Lexikographie

# Vom Papierfragebogen zur Online-Datenbank

## Das Bayerische Wörterbuch auf dem Weg ins digitale Zeitalter.

Von Edith Burkhart-Funk und Ursula Welsch

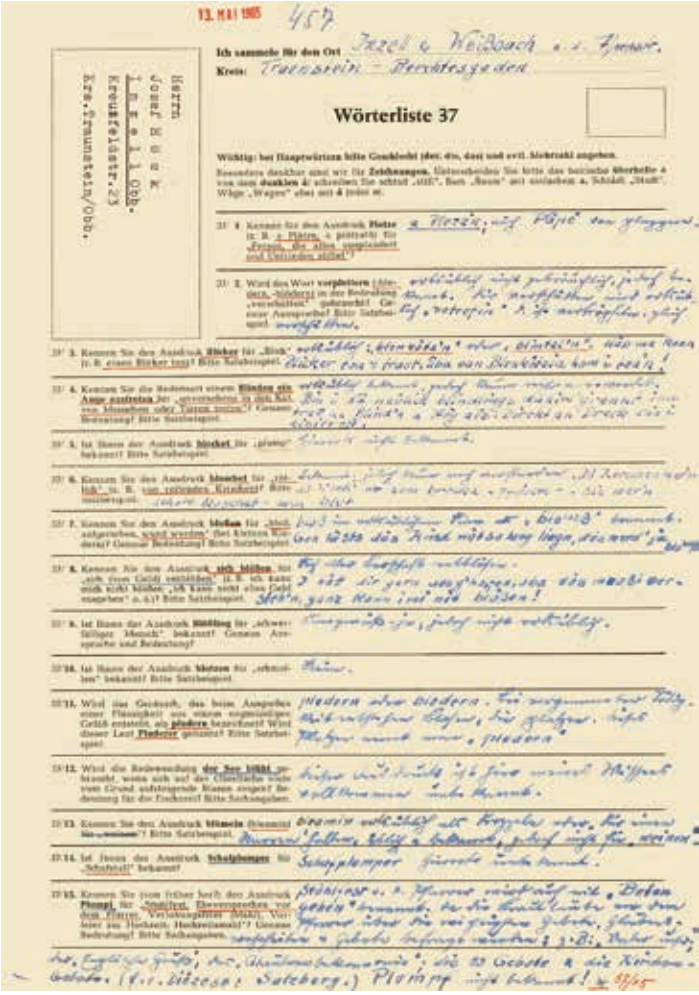



**Abb. 1: Beispiel einer Wörterliste aus der Sammlung des Bayerischen Wörterbuchs.**

DIE LEICHTIGKEIT und Unscheinbarkeit manches Wörterbuchartikels kann täuschen. Hinter ein paar Zeilen stecken oft stundenlange Exzerpier-, Sortier- und Recherchearbeiten. Moderne wissenschaftliche Mundartwörterbücher sind rein empirisch, d. h. ausschließlich beleggestützt. Die Kompetenz des Bearbeiters, seine Introspektion, darf keine Rolle spielen.

Mehr Freude als das Zusammensuchen des Materials macht den Redaktoren die anschließende Interpretation: Aus den amorphen Rohdaten werden die unterschiedlichen Bedeutungen eines Wortes herausgehoben und in eine Struktur gebracht, die dem Artikel sein Gerüst gibt. Die Gliederung eines Artikels nach unterschiedlichen Bedeutungen spiegelt also die Strukturierung und Interpretation des Materials durch den Bearbeiter wider.

Im Folgenden geht es um das Aufbereiten und Herrichten der Sprachdaten für einen Artikel des Bayerischen Wörterbuchs, und zwar zum einen um die seit Jahren praktizierte „händische" Methode und zum anderen darum, wie die Digitalisierung des gesamten Materials und seine Einrichtung in einer Datenbank die Arbeit vereinfachen und beschleunigen kann.

