

Korpus-Übersicht des SLDR (Speech and Language Data Repository)

Das französische [SLDR/ORTOLANG](#) (Speech and Language Data Repository/Outils et Ressources pour un Traitement Optimisé de la LANGue) stellt, in Zusammenarbeit mit dem [CNRTL](#) (Centre National des Ressources Textuelles et lexicales) eine Online-Plattform zur Verfügung, in der Korpora und andere Sprachdaten gesammelt, dokumentiert und der Forschungsgemeinschaft zugänglich gemacht werden.

Auf der Homepage des SLDR ist eine umfangreiche Liste mit Korpora zum Französischen, aber auch zu diversen anderen europäischen (Englisch, Deutsch, Polnisch, Italienisch, Spanisch) und nicht-europäischen Sprachen (Hindi, Arabisch, Chinesisch, Quechua) zu finden. Interessant sind auch die Sprachdaten zu Regionalsprachen, z.B. zum Okzitanischen und Frankoprovenzalischen.

Die Korpora sind detailliert in Bezug auf Forschungsbereich, Art und Format der Daten, Umfang, Erhebungsjahr, Metadaten und Zugangsvoraussetzungen beschrieben.

Der Link zur Liste

http://sldr.org/SLDR_data/Disk0/doc/Corpora-fr.html