

Researcher onboarding

AIPLA researcher guide — R1

AIPLA project team

2026-07-14

Table of contents

1	What a researcher can do	1
2	Getting access	2
3	Cross-teacher observation	2
4	Rubric experimentation	3
5	Next steps	4

For researchers

A researcher account layers **cross-teacher, read-only** views and a **rubric experimentation** workspace on top of the normal teacher surfaces. You observe across every teacher and class; you do not edit their work. Researcher access is a role an administrator grants to your account (see *Getting access* below).

1 What a researcher can do

Two capabilities beyond a normal teacher account:

1. **Observe across the whole cohort** — every teacher's classes, activities, engagement and cost, read-only.
2. **Experiment with rubrics** — author and version the judge prompts (lenses) used to score student sessions, and run a judge over a captured session.

Everything a researcher sees across teachers is **observation only** — nothing in the cross-teacher views is editable.

2 Getting access

The researcher role is a claim an administrator sets on your account (via `aiplatform users grant-researcher <uid>`). Once granted, sign out and back in so your session picks up the new role. You will then see a **Research** item in the sidebar and extra controls described below. Without the role, these surfaces show a “Researcher access required” message.

3 Cross-teacher observation

- **Research (activity scan)** — the **Research** sidebar item opens a read-only scan of *every* teacher’s activities, in every state (Draft / Private / Shared), with the owner shown. Use it to see what is being built across the cohort.
- **Research view (classes)** — on **Classes**, a **My classes** / **Research view** toggle switches the list to every teacher’s classes; pick one to drill into its sessions.
- **Insights across all teachers** — on **Insights**, a **My classes** / **All teachers** scope toggle widens the engagement comparison to the whole cohort.
- **Cost** — a researcher-only **Cost** tab under Insights breaks cross-class spend down by cohort, model, voice (speech-to-text / text-to-speech) and class, over a period you choose.

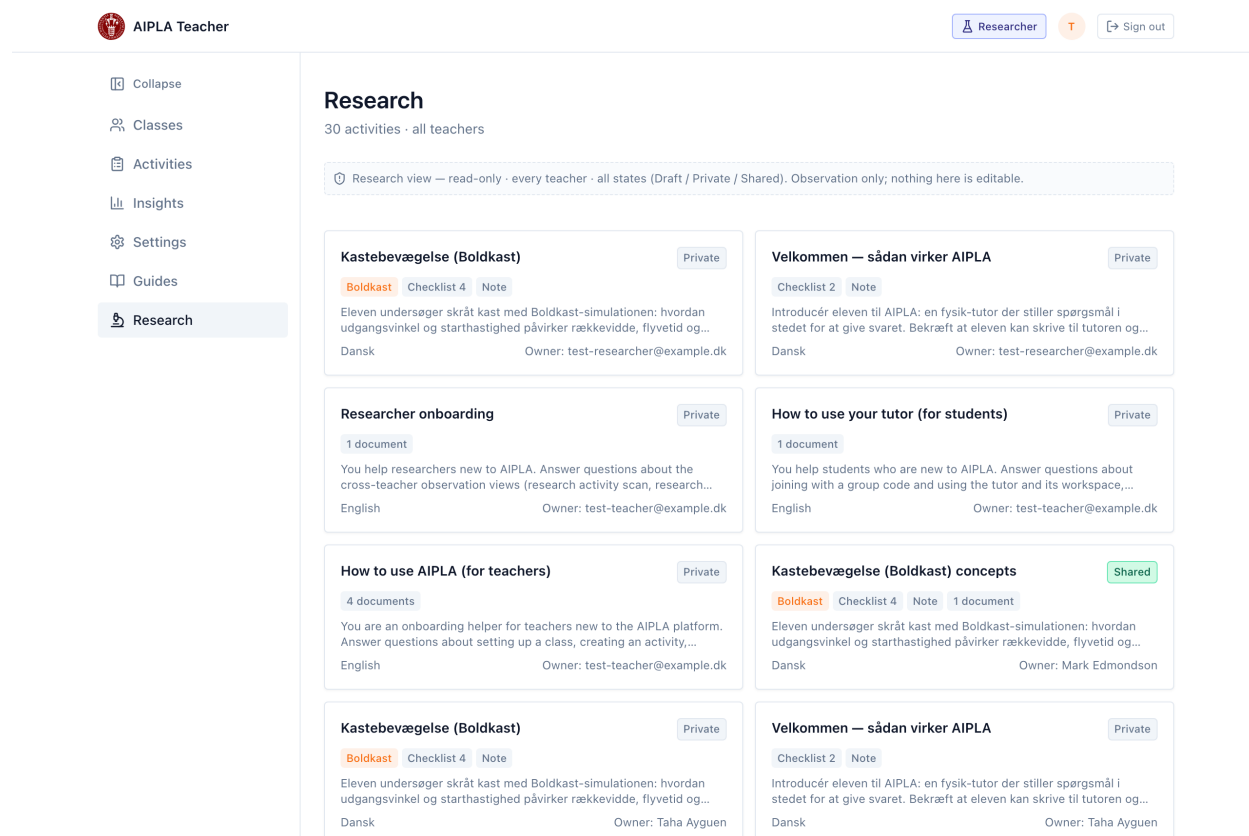


Figure 1: The cross-teacher research surfaces are read-only — observation across every teacher and class.

4 Rubric experimentation

Under **Settings**, the **Research · judge lenses** panel is where you shape and test how student sessions are scored:

- **Judge lenses** — for each lens you can enable/disable it, choose the **judge model**, inspect the **default prompt**, and write a **judge prompt override**. Saving an override **bumps its version**, so you can iterate on the wording and keep the history. **Reset to default** reverts it.
- **Experiment — score a captured session** — paste a session id, pick a lens, and **Run judge** to see the score profile that lens produces on real data (or an “Abstained” result when the lens declines to score).

Researcher

T

Sign out

score is stamped with the version that produced it.

MAPS problem solving (Docktor 2016)

maps-r2

enabled

Judge model

gemini-2.5-flash

► Default prompt — click to view

Judge prompt override (empty = the default above; saving bumps the version)

You are a physics-education research judge scoring a student's problem-solving PROCESS with the MAPS rubric. Score each category 0-5, or NA_problem (the problem doesn't elicit the category) or NA_solver (the student produced no independent evidence for it). Do NOT reward a correct final answer as such: expert problem solvers generate incorrect answers a significant fraction of the time – the categories score process quality. Apply the consistency rule: an early error used consistently afterwards is penalised once, not in every category it touches.

The override replaces only these instructions. The rubric categories, your anchor pack, the attribution, and the student's evidence are always appended automatically.

Save prompt

Reset to default

SAAR scientific abilities (Etkina 2006)

saar-r1

enabled

Judge model

gemini-2.5-flash

► Default prompt (currently in use) — click to view

Judge prompt override (empty = the default above; saving bumps the version)

Empty → the default prompt above is used. Click "Copy into editor" to start from it.

The override replaces only these instructions. The rubric categories, your anchor pack, the attribution, and the student's evidence are always appended automatically.

Save prompt

Reset to default

Experiment — score a captured session

session id

maps

Run judge

Figure 2: The rubric experimentation panel: versioned judge prompts and a judge you can run on a captured session.

This is the workflow for developing the capability-floor rubrics: change a lens prompt, run it against known sessions, compare the profiles, and promote a version when it behaves.

5 Next steps

- The teacher-facing surfaces work as documented in $T1-T4$; the researcher views sit alongside them.