DISASTER DRAW

About this edition

Disaster Draw · November 2026

© 2026 Tim O’Brien. All rights reserved. Sources accompany the cards and reading editions.

disasterdraw.com

24 content cards, two promotion cards, and this edition card. General education and discussion. Follow current local instructions.

DISASTER DRAW

About this edition

Disaster Draw · November 2026

© 2026 Tim O’Brien. All rights reserved. Sources accompany the cards and reading editions.

disasterdraw.com

24 content cards, two promotion cards, and this edition card. General education and discussion. Follow current local instructions.

The Campus

AI alignment disasters

CONSIDER TOGETHER

What does Lighthaven actually describe itself as?

Everyone answers before the reader reveals the back.

The Campus

WHAT TO NOTICE

A Berkeley campus operated by Lightcone, hosting events across AI and intellectual culture.

TRY THIS

Examine the institution's public program before drawing conclusions about it.

Lighthaven



Keep the conversation going.

Free books, more decks, and the sources behind every card.

disasterdraw.com

The Lost Apprentice

AI alignment disasters

CONSIDER TOGETHER

A hypothetical organization removes entry-level roles. How will the next generation gain judgment?

Everyone answers before the reader reveals the back.

The Lost Apprentice

WHAT TO NOTICE

Training and operational knowledge require explicit attention during restructuring.

TRY THIS

Ask who owns skill development after automation.

NIST · AI Risk Management Framework

Keep the conversation going.

Free books, more decks, and the sources behind every card.

disasterdraw.com

The Extreme Prescription

AI alignment disasters

CONSIDER TOGETHER

Yudkowsky proposed coercive international limits on large AI training runs. What evidence should support such a policy?

Everyone answers before the reader reveals the back.

The Extreme Prescription

WHAT TO NOTICE

A catastrophic forecast carries an evidentiary burden alongside the risks of its proposed enforcement.

TRY THIS

Read the original argument and identify its assumptions.

Yudkowsky essay, preserved copy

Keep the conversation going.

Free books, more decks, and the sources behind every card.

disasterdraw.com

The Basilisk

AI alignment disasters

CONSIDER TOGETHER

A speculative future punisher makes a reader feel obliged to act. What makes the argument credible?

Everyone answers before the reader reveals the back.

The Basilisk

WHAT TO NOTICE

Emotional pressure does not establish the truth of a thought experiment.

TRY THIS

Identify the assumptions before accepting an obligation.

LessWrong · Roko's basilisk

Keep the conversation going.

Free books, more decks, and the sources behind every card.

disasterdraw.com

The Greater Good

AI alignment disasters

CONSIDER TOGETHER

Does a donor's promise to support worthy causes establish the legitimacy of the money?

Everyone answers before the reader reveals the back.

The Greater Good

WHAT TO NOTICE

Bankman-Fried's case demonstrates the need to examine conduct and financial controls.

TRY THIS

Ask how an organization independently reviews major support.

US Department of Justice

Keep the conversation going.

Free books, more decks, and the sources behind every card.

disasterdraw.com

The Community Account

AI alignment disasters

CONSIDER TOGETHER

An institution describes its connection to a disgraced donor. What kind of source is that?

Everyone answers before the reader reveals the back.

The Community Account

WHAT TO NOTICE

Its own account establishes its stated position and should be read with that perspective.

TRY THIS

Compare institutional statements with primary legal records.

Centre for Effective Altruism



Keep the conversation going.

Free books, more decks, and the sources behind every card.

disasterdraw.com

Forty-five Percent

AI alignment disasters

CONSIDER TOGETHER

Veracode reports flaws in 45% of tested generated samples. Forty-five percent of what?

Everyone answers before the reader reveals the back.

Forty-five Percent

WHAT TO NOTICE

The denominator is the evaluated outputs under the study conditions.

TRY THIS

Record task set, models, date, and methodology with the statistic.

Veracode · 2025 report

Keep the conversation going.

Free books, more decks, and the sources behind every card.

disasterdraw.com

The Review Queue

AI alignment disasters

CONSIDER TOGETHER

A hypothetical team writes more code and cuts reviewers. What burden grows?

Everyone answers before the reader reveals the back.

The Review Queue

WHAT TO NOTICE

Generation speed can increase the amount needing security scrutiny.

