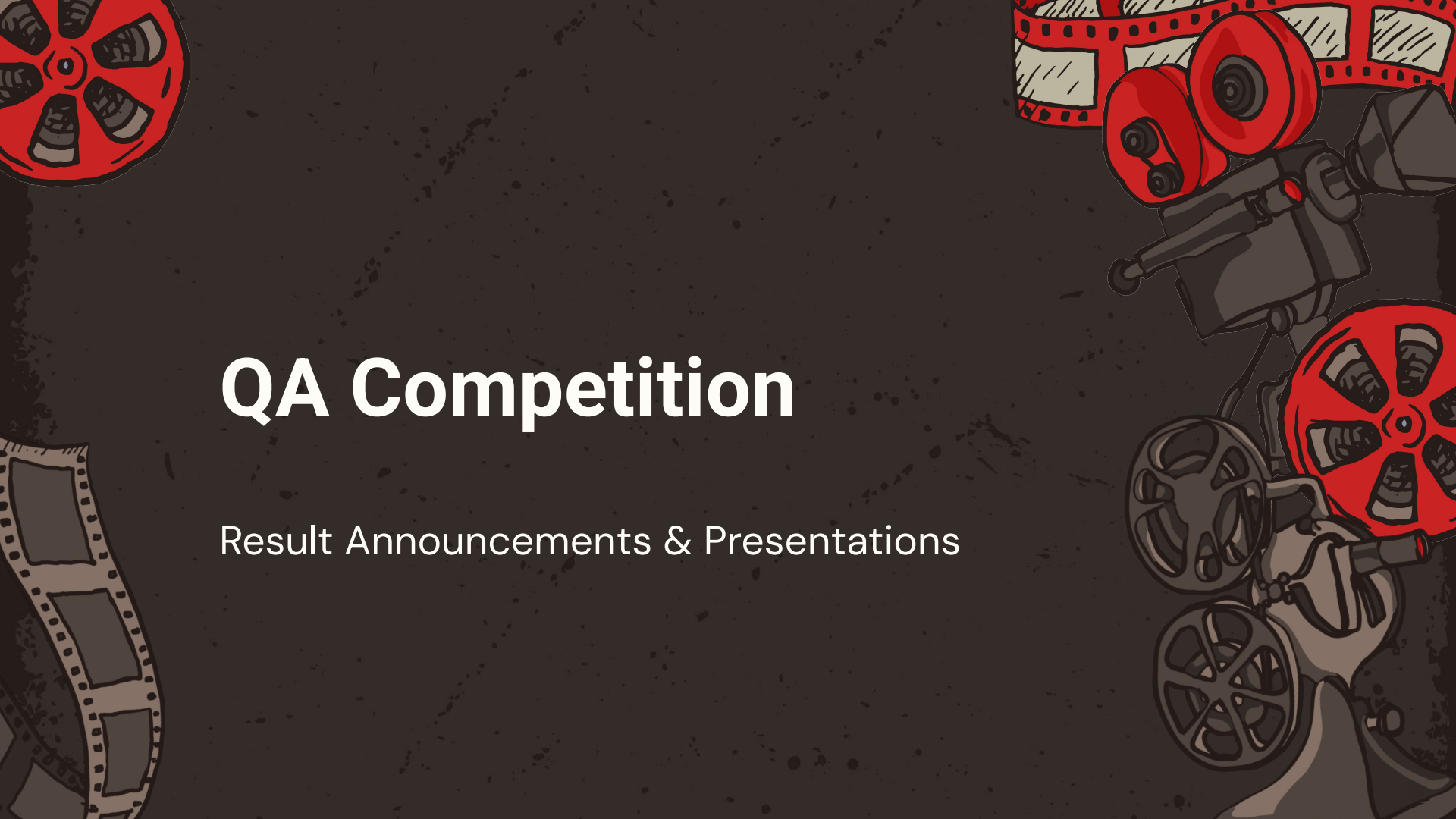


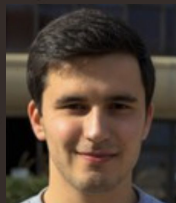
QA Competition

Result Announcements & Presentations



Short Films 20K (SF20K) Competition

Organizers:



