

Back to [Linguistic Research Infrastructure](#) (LiRI)

Corpora & Assistive Technology

The [Language Technology](#) group has expertise in handling various types of corpora. We are building tailor-made applications to explore large and structurally complex collections of language data. In particular, we are competent in:

- The design of databases to hold application-relevant data
- Generating interactive visualizations
- Efficiently querying large data collections (in particular corpora)
- Anonymisation of large data sets
- Data crawling/scraping and processing of web sources, batch download of documents
- Data extraction and conversion

From:

<https://liri.linguistik.uzh.ch/wiki/> - **LiRI Wiki**

Permanent link:

<https://liri.linguistik.uzh.ch/wiki/langtech/corpora?rev=1675169495>

Last update: **2023/01/31 12:51**

