

Title: Examining Mobs Images Using Vision-Language Models



PRESENTER
Ariana Mondiri

Background

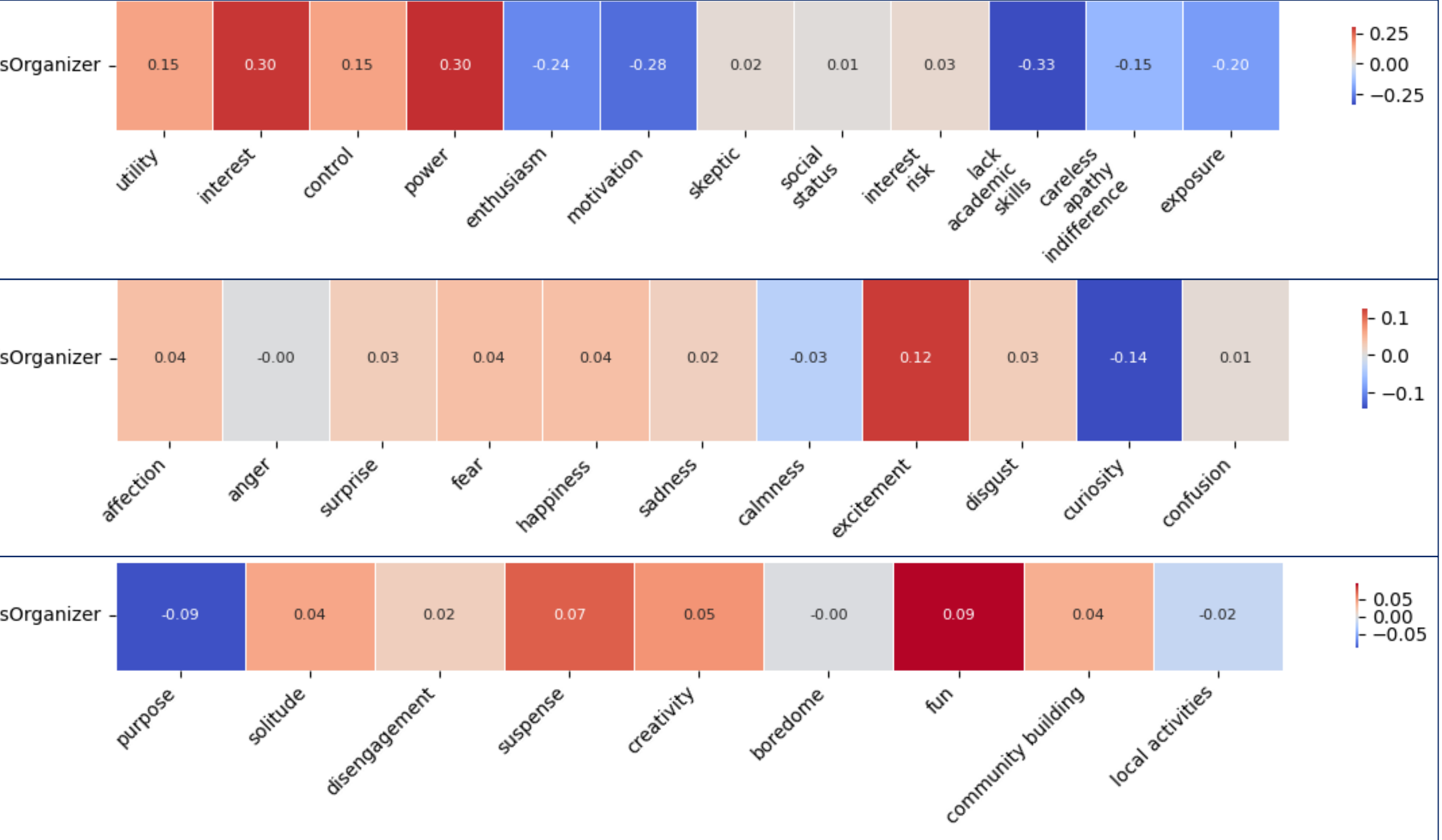
A mob is a type of event that is usually coordinated through social media, where individuals can gather, engage in an activity with a specific goal, and then disperse. Event-based social media (EBSM) platforms provide a fertile ground for groups to coordinate and organize events online, offline, or both, such as mobs. A mob has a "mob organization process" with five phases, and during the "replaying and republishing" phase, organizers and participants share various content on social media—including text, URLs, images, and videos across multiple platforms. In this research, we examine the images shared on one of the most popular EBSM platforms, [Meetup.com](https://www.meetup.com).