Ridouane Ghermi



Xi Wang



Vicky Kalogeiton



Ivan Laptev

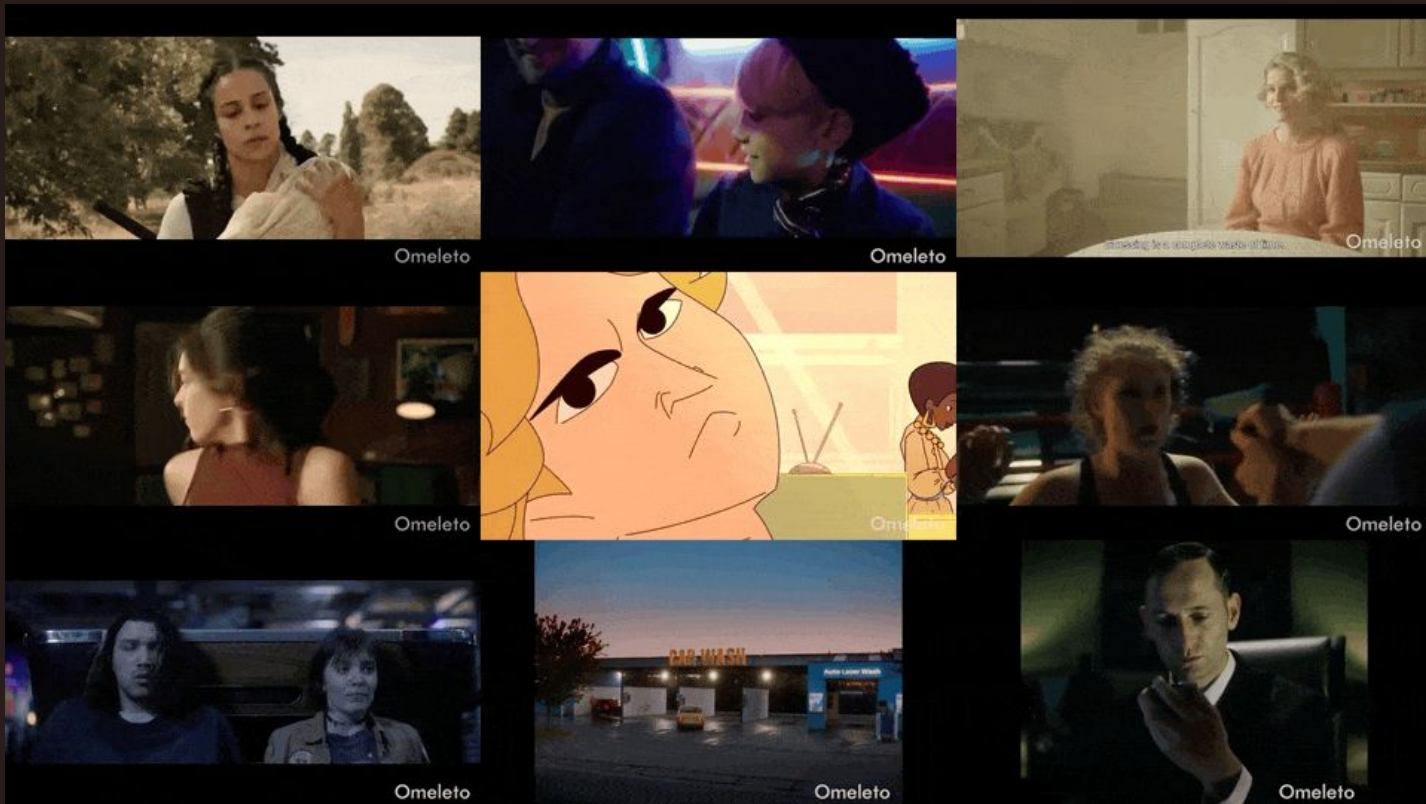


MOHAMED BIN ZAYED
UNIVERSITY OF
ARTIFICIAL INTELLIGENCE



INSTITUT
POLYTECHNIQUE
DE PARIS

What are Short Films?



