

# **Unterstützung von Forschungsprozessen in einem internationalen Forschungsprojekt: Variantengrammatik des Standarddeutschen**

Gunter Vasold, Zentrum für Informationsmodellierung – Austrian Centre for Digital Humanities (ZIM – ACDH), Universität Graz, Österreich

## ***Einleitung***

In diesem Poster wird anhand eines konkreten Forschungsprojekts eine verteilte Arbeitsumgebung vorgestellt, die es kollaborativ arbeitenden ForscherInnen ermöglicht, (Zwischen)Ergebnisse zu erfassen, iterativ anzureichern, zu verbessern und die geleistete Arbeit zu dokumentieren bzw. dokumentieren zu lassen. Bei der Entwicklung wurde besonderes Augenmerk auf die Unterstützung typisch geisteswissenschaftliche, d.h. hermeneutisch ausgerichteter Forschungsprozesse gelegt.

## ***Das Projekt***

Das Forschungsprojekt "Variantengrammatik des Standarddeutschen" ist ein durch den Schweizerischen Nationalfonds (SNF), die DFG und den österreichischen FWF gefördertes Projekt. Forschergruppen in Salzburg (ursprünglich Augsburg), Graz und Zürich untersuchen gemeinsam mit Kooperationspartnern in Liechtenstein, Belgien, Luxemburg und Südtirol systematisch nationale und regionale Unterschiede in der Grammatik der deutschen Standardsprache. Methodisch stützt sich das Projekt auf die Analyse eines Korpus mit Zeitungstexten aus allen deutschsprachigen Ländern, das mehr als 300 Millionen Tokens enthält.

## ***Herausforderungen***

Die geographische Verteilung der beteiligten ForscherInnen auf mehrere Länder erforderte Hilfsmittel zur Unterstützung kollaborativer Arbeitsweisen, die größtenteils mit bestehenden Werkzeugen abgedeckt werden konnten. So wurde der gemeinsame Zugriff auf das Korpus durch ein von Semtracks bereitgestelltes Webinterface zur Corpus Workbench (CWB) gelöst. Die projektinterne Kommunikation erfolgte in bewährter Weise über E-Mail und Skype; auch für die Literaturverwaltung standen zu Projektbeginn entsprechende netzwerkfähige Produkte zur Verfügung. Alleine für die Verzeichnung und Verwaltung bereits untersuchter grammatischer Varianten und die projektinterne Bereitstellung von Zwischenergebnissen fehlte es an fertigen Lösungen. Ein erster Ansatz, nämlich die benötigten Daten in einem via Google-Docs zentral bereitgestellten Tabellendokument zu speichern, erwies sich mit wachsender Datenmenge bald als unpraktikabel, nicht zuletzt, weil die zu speichernden Daten nur unzureichend in eine zweidimensionale Matrix abbildbar waren.

## ***Die Lösung***

Anfang 2012 wurde das Zentrum für Informationsmodellierung – Austrian Centre for Digital Humanities (ZIM – ACDH) mit der Lösung dieses Problems betraut. Es wurde auf Basis eines zuvor für ein anderes Projekt entstandenen Frameworks eine Datenbank mit Web-Interface entwickelt, über das alle ProjektmitarbeiterInnen die von ihnen untersuchten Varianten zentral erfassen und schrittweise weiter bearbeiten können. Dazu gehört die Dokumentation aller Korpusabfragen und daraus resultierender Ergebnisse ebenso wie das Speichern und Kommentieren statistischer Auswertungen, sowie das Zitieren von Belegstellen und von variantenspezifischer Sekundärliteratur. Jede Variante

kann grammatisch (Wortbildung, Flexion, Genus, Valenz/Rektion; teilweise mit mehreren Kategorieebenen) und inhaltlich (Einzelvariante, Variantenphänomen) kategorisiert und zusätzlich frei beschlagwortet werden.