### Die Sammlungen des Bayerischen Wörterbuchs

Der Bestand des Bayerischen Wörterbuchs ist ziemlich heterogen. Gesammelt wird schon seit 1913. Es wurden literarische Quellen ausgewertet, zeitgenössische und ältere Mundartliteratur, aber auch Monographien und kleinere Darstellungen der bairischen Mundarten (wenn vom Dialekt die Rede ist, schreibt man „bairisch" mit „i") exzerpiert, ebenso wie Texte aus älteren Sprachstufen, da das Bayerische Wörterbuch auch den gesamten historischen Wortschatz darstellen soll. Viele an ihrer Mundart Interessierte haben der Arbeitsstelle ihre privaten Sammlungen zur Verfügung gestellt. Ende der 1920er Jahre wurde auf so genannten „Kundfahrten" Material durch Befragung einheimischer Sprecher erhoben (Abb. 2). Der größte Teil des Bestandes basiert jedoch auf großangelegten Sammlungsaktionen mittels Fragebögen. Systematisch wurde in Altbayern und den angrenzenden bairischen Gebieten Schwabens und Frankens ein Netz von Sammlern aufgebaut, deren Anzahl zwischen ca. 200 und über 500 schwankte. In der schlechtesten Phase nach dem Ersten Weltkrieg war die Zahl auf nur noch 38 Sammler zurückgegangen. Heute haben wir etwa 430 Sammler.

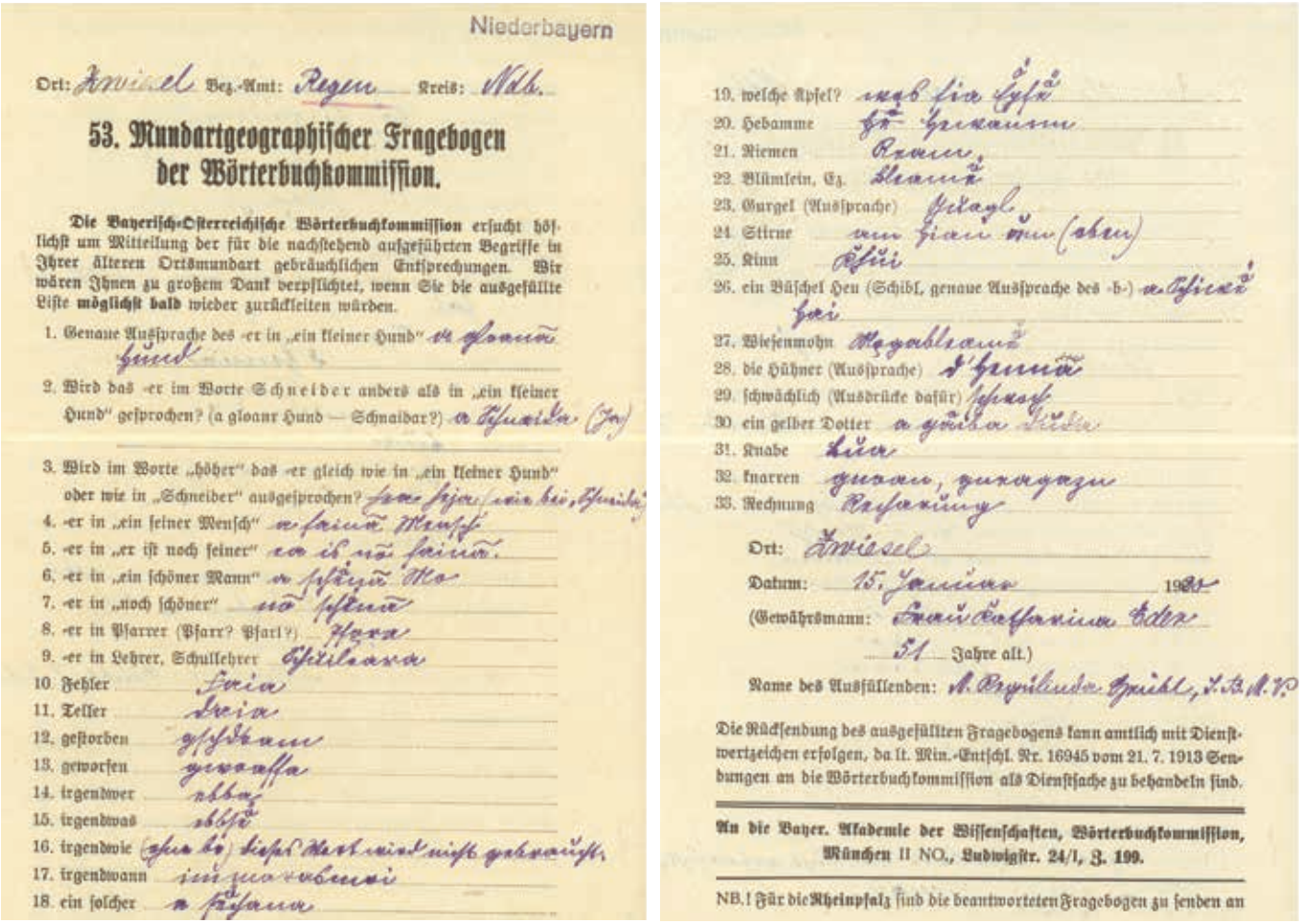
In der ersten Periode bis zur Jahrhundertmitte fragte man ausgehend von Bedeutungen nach den Bezeichnungen. Auch übertragene Bedeutungen, Wendungen und Redensarten,