Short-Films 20K (SF20K)



SF20K: ♦ new dataset with 20k movies ♦ ~12 mins / video ♦ 3.6k hours of video
♦ large variety of genres ♦ accessible online ♦ limited data leakage

Comedy



The Beast, the Phantom and the Hunchback use a dating app to find love.

Romance



A couple on a first date clash over astrology.

Horror



*A young woman trapped in a bathroom stall during an active sho*ting.*

Action



The biggest boxing fight of 1960 takes an unexpected turn.

Drama



A mother struggles with bullies who torment her disabled daughter.

Animation



A young boy gets lost in a strange forest. Then his father tries to rescue him.

Documentary



How Movie Sounds are Made. An Inside Look at the World of Foley Artist.

Experimental



A stranded soul searches an endless desert for his purpose.

Why another VideoQA benchmark?

- Focus on storytelling and personal relations



Movies



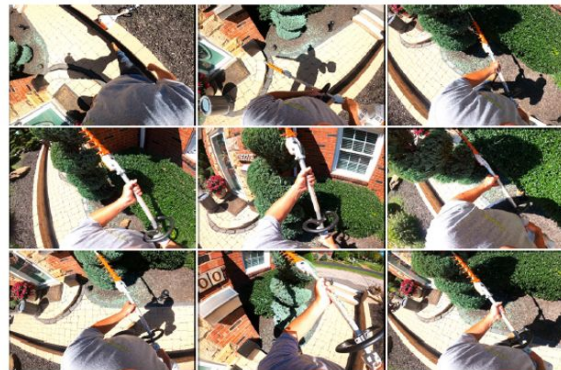
A young girl lives in a bunker, protected from the monsters. Today she goes outside...

Instructional videos



How to tie a necktie?

Egocentric videos



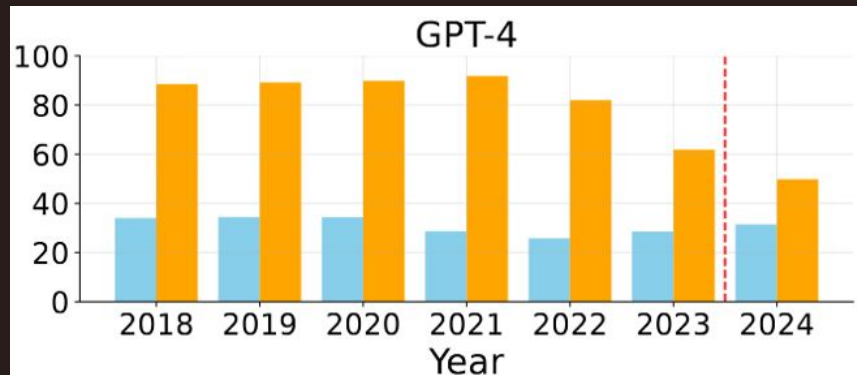
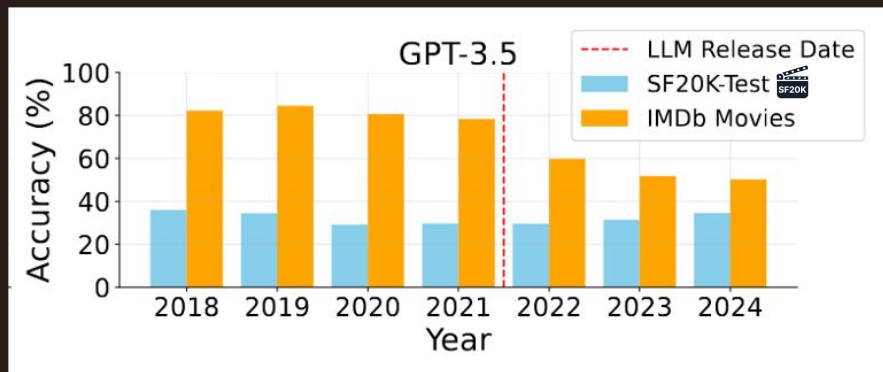
C was in the front yard, pulled the starter string
and trimmed the tree with a hedge trimmer



Why another VideoQA benchmark?

