



D:3.3.3 b. Multimodal Analysis Tools for Understanding Livestreams

One of the biggest transformations across contemporary media cultures is the proliferation of streamed audio-visual content with attached textual communication features. Online streams of both amateurs and professionals are changing the ways through which people engage with the world. While online streams have been studied increasingly over the past years, the methods employed are typically qualitative and have focused on limited samples. There is a lack of means for processing larger amounts of content. Thus, we need tools to study online streams quantitatively, for instance, to enable even thousands of hours of content in genre-profiling analyses.

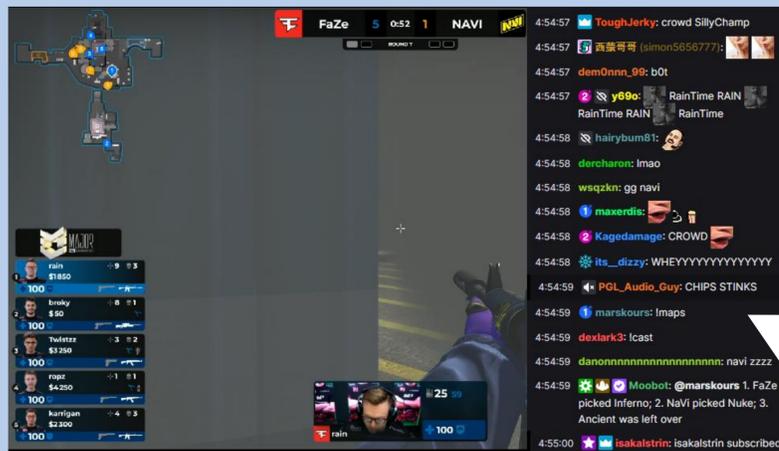
MULTIMODAL LARGE LANGUAGE MODELS FOR LIVESTREAM VIDEO UNDERSTANDING

Twitch streams offer a rich opportunity to analyze data via machine learning, and new methods will be developed for this. Our approach leverages multimodal large language models (LLMs) to understand both the visual and textual content of streams, enabling a deeper understanding of the complex interactions between viewers and streamers.

Tools will include comprehensive analysis of visual content by combining visual and textual information for comprehensive video via multimodal LLMs. These models will be trained to understand the spatial and temporal aspects of livestream videos, enabling understanding of ongoing events, peak engagement moments, and recognition of scene transitions to provide the overall narrative and context of the stream.

TEAM

Work in this WP will be conducted by the CoE in Game Culture Studies JYU team, which has already done significant work on analyzing discussions in game streams, and the TAU machine learning team of the CoE having worked on gaming social media analytics.



Example of Twitch gameplay stream and the accompanying chat

TWITCH.TV

- Started in 2011, currently owned by Amazon
- 7 million broadcasters monthly
- 26 million daily active users, 2.3 million average concurrent users
- 93 000 concurrent streams
- Categories: Games, Esports, IRL, Sports, Music, Creative

Twitch chat

- Participation requires logging in
- Sub-cultures based on games and genres, built around channels and casters
- Sub-culture specific communication
- Emotes, copy-pasta, cheers...
- Constantly evolving
- Toxic behavior and controversial ideological views