Besonderes Augenmerk wurde dabei dem prozesshaften Voranschreiten der Arbeit geschenkt. Jeder einzelne Speichervorgang wird unter Berücksichtigung von BearbeiterIn, Zeitstempel, Bearbeitungsstatus und einer allfälligen Anmerkung protokolliert. Diese Information steht bei Bedarf im Kontext jeder einzelnen Variante zur Verfügung. Zusätzlich kann gezielt nach diesen Informationen gesucht werden. Die Prozessdaten bilden das Gerüst eines „Laborbuches“ und sind wesentlich für die Zuschreibung geleisteter Beiträge. Sie sind auch nützlich für das Projektmanagement, etwa zur Identifizierung von Varianten mit Bearbeitungsbedarf oder als Maßzahlen für die Projektplanung.

Ein formulargesteuertes Suchwerkzeug erlaubt kombinierte Abfragen über alle Datenbankfelder. Die Suche liefert eine tabellarische Überblickdarstellung, aus der heraus einzelne Varianten im Detail angesehen oder direkt zum Bearbeiten geöffnet oder gelöscht werden können. Als arbeitstechnisch hilfreich hat sich die Möglichkeit erwiesen, im Suchresultat bestimmte Varianten zur genaueren Analyse auswählen zu können. Die gewählten Einträge können in einer Vorschaufunktion rasch durchgesehen und aus dieser heraus wieder deselektiert oder nachbearbeitet werden. Aus erkenntnistheoretischer Sicht ergibt sich daraus die Möglichkeit, aus im System vorhandenen Daten Erkenntnisse abzuleiten und diese in Form adaptierter oder neuer Daten unmittelbar wieder in das System zurückfließen zu lassen. Aus dem Suchergebnis heraus können gewählte Varianten auch in verschiedene Formate (Excel, PDF, HTML/XML) exportiert werden, was etwa ein Korrekturlesen auf Papier oder die Nutzung spezieller Analysesoftware ermöglicht.

Da die Datenbank nicht nur den Forschungsprozess unterstützen und dokumentieren soll, sondern vor allem die Datengrundlage für ein zum Projektende geplantes Handbuch darstellt, wurde zusätzlich eine Registersicht als Vorschau und Ausgangspunkt für die noch zu erstellende Druckvorlage implementiert.

## **Résumé**

Das Beispiel der Forschungsdatenbank zur „Variantengrammatik des Standarddeutschen“ zeigt, dass mit geringem Aufwand individuelle VRE-artige Arbeitsumgebungen realisiert werden können, die auf pragmatische Weise die Verwaltung von Forschungsergebnissen im Sinne von (vorläufigen) Erkenntnissen erleichtern. Eine spezifisch geisteswissenschaftliche, d.h. hermeneutische Herangehensweise im Sinne des von Gadamer als hermeneutischer Zirkel beschriebenen Erkenntnisprozesses erfordert, dass in einem solchen System nicht nur Ergebnisse, sondern auch deren Fortentwicklung und Veränderung abbildbar sind und entsprechende Forschungsprozesse unterstützt werden. Es geht dabei weniger um die Vorgabe genau spezifizierter Workflows, sondern um die flexible Unterstützung iterativer Prozesse und deren Dokumentation.

## **Literatur**

<http://variantengrammatik.net/>

Christa Dürscheid, Stephan Elspaß und Arne Ziegler: Grammatische Variabilität im Gebrauchsstandard – das Projekt »Variantengrammatik des Standarddeutschen«. In: Marek Konopka et al. (Hgg.), Grammar und Corpora 2009, Tübingen 2011, S. 123–140.

Hans-Georg Gadamer: Vom Zirkel des Verstehens. In: Hans-Georg Gadamer, Gesammelte Werke 2, Tübingen 1993, S. 57–65 (Zuerst erschienen in: Festschrift für Martin Heidegger zum 70. Geburtstag, Pfullingen 1959).

Jan Potthoff und Sebastian Rieger: Elektronisches Laborbuch: Beweiswerterhaltung und Langzeitarchivierung in der Forschung. In: Silke Schomburg, Claus Leggewie, Henning Lobin und Cornelius Puschmann (Hgg.) Digitale Wissenschaft, Köln 2011, S. 149–156.

Thomas Süptitz, Stephan J. J. Weis und Torsten Eymann: Was müssen Virtual Research Environments leisten? – Ein Literaturreview zu den funktionalen und nichtfunktionalen Anforderungen. In: Proceedings of the 11th International Conference on Wirtschaftsinformatik (WI2013), Leipzig 2013, S. 327–341.