TRY THIS

Track review capacity and consequential defects alongside output.

Veracode · 2026 update



Keep the conversation going.

Free books, more decks, and the sources behind every card.

disasterdraw.com

121,401

AI alignment disasters

CONSIDER TOGETHER

What does Challenger's January–October 2026 AI job-cut total count?

Everyone answers before the reader reveals the back.

121,401

WHAT TO NOTICE

Announced US cuts where employers cited AI as the reason.

TRY THIS

Keep the period and definition attached to the number.

Challenger · October 2026, Table 4



Keep the conversation going.

Free books, more decks, and the sources behind every card.

disasterdraw.com

The CEO's Forecast

AI alignment disasters

CONSIDER TOGETHER

Jassy said he expected AI to reduce Amazon's total corporate workforce over the following years. Is that a completed layoff count?

Everyone answers before the reader reveals the back.

The CEO’s Forecast

WHAT TO NOTICE

It is a management forecast about staffing.

TRY THIS

Distinguish future expectations from completed reductions.

Amazon · June 2025 message



Keep the conversation going.

Free books, more decks, and the sources behind every card.

disasterdraw.com

DISASTER DRAW

The End of the World Has a Campus

Read the companion book

AI prophets, rationalist institutions, broken software, and disappearing jobs



disasterdraw.com

Read online or download the PDF and EPUB at disasterdraw.com.

DISASTER DRAW

The End of the World Has a Campus

Read the companion book

AI prophets, rationalist institutions, broken software, and disappearing jobs



disasterdraw.com

Read online or download the PDF and EPUB at disasterdraw.com.

Disaster Draw

Seven subjects. One growing library.

Read the books. Explore a new deck. Bring the next conversation to your table.



disasterdraw.com

Scan for this deck, its companion book, and the complete collection.

Disaster Draw

Seven subjects. One growing library.

Read the books. Explore a new deck. Bring the next conversation to your table.



disasterdraw.com

Scan for this deck, its companion book, and the complete collection.

The Agent Chain

AI alignment disasters

CONSIDER TOGETHER

One assistant delegates to another. Who remains accountable for the final action?

Everyone answers before the reader reveals the back.

The Agent Chain

WHAT TO NOTICE

Delegation can blur responsibility without removing consequences.

TRY THIS

Name the owner of the complete workflow.

NIST · AI Risk Management Framework

Keep the conversation going.

Free books, more decks, and the sources behind every card.

disasterdraw.com

The Appeal Window

AI alignment disasters

CONSIDER TOGETHER

An automated decision harms someone. How can they challenge it?

Everyone answers before the reader reveals the back.

The Appeal Window

WHAT TO NOTICE

A correction channel belongs in the design of consequential systems.

TRY THIS

Provide a reachable human appeal process.

NIST · AI Risk Management Framework

Keep the conversation going.

Free books, more decks, and the sources behind every card.

disasterdraw.com

The Cheap Retry

AI alignment disasters

CONSIDER TOGETHER

An automated task repeats until a budget is exhausted. What boundary was missing?

Everyone answers before the reader reveals the back.

The Cheap Retry

WHAT TO NOTICE

A task can pursue an assigned goal beyond sensible limits.

TRY THIS

Set monitored time and spending limits.

NIST · AI Risk Management Framework

Keep the conversation going.

Free books, more decks, and the sources behind every card.

disasterdraw.com

The Evaluation Room

AI alignment disasters

CONSIDER TOGETHER

A model passes a test in a controlled setting. What does that establish about deployment?

Everyone answers before the reader reveals the back.

The Evaluation Room

WHAT TO NOTICE

A test establishes performance under its tested conditions.

TRY THIS

Document what the evaluation excluded.

International AI Safety Report 2026

Keep the conversation going.

Free books, more decks, and the sources behind every card.

disasterdraw.com

The Future Scenario

AI alignment disasters

CONSIDER TOGETHER

A speaker assigns a precise date to catastrophic loss of control. What supports that date?

Everyone answers before the reader reveals the back.

The Future Scenario

WHAT TO NOTICE

Serious future risks carry unresolved forecasting uncertainty.

TRY THIS

Distinguish observations, assumptions, and forecasts.

International AI Safety Report 2026



Keep the conversation going.