- Focus on storytelling and personal relations
- Limited data contamination

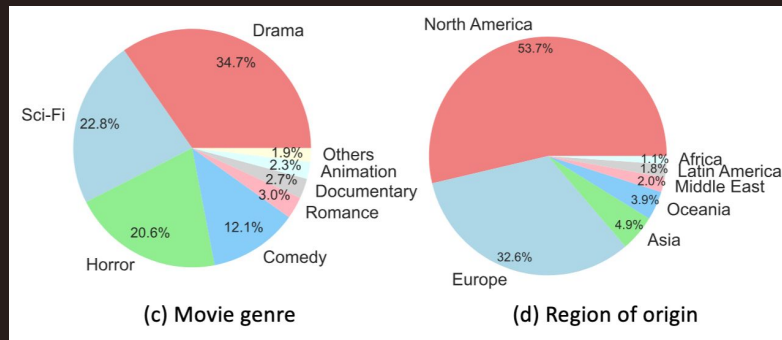
Movie question answering *from movie titles*



➡ Modern LLMs memorize common movies and can answer questions in other MovieQA benchmarks w/o watching the movies

SF20K Challenge

- Task: Open-ended video question answering
- Source: High-quality movies from SF20K
- Statistics: 95 movies, 979 question-answer pairs



After The Rain

A dog turn sheep's wool into clouds.



[4:19] The shepherd passes away.



[4:36] The meadow becomes a dry, desert-like landscape.



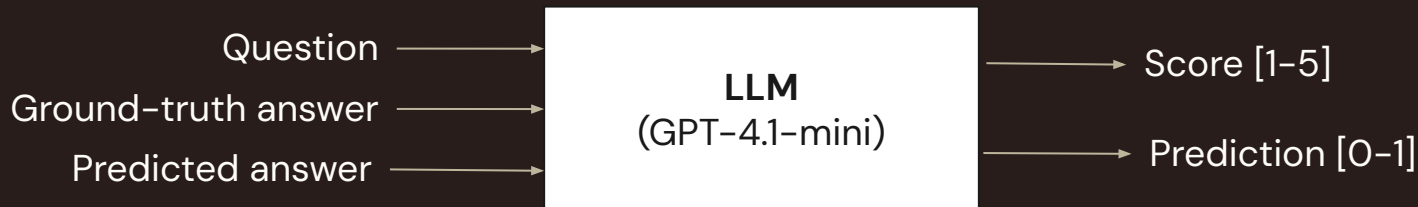
[4:58] The sheeps are not sheared.

Q: What happens to the meadow after the shepherd dies?

A: The meadow becomes dry, invaded by brambles.

Metric

LLM-QA-Eval: LLM-as-a-judge scoring the similarity between the predicted and the ground-truth answer given the question as context



Baselines

Model	#Frames	Accuracy (%)
Random	-	0.4
Question-only	0	14.7
Qwen3-VL-8B	256	50.1
Gemini-2.5-Pro	1 FPS	59.3
Human	-	91.7

- Accuracy can go from 0 to 100%
- Question-only performance is non-zero but low
- Human solve most of the task (91.7%)
- Baseline (Qwen3-VL-8B) and SOTA (Gemini-2.5-Pro) performances still lags behind (50.1% / 59.3%)

