

Activity

Chatbot Arena

Spring 2026

Fernanda M. Elliott, Ph.D.

Computer Science, Grinnell College

Instructions

1. You will read the article and discuss with four peers:

- [Why most AI benchmarks tell us so little](#)
 - a) What is the problem described in the article?
 - b) Why does it matter?

2. You will read the article and then follow the instructions on the next slide

- [The AI industry is obsessed with Chatbot Arena, but it might not be the best benchmark](#)

Chatbot Arena (Do not add any private or sensitive info!)

1. What is the chatbot arena?
2. What problems is it trying to solve?
3. You will craft three questions and test them on the [Chatbot arena](#).

Things to consider:

1. What are you testing through your questions?
2. How do your questions relate to the problem described in the previous article?
3. What do you expect the answer to look like?
4. Analyze and compare the output you got.