

# Reading guide

One technique paper, one art-and-tech project, one culture piece. For Quiz 07, reflect on any one of them.

## READING 1 / TECHNIQUE

### [Helios: Real Real-Time Long Video Generation Model](#)

PKU-YuanGroup, Peking University, with ByteDance • arXiv 2603.04379 | March 2026

**Task.** Read the abstract, then open the project page and watch two of the minute-long videos. You do not need the method sections.

Helios generates about a minute of video at roughly 19.5 frames per second on a single H100 GPU, which is close to watching it being made. Its key idea for this course: long videos fall apart because of drift, and Helios is trained on deliberately drifted video so it learns to recover.

#### Questions to keep in mind

- What is drift, in your own words?
- Watch one long example. Where, if anywhere, does it start to change?
- What could you make with a minute that you couldn't with five seconds?

This is the model behind Story 1: the real-time minute and the drift problem.

Project page with videos: <https://pku-yuangroup.github.io/Helios-Page>

## READING 2 / ART + TECH

### [Wan-Dancer: Minute-scale Coherent Music-to-Dance Generation](#)

Mingyang Huang, Peng Zhang, Li Hu, Guangyuan Wang, Ruoshi Zhang, Yi Lu, Gang Cheng, and Bang Zhang • Tongyi Lab, Alibaba • Project page | July 2026

**Task.** Watch at least three examples on the project page, including one tap dance and one Chinese classical dance. The videos are the reading.

Wan-Dancer turns a music track, one photo of a dancer and a style word into a dance video longer than a minute, on the beat. It plans the whole song first (intro, verse, chorus) and then renders the steps, much as a choreographer works.

#### Questions to keep in mind

- Which example would you call choreography, and why?
- Does it follow the music's structure, or only the beat?
- Who should be credited for one of these dances?

Tonight's Story 2 and the 'Pitch a one-take' activity.

Open weights: <https://huggingface.co/Wan-AI/Wan-Dancer-14B>

## What Happened to Sora? OpenAI Shuts Down Its AI Video App

pasqualepillitteri.it (blog) • 2026

**Task.** Read the article (about 10 minutes), then skim OpenAI's own help page on the discontinuation.

OpenAI closed the Sora app on April 26, 2026 and switched off the Sora API on September 24, 2026, telling users to export their work before it was deleted. The article traces the timeline and the reported reasons, including the cost of compute.

### Questions to keep in mind

- What would you lose if your main tool disappeared tomorrow?
- Which reasons for the shutdown are confirmed, and which are reported?
- What would you keep from a project so it outlives its tool?

Tonight's Story 4 and the 'Keep your sources' checklist.

Official source: <https://help.openai.com/en/articles/20001152-what-to-know-about-the-sora-discontinuation>

## Self Forcing: Bridging the Train-Test Gap in Autoregressive Video Diffusion

Huang et al. • Adobe Research and The University of Texas at Austin • NeurIPS 2025 | arXiv 2506.08009

**Task.** Optional, for students who want the technical side. Read the abstract and introduction and look at the first figure; skip the derivations. Not needed for Quiz 07.

Autoregressive video models are trained on real previous frames but must continue from their own imperfect frames when used, so errors pile up (exposure bias, or drift). Self Forcing closes this train-test gap by rolling the model out on its own frames during training, using a KV cache, and supervising the whole video.

### Questions to keep in mind

- In your own words, what is the train-test gap?
- Why does a KV cache make training on the model's own frames affordable?
- How is Helios's answer to drift different?

Goes with the slides “Under the hood: the KV cache” and “Under the hood: the train-test gap” .

Background on diffusion and flow matching: MIT 6.S184 course notes, <https://diffusion.csail.mit.edu/>