



Unravix
Find the knot. Watch it click.

TECHNICAL WHITE PAPER · PRODUCT THESIS

Unravix: A Misconception-First Learning System

The Confusion Graph, diagnose-first answers and verified AI on free infrastructure

"AI search hands you an answer.
Unravix finds the wrong idea behind yours."

Swanix

Version 1.0 · October 2026

unravix.pages.dev

Abstract

Large language models have made correct-sounding answers available to every student instantly, yet decades of research on misconceptions show that students can reproduce a correct answer while keeping the faulty mental model that produced their error. This white paper presents **Unravix**, a free, web-based learning system that indexes learning by misconceptions (“knots”) rather than by questions. Its central data structure, the **Confusion Graph**, links each knot to subjects, chapters and exam levels, to the prerequisite ideas it rests on, and to the knot students usually hit next. Every answer follows a **diagnose-first** protocol grounded in conceptual-change theory: a three-option probe the student must answer first, a named knot, a click-moment explanation that starts from the wrong belief, an analogy, a transfer check in a new situation, and a different-angle retry.

The system runs entirely on free infrastructure: a static site on Cloudflare Pages, a Cloudflare Worker agent that rotates across a 32-entry pool of free model providers with circuit breakers, cross-family verification with a verified-only answer cache and student flags, library and queue fallbacks, and a Supabase back end with row-level security and AI-moderated forums. We describe the literature behind each design decision, the competitive landscape, the system design and specifications, and the implementation, including a launch audit and load test. In that test cached answers returned in a median of 1.22 s, while fresh answers under concurrent load reached the 39 s time budget. We then set out a pre-registered-style evaluation plan (A/B, pre/post concept inventory, verification audit), a roadmap toward a peer “click moment” explanation economy, and the ethical limits of the approach. No learning-outcome data are reported; all outcome claims are left to the planned studies.

Keywords: misconceptions · conceptual change · diagnostic assessment · distractor-driven items · generative AI in education · LLM verification · provider failover · knowledge graph · JEE · NEET · GATE · educational technology

Contents

1	Introduction and Motivation	1
1.1	Answers are everywhere; understanding is not	1
1.2	Exam pressure rewards the memorised answer	1
1.3	The wrong mental model survives the right answer	2
1.4	Why Unravix was made	2
1.5	Contributions	2
1.6	How this document is organised	3
2	Literature Review	4
2.1	Misconceptions, preconceptions and knowledge in pieces	4
2.2	Conceptual change	4
2.3	Measuring misconceptions: inventories and distractor-driven diagnosis	5
2.4	Interactive engagement and peer explanation	5
2.5	Tutoring and the two-sigma problem	6
2.6	Retrieval, guessing first and productive failure	6
2.7	Feedback	6
2.8	Generative AI in learning: promise and failure modes	7
2.9	Synthesis	7
3	Problem Statement and Research Questions	8
3.1	Problem statement	8
3.2	Research questions	8
3.3	Scope and assumptions	9
4	Competitive Landscape	10
4.1	Four families of study help	10
4.2	Capability matrix	11
4.3	Where competitors are stronger	11
4.4	Positioning	11
5	System Design	12
5.1	Design principles	12
5.2	The Confusion Graph	12
5.3	Knot schema	14
5.4	The diagnose-first pipeline	14
5.5	Verification and trust model	16
5.6	Provider rotation and failover	17
5.7	Data model	19
5.8	Deployment architecture	20

5.9	Privacy by design	21
6	Product Specifications	23
6.1	Feature set	23
6.2	Coverage: every level, every country	24
6.3	User journey	25
6.4	Platforms	25
6.5	Performance	26
6.6	Accessibility and theming	26
6.7	Limits and quotas	26
7	Implementation	28
7.1	Technology stack	28
7.2	Evolution of the Ask agent	28
7.3	Content pipeline	29
7.4	Launch audit	29
7.5	Load test	30
7.6	Lessons from implementation	30
8	Evaluation Plan	31
8.1	Metrics	31
8.2	Studies	32
8.3	Research ethics	33
9	Goals and Roadmap	34
9.1	Goals	34
9.2	Roadmap	34
9.3	From wedge to global	35
9.4	The peer explanation economy	35
9.5	Sustainability	35
10	Ethics, Safety and Limitations	36
10.1	AI errors	36
10.2	Children and self-declared age	36
10.3	Data sent to AI providers	36
10.4	Community safety	36
10.5	Academic integrity	37
10.6	Curricular and cultural bias	37
10.7	Dependence on free infrastructure	37
10.8	Other limitations	37
11	Conclusion	38
	References	39
	Appendix A - Sample Curated Knots	41
	Appendix B - Glossary	43

List of Figures

2.1	Posner et al.'s (1982) four conditions for conceptual change and the Unravix step that meets each one.	5
4.1	Capability matrix of eight study tools	11
5.1	The curated core of the Confusion Graph: 42 hand-written knots arranged from subject to chapter to knot (data: knots.json).	13
5.2	Knot taxonomy: curated knots by chapter, coloured by subject	13
5.3	Anatomy of a curated knot	14
5.4	The diagnose-first pipeline as a swimlane	15
5.5	Verification flow: how verifier outcomes map to the Verified and Needs review badges and to the cache.	16
5.6	Request failover cascade from cache to provider tiers, library and queue, with circuit-breaker rules and the time budget.	18
5.7	The provider pool (32 entries) by model family, tier and trust role, as configured in the Worker on 18 October 2026.	19
5.8	Data model	20
5.9	Deployment architecture: static site, Worker agent, relay, database and provider pool	21
6.1	Unravix by the numbers at launch	24
6.2	Levels per category in the level catalogue (levels.json).	24
6.3	Curated knots by subject (inner ring) and exam tag (outer ring).	25
6.4	First-session user journey	25
7.1	Ask latency per request in the launch load test (work/launch/load_results.json).	30
8.1	Planned two-arm study design with pre-, post- and delayed tests	32
9.1	Roadmap ribbon	35

List of Tables

2.1	Research findings and the Unravix design decisions they support	7
5.1	Fields of a curated knot record (knots.json)	14
5.2	Fields of a diagnose-first answer object (agent v4)	15
5.3	Components and flows in Figure 5.9	21
6.1	Live features at launch (18 October 2026)	23
6.2	Measured performance at launch (launch audit, headless Chromium, 18 October 2026)	26
6.3	Service limits	26
7.1	Technology stack	28
7.2	Launch audit summary (18 October 2026)	29
8.1	Evaluation metrics	31

Introduction and Motivation

Every major AI system can now hand a student a fluent, confident answer to almost any question in seconds. This document argues that, for learning, the answer is rarely the scarce resource. What is scarce is a precise account of why a particular student got it wrong, and a reliable way to change that reason. Unravix is built around that claim. Its tagline states the thesis in one line: “AI search hands you an answer. Unravix finds the wrong idea behind yours.”

1.1 Answers are everywhere; understanding is not

Between 2023 and 2025 the general-purpose AI assistants moved into homework at scale. Their own makers acknowledged the risk that students would use them to skip thinking: OpenAI introduced a “study mode” in ChatGPT on 29 July 2025, describing it as a way to work through problems “step by step instead of just getting an answer” (OpenAI, 2025), and Google followed with “Guided Learning” in Gemini on 6 August 2025 (Google, 2025). Both are opt-in modes layered over an answer engine.

The most careful evidence so far suggests the default behaviour matters. In a pre-registered field experiment with nearly 1,000 high-school mathematics students in Turkey, access to a standard GPT-4 chat interface raised scores on assisted practice by 48%, yet when access was removed the same students scored 17% *lower* on an unassisted exam than students who never had the tool. A tutor variant whose prompt carried teacher-written solutions and common mistakes, and which was told to give hints rather than answers, largely removed that harm (Bastani et al., 2025). The authors traced the damage to students using the plain chatbot as a “crutch” (asking for and copying solutions), and they also reported that GPT-4 answered the practice problems correctly only 51% of the time when asked directly. Two lessons follow for any learning product built on large language models: the interaction design decides whether learning happens, and the model's own answers cannot be trusted without a check.

1.2 Exam pressure rewards the memorised answer

High-stakes entrance examinations sharpen the problem. Unravix launches first for students preparing for India's engineering, medical and postgraduate engineering entrance examinations (JEE, NEET and GATE), where months of practice on past-paper patterns are the norm. The same pattern exists for the SAT and AP in the United States, A-Levels and GCSEs in the United Kingdom, the IB Diploma and China's Gaokao. Under time pressure it is rational to memorise solution templates. The cost appears later: a student who has memorised that “at the top of a throw the ball stops” can reproduce the textbook answer about velocity and still believe that the acceleration there is zero, and that belief resurfaces in every problem that touches it.

1.3 The wrong mental model survives the right answer

Physics education research documented this decades ago. Halloun and Hestenes (1985) showed that students arrive at university mechanics with strongly held commonsense beliefs about motion, and that conventional instruction barely changes them. The Force Concept Inventory, built from those beliefs, made the gap measurable across thousands of students (Hestenes et al., 1992; Hake, 1998). Later work explained why some misconceptions are so robust: they are not isolated factual errors but parts of a coherent, wrongly categorised mental model (Chi, 2005; Vosniadou, 1994). Re-reading the correct explanation does not touch the faulty model; the student has to see the model fail and replace it.

This is the gap Unravix targets. Students do not usually lack the correct sentence. They hold a specific wrong idea, a *knot*, that keeps producing wrong answers in new situations. A system that answers questions treats each question as new. A system that diagnoses knots can recognise that a hundred different questions are symptoms of the same few beliefs.

1.4 Why Unravix was made

Unravix began as an entry in its creator's idea log on 18 October 2026, under the working title “Click: the Confusion Graph”. An earlier idea from the same day, an AI oral viva that tests whether students understand their own work, was rejected after a scan found at least nine products already in that space. The Confusion Graph idea survived the same scan: index learning by *stuck points* (misconceptions) rather than by questions; when a student finally gets it, capture a short “what made it click” explanation; route the next student stuck at the same point to the explanations that measurably unstuck people like them; and credit the authors each time their explanation works. The long-term asset is a map of where humans get stuck and what fixes it, which no answer engine collects.

An earlier candidate name was rejected before *Unravix* was chosen, and the product was built and launched in October 2026 as a free web application at `unravix.pages.dev`. A decision recorded in the idea log shaped the order of work: launch on the web first and earn traffic, and build the mobile app later. The site now announces Android and iOS apps for 2027.

1.5 Contributions

This document makes five contributions:

1. A **product thesis**. It frames AI-era learning as a misconception problem rather than an answer problem and grounds that framing in conceptual-change research (Chapters 2 and 3).
2. **The Confusion Graph**. It defines a data model that treats misconceptions as first-class nodes, with a concrete schema already in production (Chapter 5).
3. A **diagnose-first answer protocol**. Every answer is a structured object (probe, knot, click-moment explanation, analogy, transfer check, retry, roots, next knot) rather than free text (Chapter 5).
4. A **trust architecture on free infrastructure**. It combines cross-family verification, a verified-only cache, student flags, and a 32-entry rotation over free model providers with circuit breakers and graceful fallbacks (Chapters 5 and 7).
5. An **evaluation plan**. It specifies metrics and a pre-registered-style A/B and pre/post concept-inventory study. No outcome data exist yet, and nothing in this document should be read as a learning-gain result (Chapter 8).

1.6 How this document is organised

Chapter 2 reviews the research on misconceptions, conceptual change, diagnostic testing, tutoring, retrieval practice and feedback. Chapter 3 states the problem and research questions. Chapter 4 maps the competitive landscape. Chapters 5 to 7 describe the system's design, specifications and implementation. Chapter 8 sets out the evaluation plan, Chapter 9 the goals and roadmap, and Chapter 10 the ethics, safety measures and limitations. Chapter 11 concludes. The appendices give a sample of the 42 hand-written knots and a glossary.

Evidence policy. Facts about Unravix come from its source code, configuration and launch records as of 18 October 2026. Research claims are cited to the original peer-reviewed sources, each of which was checked against its DOI record. Anything planned is labelled as planned.

