

Layers of information

WhatsApp messages are built up in a hierarchy: a chat contains messages that contain tokens that contain characters. A corpus of WhatsApp chats should allow for all these layers to be queried. Additionally, there is meta-data about the chats (e.g. number of messages) and about the messages (e.g. the timestamp when it was written) and about the informant (e.g. his/her age) and about the tokens (e.g. part of speech). This makes our corpus a rather challenging and complex endeavor.

From:

<https://whatsup.linguistik.uzh.ch/> -

Permanent link:

https://whatsup.linguistik.uzh.ch/02_browsing/02_layers?rev=1573030033

Last update: **2022/06/27 07:21**