Ablations

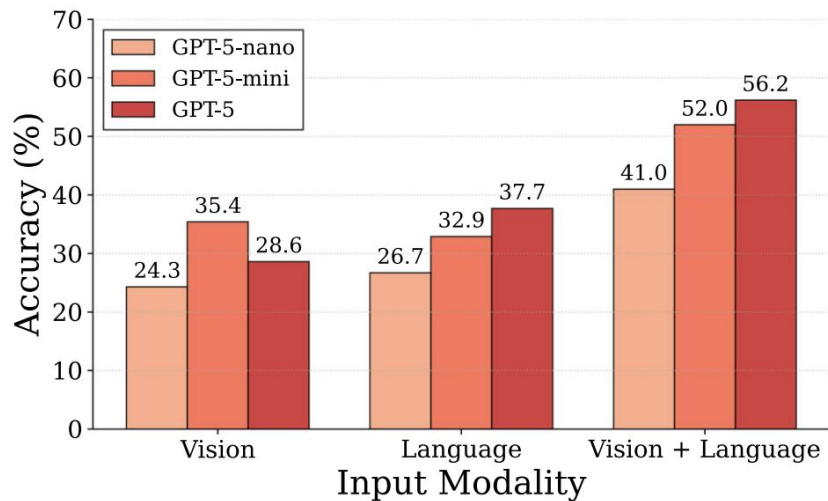


Fig. 7 Accuracy vs. Input Modality of GPT-5 variants on SF20K-Test

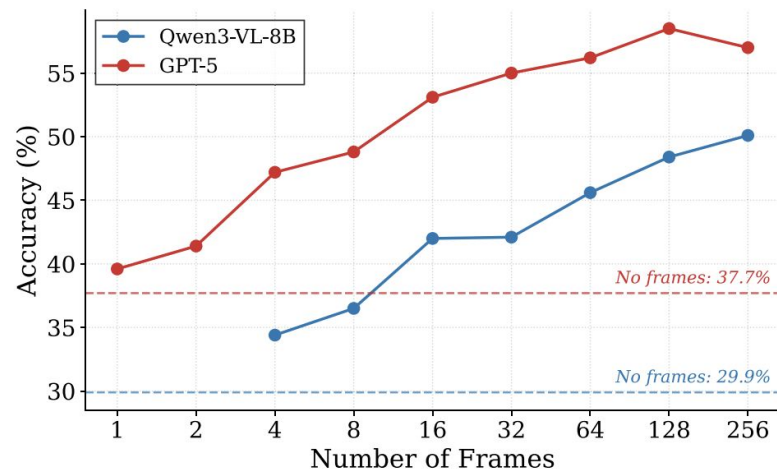


Fig. 8 Accuracy vs. Number of Frames of GPT-5 and Qwen3-VL-8B on SF20K-Test

Competition results

- Competition dates: July 1 - August 31
- Two tracks: Main Track (unlimited model size) and Special Track (<10B params models)
- Two phases: Public Test Set and Private Test Set
- **11 active teams, 660 submissions**

Private leaderboard - Main Track

Rank	Team ID	Acc.
1	Turismo	87.76
2	Math	67.15
3	ZFTurbo	65.99
4	ucas_xlab	64.40
5	Blueberry	60.54
6	SEAMiLab	56.69

Private leaderboard - Special Track

Rank	Team ID	Acc.
1	l1d1w1	58.50
2	Nathaniel	57.14
3	Blueberry	52.61
4	ZFTurbo	51.70
5	ucas_xlab	51.02
6	SEAMiLab	39.46

Competition winners

Main Track	
	Team Turismo
	Team Math

Special Track	
	Team l1d1w1
	Team Nathaniel

Prizes

Coud GPU credits:

\$1,000

Sponsored by:



MOHAMED BIN ZAYED
UNIVERSITY OF
ARTIFICIAL INTELLIGENCE

Congratulations!

Presentation of the Winners' Team

Team Math

Second place in the Main Track.

Presentation of the Winners' Team

Team l1d1w1

First place in the Special Track.

A top-down view of a dark brown wooden surface. In the top left, a large white film reel is partially unspooled, with its black and white film strip winding across the frame. Scattered throughout are numerous small, fluffy, yellow popcorn kernels. In the top right, a white ceramic cup of coffee sits on a matching saucer, featuring a latte art design of a heart with symmetrical leaf-like patterns. In the bottom right, a black clapperboard is positioned diagonally, showing the words 'SCENE', 'TAKE', and 'ROLL' in white. In the bottom left, another film reel is visible. In the center, two overlapping tan-colored tickets with scalloped edges are placed; the top ticket reads 'SLoMO workshop' and the bottom one reads 'workshop'.

Coffee Break, come back at 15:40

SLoMO
workshop

workshop

SCENE

TAKE

ROLL