Literature Review

Unravix borrows almost nothing new from cognitive science. Its design choices were arrived at by building and testing, but each one lines up with a long research tradition that predates generative AI by decades. This chapter reviews that tradition and ends with a table linking each finding to the design decision it supports.

2.1 Misconceptions, preconceptions and knowledge in pieces

The term *misconception* covers ideas that learners hold before or despite instruction and that conflict with accepted science. Halloun and Hestenes (1985) measured the “initial knowledge state” of college physics students and found commonsense beliefs about force and motion, close to pre-Newtonian impetus ideas, that were widespread, stable, and only weakly affected by conventional lectures. Their diagnostic test became the ancestor of the Force Concept Inventory.

Not everyone agrees that misconceptions are coherent theories. diSessa (1993) argued that intuitive physics consists of many small, self-explanatory pieces (“phenomenological primitives”, such as *more effort gives more result*) that are activated by context. Smith, diSessa and Roschelle (1994) went further: misconceptions are not simply errors to be swapped out but raw material that expert understanding grows from, so instruction should refine them rather than attack them. For a product, both views matter. The coherent-theory view says a single knot can drive many wrong answers, so diagnosing the knot pays off. The knowledge-in-pieces view says the same student may hold a knot in one context and not in another, so diagnosis has to be checked again in a new situation. Unravix's transfer check (Chapter 5) exists for that reason.

2.2 Conceptual change

Posner, Strike, Hewson and Gertzog (1982) proposed that replacing a conception requires four conditions: the learner must be *dissatisfied* with the existing conception, and the new one must be *intelligible*, *plausible* and *fruitful*. The model has been criticised as too rational, but it remains a useful checklist for designing an intervention, and Unravix's answer protocol follows it closely (Figure 2.1).

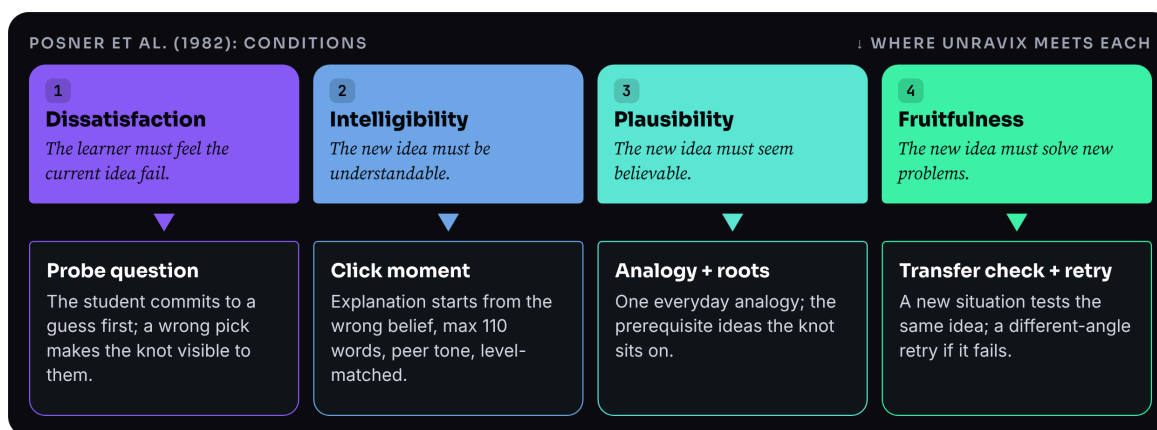


Figure 2.1. Posner et al.'s (1982) four conditions for conceptual change and the Unravix step that meets each one.

Vosniadou and Brewer (1992) studied children's mental models of the Earth and found *synthetic models*, such as a hollow Earth people live inside or a flattened disc, that blend instruction (the Earth is round) with an underlying framework (things fall down; the ground is flat). Vosniadou (1994) generalised this into a framework-theory account: misconceptions arise when new information is assimilated into constraining presuppositions that the learner is not aware of. The implication for design is that an explanation must start from what the learner currently believes, not from the correct statement. Unravix's generation prompt encodes this literally: the click-moment answer “starts from the student's likely wrong belief”.

Chi, Slotta and de Leeuw (1994) located the hardest misconceptions in category errors, where a process such as heat flow or current is treated as if it were a substance. Chi (2005) showed that misconceptions about emergent processes are especially robust because learners assign them to the wrong ontological category, so no amount of detail within that category fixes them. Several of Unravix's curated knots are of this kind, for example “current gets used up in a bulb” and “a reaction stops at equilibrium”. For these, the protocol asks for an analogy and a retry from a *different* angle rather than a reworded repeat.

2.3 Measuring misconceptions: inventories and distractor-driven diagnosis

The Force Concept Inventory (Hestenes, Wells & Swackhamer, 1992) is a multiple-choice test whose wrong options are the commonsense beliefs found by earlier interview studies. A wrong answer therefore names a belief, not just an error. Treagust (1988) described a procedure for building such diagnostic instruments in any science topic, including two-tier items that ask for an answer and then a reason. Sadler (1998) analysed responses to distractor-driven items with item-response models and showed that the attractiveness of particular misconception distractors follows systematic patterns as ability grows. These items reveal something that open-ended marking misses.

The same principle now runs at classroom scale in commercial tools. Eedi, a UK mathematics platform, builds diagnostic questions in which “every wrong answer is engineered ... to reveal a specific misconception” (Eedi, n.d.-a), and an independent randomised trial reported additional mathematics progress for students using it (Eedi, n.d.-b). Unravix's 42 curated knots (sampled in Appendix A) use the same design: three options, one correct, and at least one distractor that is the misconception (Figure 5.3 in Chapter 5).

2.4 Interactive engagement and peer explanation

Hake (1998) pooled pre- and post-test inventory data from 62 introductory physics courses with 6,542 students. Courses that relied on traditional lectures reached an average normalised gain of 0.23 ± 0.04 , while courses that made substantial use of interactive-engagement methods reached 0.48 ± 0.14 , roughly two stan-

dard deviations apart. The normalised gain, the share of the possible improvement that a class actually achieves, became the standard metric for concept inventories, and Unravix's evaluation plan adopts it (Chapter 8).

One of the most studied interactive methods is Peer Instruction, in which students answer a conceptual question, discuss it with a neighbour, and answer again. After ten years of use at Harvard, Crouch and Mazur (2001) reported consistent conceptual gains. Peer explanations work partly because a student who has just understood an idea remembers what the confusion felt like. Self-explanation research supports the same mechanism from the other side: students who explain worked examples to themselves learn more than students who only read them (Chi, Bassok, Lewis, Reimann & Glaser, 1989). The Confusion Graph's long-term vision, capturing a short explanation from a student at the moment it clicks and serving it to the next stuck student, combines these two findings. That vision is not yet live (Chapter 9).

2.5 Tutoring and the two-sigma problem

Bloom (1984) reported that students taught one-to-one by a tutor performed about two standard deviations above students in a conventional class, so that the average tutored student outperformed 98% of the comparison group. He posed the “2 sigma problem”: find group methods that come close. VanLehn (2011) re-examined the evidence and found that the effect of human tutoring was much smaller than commonly believed ($d = 0.79$), and that step-based intelligent tutoring systems were nearly as effective ($d = 0.76$). The useful lesson is that the benefit comes from fine-grained interaction (diagnosing and responding to each step) rather than from the human as such. That is encouraging for software, provided the software diagnoses rather than merely answers.

2.6 Retrieval, guessing first and productive failure

Roediger and Karpicke (2006) showed that taking a test enhances later retention more than restudying the same material: after five minutes restudy looked better, but after two days and one week prior testing produced substantially greater retention, even though restudy made students more confident. Kornell, Hays and Bjork (2009) found that even *unsuccessful* retrieval attempts, where the learner guesses wrong before seeing the answer, enhance subsequent learning. Kapur (2008) reported that students who first struggled with complex, ill-structured problems, failing in the process, later outperformed peers on well-structured and transfer problems. In their review of learning techniques, Dunlosky, Rawson, Marsh, Nathan and Willingham (2013) rated practice testing as one of only two techniques with high utility.

Together these findings justify the most visible and least intuitive feature of Unravix: the student must commit to a guess on a probe question before any explanation appears, even when they asked a direct question.

2.7 Feedback

Feedback is powerful but not reliably positive. Kluger and DeNisi (1996), in a meta-analysis of 607 effect sizes, found that feedback interventions improved performance on average ($d = 0.41$) but that over one-third of them *decreased* performance. Hattie and Timperley (2007) proposed a model in which effective feedback answers three questions (Where am I going? How am I going? Where to next?) and works at four levels (task, process, self-regulation and self), with praise directed at the self being least useful. Black and Wiliam (1998) reviewed formative assessment and concluded that feedback which tells students how to improve, rather than a mark, raises achievement. Unravix maps onto Hattie and Timperley's three questions: the knot names where the learner stands, the click moment and analogy work at the process level, and the *next knot* and *roots* fields answer “where to next?”. The Trap Report at the end of a Knot Scan names specific traps rather than giving a score.

2.8 Generative AI in learning: promise and failure modes

Large language models make it possible to generate a diagnosis, explanation and check for any question at any level. They also hallucinate: they produce fluent statements that are unsupported or false, a failure catalogued across tasks by Ji et al. (2023). In education the risk is compounded because the reader, by definition, cannot yet judge correctness. Bastani et al. (2025) add the behavioural risk: without guardrails students outsource the thinking. Two design consequences follow. First, generation must be constrained into a pedagogical structure that forces engagement, so that the student guesses, then reads, then is checked. Second, every generated answer needs an independent check before it is trusted or reused. Unravix implements both, and is explicit that its automated “Verified” badge is not a substitute for a teacher (Chapter 10).

2.9 Synthesis

Table 2.1 summarises how the research translates into product decisions.

Table 2.1. Research findings and the Unravix design decisions they support

Finding	Source	Design decision in Unravix
Conceptual change needs dissatisfaction, intelligibility, plausibility, fruitfulness	Posner et al. (1982)	Probe → click moment → analogy and roots → transfer check
Misconceptions are often coherent, framework-level beliefs	Vosniadou (1994); Chi (2005)	Explanations start from the student's wrong belief; the retry must take a new angle
Intuitive knowledge is context-dependent	diSessa (1993); Smith et al. (1994)	The check always uses a new situation
Distractors built from misconceptions diagnose beliefs	Hestenes et al. (1992); Treagust (1988); Sadler (1998)	Every probe and curated knot has three options, one of which is the misconception
Interactive engagement roughly doubles normalised gain	Hake (1998)	No passive reading: guess, read, then check
Peers and self-explanation help understanding	Crouch & Mazur (2001); Chi et al. (1989)	Planned click-moment capture and peer ranking
Fine-grained tutoring approaches human tutoring	Bloom (1984); VanLehn (2011)	Per-knot diagnosis instead of whole-answer delivery
Testing and even failed guesses improve retention	Roediger & Karpicke (2006); Kornell et al. (2009); Kapur (2008)	Commit-before-explain probe; Knot Scan
Feedback can hurt unless it says where to next	Kluger & DeNisi (1996); Hattie & Timperley (2007)	Trap Report names traps, not scores; next_knot and roots fields
Unguarded AI access can harm learning; LLMs hallucinate	Bastani et al. (2025); Ji et al. (2023)	Structured protocol plus cross-family verification and a verified-only cache

03

Problem Statement and Research Questions

3.1 Problem statement

Students at every level (school, entrance examinations, university and job interviews) now have instant access to AI-generated answers, but answers do not repair the specific mistaken ideas that cause repeated errors. Existing tools are organised around *questions* (Q&A sites, solvers, chat assistants) or *content* (courses, videos), not around *misconceptions*. As a result:

1. the same misconception is explained thousands of times, each time from scratch and usually starting from the correct statement rather than from the student's wrong belief;
2. students are rarely made to commit to a prediction, so they cannot notice that their current model fails;
3. whether the explanation worked is almost never checked in a new situation;
4. AI-generated explanations reach students without an independent accuracy check; and
5. no shared dataset records where learners get stuck, in what order, and what fixes it.