ALLE ABB.: BWB

Sprichwörter etc., in denen ein Wort vorkommen kann, sowie Volkskundliches zu einem Begriff wurden erhoben. Ein Beispiel: „Welche Bezeichnungen verwenden Sie für das Schaf? Für das männliche, weibliche? Nennen Sie Schimpfwörter, Redensarten mit diesen Bezeichnungen, Bezeichnungen für unterschiedliche Altersstufen, unterschiedliche Lämmerrassen." Um die Jahrhundertmitte wurde das gesammelte Material gesichtet, um Schwachstellen und Lücken zu ermitteln. Von nun an konnte man umgekehrt vorgehen und fragte nicht mehr nach dem Wort für eine Sache, sondern in alphabetischer Reihenfolge gezielt nach dem Vorkommen, der Bedeutung, der Verwendung und der Aussprache von bestimmten Bezeichnungen. Die Sammler wurden gebeten, immer auch ein Satzbeispiel anzuführen. Das Material dieser sogenannten „Wörterlisten" (Abb. 1) ist die umfangreichste der Sammlungen, sie umfasst derzeit ca. 108.000 Listen mit je 60 Fragen, die wegen der regionalen sprachlichen Unterschiede freilich selten alle beantwortet wurden. Die Befragungen mit solchen „Wörterlisten" werden bis zum Ende des Alphabets durchgeführt. Die Anzahl der Belege in den unterschiedlichen Sammlungen ist schwer zu schätzen, liegt aber auf jeden Fall bei mehreren Millionen. Aufgrund der mittlerweile sehr weit verbreiteten digitalen Technologien können Sammler die Fragebögen auch als Dateien einreichen. Ein Online-Eingabe-Tool ist in Planung – dann können die Fragen direkt in einem Bildschirmformular beantwortet werden.

## Bearbeitung des Sprachmaterials vor dem Verfassen eines Wortartikels

Die Exzerpte der literarischen Quellen, die Sammlungen der Laien, der Kundfahrtenkatalog sowie die Ergebnisse der ersten schriftlichen Erhebungen sind in der alphabetisch geordneten Zentralkartei abgelegt (Abb. 3). Diese Belege sind sofort zur Hand, wenn es darum geht, einen bestimmten Wortartikel zu schreiben. Mehr Arbeit machen die Fragelisten. Aufgrund der Fragestellung weiß man, in welcher Liste ein bestimmtes Wort zu suchen ist. Nun kann man darangehen,



Niederbayern

Ort: Zwiesel Bez.-Amt: Regen Kreis: Ndb.

**53. Mundartgeographischer Fragebogen der Wörterbuchkommission.**

**Die Bayerisch-Österreichische Wörterbuchkommission** ersucht höflichst um Mitteilung der für die nachstehend aufgeführten Begriffe in Ihrer älteren Ortsmundart gebräuchlichen Entsprechungen. Wir wären Ihnen zu großem Dank verpflichtet, wenn Sie die ausgefüllte Liste **möglichst bald** wieder zurückleiten würden.

1. Genaue Aussprache des -er in „ein kleiner Hund"
2. Wird das -er im Worte Schneider anders als in „ein kleiner Hund" gesprochen? (a gloanr Hund — Schnaidar?)
3. Wird im Worte „höher" das -er gleich wie in „ein kleiner Hund" oder wie in „Schneider" ausgesprochen?
4. -er in „ein feiner Mensch"
5. -er in „er ist noch feiner"
6. -er in „ein schöner Mann"
7. -er in „noch schöner"
8. -er in Pfarrer (Pfarr? Pfarl?)
9. -er in Lehrer, Schullehrer
10. Fehler
11. Teller
12. gestorben
13. geworfen
14. irgendwer
15. irgendwas
16. irgendwie
17. irgendwann
18. ein solcher

19. welche Äpfel?
20. Hebamme
21. Riemen
22. Blümlein, Ez.
23. Gurgel (Aussprache)
24. Stirne
25. Kinn
26. ein Büschel Heu (Schibl, genaue Aussprache des -b-)
27. Wiesenmohn
28. die Hühner (Aussprache)
29. schwächlich (Ausdrücke dafür)
30. ein gelber Dotter
31. Knabe
32. knarren
33. Rechnung

Ort: Zwiesel

Datum: 15. Januar 1935

(Gewährsmann: Franz Xaver Eder, 51 Jahre alt.)

