

New experiment

new-experiment.html · 1060x724

New experiment

Task & prompts

Enrollee experience

Scoring

Run defaults

2 Identity (Staff-Facing)

Experiment name

e.g. Terse-prompt study

Description

What question does this study answer? (staff only)

5 What the Enrollee Sees

6 Enrollee title

7 e.g. Chat session

8 Scenario

9 The framing the enrollee reads above the chat.

3 System prompt

4 The model's standing instruction — the base of the prompt cascade.

10 Opening message

10 The assistant's opening line to the enrollee.

Context Documents

11 Grounding corpus

12 + Attach documents

reference-guidelines.pdf

READY

market-primer.docx

READY

Attach none and groundedness is not scored — the composite is the weighted mean of whichever factors apply.

13 Cancel

14 Create experiment

Edits apply to future runs; active runs are frozen.

Annotations

- 1

A frame, not a run The reusable experiment frame holds only what is **constant across every run**. The per-run variables (model configuration, nudge, cohort) are chosen at **launch**. Opened as a modal over the Experiments list — the list stays put behind it.
- 2

Staff-facing name and description are shown to staff only, so they can describe the study's intent. Enrollees never see them — they see the enrollee-safe fields on the right.
- 3

The system prompt is hidden — enrollees never see it.
- 4

Prompt cascade base `system_prompt` — the model's standing instruction and the **base** of the cascade, and it belongs to the study — there is no per-run override. The one override that exists is per **model**: `model_config_model.system_prompt`, authored on the model configuration.
- 5

Enrollee-safe These three are the only study text an enrollee sees. Keeping them separate from the staff name/description is what lets the study stay **blinded**. The nudge (per-run guidance) is chosen at launch, not here.
- 6

An enrollee-safe title shown in place of the staff name, so the study's intent stays hidden.
- 7

Enrollee-safe title `enrollee_title` — shown to enrollees instead of name. The app prefills a head-started value you can edit.

-
- 8 Enrollees read this text above the chat for the whole session.
-
- 9 **Enrollee framing** scenario — the framing the enrollee reads (a persistent strip above the chat).
-
- 10 **First message** opening_message — the assistant's opening line to the enrollee.
-
- 11 **Attached to the study, not to a run** experiment_context_file joins this experiment to the document rows its answers are grounded in. They are attached **here** and nowhere else: a run varies only the nudge, the model configuration and the cohort, so run-level context would change what is being measured between siblings that are supposed to differ in one thing. Every run of this study therefore retrieves from the same corpus. Only documents whose extraction is ready can be attached, since an unextracted file has nothing to embed. The set is frozen once the study has runs. Attaching none is allowed and the composer is warned that **groundedness will not be scored** — there is nothing to ground against, so the factor is N/A rather than zero. Reference documents the assistant's answers are grounded in. They belong to the study, so every run of it uses the same set.
-
- 12 Attach a ready document to this study.
-
- 13 Discard changes and close.
-
- 14 Create this new experiment.