Unravix is designed to address all five while running entirely on free infrastructure, so that it can be offered at no cost to students worldwide.

3.2 Research questions

The project is framed by four research questions. RQ1 and RQ2 concern design and are answered descriptively in this document. RQ3 and RQ4 are empirical, and Chapter 8 sets out how they will be answered.

RQ	Question	How it is addressed
RQ1	Can learning content be organised around misconceptions (knots) as first-class objects, across subjects and levels from Class 1 to PhD?	Confusion Graph data model; 42 curated knots; AI-generated library across 329 levels (Ch. 5–7)
RQ2	Can a generative-AI pipeline produce diagnose-first answers (probe, knot, click moment, check, retry) reliably enough to show students, on free infrastructure?	Shape check, cross-family verification, verified-only cache, fallbacks; launch load test (Ch. 5, 7)
RQ3	Do diagnose-first answers produce larger and more durable conceptual gains than answer-first explanations of the same content?	Planned randomised A/B with pre/post concept inventory and delayed test (Ch. 8)
RQ4	Do peer “click moment” explanations, ranked by measured unsticking, outperform AI explanations for the same knot?	Planned once click capture ships (Ch. 8–9)

3.3 Scope and assumptions

This document covers the web product as launched on 18 October 2026. It assumes that students will use Unravix voluntarily and briefly, often on mobile phones and shared networks, so every interaction must work without an account and degrade gracefully when free AI quotas run out. The product is global; India's JEE, NEET and GATE students are the launch wedge, not the boundary. Outcome claims are explicitly out of scope until the studies in Chapter 8 report.

04

Competitive Landscape

The market for study help is crowded, but it is crowded around answers. This chapter groups the main alternatives by what they organise learning around, scores them against the capabilities that matter for misconception repair, and is explicit about where competitors are stronger.

4.1 Four families of study help

General AI assistants with learning modes. ChatGPT's study mode (launched 29 July 2025) guides users step by step, asks questions and checks understanding, and is available across plans on web and mobile (OpenAI, 2025). Gemini's Guided Learning (launched 6 August 2025) breaks problems down, adapts explanations, and can generate quizzes and study guides from uploaded material (Google, 2025). Both cover every subject and have very large audiences. In both, diagnosis is optional behaviour within a general chat; neither product exposes the misconception as an output, and neither publishes a per-answer verification signal.

Purpose-built AI tutors. Khanmigo, from the non-profit Khan Academy, is built into Khan Academy's content library and explicitly “doesn't just give answers”; it guides learners to find answers themselves. It is free for teachers and costs US\$4 a month or US\$44 a year for learners and parents, and access for under-18s requires a parent or a partnered school district (Khan Academy, n.d.). Its strengths are curated content and a trusted brand.

Solvers and Q&A platforms. Photomath scans a maths problem with the phone camera and shows step-by-step solutions from elementary maths through calculus; it joined Google in 2023 (Photomath, n.d.; European Commission, 2023). Brainly began in 2009 as a peer-to-peer Q&A community and added an AI Tutor and AI-powered answers in 2023, built on a knowledge base it described as more than 250 million answers, alongside human “Expert-Verified” answers (Brainly, 2023a, 2023b). Doubtnut, an Indian app, matches a photographed question against a repository of answered questions and returns a video solution. Allen Career Institute acquired it in December 2023, when it reported a monthly reach of 32 million students (Business Standard, 2023). These tools are fast and familiar. They organise help around the question.

Misconception diagnostics. Eedi is the closest conceptual neighbour. Its diagnostic questions are multiple-choice items in which each wrong option is designed to reveal a specific misconception, followed by a misconception explanation, an intervention and a check-out question (Eedi, n.d.-a, n.d.-c). It targets mathematics for roughly ages 9 to 16 and is sold through schools and, increasingly, as infrastructure for other platforms. An independent randomised trial is cited on its site (Eedi, n.d.-b). Eedi shows that misconception-first assessment works and has a market. Unravix differs in scope (every subject and level, including exams and job interviews), in being student-led and free with no account required to ask, and in generating diagnoses on demand for questions students actually type, speak or photograph.

4.2 Capability matrix

Figure 4.1 scores eight products on seven capabilities. Scores come from each product's public description. Anything not publicly documented is scored “No”, which may understate some products.

	Names the misconception	Guess before explanation	Transfer check after	Per-answer verification signal	Beyond maths	Peer explanations ranked by unsticking	Misconception data as core asset
ChatGPT Study mode	Partial	Partial	Partial	No	Yes	No	No
Gemini Guided Learning	Partial	Partial	Partial	No	Yes	No	No
Khanmigo	Partial	Yes	Partial	No	Yes	No	No
Photomath	No	No	No	n/a	No	No	No
Brainly	No	No	No	Partial	Yes	Partial	No
Doubtnut	No	No	No	No	Yes	No	No
Eedi	Yes	Yes	Yes	n/a	No	No	Yes
Unravix	Yes	Yes	Yes	Yes	Yes	Planned	Partial

Partial = an optional mode, or not the primary output · **n/a** = solver or expert-authored content, not generated per request · **Planned** = Unravix vision, not live · not publicly documented = No. Verification signal: Brainly marks some answers Expert-Verified (human); Unravix runs an automated second model on every generated answer.

Figure 4.1. Capability matrix of eight study tools. Unravix's peer ranking is planned, not live, and its misconception data asset is early-stage, so both are scored accordingly.

4.3 Where competitors are stronger

- **Distribution and trust.** ChatGPT, Gemini and Photomath reach very large audiences, and Khan Academy and Eedi carry institutional trust and published evidence. Unravix launched with no audience and no outcome data.
- **Content quality control.** Khanmigo is grounded in curated lessons and Eedi in expert-authored questions. Only 42 of Unravix's knots are hand-written; the rest are AI-generated and labelled *Needs review*.
- **Determinism.** For symbolic maths, a solver such as Photomath is more reliably correct than any generated explanation.
- **Fast followers.** The large AI labs can add misconception diagnosis to their study modes. Unravix's defensible position therefore cannot be the prompt. It has to be the accumulated graph of knots, the verified answer cache, and, once built, a network of peer explanations ranked by what actually unsticks students.

4.4 Positioning

Unravix positions itself as *the misconception layer*: a search engine runs on questions and returns an answer, while Unravix takes a question, returns the wrong idea behind it, checks that it has been repaired, and records what worked. The product's own comparison states the difference in student language: “Answers are everywhere. The knot map is not.”

05

System Design

This chapter describes the design of Unravix as built: the Confusion Graph and its knot schema, the diagnose-first answer pipeline, the verification and trust model, the provider rotation that keeps it running on free tiers, the data model, the deployment architecture, and the privacy decisions built into each part.

5.1 Design principles

1. **Diagnose before explaining.** The first thing a student sees after asking is a probe question, never the explanation.
2. **Structure over prose.** Every answer is a typed JSON object that the interface renders and the verifier can check field by field.
3. **Trust is earned per answer.** Nothing generated is cached or labelled *Verified* until a model from a different family has checked it.
4. **Never a dead end.** When providers fail, the system falls back to the library, then to a queue, rather than showing an error.
5. **Zero cost, zero login to learn.** The core loop runs on free tiers and needs no account.
6. **Collect the minimum.** No email is needed to ask or to post, and only hashed IP addresses are stored for rate limits.

5.2 The Confusion Graph

The Confusion Graph is the central data structure. Its nodes are *knots*: specific, testable wrong beliefs, each written in the student's voice ("If velocity is zero at an instant, acceleration must be zero too"). A knot belongs to a subject, a chapter and one or more exam levels, and it carries a diagnostic item. Generated answers add two kinds of edges: *roots*, the deeper or prerequisite ideas a knot sits on, and *next_knot*, the knot students usually hit next. Answers in the cache and items in the knot library are linked to knots through full-text search, so a question in new words can still reach a known knot. The planned peer layer adds a third kind of edge: from a knot to the click-moment explanations that resolved it, weighted by measured un-sticking (Chapter 9).

Figure 5.1 draws the curated core of the graph, the 42 hand-written knots, as a radial hierarchy from subject to chapter to knot. Physics accounts for 20 of them, chemistry 10, mathematics 7 and biology 5. Figure 5.2 shows the same knots by chapter.



Figure 5.1. The curated core of the Confusion Graph: 42 hand-written knots arranged from subject to chapter to knot (data: knots.json).

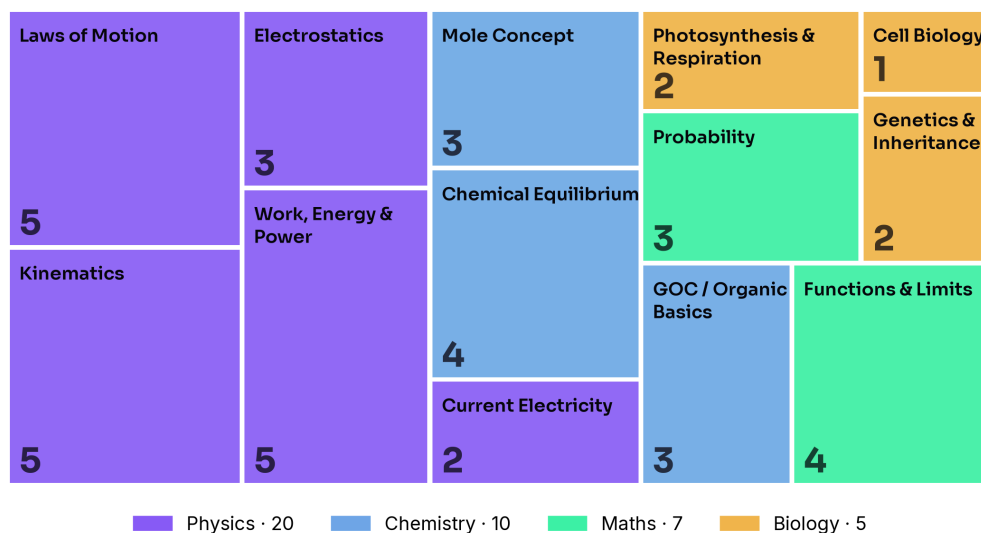


Figure 5.2. Knot taxonomy: curated knots by chapter, coloured by subject. Area is proportional to the number of knots.

5.3 Knot schema

A curated knot is a compact record (Figure 5.3; Table 5.1). Its probe is written so that the attractive wrong option is the misconception, following the distractor-driven tradition reviewed in Chapter 2. In the database each knot is stored with its category and level, a curated flag that is true for the 42 seeds, and a generated full-text vector over subject, chapter, label, knot, question and fix, indexed for search.

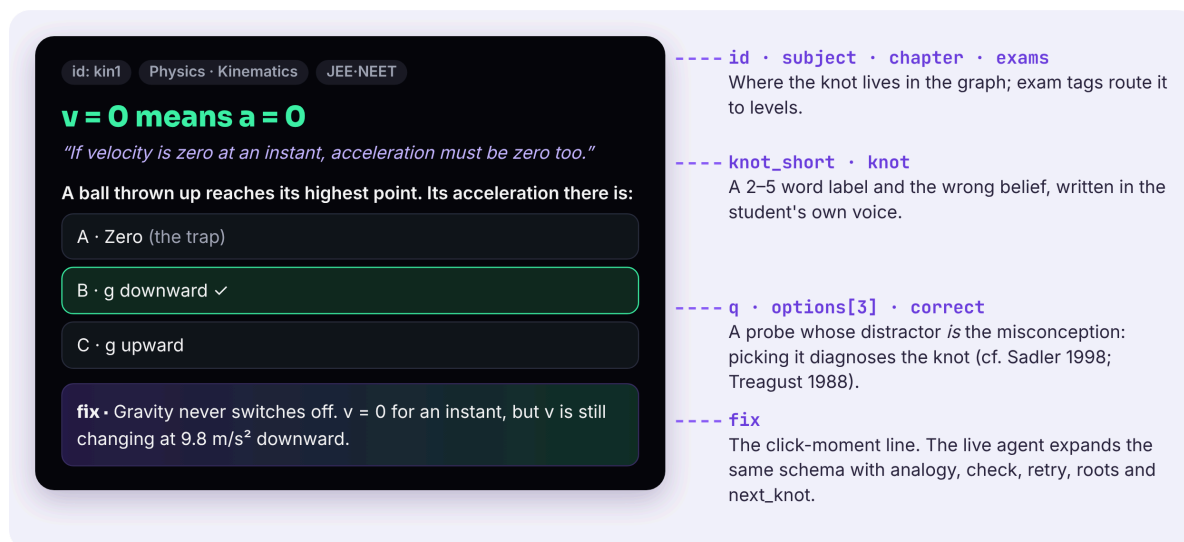


Figure 5.3. Anatomy of a curated knot. The attractive wrong option is the misconception itself.

