

Lecture 02 · Reading guide

1. An Alien Mind

Jakub Pachocki / OpenAI · 6 September 2026

[Read the original article](#)

Research perspective and policy argument

Pachocki argues that increasing AI capability raises difficult questions about alignment and monitoring. He distinguishes accomplishing an assigned goal from generalizing human values in unfamiliar situations. Read this as a research leader's position about the direction and governance of AI development, including claims and forecasts that deserve scrutiny.

Read the opening, then focus on “Teaching machines to love” and “Monitoring generalization.” Skim the remaining sections to understand the author's larger argument.

Slides 11–12, 37–40, 101–104, 111 and 120

The analogy to art direction is our classroom application. It does not establish that an image editor has the agency or risk profile of an autonomous research system.

Discuss

- How does the author distinguish goal alignment from value alignment? Explain using an example of your own.
- Which statement in the article is an observation, which is a forecast, and which is a recommendation? What evidence would you seek for each?
- If a creative agent achieves a requested visual effect while undermining the artist's purpose, what should the artist inspect or change?

Discuss further

Which task could an exhibition team delegate to AI, which artistic decision should remain human, and what evidence would justify acceptance?

2. How to fix your entire life in 1 day

Dan Koe · 23 December 2025 (author newsletter version)

[Read the original article](#)

Personal-development essay, selected by the instructor

Koe proposes reflecting on identity and goals, interrupting habitual behavior during the day, and turning the evening's reflections into concrete actions. Use the essay as a prompt to examine creative habits and intentions. Its promise is a proposition to evaluate, rather than an established guarantee of transformation.

Read the opening and section VI, which presents the day-long reflection process. Follow the argument around those sections as time allows. Completing the full personal exercise is optional.

Slides 49–72, 109–111, 121 and 123

This is the instructor-selected companion on human agency to the AI readings. It is not an AI research article. Students need not disclose private struggles or accept every psychological claim.

Discuss

- Which part of the proposed process could help an artist decide what to make, rather than only produce more?
- Identify a claim you find useful and one you would challenge. What experience or evidence would change your view?
- How could abundant AI-generated options reinforce an unhelpful habit, or support a deliberate experiment? Make the mechanism specific.

Discuss further

How could a small creative experiment clarify an intention, what observation would make it useful, and where might the essay's advice fail?

3. Introducing Aleph 2.0 and Edit Studio

Runway · 21 May 2026

[Read the original article](#)

Public AI product announcement with visual examples

Runway describes a video-editing workflow in which an edited image establishes the desired appearance before the change is applied through a clip. The announcement presents localized edits and preservation as product capabilities. Read the examples alongside the tests you would require for your own shot.

Read the sections describing image-level control and what the workflow makes possible. Inspect one visual example, then separate what the page shows from what it asserts.

Slides 73–76, 95–104, 112 and 122

This is a vendor announcement, not an independent comparative evaluation. Selected demonstrations do not establish reliability on every clip, and the page does not disclose the complete model architecture.

Discuss

- What decision can the edited reference frame settle, and what remains uncertain until the full clip is reviewed?
- Choose one preservation claim. What input clip and failure event would make a useful test?
- How could this workflow support the AFTER RAIN concept without letting a product demonstration determine the artwork's purpose?

Discuss further

What should an edited keyframe establish, which motion event is most likely to reveal failure, and what evidence should be reviewed before acceptance?