Name des Ausfüllenden: A. Brunhilde Spiehl, z. Zt. M. P.

Die Rücksendung des ausgefüllten Fragebogens kann amtlich mit Dienstwertzeichen erfolgen, da lt. Min.-Entschl. Nr. 16945 vom 21. 7. 1913 Sendungen an die Wörterbuchkommission als Dienstsache zu behandeln sind.

**An die Bayer. Akademie der Wissenschaften, Wörterbuchkommission, München II NO., Ludwigstr. 24/I, Z. 199.**

NB.! Für die Rheinpfalz sind die beantworteten Fragebogen zu senden an

**Abb. 2: Mundartgeographischer Fragebogen (1920er und 1930er Jahre), hier aus dem Landkreis Regen (Vorder- und Rückseite).**

die entsprechenden Listen durchzugehen und zu exzerpieren. Bei z. T. über 500 Listen nimmt diese oft recht monotone und unangenehme Arbeit aber sehr viel Zeit in Anspruch. Sie ist auch kaum von Hilfskräften zu leisten, da das Material bereits beim Exzerpieren einer sprachwissenschaftlichen Interpretation bedarf. Man muss schon die bairischen Lautgesetze im Kopf haben, um sicher zu wissen, dass eine Lautform *bema* für „mit einem Fußboden belegen" zum Stichwort „bödnen" gehört. Da ein und dasselbe Wort aufgrund seiner unterschiedlichen Bedeutungen und Verwendungsweisen in den Listen häufig an mehreren Stellen abgefragt wurde, dauert das Exzerpieren oft wesentlich länger als das Verfassen des Artikels.

Nachdem die Wörterlisten und die Original-Fragebögen der verschiedenen Fragerunden mittlerweile komplett digital als Scans zur Verfügung stehen, können sie bei der Suche nach passenden Belegen bequem am Bildschirm durchgeblättert werden. Aktuell entwickelt die IT-Abteilung der Bayerischen Akademie der Wissenschaften ein Redaktionssystem, das es außerdem erlaubt, sich auf einzelne Fragen zu konzentrieren und die verschiedenen Antworten unmittelbar am Bildschirm miteinander zu vergleichen (Abb. 4). Mit einem einfachen Mausklick können sie markiert und z. B. nach unterschiedlichen Bedeutungen gruppiert werden. Damit wird die Arbeit des Vergleichens und Strukturierens der relevanten Belege sehr erleichtert und beschleunigt.

In diesem Redaktionssystem ist auch eine spezielle Suchmöglichkeit implementiert, die im Bestand der Dialektwörter unterschiedliche Schreib- und Lautungsvarianten findet, und zwar unabhängig von der Region, aus der diese stammen. Zudem können neben den bairischen Belegen auch diejenigen des Fränkischen und des Schwäbischen Wörterbuchs gefunden und auf diese Weise miteinander vernetzt werden. Das Fränkische Wörterbuch ist ebenfalls ein Projekt der Bayerischen Akademie der Wissenschaften.