Free books, more decks, and the sources behind every card.

disasterdraw.com

The Human Shortcut

AI alignment disasters

CONSIDER TOGETHER

An operator bypasses a control to meet a deadline. Which incentive needs attention?

Everyone answers before the reader reveals the back.

The Human Shortcut

WHAT TO NOTICE

Organizational pressure can weaken a technical safeguard.

TRY THIS

Make safe escalation possible within the workflow.

NIST · AI Risk Management Framework

Keep the conversation going.

Free books, more decks, and the sources behind every card.

disasterdraw.com

The Incident Record

AI alignment disasters

CONSIDER TOGETHER

The output disappears after a failure. What can investigators still reconstruct?

Everyone answers before the reader reveals the back.

The Incident Record

WHAT TO NOTICE

Investigation depends on retained evidence and appropriate records.

TRY THIS

Define privacy-conscious incident logging.

NIST · AI Risk Management Framework



Keep the conversation going.

Free books, more decks, and the sources behind every card.

disasterdraw.com

The Independent Check

AI alignment disasters

CONSIDER TOGETHER

Two assistants agree. What would make that agreement less reassuring?

Everyone answers before the reader reveals the back.

The Independent Check

WHAT TO NOTICE

Similar systems can share training patterns and blind spots.

TRY THIS

Use independent evidence and qualified human judgment.

NIST · AI Risk Management Framework

Keep the conversation going.

Free books, more decks, and the sources behind every card.

disasterdraw.com

The Irreversible Step

AI alignment disasters

CONSIDER TOGETHER

A workflow is about to make a costly irreversible change. What should happen first?

Everyone answers before the reader reveals the back.

The Irreversible Step

WHAT TO NOTICE

Consequences should shape the required authority and review.

TRY THIS

Require a specific responsible approval.

NIST · AI Risk Management Framework



Keep the conversation going.

Free books, more decks, and the sources behind every card.

disasterdraw.com

The Persuasive Answer

AI alignment disasters

CONSIDER TOGETHER

An answer becomes more convincing after criticism. Has it become more accurate?

Everyone answers before the reader reveals the back.

The Persuasive Answer

WHAT TO NOTICE

Persuasion and evidence are different qualities.

TRY THIS

Ask for verifiable evidence before accepting the revision.

NIST · AI Risk Management Framework

Keep the conversation going.

Free books, more decks, and the sources behind every card.

disasterdraw.com

The Public Consequence

AI alignment disasters

CONSIDER TOGETHER

A deployment benefits the buyer while shifting harms to others. Who had a voice?

Everyone answers before the reader reveals the back.

The Public Consequence

WHAT TO NOTICE

Affected people may be absent from procurement decisions.

TRY THIS

Include affected groups in risk assessment.

NIST · AI Risk Management Framework

Keep the conversation going.

Free books, more decks, and the sources behind every card.

disasterdraw.com

The Research Claim

AI alignment disasters

CONSIDER TOGETHER

A lab demonstration is described as an inevitable worldwide outcome. What changed in the retelling?

Everyone answers before the reader reveals the back.

The Research Claim

WHAT TO NOTICE

A demonstration and a forecast answer different questions.

TRY THIS

Record the demonstrated conditions and open uncertainties.

International AI Safety Report 2026



Keep the conversation going.

Free books, more decks, and the sources behind every card.

disasterdraw.com

The Unusual Case

AI alignment disasters

CONSIDER TOGETHER

A system performs well on average but fails for a small group. Who notices?

Everyone answers before the reader reveals the back.

The Unusual Case

WHAT TO NOTICE

Aggregate performance can hide uneven harms.

TRY THIS

Measure results for relevant affected groups.

NIST · AI Risk Management Framework



Keep the conversation going.

Free books, more decks, and the sources behind every card.

disasterdraw.com

The Vendor Promise

AI alignment disasters

CONSIDER TOGETHER

A contract says the system is safe. What evidence will your team request?

Everyone answers before the reader reveals the back.

The Vendor Promise

WHAT TO NOTICE

A claim needs relevant evaluations and an incident process.

TRY THIS

Ask for documented limitations and change notices.

NIST · AI Risk Management Framework



Keep the conversation going.

Free books, more decks, and the sources behind every card.

disasterdraw.com