1.2.2 Data without permission

During the data collection process, not all communication partners in all the chats gave their permission for their texts to be used. In chats where we did not get the permission of all participants, we still used the messages for which we had the permission and disguised those without.

In order to keep the chats more or less readable, we decided to give some information on the length of disguised messages, which are accordingly marked as e.g. `redactedQ12tokens55characters`. This annotation lets you know that the original message consisted of 55 characters that formed 12 tokens.

From:
https://whatsup.linguistik.uzh.ch/ -

Permanent link:
https://whatsup.linguistik.uzh.ch/01_corpus/02_preprocessing/02_without_permission

Last update: **2022/06/27 09:21**