Goals

1. Generate accurate captions for images shared during the "replaying and republishing" phase of a mob.
2. Detect emotions and sentiments from the shared images.
3. Explore whether these images reveal anything about who posted them.

Results

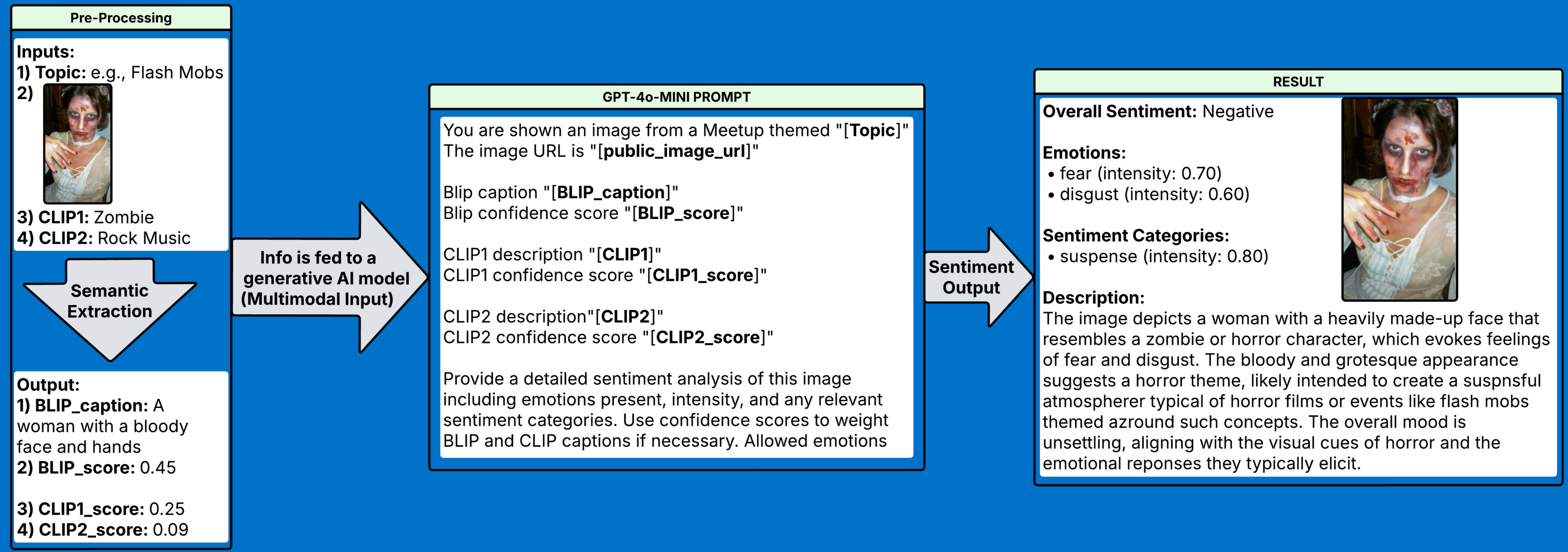
1. Results show that some images can have a poor caption while others have a very good one.
2. Using a visual-textual multi-modal approach performed very well in detecting the sentiments and emotions expressed in the images.
3. Images shared by event organizers seem to have distinct characteristics from those shared by regular mobbers.



