A briefing for

STEM Faculty

An AI learning platform engineered to strengthen the reasoning habits of students in STEM coursework and provide faculty defensible evidence of where understanding breaks down.

Field	Detail
Audience	STEM faculty, course coordinators, and curriculum committees
Approach	Socratic dialogue, adaptive diagnosis, retrieval practice, spaced review, bounded to your course materials
Origin	A Harvard Innovation Labs venture
Founder	Kuljit Takhar

INTRODUCTION

A note to STEM faculty

You have spent the last two years watching something subtle happen to problem sets. The work submitted is often correct. It is sometimes more polished than what the same students produced a year ago. But the questions that arrive in office hours have flattened, and the moments of genuine confusion that used to lead to understanding have grown rarer.

Exam performance, lab reports, and weekly problems are giving noisier signals than they used to. Students can derive the right answer without having internalized the structure that makes the derivation possible. This shows up most acutely in courses where downstream success depends on the reasoning being load-bearing: organic chemistry, linear algebra, thermodynamics, algorithms. Anything where this week's concept is next month's prerequisite.

Classy AI was built in response to that drift. It sits at the intersection of two facts: Bloom's 1984 finding that one-to-one tutoring produces roughly two standard deviations of gain over conventional instruction, and the simple economic constraint that has kept that level of personalization out of reach until now. It is an attempt to deliver one-to-one tutoring without sacrificing the cognitive rigor that makes STEM learning transfer to new problems.

> *The friction is the feature. Productive struggle is what converts information into transferable understanding.*

2σ

Bloom's tutoring effect over classroom instruction

1:1

Student-to-tutor ratio, delivered at software economics

0

Times Classy hands a student the answer on demand

01 · THE PRODUCT

What Classy AI is

An infinitely patient tutor available 24/hrs a day.

Classy AI is a personalized learning platform that coaches students through their own reasoning rather than producing the reasoning for them. It is built on five pillars, each drawn from cognitive science and chosen because it has survived decades of replication.

Pillar	What it does
Socratic questioning	Guides students to construct ideas through dialogue, modeling the kind of one-to-one conversation a skilled TA or faculty mentor would lead at the whiteboard.
Adaptive scaffolding	Diagnoses what the student actually knows and rebuilds from the first broken link forward.
Retrieval practice	Closes each session with a dynamically generated, conversational quiz calibrated to demonstrated level.
Spaced repetition	Returns concepts at the intervals research shows are optimal for long-term retention.
Bounded learning	Grounds every response in your course materials rather than the open internet.

What a session looks like

A student in your linear algebra course opens Classy and types something simple: I don't understand why eigenvectors matter. Classy asks the student a question. The student answers. Classy asks another.

Within a few exchanges, it has located the actual point where understanding broke down. Perhaps a missing intuition about linear transformations from two semesters earlier, or a brittle grasp of basis change. It begins rebuilding from there.

After the dialogue, the student is offered a short, conversational quiz calibrated to the level they have just demonstrated. If they hesitate on a question, that hesitation is recorded. The next session begins where the last one ended.

02 · WHY THIS MATTERS FOR STEM

The cognitive stakes are higher in your courses

General-purpose generative AI is a particular problem for STEM in a way it is not for many other disciplines. The reason is structural: STEM curricula are stacked. Mastery of week 4 is the working memory of week 9. A student who outsources the derivation in October arrives in February without the schemas they need to even parse the new material, let alone reason about it.

Students using general-purpose AI to complete STEM coursework can produce correct artifacts while developing brittle, non-transferable understanding. The artifact looks like learning but it is far from it.

> *In STEM, the cost of a shortcut surfaces two semesters later, when the foundation that should have been built isn't there.*

Where Classy is architecturally different

Classy AI refuses to do the work for the student. This refusal is the design center of the product.