Table 5.1. Fields of a curated knot record (knots.json)

Field	Type	Purpose	Example (kin1)
id	string	Stable identifier	kin1
subject / chapter	string	Position in the graph	Physics / Kinematics
exams	string	Level routing	JEE-NEET
knot_short	string, 2–5 words	Label shown on chips and reports	v = 0 means a = 0
knot	string	The wrong belief, in the student's voice	If velocity is zero at an instant, acceleration must be zero too.
q	string	Diagnostic probe	A ball thrown up reaches its highest point. Its acceleration there is:
options	string[3]	One correct option and misconception distractors	Zero · g downward · g upward
correct	0–2	Index of the correct option	1
fix	string	One-line click moment	Gravity never switches off...

5.4 The diagnose-first pipeline

When a student asks a question (typed, spoken or photographed) the agent returns one structured object rather than a paragraph. The generator is instructed that Unravix “is NOT a search engine” and that its job is to diagnose the exact misconception behind the question, make the student commit to a guess before ex-

plaining, and then verify that they are truly unstuck. It must match the student's level: simple words and small numbers for young children, rigorous treatment for postgraduates. Table 5.2 lists the fields; Figure 5.4 shows the flow across the student, the agent and the trust layer.

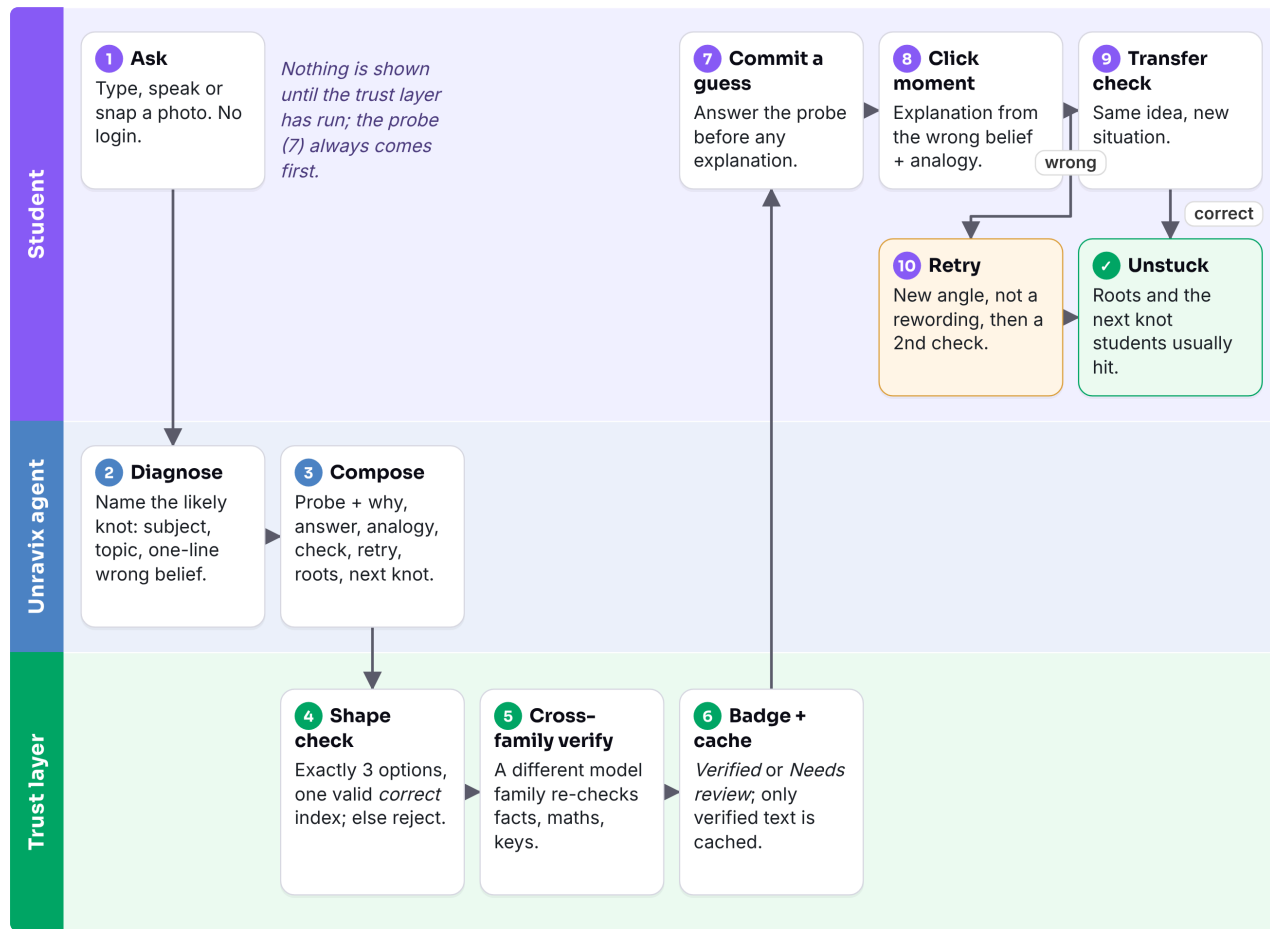


Figure 5.4. The diagnose-first pipeline as a swimlane. The trust layer runs before the student sees anything; the probe always comes before the explanation.

Table 5.2. Fields of a diagnose-first answer object (agent v4)

Field	Content	Constraint enforced
subject, topic	Where the question sits	Free text
knot, knot_short	The diagnosed misconception and its label	Required
probe	Question that reveals the knot, with 3 options, correct index and a reason per option	Exactly 3 non-empty options; correct $\in \{0,1,2\}$; why padded to 3
answer	Click-moment explanation starting from the wrong belief	Required; at most 110 words (prompt)
analogy	One everyday analogy	At most 30 words (prompt)
check	New-situation question with explanation	Exactly 3 options; valid index
retry	Different-angle explanation plus a second check	Exactly 3 options; valid index
roots	2–3 prerequisite ideas	Trimmed to at most 3
next_knot	The knot students usually hit next	—
factcheck	Verdict, claim, verifier, issues, corrected flag	Set by the trust layer

Field	Content	Constraint enforced
sources	Real textbook chapters	At most 5; doubtful ones removed by verifier
verified, cached, fallback, model	Trust and provenance metadata	Set by the Worker

Three properties of the pipeline matter most. First, the student commits before reading: the probe comes first and each of its options carries a one-line reason, so a wrong pick is explained at the level of the belief. Second, the check uses a new situation, which follows the knowledge-in-pieces caution that a student may hold a knot in one context only. Third, the retry is required to be a *different* explanation from a new angle, not a rewording. This answers Chi's (2005) observation that robust misconceptions need a change of category, not more detail. Before anything is shown, a deterministic shape check rejects any output that lacks a knot or an answer, or whose probe, check or retry check does not have exactly three non-empty options with a valid correct index.

5.5 Verification and trust model

Every generated answer goes to a strict fact-checker prompt at temperature 0. The checker reviews every factual claim, recomputes every number, unit and formula, confirms that each multiple-choice item has exactly one correct option at the stated index, and removes doubtful textbook references. It returns either *ok* or a list of issues with a fully corrected answer object. Figure 5.5 shows how the outcomes map to what the student sees.

TRUST RULES

- The verifier comes from a **different model family** first.
- Keyless, low-trust generators may only be verified by **strong** models; keyless models never verify.
- Two student flags stop a cached answer from being served.

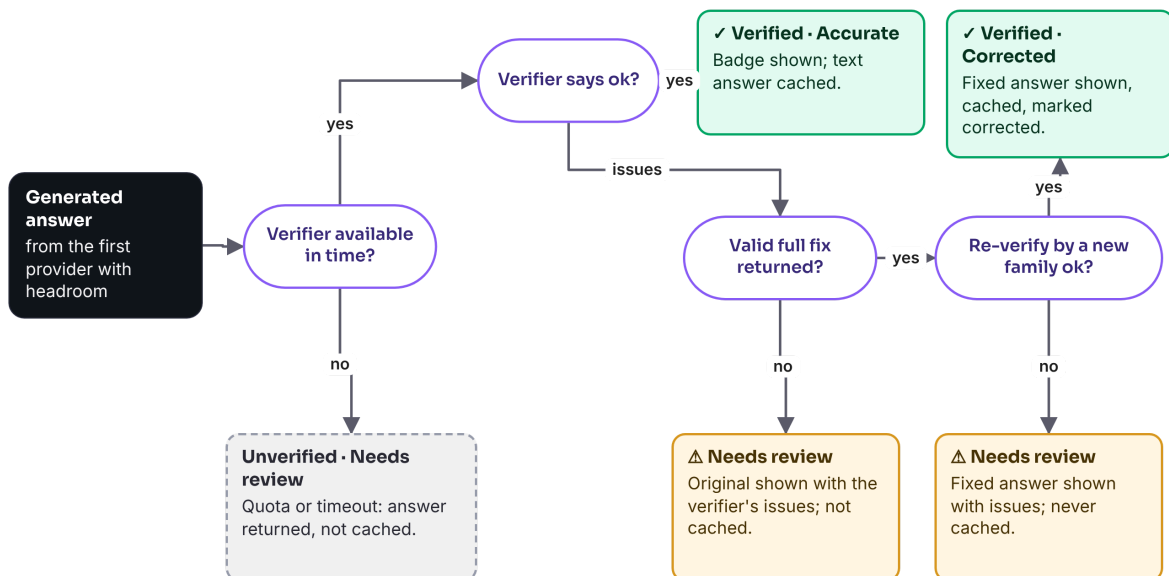


Figure 5.5. Verification flow: how verifier outcomes map to the Verified and Needs review badges and to the cache.

The trust model has five rules:

1. **Cross-family verification.** The verifier is drawn from a different model family than the generator; Gemini and Gemma count as one family, as do all Groq-hosted models. A different model from the same family is used only as a last resort.
2. **Strong verifiers for low-trust generators.** Answers from keyless, low-trust providers count as verified only if one of the designated strong models has checked them. Keyless providers never verify.
3. **Verified-only cache.** A text answer enters the shared cache only if it is verified. Unverified answers, answers to photos and voice notes, and answers for which no verifier was available in time are returned once and never cached.
4. **Student flags.** Every answer carries a “This looks wrong” button. Two flags on a cached answer stop it from being served, and a fresh verified answer replaces it.
5. **Visible status.** The interface shows a *Verified* badge or a *Needs review* label with the checker's issues. The Terms of Use state plainly that a *Verified* badge means an automated check passed, not that a teacher reviewed the answer.

The cache key is a SHA-256 hash of the normalised level and question (lower-cased, whitespace collapsed, punctuation trimmed at both ends). Only an exact match is served from the cache. Near matches are handled by the library fallback below, which marks itself as such.

5.6 Provider rotation and failover

Unravix runs on free model quotas, which are small and unpredictable. Rather than depend on one provider, the Worker holds a configuration-driven pool of 32 entries from 12 model families, arranged in four tiers. Within a tier it rotates round-robin from a random starting point among entries that still have headroom. An entry has headroom if it is not cooling down, is under its daily cap and any shared group cap, and is under its requests-per-minute limit. A circuit breaker cools an entry for a period that depends on how it failed, and the state is shared across isolates through the database every 20 seconds. Adding an OpenAI-compatible provider takes one configuration entry and one secret. Figure 5.6 shows the request cascade, and Figure 5.7 shows how the pool divides by family, tier and trust role.