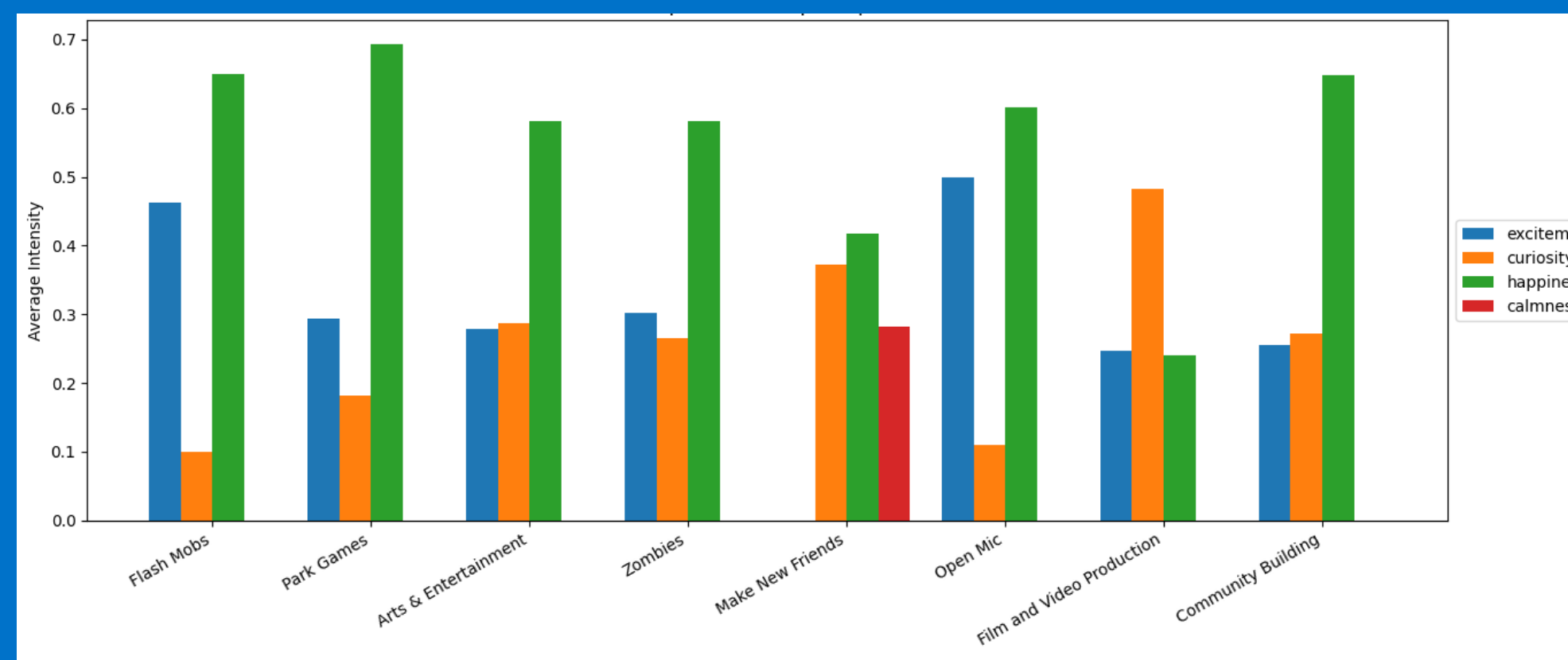
- **Emotions:** Immediate, short-term, e.g., happiness, anger, fear, surprise, disgust, sadness.
- **Sentiments:** Longer-lasting, generalized attitudes, e.g., positive, negative, neutral.