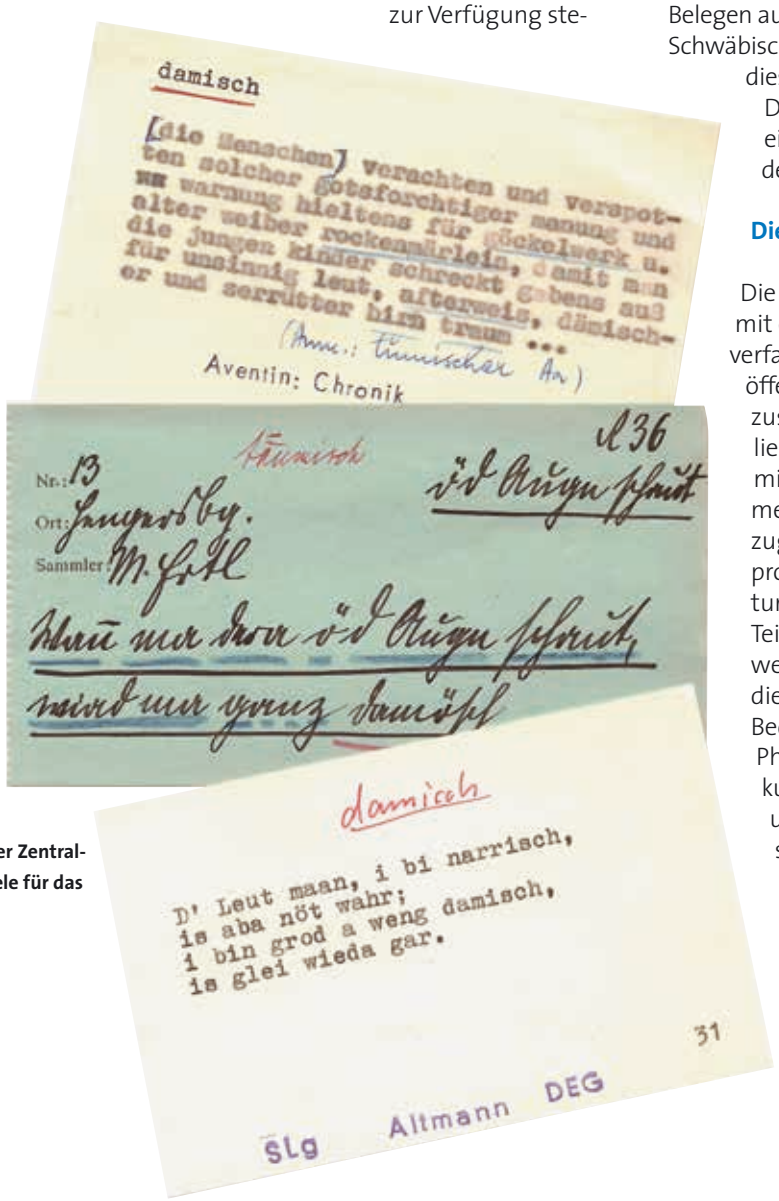



**Abb. 3: Zettel aus der Zentralkartei: Belegbeispiele für das Wort „damisch".**

## Die Wörterbuchartikel

Die Wörterbuchartikel werden bislang mit einem Textverarbeitungsprogramm verfasst. Da sie in Zukunft online veröffentlicht werden sollen, und zwar zusammen mit dem ihnen zugrundeliegenden Belegmaterial, wird gerade mit einer Erfassungsmethode experimentiert, die es ermöglicht, die Artikel zugleich in Print- und in Online-Form zu produzieren. Die Texte werden strukturiert erfasst, um ihre verschiedenen Teile explizit zu kennzeichnen. Damit werden die unterschiedlichen Angaben, die in einem Artikel stecken, wie z. B. die Bedeutungen, die Verbreitungsgebiete, Phraseologismen, sachliche und volkskundliche Hinweise, die Etymologie und die unterschiedlichen Lautformen, sichtbar und recherchierbar gemacht (Abb. 5).

Bislang sind vom Bayerischen Wörterbuch zwei Bände und drei Hefte des dritten Bands (von insgesamt 10 Bänden) in Buchform publiziert. Demnächst soll neben die

