AI SAFETY SANDBOX · AI POLICY

Sort 20 AI news stories with the MIT AI Risk Repository

Starter · A weekend · Free · Step-by-step guide

Online version: https://ai-safety-sandbox.vercel.app/projects/sort-risks-with-the-mit-repository/

Summary

The MIT AI Risk Repository organises over a thousand AI risks into a shared taxonomy. Use it to classify 20 recent news stories, then look at what the pattern tells you.

Why it matters

Shared categories let researchers and policymakers compare notes. Applying a taxonomy to real events also reveals where it fits badly.

What you will have at the end

A spreadsheet of 20 classified stories and a short analysis.

Steps (tick each box as you go)

- [ ] **1. Learn the taxonomy (45 min)**
  Explore the MIT AI Risk Repository and read how its causal and domain taxonomies work.
- [ ] **2. Collect 20 stories (1 hr)**
  Gather recent AI news stories involving harm or near-misses. Mix topics: jobs, misinformation, security, safety.
- [ ] **3. Classify them (2 hr)**
  Put each story into the taxonomy. Note when it is ambiguous or fits nowhere.
- [ ] **4. Look for patterns (1 hr)**
  Which domains dominate the news? Which risks are missing from coverage entirely?
- [ ] **5. Write it up (1 hr)**
  Share your spreadsheet and three findings, including one place the taxonomy struggled.

What you need

- A spreadsheet
- A weekend

What you will learn

- The landscape of AI risks
- Using a research taxonomy

- Finding gaps in public attention

Source and links

MIT AI Risk Repository
https://airisk.mit.edu/

When you have finished

Write a few paragraphs on what you did, what you found and what you would try next, and share it with a friend, your society or online.

Try next: Investigate a real AI incident
https://ai-safety-sandbox.vercel.app/projects/investigate-an-ai-incident/

Courses, mentors, small grants and communities:
https://ai-safety-sandbox.vercel.app/next-steps/

Questions? Ask in AI Safety Cardiff University (AISCU): https://aiscu.org/

Notes