Figure 5.6. Request failover cascade from cache to provider tiers, library and queue, with circuit-breaker rules and the time budget.

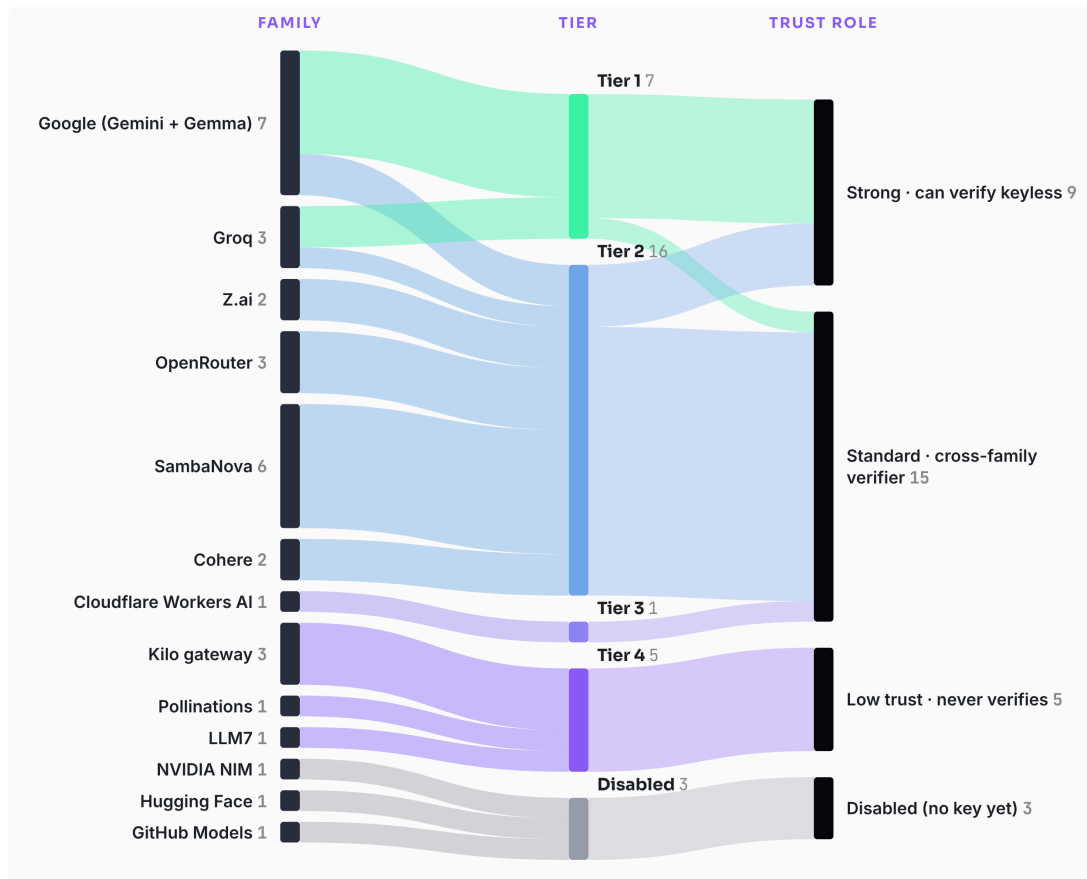


Figure 5.7. The provider pool (32 entries) by model family, tier and trust role, as configured in the Worker on 18 October 2026.

If every provider fails, a library search runs a full-text query over verified cache rows and curated knots. It requires at least two matched terms and at least half of the question's terms (60% for cache rows), and returns the normal answer shape labelled “From the Unravix library”. If nothing matches, the question is stored in a queue and the student gets a ticket. A cron job every five minutes answers up to four queued items through the same pipeline, and the page polls the ticket until the answer is ready. Tickets expire after 48 hours. The student therefore always gets an answer, a labelled library match, or a promise that is kept.

Time is budgeted as tightly as quota. Generation attempts stop at 28 seconds, verification starts only if less than 30 seconds have passed, re-verification only before 33 seconds, and the request ends at about 39 seconds. An answer that misses the verification window is returned as *Needs review* and not cached.

5.7 Data model

All persistent state lives in one Postgres database with row-level security enabled on every table (Figure 5.8). Knowledge tables hold the knots and verified answers; reliability tables hold the ask queue and provider state; community tables hold profiles, forums, threads, posts and the moderation log; growth tables hold the waitlist, feedback and the country catalog. Anonymous clients can read only approved forum content and public catalog and knot data. They can insert into the waitlist and feedback tables only through column-limited grants guarded by a per-IP rate-limit trigger. All other access goes through server functions holding the service role.

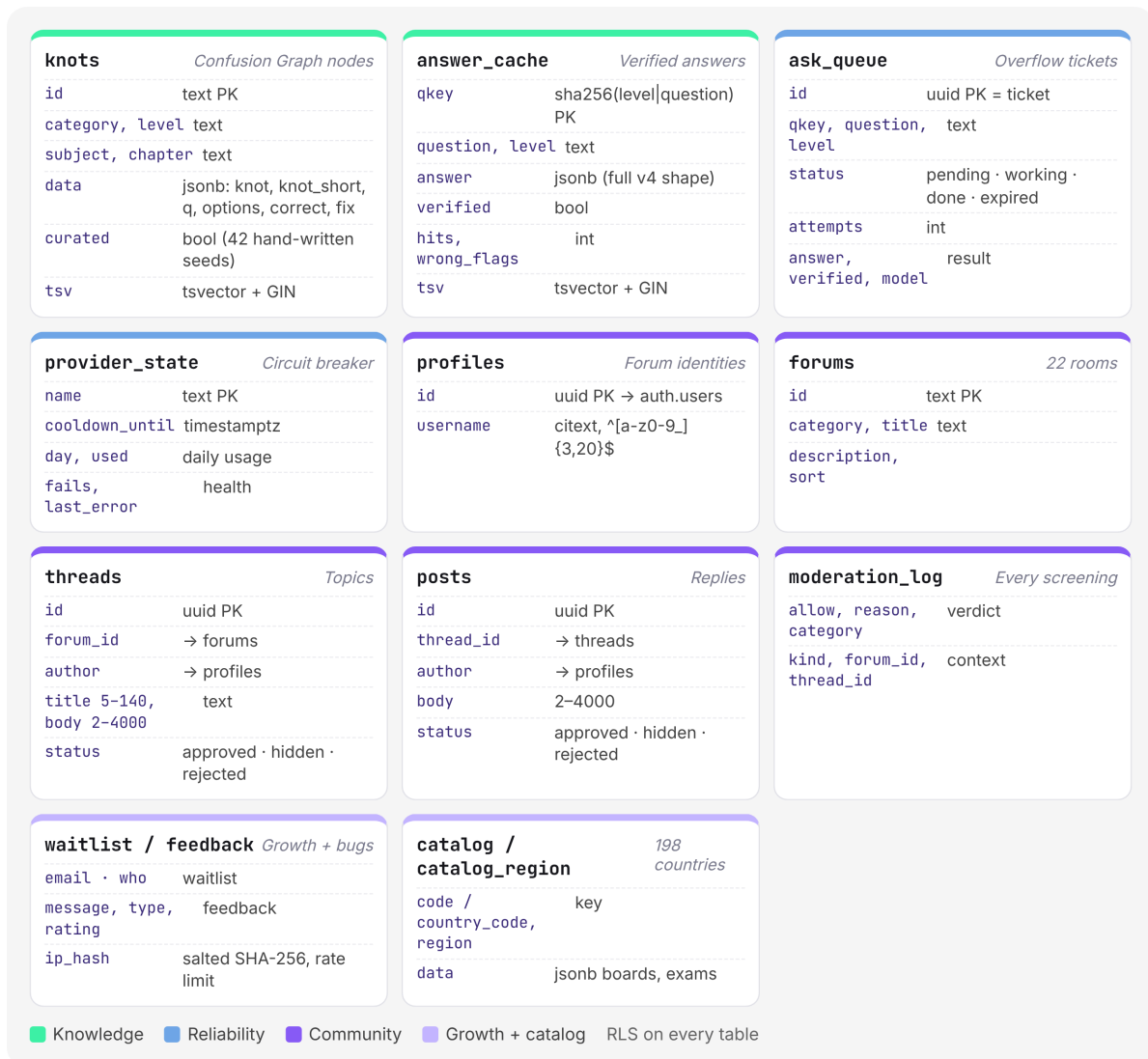


Figure 5.8. Data model. Every table has row-level security; colours group tables by purpose.

5.8 Deployment architecture

Figure 5.9 shows the deployed system. The site is static and served from Cloudflare Pages with strict security headers and a hashed-script content security policy. The Ask agent is a single Cloudflare Worker exposing modes for asking, scanning, flagging, ticket polling, moderation and catalog generation, with a cron trigger that drains the queue. Because some Indian mobile networks were found to block the *.workers.dev domain, student requests reach the Worker through a small relay edge function hosted by Supabase in the Mumbai region. The same Supabase project provides Postgres, authentication for forum accounts, and a forum edge function that sends each post to the Worker for AI moderation.

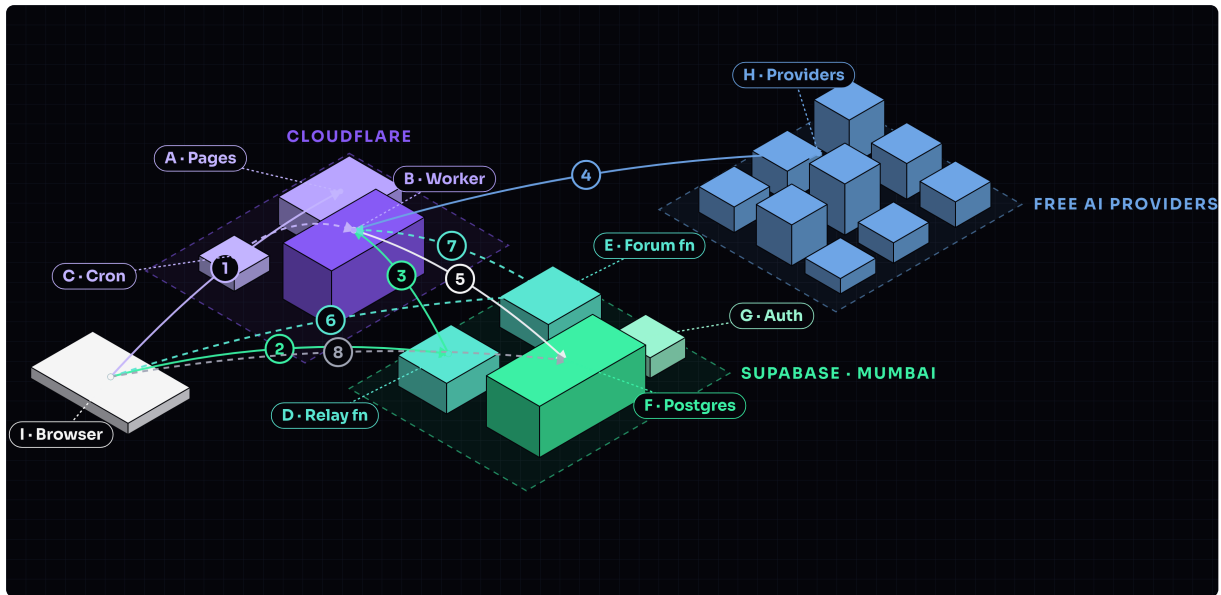


Figure 5.9. Deployment architecture: static site, Worker agent, relay, database and provider pool. Letters and numbered flows are explained in Table 5.3.