- No on-demand answers. Classy will not produce a solution to a problem the student has been assigned. It works with the student on the structure of the problem, the relevant principles, and the next defensible step.
- Prerequisite-first diagnosis. Before instruction begins, Classy probes prerequisite understanding through dialogue. A student asking about Lagrangian mechanics without a working grasp of partial derivatives gets the partial derivatives first.
- Retrieval as part of the loop. Quizzes are dynamically generated against the concept just discussed, calibrated to demonstrated level.
- Spaced return. Concepts the student has touched return days and weeks later at intervals the spacing literature identifies as productive.
- Bounded to your materials. Classy draws from the materials you have chosen for the course, limiting hallucinations and staying aligned with your lesson plan. The tutor speaks your course's notation and conventions, citing the section it is drawing from.

03 · BOUNDED LEARNING

The tutor is bounded to your materials.

Refusing to give the answer is one half of what makes Classy AI architecturally different. The other half is what the tutor draws on when it does respond. A general-purpose LLM has been trained on most of the public internet, which gives it broad coverage but introduces two specific problems in a course context. It has no commitment to your course's notation or framing, and the completeness of what it generates cannot be guaranteed against any particular reference.

Bounded learning is the architectural response. The tutor is given the materials you have chosen for the course, retrieves the passages relevant to the student's question, and is prompted to draw from those passages first. The pre-trained knowledge of the underlying model remains available as a fallback when the corpus does not cover the question, but it is no longer the primary source. The approach is drawn from the SocratiQ research framework (Jabbour et al., 2025), which documents the design choice directly: "relying solely on the pre-trained knowledge of the language model introduces the risk of generating content that is inconsistent or not related to the course content curated."

Why this matters in STEM

In a STEM course, the difference between generally correct and course-correct is the entire point. A linear algebra course that uses Strang's framing of eigenvectors as invariant directions will scaffold differently than one that uses Axler's coordinate-free approach. A thermodynamics course that derives entropy from statistical mechanics will scaffold differently than one that derives it from the Clausius inequality. A bounded tutor speaks the language of your lecture; an unbounded tutor speaks the language of the average textbook on the internet.

How the binding works

Commitment	What it does
Corpus scoping	Only documents shared with the class are retrievable. Solutions handouts, the answer key to next week's problem set, instructor notes — if they are not in the shared corpus, the tutor cannot draw from them.

In-context grounding	Retrieved passages are passed to the tutor as the primary source. The system prompt instructs the model that course materials take precedence, and that conflicts should be resolved in favor of the corpus.
Citation	When the tutor draws from a specific section, it names the section. Students can verify; instructors can audit.
Graceful fallback	When the curated corpus does not cover the question, the tutor falls back to general knowledge but signals that it is doing so.

> *Bounded learning is the architectural commitment that the tutor speaks the same language as your lecture.*

This is the architectural answer to the question faculty often ask first: how do I know the tutor won't teach my students the wrong thing? The answer is that you choose what it teaches from.

04 · THE EDUCATOR DASHBOARD

Defensible evidence of where understanding breaks down

The instrument many faculty have wanted for years.

Every Classy AI course comes with an educator dashboard built for instructors and institutions. It is a diagnostic instrument, engineered to feel like a tool a department chair could put in front of a curriculum committee. Four views, sequenced so that an instructor opening it for the first time lands on a view that orients them.

View	What it shows
Class overview	Stat cards for active members, average proficiency, total quizzes taken, active-this-week. Top five struggling concepts on the left, recent activity feed on the right.

Concepts	A single-column heatmap of every section across the course's shared documents, sorted by struggle score. Click any section to expand: every student who has engaged with it, sorted strugglers-first, alongside a Bloom's-taxonomy decomposition of how the class is engaging.
Students	Per-student summary stats with click-through to per-section progress, quiz performance, and session history. Used when exam performance has raised a question and you want a finer-grained read than the gradebook supports.
Settings	Class controls: rename, rotate join code, manage shared documents, manage roster, archive.

