

Responsible NLP Checklist

Paper title: *EgoMemory: Memory-Augmented Personalized Retrieval for Long-Context Egocentric Video*
Authors: *Yuanmin Tang, Jue Zhang, Xiaoting Qin, Jing Yu, Meikang Qiu, Gaopeng Gou, Gang Xiong, Qingwei Lin, Saravan Rajmohan, Dongmei Zhang, Qi Wu*

How to read the checklist symbols:

- the authors responded 'yes'
- the authors responded 'no'
- the authors indicated that the question does not apply to their work
- the authors did not respond to the checkbox question

For background on the checklist and guidance provided to the authors, see the [Responsible NLP Checklist](#) page at ACL Rolling Review.

A. Questions mandatory for all submissions.

A1. Did you describe the limitations of your work?

This paper has a Limitations section.

A2. Did you discuss any potential risks of your work?

See "Ethical Considerations" section (page 9), which discusses surveillance and manipulation risks of personalized egocentric retrieval and the safeguards required for responsible deployment (access control, encryption, data minimization, user-facing transparency, fairness/misuse monitoring).

B. Did you use or create scientific artifacts? (e.g. code, datasets, models)

B4. Did you discuss the steps taken to check whether the data that was collected/used contains any information that names or uniquely identifies individual people or offensive content, and the steps taken to protect/anonymize it?

Discussed in "Ethical Considerations" (page 9): egocentric video can expose sensitive information about users and bystanders. Our benchmark inherits Ego4D's original consent and access protocols; we recommend safeguards beyond the scope of this work (strong access control, encryption, data minimization, user-facing transparency).

B6. Did you report relevant statistics like the number of examples, details of train/test/dev splits, etc. for the data that you used/created?

Section 3.1 "The EgoMemory Benchmark" (Benchmark Construction & Filtering, Dataset Composition): 165,795 user-specific object annotations across 245 videos from 45 participants (~64.25h total), yielding 639 high-quality personal queries (~91.6% personal). Per-user candidate sizes range 961 clips, clip durations 4300s. Further distribution statistics in Appendix B.1.

C. Did you run computational experiments?

C2. Did you discuss the experimental setup, including hyperparameter search and best-found hyperparameter values?

Section 4 "Implementation Details": retrieval experiments on 4 NVIDIA V100 GPUs (32GB) with GPT-4o backbone and EgoVLPv2 text encoder; CLIP/BLIP video features averaged over 15 uniformly sampled frames. Training-free pipeline no hyperparameter search required. Annotation cost ~\$670 via GPT APIs (one-time). Additional configuration details in Appendix F.1.

The Responsible NLP Checklist used at ACL Rolling Review is adopted from NAACL 2022, with the addition of ACL 2023 question on AI writing assistance and further refinements based on ARR practice. ACL 2026 used a subset of ARR checklist form.

- C3. Did you report descriptive statistics about your results (e.g., error bars around results, summary statistics from sets of experiments), and is it transparent whether you are reporting the max, mean, etc. or just a single run?

Section 4 "Evaluation Metrics": we report mean Recall@K computed per user and macro-averaged across all 45 users to mitigate per-user query-count bias. Single-run evaluation for the training-free deterministic retrieval pipeline (no stochastic training).

D. Did you use human annotators (e.g., crowdworkers) or research with human subjects?

- D1. Did you report the full text of instructions given to participants, including e.g., screenshots, disclaimers of any risks to participants or annotators, etc.?

Annotation and verification performed by paper authors and lab collaborators following internal protocols (described qualitatively in Section 3.1 Human Verification: reviewing 20 additional short clips per query for the same user/object class, labeling personal/uncertain by 90%/75% match thresholds). No third-party crowdworker instructions or screenshots applicable.

- D2. Did you report information about how you recruited (e.g., crowdsourcing platform, students) and paid participants, and discuss if such payment is adequate given the participants' demographic (e.g., country of residence)?

Annotation conducted by paper authors and lab collaborators; no external recruitment or per-task compensation.

- D3. Did you discuss whether and how consent was obtained from people whose data you're using/curating (e.g., did your instructions explain how the data would be used)?

"Ethical Considerations" section (page 9): the benchmark is derived from Ego4D (Grauman et al., 2022); all use complies with Ego4D's original data licenses, consent protocols, and access restrictions.

- D4. Was the data collection protocol approved (or determined exempt) by an ethics review board?

No new human-subjects data collection in this work. We re-annotate existing Ego4D videos that were collected under Ego4D's original IRB-approved protocols (Grauman et al., 2022). Our annotation step adds object-attribute labels to already-collected, already-consented video data.

E. Did you use AI assistants (e.g., ChatGPT, Copilot) in your research, coding, or writing?

- E1. If you used AI assistants, did you include information about their use?

AI assistants used in this work fall into three categories, all disclosed: (i) GPT-4o is a method component used for memory bank construction and reflective Chain-of-Thought reasoning at inference (Section 4 "Implementation Details", Section 3.2). GPT-4o is also used for query pre-screening during dataset filtering (Section 3.1). (ii) GPT-4o-mini, LLaVA, and Qwen2.5-VL are evaluated as alternative MLLM backbones in ablations (Section 4.2, Table 3 rows 1315). (iii) Claude Code (Anthropic) was used to assist with camera-ready preparation: LaTeX edits, page-budget verification, and build pipeline. All experimental design, data analysis, and scientific claims were conducted and verified by the human authors.