ausgefüllte Fragebögen
  - > 6.000.000 halbautomatisch erstellte Schnipsel
  - Filtermöglichkeiten für Belege:
    - fertig bearbeitete Antworten
    - auskommentierte, unbrauchbare Antworten (z. B. leere oder solche, die lediglich „unbekannt“ oder „nein“ etc. beinhalten)
  - Suchoptionen zur geographischen Filterung (z. B. Regierungsbezirk, Stadt)
  - einfache Zuweisung eines Lemmas zum Beleg
  - PDF-Export der Suchergebnisse oder ganzer Lemmalisten samt Schnipsel
  - Erfassung neuer Fragebögen in nativ digitaler Form
- ✓ schnelles Auffinden von Belegen
  - ✓ Vorbereitung der Artikelerstellung durch schnelle Lemmatisierung der Belege
  - ✓ direkte Veröffentlichung der bearbeiteten Belegmaterialien online und gedruckt (via XML-Export und Weiterbearbeitung in Oxygen).

<http://www.bwb.badw.de/bwb-digital/datenbank/>

## Beispiel: Fränkisches Wörterbuch

STARTSEITE LEMMALISTE BENUTZER-VERWALTUNG ADMINISTRATION LOGOUT KONTAKT / HILFE

Bitte geben Sie hier die gewünschten Suchparameter bzw. einen Suchbegriff ein.  
Die Eingabe mehrerer Parameter bewirkt eine UND-Verknüpfung:

Fragebogen: 60  
Fragebogennummer: 4  
Fragebogen: Frage 4: Kennen Sie das Wort Tiefet (Diafat, Duifat)? Was verstehen Sie +  
Fragebogen (vollst.):  
Fragebogen (nur Lemma):  
zugeordnetes Lemma:

Bezirk: Niederbayern  
Kreis:  
Ort:  
Sortieren: Ort

Folgende Aktionen wirken sich auf alle Suchergebnisse aus, nicht nur auf die dargestellten 10!  
Status aller Suchergebnisse ändern auf: nicht erledigt  
Alle Suchergebnisse einem Lemma zuordnen in ein PDF drucken

Ihre Suche ergab 25 Treffer:

WL-060-Abbach-Bad-43H-01.jpg

60/ 4. Kennen Sie das Wort Tiefet (Diafat, Duifat)?  
Was verstehen Sie darunter? Genaue Aussprache? Bitte Satzbeispiel.

Met mir als Eigenschaftenstort / dem Thoms  
Geo Mo schteckt 2tweifst em Dreck drin

Status ändern auf: irrelevant Status ändern auf: erledigt Schnipsel ändern / Scan anzeigen einem Lemma zuordnen

WL-060-Aidenbach-VDF-01.jpg

60/ 4. Kennen Sie das Wort Tiefet (Diafat, Duifat)?  
Was verstehen Sie darunter? Genaue Aussprache? Bitte Satzbeispiel.

Gib acht, daß die nicht jeds  
Tiafat einikommst (in das tiefe  
Bass)

Status ändern auf: irrelevant Status ändern auf: erledigt Schnipsel ändern / Scan anzeigen einem Lemma zuordnen

WL-060-Altenbuch-LAN-01.jpg

60/ 4. Kennen Sie das Wort Tiefet (Diafat, Duifat)?  
Was verstehen Sie darunter? Genaue Aussprache? Bitte Satzbeispiel.

hier für Tiefet  
Du bist aqanza Doife

Status ändern auf: irrelevant Status ändern auf: erledigt Schnipsel ändern / Scan anzeigen einem Lemma zuordnen

WL-060-Altenhausen-LA-01.jpg

60/ 4. Kennen Sie das Wort Tiefet (Diafat, Duifat)?  
Was verstehen Sie darunter? Genaue Aussprache? Bitte Satzbeispiel.

ja man versteht einen Körper.  
Der duifat den 99. Tag so dahin der mag

Status ändern auf: irrelevant Status ändern auf: erledigt Schnipsel ändern / Scan anzeigen einem Lemma zuordnen

- basierend auf Excel-Listen
- Zuweisung jeder Zeile zu Bogen und Frage der entsprechenden Gewährsperson
- Such- und Anzeigefilter
- Export nach REDE<sup>1</sup> zur halbautomatischen Kartierung (künftig: vollautomatisch)
- automatische Kompilierung und Onlinestellung des zitierfähigen Lemma-Artikels
- Anbindung der Belegmaterialien an das Wörterbuchnetz (künftig)
- CSV-Export der Suchergebnisse für individuelle Weiterbearbeitung / Forschung
- Sachgruppenzuordnung (nach Post) und Suche über diese

- ✓ schnelles Auffinden von Belegen, die in einem Ort bereits erfasst wurden und somit nicht erneut erfasst werden müssen (→ Zeitersparnis)
- ✓ Zugriff auf komplettes Belegmaterial; Suchfilter für entsprechend exakte Suchanfragen
- ✓ automatische Artikelerstellung

<http://www.wbf.badw.de/wbf-digital/wörterbuch/>

<sup>1</sup> SprachGIS (<https://www.regionalsprache.de/SprachGIS/Map.aspx>) zur individuellen Kartenerstellung