What gets measured, and the struggle score

Three signals drive everything on the dashboard: question volume per section (how much attention a piece of material is drawing), demonstrated proficiency per section (a decayed mean over recent quiz attempts, bounded to [0,1], with null for unmeasured rather than zero), and Bloom's-taxonomy level of engagement (whether the class is remembering, understanding, applying, analyzing, evaluating, or creating). The struggle score combines the first two:

```
struggle_score = (1 − avg_proficiency) × 0.6 +
normalized_question_volume × 0.4
```

Sixty percent on "how poorly is the class performing," forty percent on "how much attention is this drawing." The formula is pinned across phases so the dashboard does not silently change meaning between semesters. A faculty member can look at any number on the dashboard and trace it back to the quizzes and questions that produced it. (Full methodology, including what the architecture defensibly claims and what it does not yet claim, is available on request.)

> *A percentage grade tells you a student got 73%. A Bloom-level decomposition tells you which 27% they are missing and whether they are missing the easy part or the hard part.*

Why instructors and institutions care

- Lecture revision. Identify where the class is collectively stuck, by concept, before the next lecture. The struggle-sorted concept view answers one question fast: what should I cover differently on Tuesday?
- Targeted remediation. Two students with the same midterm grade can have very different gap profiles. The per-student view, decomposed by Bloom level, targets office-hour time where it will move the needle.
- Curriculum committee evidence. Concept-level engagement data across a cohort, semester over semester, gives departments a defensible empirical basis for curriculum redesign.
- Insight, not surveillance. The dashboard is read-only from the instructor side. No keystroke logging, no time-on-page, no individual chat transcripts surfaced to instructors. An instrument for the instructor, not a behavior-modification system for the student.

05 · IN YOUR COURSE

How Classy fits the work you are already doing

Classy AI works in concert with faculty Lecture, lab, recitation, and office hours remain the center of gravity. Classy meets students in the moments between — at 11pm before a problem set is due, on the bus, the morning of a midterm — when a question would otherwise go unasked.

Concrete instructor workflows

Use case	How it works
Pre-class preparation	Assign a short Classy session before a flipped-classroom discussion or problem session. Students arrive with prerequisite gaps already surfaced and partially closed.
Supplemental problem-set support	Students working through a problem set use Classy in place of an answer-key shortcut. The dialogue is structured to keep the reasoning load on the student.

Targeted remediation	Identify students whose exam performance suggests a specific gap; route them to a focused Classy sequence without pulling them from class or singling them out.
Independent study	Advanced students explore topics beyond the syllabus at their own pace, with the same Socratic discipline.

The research, briefly

Finding	Source
One-to-one tutoring produces ~2σ of gain over conventional instruction.	Bloom (1984)
Active retrieval, even unsuccessful, produces more durable learning than re-reading.	Roediger & Karpicke (2006)
Distributing review across time substantially increases long-term retention.	Cepeda et al. (2006)
Learning conditions that feel harder in the moment often produce stronger long-term learning.	Bjork (desirable difficulty)
Bounded, in-context grounding keeps an AI tutor anchored to curated course material.	Jabbour et al. (2025), SocratiQ (arXiv:2502.00341)

Built on a foundation of proven research, Classy AI scales high-impact instruction to reach every student in the classroom.

> *Classy AI serves as a force multiplier for the instructional team. While the professor remains the center of gravity for the course, the AI acts as a 24/7 expert mentor that democratizes high-level academic support.*

INVITATION

There is a third path

If you have read this far, you are likely one of the faculty we built Classy AI for: someone who refuses to treat the current moment as a forced choice between adopting tools that hollow out STEM learning and rejecting AI altogether. There is a third path. Classy AI is our attempt to walk it.

We are particularly interested in partnerships with STEM instructors willing to pilot Classy AI in a course and contribute to its evolution. A pilot can start at the scale of a single section, with the option to expand.

GET IN TOUCH

Website	stayclassy.ai
Incubated	Harvard Innovation Labs
Founder	Kuljit Takhar
Email	kt@stayclassy.ai
For faculty	Pilot programs and course partnerships available. Detailed methodology, security documentation, and FERPA materials on request.

> *Classy AI is on a mission to democratize the power of personalized STEM instruction for every student, everywhere.*