



AGREEMENT ON USE OF FaceCaption-15M

I Agree

- This dataset should be cited in your papers that utilize it. The ref is as follows:

[1] Dai D, Li Y T, Liu Y G, et al. 15m multimodal facial image-text dataset[J]. arXiv preprint arXiv:2407.08515, 2024..

[2] Dai D, Xu L, Li Y, et al. Humanvlm: Foundation for human-scene vision-language model[J]. Information Fusion, 2025: 103271.
- FaceCaption-15M is provided exclusively for non-commercial research purposes.
- NOT to reproduce, duplicate, copy, sell, trade, resell or exploit for any commercial purposes, any portion of the images and any portion of derived data.
- NOT to further copy, publish or distribute any portion of the database.
- To Click “ like ” on Huggingface and “ Starred ” on GitHub

NOTE: The Lab Director or Research Investigator should fill out this form:

Name: _____

Title: _____

Institution: _____

Date: _____

Email Address
(Organizational &
Non-student): _____

Signature:
(Handwritten &
Non-student) _____

Dawei Dai

Chongqing University of Posts and Telecommunications, Chongqing, China

ddwteam8@gmail.com