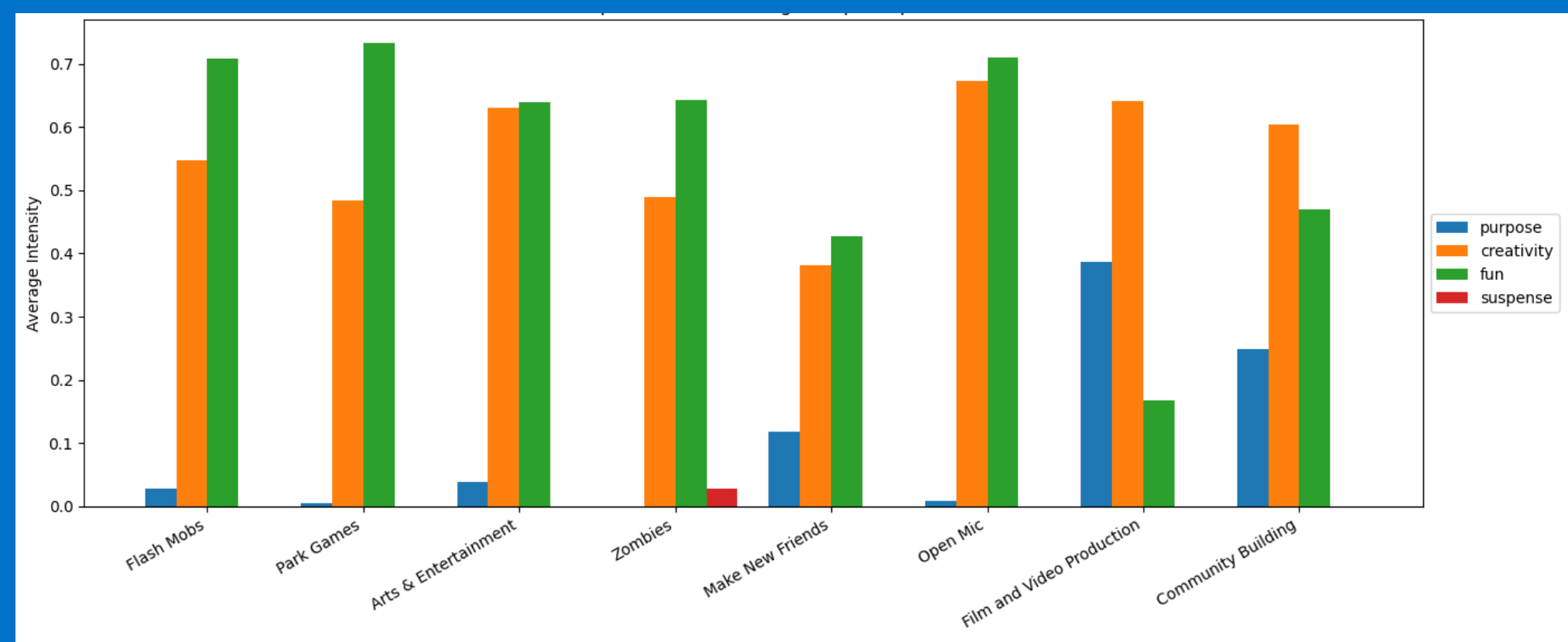
Vision-Language Models (VLMs) deliver accurate captions while detecting sentiments and emotions with precision.



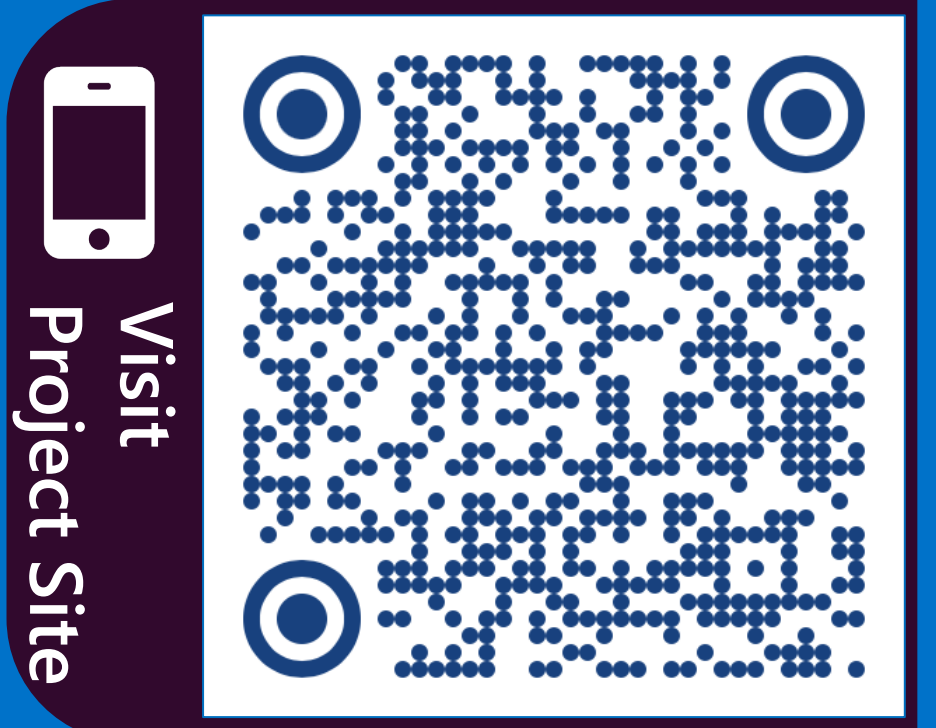
Pipeline of Vision-Language Models (VLMs): Multimodal AI systems used to generate image captions and detect sentiments and emotions.



Top 3 emotions per mob topic



Top 3 sentiments per mob topic



Ariana Mondiri & Samer Al-khateeb



Awards#: FA9550-22-1-0332 & W911NF2510147.