

ESCOW14 (Spanish web corpus)

Umfangreiches Korpus zum Web-Spanisch

Sprache	Spanisch
Varietät	nicht erfasst
Sprachliche Realisierung	schriftlich/web
Umfang	Basiskorpus ~7,3 Mrd. Token/~330 Mio. Sätze; offen zugängliches Subkorpus ~3,7 Mrd. Token/~150 Mio. Sätze
Universität	Freie Universität Berlin, Deutschland
Form der Daten	XML mit inline VRT, UTF-8
Format	Colibri2 User-Interface, Export möglich
Annotation	Wort/Token, Lemma, Morphologie (FreeLing automatische Annotation)
Quelle/Herausgeber	Roland Schäfer, 2015
Link	https://webcorpora.org/