

Experiments

experiments-home.html · 1360×820

ChatMaestro 1/2026.08 2 PN priya.n

WORK

- Dashboard
- Experiments**
- Nudges
- Model Configs
- Cohorts

MONITOR

- Live Monitor 3
- Results
- Ask
- Saved Questions

COMMS

- Messages 2
- Documents

OVERSIGHT

- Audit Trail
- Issue Tracker

ADMIN

- Experimenters
- Enrollees
- Model Catalog

Experiments · 9 experiments

3 Import + New experiment

4 Search experiments 5 Cohort: All 6 Activity: All 7

Experiment	Task	Cohort	Runs	Active	Composite	Cost	Actions
☐ Terse-prompt study Open-topic chat	llm_chat	Pilot-20	4	YES running	73.9	\$2.52	
☐ Politeness · SLM ensemble Q&A session	llm_chat	Pilot-20	2	YES running	71.4	\$2.51	
☐ Redundant-prompt · Opus · hot Explainer chat	llm_chat	Group-C	7	YES paused	68.2	\$2.10	
☐ CoT-nudge study No runs yet	llm_chat	—	0	NO	—	—	
☐ Temp sweep · Haiku 4.5 Warm-up chat	llm_chat	Batch-7	11	NO	66.9	\$0.31	

Showing 1–9 of 9 Rows per page 25 1

Runs — Terse-prompt study 4 runs

Search runs Launch run

Run	Status	Started	Composite	Cost	Enrollees	Actions
Pilot-20 · nudge A	RUNNING	14:02 · 23m	73.9	\$2.52	20	
Pilot-20 · nudge B	DONE	Jul 16	71.2	\$2.51	20	
Group-C · nudge A	DONE	Jul 14	68.4	\$2.10	18	
Pilot-20 · baseline	ABORTED	Jul 12	—	\$0.31	6	

Annotations

- Version pill (right-aligned)** Dimmed build pill sits on the RIGHT, beside the account menu. Hover reveals the build sha; on phones it folds into the account menu.
- Account menu** Avatar + name + chevron open the account menu (profile, settings, sign out). On narrow screens the feedback / bug / dark-mode icons and version pill collapse into here.
- Import** Bulk-import experiments from CSV / XLSX / JSON — the shared ImportButton affordance. Primary (filled navy).
- Master-detail split (drag the divider)** Selecting an experiment fills the lower pane with ITS runs — a proper DataTable with its own search and per-run pause / abort. The bar between the panes drags up or down, so you size the list against the runs. This replaced the old right-hand runs drawer.
- Cohort filter** Filter experiments by the cohort a run targets. Added alongside the status filter so an experimenter can narrow to one enrolled group.
- Clear search (funnel)** A funnel-with-slash icon (IconFilterOff) resets the search box and every filter dropdown at once. Disabled when nothing is filtered, so it never gives a ‘did anything happen?’ non-response.
- Select columns** The eye icon opens a column-visibility menu (toggle, reorder, reset to default). Column choices persist per-user.
- Confined horizontal scroll** The scrollbar belongs to THIS viewport — only the middle columns scroll. The Select column (left) and Actions column (right) stay pinned; the Experiment column is pinned-left too, so the row's identity never scrolls away.

-
- 9 **Active — a DERIVED boolean** Active = the experiment has a run that is RUNNING or PAUSED. This replaces the earlier (wrong) per-experiment 'Status' column — status belongs to a run, not to the reusable experiment package.
-
- 10 The run's composite score (0–100) — a weighted blend of the rubric's quality factors, shown here for the experiment's latest run.
-
- 11 **Clear a run from its row** A finished or aborted run is cleared from here, without opening it: **Clean up** (recycle) reclaims the bulk and keeps the run and its `run_result` scorecard; **Purge** (flame, admin only) removes the run and the scorecard with it. Both open the same confirmations as the Run-detail Screen, both archive-first, and both are disabled while the run is running or paused. A run that has already been cleaned up shows Clean up disabled — its exchanges are gone (no `chat_round` rows), only the scorecard remains — while Purge stays available (ALGORITHMS §13). Live runs: pause, resume, abort. Finished runs: clean up (keep the scorecard), purge (admin only), open.