



UpLORD identified several gaps

in the current situation in Switzerland in the context of the requirements of Open Science, on the one hand, and of that of more challenging data sets (e.g., sensitive or copyright protected data), on the other hand, and aims to address them from different perspectives:



1. Improve Platform Interoperability and Corpora Quality
 Develop new software applications to improve the quality of existing corpora and enhance the interoperability of their platforms. Improve discovery, access, integration, usability and reusability of corpora according to FAIR principles, as well as simplify re-formatting, assembling, and standardizing the (meta)data:



More on VIAN-DH
 VIAN-DH is a web application for viewing, querying, automatic annotation and analysis of multimodal interaction and language use in videos. The project aims at facilitating creation and processing of multi-modal corpora by automating annotation and transcription tasks.

The Linguistic Corpus Platform **LCP** is a system for handling and querying large corpora as a replacement for terminal-based corpus software. LCP can serve the interests of both newcomers to corpus linguistics and experienced users.



Both applications are developed by LiRI – the Linguistic Research Infrastructure at UZH.

2. Increase FAIR Adherence of SWISSUbase
 Upgrade workflows and interoperability of existing infrastructure services by specifying and adapting the already implemented modular linguistic metadata schema of LaRS/SWISSUbase to the needs that are not currently satisfied. To increase the usability of SWISS-Ubase, it should also be possible to provide a workflow via an API.



More on LaRS@SWISSUbase

3. Standardize Metadata
 National CLARIN-CH working groups develop specific metadata for various types of linguistic data (in, e.g., sociolinguistics, experimental psycholinguistics, neurolinguistics, conversational analysis, lexicography, computational linguistics and including multimodal data) that satisfy the FAIR principles.



More on CLARIN-CH working groups

3. Document and Share Best Practices
 Document showcases, best practices and tutorials for data management according for FAIR principles (e.g., planning and writing data management plans, using sustainable file formats, setting solid file-naming conventions, finding data storage backup schemes, creating informative metadata) and raise awareness about ORD practices in the context of teaching, research and publishing. Special attention given to the management of personal and sensitive data.

UpLORD is a swissuniversities ORD-funded 2-year project (2023-2024) hosted by UZH and supported by the Zurich University Library and the CLARIN-CH consortium.

Since 2018, a consortium of partners has been working on building a national ecosystem of infrastructures, which covers the whole linguistic data lifecycle according to ORD requirements (FAIR principles) from data generating, processing and analyzing to data sharing and archiving. In the long-term, this project will significantly contribute to a strong foundation for a sustainable ORD strategy for linguistic data in Switzerland.

UpLORD

Upgrading the Linguistic ORD-ecosystem

Are you a member of the Swiss scientific community working with language resources and would like to get involved in the project? Please contact cristina.grisot@uzh.ch or read more here:

