

Annotation Tool for Dataset Creation

Patrick Giedemann and **Pius von Däniken** and **Jan Milan Deriu**
gied@zhaw.ch

Abstract

The “Annotation Tool for Dataset Creation”, developed as part of the HAMiSoN project (Holistic Approach to Counter Disinformation), represents a significant advancement in the fight against the spread of disinformation on social networks. This web application is specifically designed to facilitate the annotation of transcripts derived from short videos found on social media platforms, which are suspected of containing disinformation. Aimed at empowering annotators, the tool is equipped with a suite of interactive features that simplify the process of navigating, annotating, and understanding the content that requires attention.

The primary objective of the tool is to enhance the annotation process, making it more intuitive and efficient for users. By providing real-time insights into the dataset creation process and the characteristics of the dataset itself, the tool not only enhances the productivity of annotators but also contributes valuable data for researchers involved in the HAMiSoN project. Its ability to handle multimodal data, including video, text, and audio, underscores the comprehensive approach taken by the developers to address the multifaceted nature of disinformation.

Moreover, the “Annotation Tool for Dataset Creation” has been thoughtfully designed with scalability and adaptability in mind, allowing for its application beyond the initial scope of the HAMiSoN project. This flexibility makes it a valuable resource for future research endeavors aimed at curating datasets across various domains. As such, the tool not only serves the immediate needs of the HAMiSoN project but also lays the groundwork for broader applications in dataset creation.