Table 5.3. Components and flows in Figure 5.9

Key	Component	Role
A	Cloudflare Pages	Static site, 56 knot pages, privacy and terms, CSP headers
B	Worker unravix-ask	Modes: ask, scan, flag, ticket, moderate, catalog, region
C	Cron trigger	Every 5 minutes: drains the ask queue through the same pipeline
D	Relay edge function	Carries student requests on networks that block *.workers.dev
E	Forum edge function	Signup, posting, pending re-screen with pg_cron
F	Postgres (RLS)	answer_cache, knots, ask_queue, provider_state, forums, waitlist, feedback, catalog
G	Auth	Username and password accounts, no email
H	Provider pool	32 entries, 12 families, 4 tiers, circuit breaker
I	Student browser	Light/dark UI; voice and photo input; no login to Ask
1-8	Flows	1 static pages · 2 Ask/Scan/flag · 3 relay to Worker · 4 generate + verify · 5 cache, queue, knots, provider state · 6 forum post (login) · 7 AI moderation · 8 public reads and rate-limited inserts

5.9 Privacy by design

The privacy decisions are architectural rather than added on afterwards:

- **No account to learn.** Ask, Knot Scan and the Library work without signing in.
- **Accounts without email.** Forum accounts are a username and a password. Usernames map internally to a non-routable pseudo-address, and users are told not to use their real names.
- **Hashed identifiers.** For the waitlist and feedback forms only a salted SHA-256 hash of the IP address is stored, solely for rate limiting.

- **Ephemeral media.** Photos and voice notes are sent to a model to produce the answer and are never written to the answer library.
- **Anonymous reuse.** Text questions and their verified answers can enter the library without any identity attached.
- **Moderation before publication.** A local pre-filter rejects phone numbers, email addresses, chat-group invite links and posts with more than three links. An AI moderator then screens every forum post before it appears.
- **Secrets stay server-side.** The launch audit confirmed that the only key in client code is the database's public anonymous key, with row-level security enforcing access.

The published Privacy Policy and Terms of Use address the EU and UK GDPR, India's Digital Personal Data Protection Act 2023 and COPPA-style protections for children: no accounts under 13, and parental permission for under-18s.

06

Product Specifications

This chapter specifies Unravix as a product: what each feature does, which platforms it runs on, how fast it is, how it handles accessibility and theming, and the limits that protect a free service from abuse.

6.1 Feature set

Table 6.1. Live features at launch (18 October 2026)

Feature	What it does	Account
Knot Scan	Pick a level (or type your own) and optionally a topic or hand-checked chapters; answer 8 quick questions, each targeting one known misconception, with a "Not sure" option	No
Trap Report	Shows the traps the student fell into, marks curated items as hand-checked, can be shared, and adds the traps to the student's Confusion Graph view	No
Knot Library	Browse common traps for any of 329 levels in 23 categories; search; 56 static knot pages for search engines	No
Ask	Text, voice or photo question; diagnose-first answer with Verified / Needs review badge, "This looks wrong" flag, library and queue fallbacks	No
Confusion Graph view	Personal map of knots found and untied, with prerequisite links	No
Forums	22 forums in 7 categories (subjects, entrance exams, postgrad and abroad, government exams, school, college and careers, community); every post AI-moderated	Username + password
Catalog	Boards and exams by country (198) and state or province	No
Waitlists	Early access and mobile-app waitlist (email and student type)	Email only
Feedback	Bug, idea or feedback with optional rating, name and email	No
Theme and legal	Light and dark themes; Privacy Policy and Terms of Use	—

The landing page also contains an interactive "Watch a knot unravel" demonstration that shows how ranked peer explanations will look. It is labelled on the page as a demo with illustrative numbers, and the student names in it are samples. It shows the vision and is not live data.

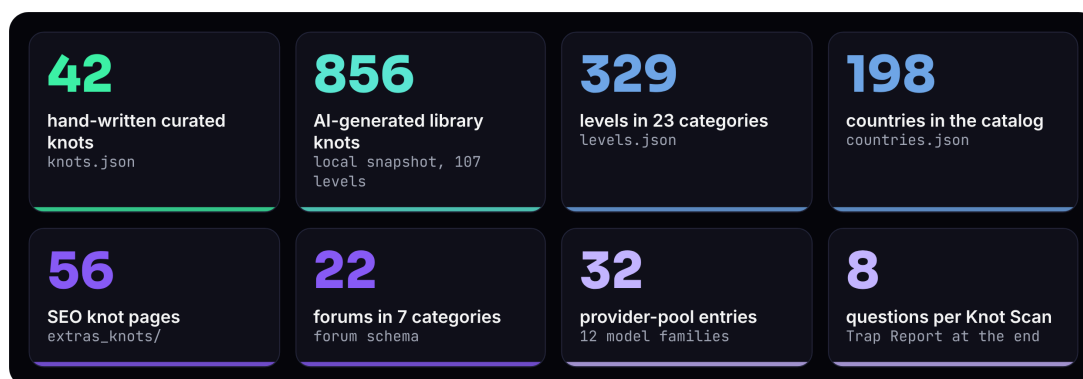


Figure 6.1. Unravix by the numbers at launch. Counts are taken from the repository files named under each figure.

6.2 Coverage: every level, every country

Unravix is designed for learners from Class 1 to PhD, including job seekers. The level catalogue contains 329 levels in 23 categories (Figure 6.2). These range from school classes and boards (CBSE, ICSE, IB, Cambridge IGCSE and A-Levels, GCSE, AP, WAEC, KCSE and others) through engineering, medical, law, design and management entrance examinations, civil-service, banking and railway examinations, study-abroad tests, professional certifications, college subjects and doctoral research, to 36 kinds of job interview. A separate catalog generated for 198 countries, with state and province detail for large federal countries, lets a student pick their own board or examination. If a student's level is missing, they can type it and the Knot Scan builds a scan on demand.

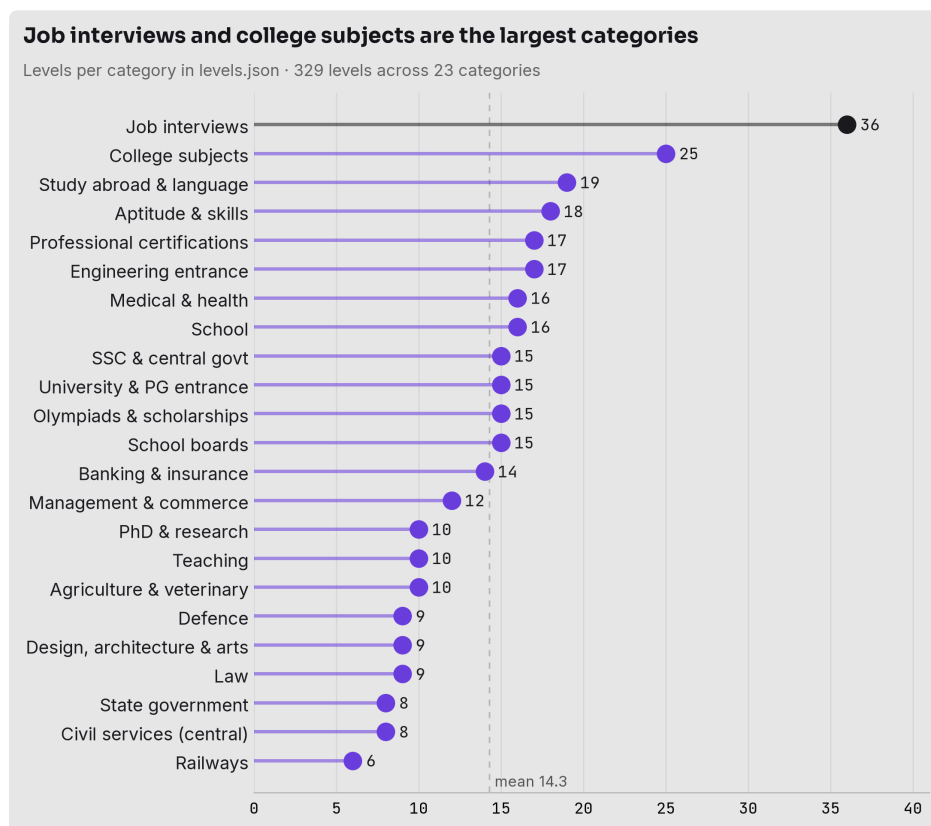


Figure 6.2. Levels per category in the level catalogue (levels.json).

The curated core is deliberately narrower: the 42 hand-written knots cover the launch wedge of JEE and NEET physics, chemistry, mathematics and biology (Figure 6.3).

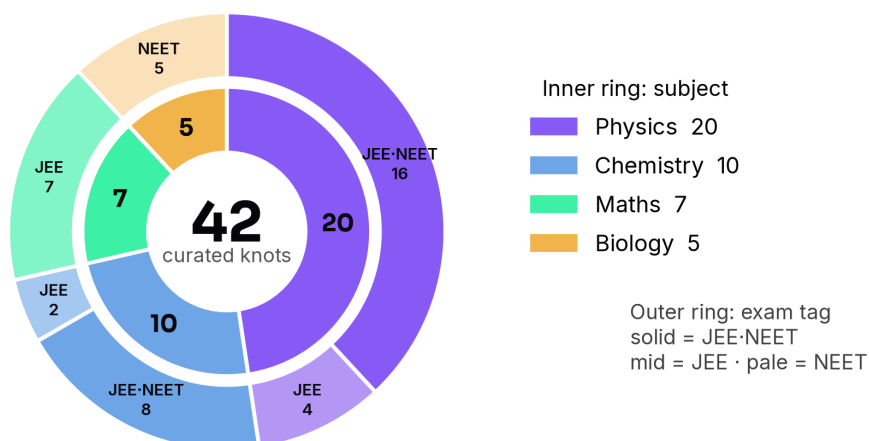


Figure 6.3. Curated knots by subject (inner ring) and exam tag (outer ring).

6.3 User journey

Figure 6.4 maps the intended first session. Its emotional arc is a design target, not a measurement: the scan deliberately produces a moment of surprise (“wait... what?”) because, following Posner et al. (1982), dissatisfaction with the current idea comes before change. The Ask loop then resolves it.

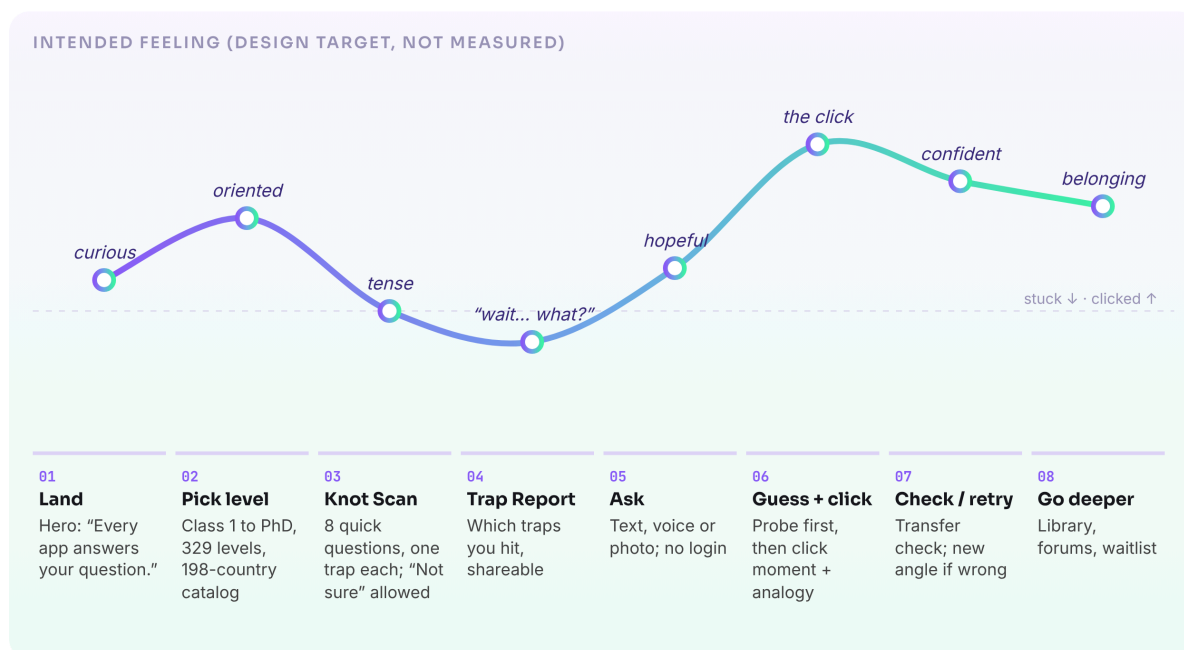


Figure 6.4. First-session user journey. The emotional curve is a design target, not measured data.