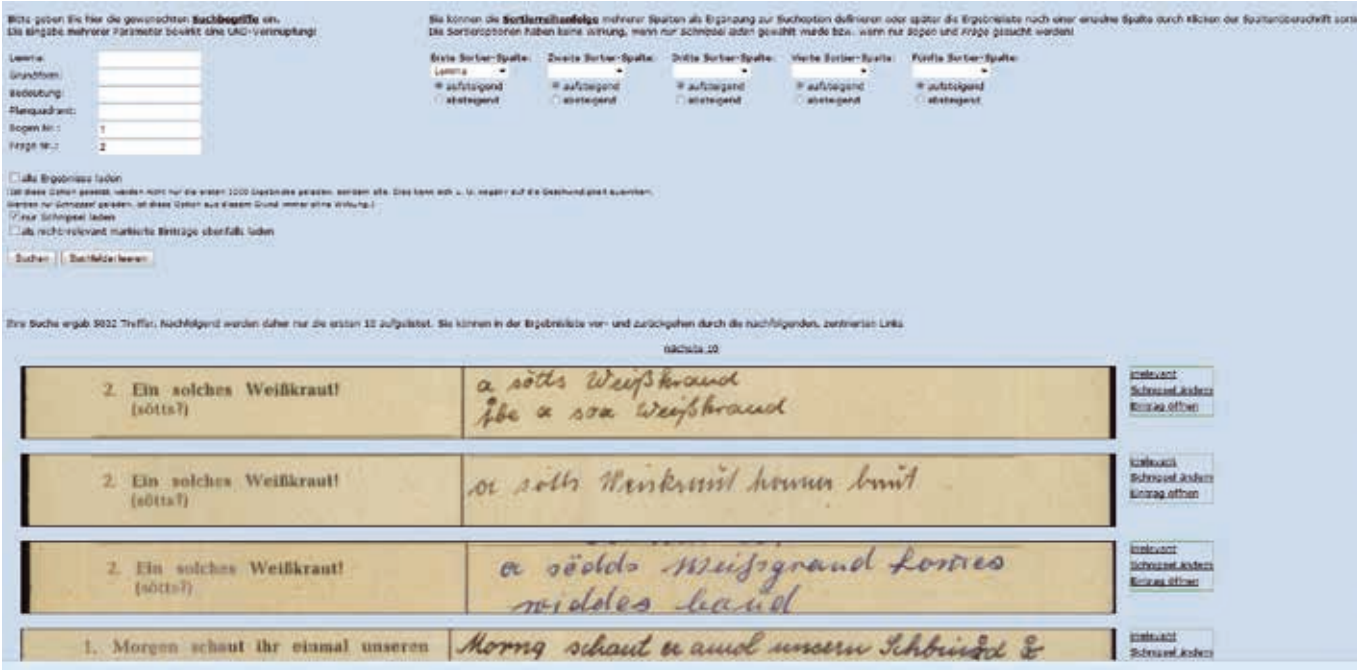



**Abb. 4: Screenshot des Abgleichs einzelner Fragen über mehrere Fragebögen hinweg.**

**Abb. 5: Screenshot der experimentellen Datenaufbereitung (XML).**

Printausgabe eine Online-Publikation treten. Natürlich werden auch die bisher erschienenen Bände online gestellt. Bis dahin ist es allerdings noch ein gutes Stück Weges: denn die nachträgliche Rückführung der Printfassung in eine Form, wie sie für die Online-Fassung neu erstellt wird, ist mit einigem Aufwand verbunden wie z. B. dem Einfügen der Links, der Auflösung von Abkürzungen und sonstigen Verdichtungen sowie der Nachstrukturierung der Inhalte in die Bestandteile, wie sie oben beschrieben sind.

Der interessierte Laie kann dann künftig sowohl in den Originalbelegen stöbern als auch sich durch die Interpretationen des Materials in Form der Artikel klicken. Es ist dann möglich, nach den Belegen einzelner Orte oder Regionen zu suchen oder etwa nach allen Belegen für einen bestimmten Begriff, ein Spiel, ein Werkzeug oder ein Tier, eine Pflanze.

Mit Hilfe der Digitalisierung wird dieses Wissen bewahrt und einer breiten Öffentlichkeit zugänglich gemacht. Die Online-Präsentation des Wörterbuchs kann jederzeit erweitert, aktualisiert und verbessert werden, da die Artikel korrigiert und erneut publiziert werden können, auch wenn sie schon geschrieben waren. Ein weiterer Vorteil ist, dass jeder Artikel publiziert werden kann, sobald er fertig ist. Man muss nicht mehr auf den Abschluss eines Heftes oder Bandes warten.

## Fazit

Mit dem digitalisierten Belegmaterial, dem Redaktionssystem für die Analyse des Materials sowie der Aufbereitung der Wörterbuchartikel und ihrer Basisbelege für eine Online-Präsentation ist das Bayerische Wörterbuch auf dem Weg zu einem modernen Wörterbuch und wissenschaftlichen Grundlagenwerk, das von interessierten Laien und Wissenschaftlern gleichermaßen sowohl digital als auch konventionell genutzt werden kann – unter Wahrung der tradierten hohen philologischen Qualitätsansprüche.

**DIE AUTORINNEN**

***Dr. Edith Burkhart-Funk ist wissenschaftliche Mitarbeiterin des Bayerischen Wörterbuchs, das an der Bayerischen Akademie der Wissenschaften entsteht.***
***Ursula Welsch ist in der IT-Abteilung der Akademie zuständig für Digital Humanities.***