

Digital Dictionary of the German Language: the evolution of its online platform since 2003

In 2003, the Digital Dictionary of the German Language (DWDS) started as an online platform under the domain www.dwds.de. The resources of the site at that time were the Dictionary of Contemporary German (WDG, [1]) compiled at the Academy of Sciences of the GDR, and a balanced reference corpus of the German language of the 20th century.

However, the aim of the DWDS was not only to digitally preserve the legacy resources, but to build a comprehensive, reliable and up-to-date general language dictionary of contemporary German [2]. To this end, a technical and a lexicographic work plan was conceived with the WDG as the lexicographic basis for the new dictionary. Two phases were foreseen: a planning and an entry compilation phase. In the planning phase the technical prerequisites had to be implemented, including the dictionary writing system, a corpus search interface and a corpus management system. At the same time the lexicographic process of revising the legacy dictionary had to be specified. Moreover, a corpus-corpus based candidate list of new words to be included had to be identified. It was only in 2013 with a more comfortable funding by the Academies' Programme when the entry compilation could start on a larger scale. In 2015, it was possible to license some 70,000 entries from Duden [3], which are complementary to the vocabulary of the WDG. Since 2019, the DWDS has been part of the newly installed Centre for Digital Lexicography of the German Language [4]. Within this framework, the lexicographic programme was further expanded. In two further sub-projects, the corpus-based description of the most important common multi-word expressions in German [5] and the revision of all headwords marked as regional according to an areal classification of the German-speaking regions [6] are being carried out. For the latter purpose, large regional corpora were created, which can also be searched on the DWDS website together with a map display. The entire DWDS system currently comprises approx. 240,000 contemporary dictionary articles and a corpus base of almost 50 billion text words.

A specific task was dedicated to increase the outreach of the website. Through targeted SEO measures in 2020, the number of page views of the dwds.de website grew from around 2.5 million (January 2019) to 9.7 million (January 2023). Individual dictionary articles, including those corresponding to trends, thus quickly reach monthly page views of 10,000 or more. The monitoring of the page views also orients the priorities in the revision of outdated dictionary

articles or the creation of dictionary articles for neologisms. Complementary to these measures, the DWDS blog went online at the end of 2022 with the aim of thematically summarising individual dictionary articles and thus creating references between the individual articles. The blog is divided into different sections: Trend words and neologisms, word histories, regionalisms, multi-word expressions and orthography.

References:

[1]: Klappenbach, Ruth; Malige-Klappenbach, Helene (1980). Das Wörterbuch der deutschen Gegenwartssprache. Entstehung, Werdegang, Vollendung. In: Abraham, Werner: Studien zur modernen deutschen Lexikographie. Amsterdam. John Benjamins. p. 3–58.

[2]: Klein, Wolfgang, Geyken Alexander (2010). Das Digitale Wörterbuch der Deutschen Sprache (DWDS). *Lexicographica* 26, p. 79–96.

[3] Duden, Das große Wörterbuch der deutschen Sprache, Mannheim 1999.

[4] <http://www.zdl.org>

[5] Ermakova, Maria; Geyken Alexander; Lemnitzer, Lothar; Roll, Bernhard (2022). Integration of multiword expressions into the Digital Dictionary of German Language (DWDS). Towards a lexicographic representation of phraseological variation. In: Annette Klosa-Kückelhaus, Stefan Engelberg, Christine Möhrs, Petra Storjohann (eds.): *Dictionaries and Society. Proc. of the XX EURALEX international Congress*. Mannheim, 12–16 Juli 2022, p. 851–861.

[6] Nolda, Andreas; Barbaresi, Adrien, Geyken, Alexander (2023). Korpora für die lexikografische Beschreibung diatopischer Variation in der deutschen Standardsprache: Das ZDL-Regionalkorpus und das Webmonitor-Korpus. In: Deppermann, Arnulf u. a. (Hrsg.): *Korpora in der germanistischen Sprachwissenschaft: Mündlich, schriftlich, multimedial* (= Jahrbuch des Leibniz-Instituts für Deutsche Sprache 2022). Berlin 2023, p. 29–52