6.4 Platforms

- **Web (live).** A single responsive web application that works in modern desktop and mobile browsers. The launch audit found no horizontal scroll at a 390-pixel viewport, and voice and photo input use the browser's microphone and camera permissions, requested only when the student taps the control.

- **Mobile apps (announced).** The site announces Android and iOS apps for 2027 and runs an app waitlist that records the student type (school, JEE, NEET, other entrance, competitive examination, college or PhD, job seeker).

6.5 Performance

Table 6.2. Measured performance at launch (launch audit, headless Chromium, 18 October 2026)

Measure	Value
First contentful paint	0.41–0.53 s (sandbox network)
Load event	0.8–0.9 s
Transferred on first load	≈ 298 KB (771 KB decoded)
Requests on load	6 page requests + 8 database reads
Console errors	0
Ask, cache hit (n = 11)	p50 1.22 s · p95 1.51 s · max 1.65 s
Ask, fresh answer (n = 6, 5 concurrent)	p50 38.2 s · p95 39.2 s; 1 of the 5 concurrent verified
Provider health check	1.1–1.5 s; 17 entries available
Client timeout for Ask	45 s

6.6 Accessibility and theming

- **Structure.** At the launch audit every image had alternative text, the page had a single h1 and a clean heading hierarchy, and every in-page anchor resolved.
- **Multiple input modes.** Students can type, speak (with a “We heard: ...” transcript shown back for confirmation) or photograph a question, which helps those who struggle with typing mathematical notation.
- **Plain notation.** Answers use plain-text symbols (^{2} , $\times$, $\approx$, $\sqrt{}$, θ) instead of LaTeX, so they render the same everywhere and need no maths renderer.
- **Level-matched language.** The generator adapts vocabulary and number size to the selected level, from young children to postgraduates.
- **Keyboard.** Enter sends a question and Shift+Enter adds a new line; Ctrl+Enter sends feedback.
- **Light and dark themes.** The default is the brand's dark “ink” theme (#06070B with Knot Green #3CF2A6 and Thread Violet #8B5CF6), and a light mode is available from the navigation. A darker green (#00A665) is used on light backgrounds to keep contrast.
- **Known gap.** No formal WCAG conformance audit or screen-reader test has been done yet. Both are planned (Chapter 10).

6.7 Limits and quotas

Because the service is free and runs on free tiers, limits protect it from abuse and from one heavy user exhausting shared quota. Per-IP limits are deliberately generous because many students in India share one public IP through mobile carrier-grade NAT or hostel Wi-Fi.

Table 6.3. Service limits

Limit	Value
Asks + scans per IP	20 per hour

Limit	Value
"Looks wrong" flags per IP	20 per hour
Ticket polling per IP	300 per hour (not counted against asks)
Queued question lifetime	48 hours
Forum posts per user	10 per hour, 40 per day (rejected attempts count)
Thread title / post body	5-140 / 2-4,000 characters
Username / password	3-20 characters [a-z0-9_] / 8-72 characters
Waitlist / feedback per IP	20 per hour / 15 per hour (hashed IP)
Feedback message	Up to 2,000 characters
Answer explanation / analogy	≤ 110 words / ≤ 30 words (prompt)

07

Implementation

Unravix was built and shipped by a single developer in October 2026 using only free tiers. This chapter records how the system is implemented, how it evolved, what the launch audit found, and how it performed under a first load test.

7.1 Technology stack

Table 7.1. Technology stack

Layer	Technology	Notes
Front end	Single-page static HTML, CSS and JavaScript; three.js for the hero	Served by Cloudflare Pages with security headers and CSP
Agent	Cloudflare Worker (JavaScript modules) + cron trigger	Modes: ask, scan, flag, ticket, moderate, catalog, region; admin tools
Relay	Supabase Edge Function (Deno, TypeScript), Mumbai	Reaches the Worker on networks that block its domain
Database	Supabase Postgres with RLS, GIN full-text indexes, pg_cron	RPCs for cache hits, flags, provider usage, library search
Auth	Supabase Auth (username mapped to pseudo-address)	No email confirmation; refresh tokens rotate
Forums	Supabase Edge Function + Worker moderation	Pending posts re-screened every 10 min
Models	32-entry pool: Gemini and Gemma, Groq (gpt-oss, Qwen, Whisper), OpenRouter, SambaNova, Cohere, Z.ai, Workers AI Llama 3.3 70B, keyless gateways	Free tiers only
Content tooling	Python scripts for the catalog, knot library and SEO pages	Use Worker admin modes

7.2 Evolution of the Ask agent

The agent went through three versions in a single day of hardening, 18 October 2026:

- **v2: diagnose-first JSON.** One generator with a fixed answer shape (probe, knot, answer, analogy, check, retry, roots, next knot, fact-check, sources) and a Workers AI fallback. A relay edge function made it reachable on networks that block the Worker domain.
- **v3: verification and cache.** This version added the deterministic shape check, cross-model verification with a corrected-answer path, the verified-only answer cache keyed by a SHA-256 hash of the normalised question, the “looks wrong” flag mode, and explicit quota errors.

- **v4: rotation and resilience.** This version replaced the two-model setup with the configuration-driven pool of 32 entries, the circuit breaker with shared state, round-robin rotation, robust JSON extraction (which strips reasoning traces and code fences and picks the largest parseable object), library fallback, the ask queue with cron drain, and Groq Whisper transcription as a fallback for voice when Gemini audio is unavailable.

The same Worker also serves the Knot Scan (eight items, each targeting one well-documented misconception at the chosen level, with a library fallback of eight known knots) and the forum moderation mode, and it generated the country catalog and the AI knot library through admin-only bulk modes. To protect student-facing quota, these bulk modes use Workers AI unless explicitly allowed to use the pool.

7.3 Content pipeline

The 42 curated knots in `knots.json` were written by hand and are marked curated in the database, so library fallbacks can label them *Verified*. A batch script then asked the Worker's scan mode to produce knots for each level in the catalogue. A local snapshot of that output contains 856 AI-generated knots across 107 of the 329 levels, and the build log records levels for which generation failed. These knots are deliberately labelled *Needs review*: spot checks show that some generated items are shallow or arguable, which is why curated and generated knots are kept apart. Finally, 56 static, search-indexable knot pages (individual knots, chapter hubs and a library index) were generated for the launch wedge and submitted through the sitemap.

7.4 Launch audit

A structured launch audit on 18 October 2026 checked security, compliance, search basics, performance and failure modes. Table 7.2 summarises its findings.

Table 7.2. Launch audit summary (18 October 2026)

Area	Finding	Status at audit
Secrets	No admin, service or provider keys in client code; only the public anonymous database key	Done
Waitlist	Forms were not saving; anonymous insert grant, rate-limit trigger added and tested	Database done; page wiring pending
Privacy and terms	Global policy (GDPR / UK GDPR, DPDP, COPPA-style)	Done
Age gate	No age gate at signup; interim self-declared confirmation	Open
Security headers	CSP with hashed scripts; validated headless	Done
Search basics	robots.txt, sitemap, 404 page; meta, canonical and social tags prepared	Done / page edit pending
Feedback hardening	Insert-only column grants, 15 per hour per IP	Done
Fresh-answer latency	Under 5 concurrent requests: 38–39 s, 4 of 5 unverified	Open (needs provider headroom)
Retention	Policy periods need scheduled purge jobs	Open
Global defaults	Default region is India on first load	Open

7.5 Load test

The audit sent 17 scripted Ask requests through the relay. All returned HTTP 200. The 11 cache hits had a median of 1.22 s and a 95th percentile of 1.51 s. The six fresh answers were the weak point: one finished verified in 17.5 s and one (a sequential repeat) in 3.8 s, but four of the five fired concurrently reached the roughly 39-second hard end and were returned as *Needs review* (Figure 7.1). A browser test of “What is Newton’s second law of motion?” resolved through the library fallback in 28.6 s, still with a *Verified* curated answer. The conclusion is that the cache and fallbacks work, while fresh generation under concurrency needs more provider headroom. That is the top open item in Chapter 9.

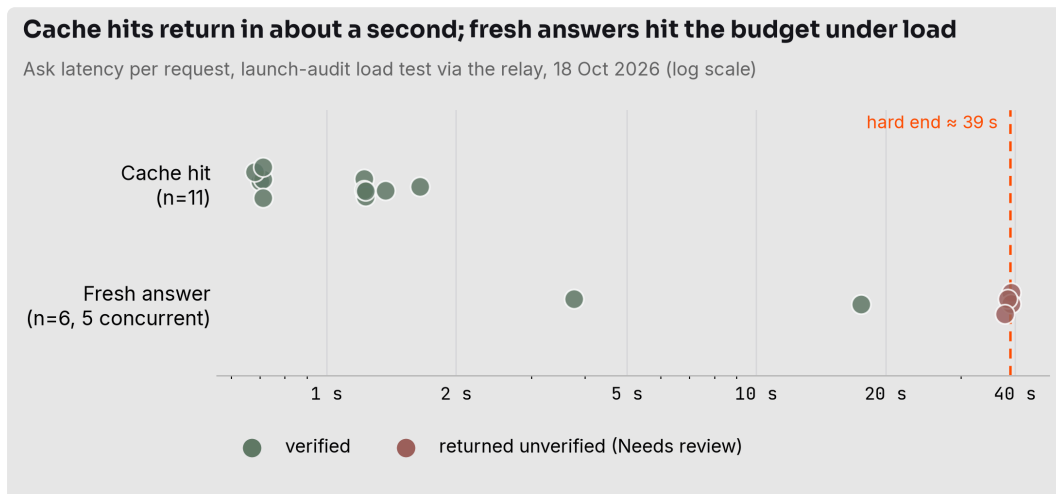


Figure 7.1. Ask latency per request in the launch load test (work/launch/load_results.json).

7.6 Lessons from implementation

- **Free quotas are volatile.** On launch day four of five Gemini text models hit their daily quota, one provider’s entire free tier turned out to require a card on file (so it now cools for six hours at a time), and a keyless provider allowed one reply per 30 seconds per IP. Treating providers as interchangeable, health-tracked entries was essential.
- **Verification costs time.** Each answer needs at least two model calls, so the time budget, not the token budget, is the binding constraint.
- **Networks differ.** A blocked platform domain on a major Indian mobile network would have silently broken the product for a large share of the launch audience. The relay fixed this at no extra cost.
- **Structure makes checking possible.** Because answers are typed objects, the shape check and the verifier can test each multiple-choice key mechanically. That would not be possible with free text.

08

Evaluation Plan

Status. Everything in this chapter is planned. No learning-outcome data have been collected, and no result in this document should be cited as evidence that Unravix improves learning.

Unravix makes a testable claim: diagnosing and repairing a specific misconception produces more durable understanding than delivering a correct explanation. This chapter defines the metrics that will test that claim and the studies that will collect them.

