

Digitale Lexikographie in den altanatolischen Sprachen

Im Zuge der fortschreitenden Digitalisierung der modernen geisteswissenschaftlichen Forschungs- und Arbeitsfelder ergeben sich auch im Bereich der Lexikographie völlig neue Möglichkeiten für die Konzeption, Gestaltung und Umsetzung digitaler Lexikonprojekte. Erfolgt die Planung von Anfang an als digitales Verbundprojekt, so lassen ehemals getrennte Bereiche wie vollständig annotierte Textkorpora, Bibliographien und die konkreten Lexikoneinträge effektiv zu einem dynamischen Ganzen verbinden. Der Vortrag möchte eines dieser digitalen lexikographischen Verbundprojekte vorstellen, welches seit Oktober 2014 als DFG-Projekt an den Universitäten München und Marburg erstellt und technisch von der IT-Gruppe Geisteswissenschaften (LMU) konzipiert und entwickelt wird: Das „Digital Philological-Etymological Dictionary of the Minor Ancient Anatolian Languages“, kurz „eDiAna“ (Miller et al. 2017).

Das Lexikon, fachwissenschaftlich im philologisch-linguistischen Überschneidungsfeld der Altanatolistik, Hetthitologie und Indogermanistik angesiedelt, hat die sogenannten altanatolischen Kleinkorpussprachen zum Gegenstand, wobei die Erfassung des Formenbestandes sowie die Rekonstruktion der Etymologie im Zentrum des philologisch-sprachwissenschaftlichen Interesses steht. Neben den konkreten Lexikoneinträgen, von denen zum aktuellen Zeitpunkt 36 öffentlich zugänglich sind, wird ein funktional sowie kategorial vollständig annotiertes Textkorpus aller erhaltenen Textzeugnisse erstellt, welches dynamisch mit dem Kernlexikon verbunden wird. Zum aktuellen Zeitpunkt umfasst das Korpus 929 Texte (ca. 38.000 Wörter) in den Sprachen Keilschrift- und Hieroglyphenluwisch, Pisidisch, Palaisch, Milyisch, Lydisch sowie Lykisch. Auch an einer umfangreichen bibliographischen Datenbank, mit der Möglichkeit, Excerpte zu den einzelnen bibliographischen Einträgen abzuspeichern und dynamisch zu suchen, wird fortlaufend gearbeitet. Die bibliographische Datenbank umfasst dabei aktuell etwa 3.000 Einträge und über 26.000 Excerpt-Datensätze.

Technisch realisiert wird das Lexikon als SQL-XML-Hybrid mit einem dynamischen Web-Interface (PHP, JavaScript, MySQL) im Kern, wobei Lexikon, Bibliographie und Korpus sowohl getrennt voneinander als auch in einer integrierten Version bedient werden können. Lexikoneinträge werden durch das Web-Interface im Moment des Aufrufes on the fly aus manuell verfassten sowie automatisch generierten Lexikon-Bestandteilen komponiert, linearisiert und zu den sichtbaren Lexikoneinträgen hin formatiert. Diese Technik erlaubt sowohl eine hohe Performance des Interfaces als auch eine Anpassung der Lexikoneinträge an das individuelle Nutzungsverhalten der Besucher, welche die dargestellte Information an ihre jeweiligen Bedürfnisse anpassen können. Die Entwicklung vielseitiger und effizienter Suchalgorithmen für die Abfrage der verschiedenen Datenbanken (Textkorpora, Lexika und Bibliographie/Excerpte) ist ebenfalls zentraler technischer Projektbestandteil: Die teils komplexen Transkriptions- und Formatierungsrichtlinien für die Sprachdaten erschweren die Benutzung sowie die Zugänglichkeit vergleichbarer Lexika erheblich, eDiAna hingegen dämmt diese Probleme durch umfangreiche automatische Hilfe- und Vorschlagsfunktionen erheblich ein, wodurch es viele der fachwissenschaftlichen Erkenntnisse erst für einen größeren Interessentenkreis öffnet. Die Ausgabe im Web-Interface wird ergänzt um die Möglichkeit einer Print-on-Demand Funktion (LaTeX), eine API-Schnittstelle (JSON) für die Interaktion mit verwandten Forschungsprojekten sowie die Möglichkeit zum Export in gewünschte XML-Formate (zum Beispiel TEI-konforme Strukturen).

Im Rahmen des Vortrages wird zunächst auf inhaltliche Aspekte des konkreten Projektes eingegangen sowie auf die Veränderungen, die sich bei der Realisierung des Lexikons als Digital Humanities Projekt im Gegensatz zu vergleichbaren analogen Projekte ergeben. Im Anschluss daran

werden die Herausforderungen für die technische Konzeption und Umsetzung des Projektes erläutert. Der Präsentation des dynamischen Web-Interfaces wird im Zuge dessen ebenfalls kurz Raum gegeben.

Literaturverzeichnis:

Miller, Jared & Rieken, Elisabeth & Hackstein, Olav & Yakubovich, Ilya (eds.) 2017: eDiAna.
<https://www.ediana.gwi.uni-muenchen.de>