

O Corpus do Português

45 Mio Wörter aus unterschiedlichen Textsorten, auch aus gesprochener Sprache, vom 14. bis ins 20. Jh.

Sprache	Portugiesisch
Sprachliche Realisierung	schriftlich, mündlich
Umfang	ca. 45 Mio. Wörter, 57.000 Texte
Medium	diverse Textsorten (Academic, News, Fiction, Oral) vom 14. bis zum 20. Jahrhundert
Geographischer Ursprung	Portugal, Brasilien
Zeitliche Einordnung	14.-20. Jh.
Form der Daten	digitalisierte, online durchsuchbare Texte und Transkripte
Format	HTML
Annotation	lemmatisiert, part-of-speech-annotiert
Mögliche Suchabfragen	Suche nach Wörtern, Wortfolgen, Lemmata, Wortarten, Kollokationen und Key Words in Context (KWIC) möglich. Zudem kann nach Genre, Land und Jahrhundert selektiert werden.
Quelle /Herausgeber	Mark Davies (Brigham Young University), Michael J. Ferreira (Georgetown University)
Nutzungsvoraussetzungen	Zugang frei
Link	http://www.corpusdoportugues.org/