



Swissdox@LiRI

Swissdox@LiRI is a tool that allows you to retrieve large quantities of Swiss media data for research purposes.

LiRI cooperates with SMD (Schweizer Mediendatenbank AG) to make the Swissdox database easily accessible to researchers. The Swissdox@LiRI database includes about 23 million media articles from a wide range of Swiss media sources covering many decades, and is updated daily with about 5'000 to 6'000 new articles mainly from the German and French speaking parts of Switzerland.

Features

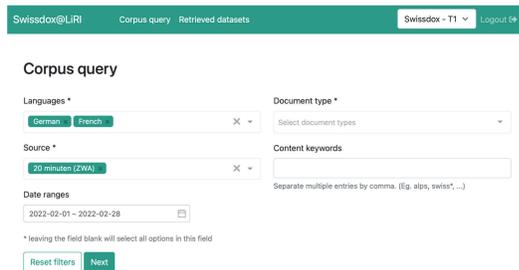
With Swissdox@LiRI you can retrieve articles by filtering for language, time interval, keywords, sources and document types. More complex query options are offered by demand. After submitting a query, Swissdox@LiRI compiles the corresponding dataset and notifies the user when it becomes available.

The data can then be further processed by the user as desired, e.g. with programming languages such as Python or R, or to apply methods such as topic modelling, word embeddings, etc.

Web Interface

The user can define a query in order to retrieve articles matching his particular research interests. A list of attributes that the database can be filtered for:

- Languages
- Document type
- Source
- Content keywords
- Time intervals



Result datasets are available in a compressed TSV (tab-separated values) format. An API will be provided soon.

Access

Access to the database is free for all members of the supporting institutions: University of **Basel**, University of **Bern**, **ETH Zurich**, **ZHAW Zurich**, University of Applied Sciences, University of **Zurich** (as of June 2022). Various subscription types are offered for other users.

If your institution is not yet a supporter, make it aware of the offer and express your interest in using Swissdox@LiRI.

Also access for **commercial purposes** are offered. Please contact Swissdox (<https://swissdox.ch>) if you are interested.

Background

Swissdox@LiRI has been initiated by Prof. Dr. Noah Bubenhofer, Prof. Dr. Fabrizio Gilardi (UZH) and Roberto Nespeca (SMD) and is funded by the University of Zurich UZH (Technology Platform Commission) and the following supporters: Zurich University of Applied Sciences (Department of Applied Linguistics), University Basel / University Library Basel, ETHZ Library, University Library Bern.



Scan for more information



Winner of SwissNLP Award 2022

more than 23 million media articles

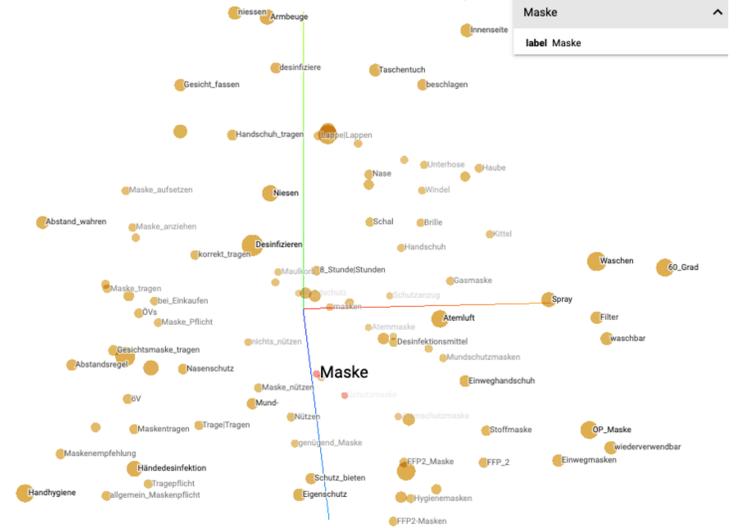
5000 to 6000 new articles daily

250+ sources

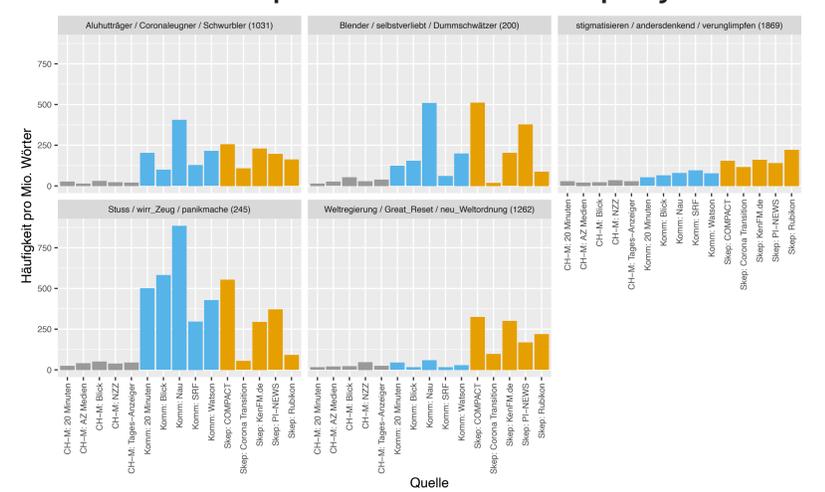
CH Media, NZZ media group, Ringier, TX Group, SRF/SRG, Wochenzeitung etc.

Example

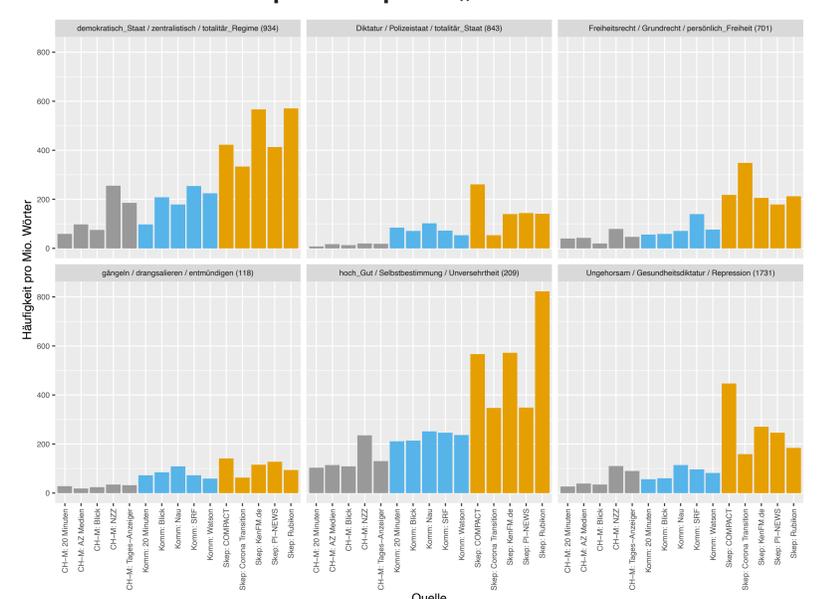
Semantic Space of „Maske“ („mask“) during COVID-19 Pandemic



Clustered Semantic Spaces: Distribution of Conspiracy Stories



Clustered Semantic Spaces: Topoi of „Totalitarian State“



Weiterführende Informationen:

Bubenhofer, Noah (2022): Exploration semantischer Räume im Corona-Diskurs. In: Kämper, Heidrun / Plewnia, Albrecht: Sprache in Politik und Gesellschaft. Jahrbuch des Instituts für Deutsche Sprache, 2021. Mannheim: 2022, S. 197-216. DOI: 10.1515/9783110774306-013.