

VIAN-DH

VIAN-DH is web-based software for multimodal corpora. It is a part of LiRI Corpus Platform (LCP), specializing in audio-visual language resources. The main components of VIAN-DH are:

- Corpus viewer with media player and a timeline preview of annotations
- Multimodal corpus query, including a specific query language and multimodal concordance preview
- Automatic annotation pipeline for speech recognition and pose estimation

Corpora stored in VIAN-DH containing different types of annotations. To name some of them, annotations can include speech transcript, part-of-speech annotation and other kinds of labels marking embodied communication, such as gestures or facial expressions, or document segmentation such as scene for film, turn for conversations, or experiment phase for various experimental settings. This is not an exhaustive list, and the scope of annotations varies from one corpus of another.

Data structure

VIAN-DH focuses on corpora based on video or audio source materials. They can include transcripts of speech and further text processing or tagging, annotation of embodied communication or interaction, annotation of persons or objects, various types of segmentation, depending to the corpus.

VIAN-DH allows for different types of alignment. Primary types are:

- time-frame alignment for any annotations linked to the time or frame stream
- text-alignment referring to continuous stream of characters for annotations referring to textual content, such as part-of-speech, named-entity-recognition, etc.

Any annotation can have multiple alignment to both of these streams by marking its start and end position/index, e.g., a sentence can have a start and end time and a start and end character. See the Fig. 1 below for a graphical representation of data being linked to multiple alignment streams. Note how each annotation track (gestures, utterances, tokens, named entities and constituents) are linked to both time and text stream (utterances), or either text stream (tokens, named entities and constituents), or time stream (gestures). In case of video, time stream also refers to frame stream - since each video has a specific frame rate.

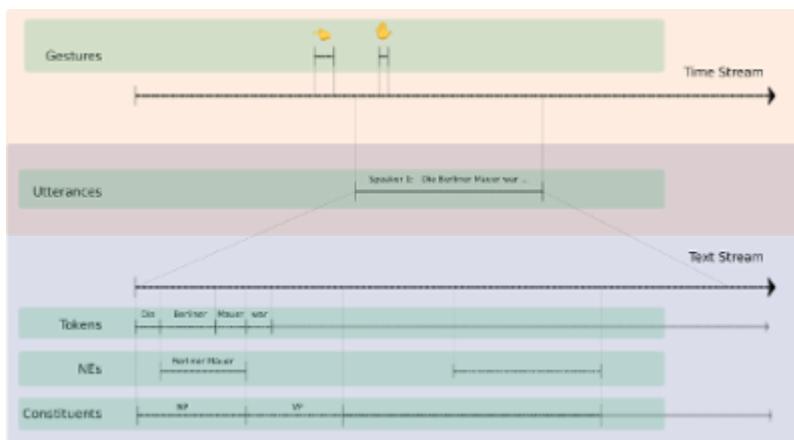


Fig 1. Data representation on virtual alignment streams

Access

VIAN-DH web-app is currently in beta-testing mode. The test application can be accessed via this [link](#).

The application is accessed through a Shibboleth login form, see Fig 2.

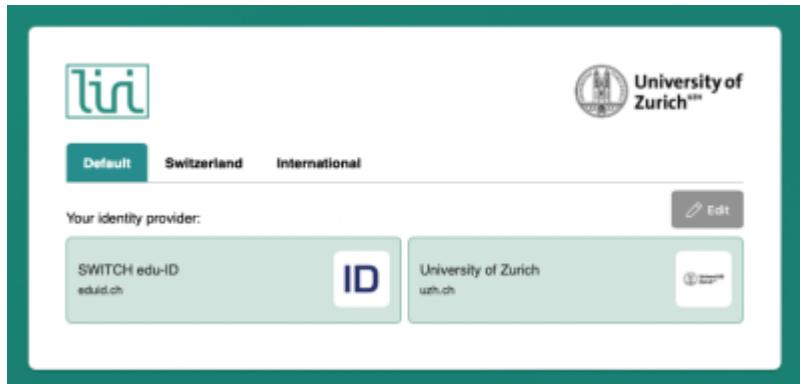


Fig 2. VIAN-DH login form

The application can be accessed by anyone with a EduID login. If you are not affiliated to the University of Zurich, please use the EduID access (left on Fig 2.) and do not try to access by searching for your home institution.

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



Permanent link:
<https://liri.linguistik.uzh.ch/wiki/langtech/vian-dh>

Last update: **2023/07/13 20:24**