8.1 Metrics

Table 8.1. Evaluation metrics

Metric	Definition	Use
Commit rate	Share of Ask sessions in which the student answers the probe before reading	Is guess-first accepted?
Unstick rate	Among sessions where the probe exposed the knot, share that pass the transfer check (first check or after retry)	Primary in-product outcome
Retry rescue rate	Share of first-check failures that pass after the retry	Value of the second angle
Relapse rate	Share of unstuck students who fail a delayed item on the same knot	Durability
Trap Report accuracy	Per-knot precision and recall of scan diagnoses against an independent measure	Validity of the scan
Verified-answer rate	Verified answers ÷ generated answers (also corrected rate and needs-review rate)	Trust-layer coverage
Badge precision	Share of Verified answers that experts grade as fully correct	Trust-layer accuracy
Flag rate	Flags per 1,000 answers served, and share of flags upheld by experts	Error discovery
Cache hit rate	Answers served from cache ÷ all answers	Cost and latency
Fallback and queue rates	Share of library and queued answers; time to ticket completion	Resilience
Latency	p50 and p95 time to answer, by path	Experience
Normalised gain (g)	$(\text{post} - \text{pre}) \div (100 - \text{pre})$ on targeted inventory items	Study 2 primary outcome

8.2 Studies

Study 1: in-product A/B test of diagnose-first versus answer-first

Sessions will be randomised to receive the same generated answer object rendered in one of two orders: *diagnose-first* (probe, then click moment, analogy, check and retry, as built) or *answer-first* (explanation first, no probe, check optional). Holding the content constant isolates the effect of the protocol. The primary outcome is performance on a delayed transfer item for the same knot, served when the student next returns. Secondary outcomes are the unstick rate, the relapse rate and time on task. Assuming a small-to-moderate effect ($d = 0.3$), a two-sided α of 0.05 and power of 0.80, a simple two-group comparison needs about 175 returning students per arm. This is a planning estimate to be refined once real variance is known.

Study 2: pre/post concept inventory with partner classes

With consenting teachers and coaching institutes, students will take a pre-test of concept-inventory items matched to the targeted knots, use Unravix (arm A) or an answer-first control (arm B) for four weeks, then take a parallel-form post-test and an eight-week delayed test without tool access (Figure 8.1). The primary outcome is Hake's (1998) normalised gain on the targeted knots. The delayed test addresses the risk documented by Bastani et al. (2025), where gains with AI assistance disappeared when the tool was removed. For mechanics, Force Concept Inventory items are the natural instrument (Hestenes et al., 1992), subject to the publisher's terms of use. For other subjects, two-tier items will be built following Treagust (1988).

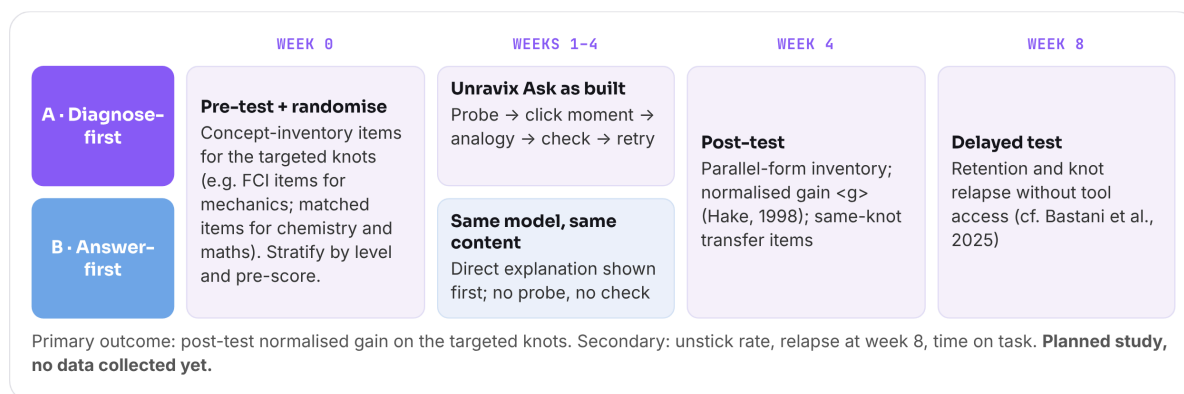


Figure 8.1. Planned two-arm study design with pre-, post- and delayed tests. No data have been collected.

Study 3: verification audit

Subject experts will blind-grade a stratified random sample of answers (verified, corrected and needs-review) for factual accuracy, answer-key correctness, level appropriateness and whether the identified knot is a real, documented misconception. This gives the precision of the *Verified* badge, the false-alarm rate of *Needs review*, and the share of errors that student flags catch. The same protocol will review the 856 AI-generated knots so that the best can be promoted to curated status.

Study 4: Trap Report validity

For a subsample of students, Knot Scan diagnoses will be compared with an independent measure: a short think-aloud interview or a second, differently worded two-tier item for the same knot. Agreement gives per-knot precision and recall for the Trap Report.

Study 5: peer click moments versus AI explanations (after launch of capture)

Once click-moment capture ships, students stuck on the same knot will be randomised to a top-ranked peer explanation or the AI click moment, using the same transfer outcome as Study 1. This answers RQ4 and tests the central bet behind the peer explanation economy.

8.3 Research ethics

All studies will use informed consent, with parental consent and age-appropriate assent for minors, will collect no more personal data than the outcome requires, and will be pre-registered before data collection. Partner institutions will be asked to obtain ethics approval where their rules require it. Students in the control arms will get the full product after the study.

09

Goals and Roadmap

Unravix's strategy is to win a narrow wedge with a product that is clearly different, then widen it. The wedge is India's JEE, NEET and GATE students. The destination is every learner in every country, with misconception data and peer explanations as the lasting asset.

9.1 Goals

Horizon	Goal	Indicator
Now	Get the first real students through the loop, from scan or ask to a passed check	Weekly active students; share of sessions that answer the probe; unstick rate
Now	Close the open launch-audit items	Age gate live; retention purge jobs scheduled; fresh-answer verified rate under load
Next	Prove the protocol	Study 1 and Study 3 complete and published, including null results
Next	Turn students into authors	Click moments captured per knot; share of stuck students served a peer explanation
Later	Mobile apps for Android and iOS (announced for 2027)	App waitlist conversion; retention
Vision	A global map of confusion with a peer explanation economy	Knots with measured fixes; authors credited per unstick

9.2 Roadmap

Figure 9.1 places these goals on a timeline. Only the shipped web launch and the announced 2027 app target carry dates. The other phases are ordered by priority rather than scheduled, because a single developer's capacity and free-tier limits set the pace.

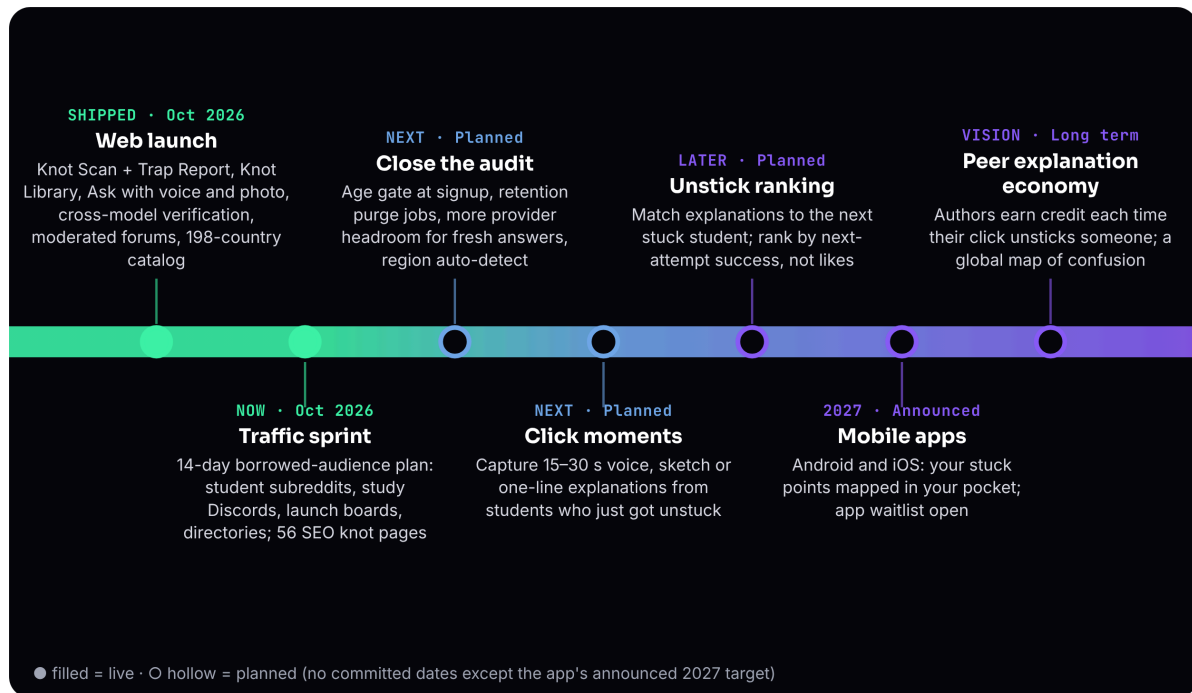


Figure 9.1. Roadmap ribbon. Filled markers are live; hollow markers are planned.

9.3 From wedge to global

Several design choices already serve the global goal: the 329-level catalogue spans boards and examinations on every continent, the country catalog covers 198 countries, the generator adapts to any level, and the policies address EU, UK, Indian and US-style children's privacy rules. The marketing plan follows the same order. A 14-day sprint borrows existing audiences (student subreddits for JEE and NEET, SAT, AP, IB, A-Level and GCSE, study Discords, launch boards and directories) rather than paying for reach. Public posts are drafted for the founder's approval first. Two launch-audit items still lean towards India: the default region on first load and the NCERT-first instruction in the generator. Both are scheduled to become locale-aware.

9.4 The peer explanation economy

The most ambitious part of the roadmap turns the Confusion Graph from an AI-built map into a human-built one. The recorded design is as follows. When a student passes the transfer check, they are invited to record a 15 to 30 second “what made it click” moment as a voice note, a sketch or a single line. When the next student is stuck at the same knot, they are served the explanations that most often led students like them to pass their *next* attempt, ranked by results rather than likes. Authors earn credit and reputation each time their explanation unsticks someone (the idea log notes that this credit could optionally be recorded on-chain), and a shareable proof card records what they untangled. This would combine the peer-instruction and self-explanation effects reviewed in Chapter 2 with the outcome measurement that the protocol already collects. None of it is live, and it will ship only together with the moderation, privacy and minor-safety work that user-generated voice content requires.

9.5 Sustainability

Unravix is free for students and runs on free tiers. The Privacy Policy commits to no advertising, no advertising profiles and no sale of data. No revenue model has been chosen. Options consistent with those commitments, such as institutional dashboards for schools and coaching centres or paid capacity for high-volume partners, are open questions for after the evaluation, not assumptions in this document.

10

Ethics, Safety and Limitations

A study tool used by minors and powered by fallible models carries real obligations. This chapter sets out the main risks, what Unravix does about each, and what remains unsolved.

10.1 AI errors

Generated answers can be wrong. Mitigations: the structured format, the deterministic shape check, cross-family verification with a corrected-answer path, the verified-only cache, the student flag that withdraws a cached answer after two reports, and visible *Verified* and *Needs review* labels. The Terms of Use state that the badge means an automated check passed, not that a teacher reviewed the answer, and they ask students to check important answers against their textbook, teacher or examination body. **Residual risk:** two models can share a misconception, especially on niche or advanced topics; verifier coverage drops under load (four of five concurrent fresh answers in the load test went out unverified); and the true precision of the badge is unknown until Study 3.

10.2 Children and self-declared age

Ask and the library need no account, so no age data are collected there; the Terms advise children under 13 to use Ask with a parent, guardian or teacher. Forum accounts and the waitlist require users to be at least 13, with parental permission under 18 or under the local age of digital consent. **Limitation:** age is self-declared. At the time of the launch audit, signup relied on an interim self-declared confirmation, and a proper age gate was an open item. Self-declaration can be falsified, and stronger verification would conflict with the minimal-data design. This trade-off is acknowledged, not solved.

10.3 Data sent to AI providers

Questions, photos and voice notes are processed by third-party model providers, which the Privacy Policy discloses. Students are told not to include names, phone numbers or other personal details. Photos and voice notes are not stored in the library. Because the rotation spans many providers, each with its own terms, the set of processors changes over time. Keeping the disclosure accurate as providers are added is an ongoing duty.

10.4 Community safety

Forum posts pass a local pre-filter (phone numbers, email addresses, chat-group invite links, more than three links) and then an AI moderator that rejects spam, abuse, sexual content, doxxing, cheating and leaked-paper requests, scams, self-harm encouragement, illegal content and gibberish. If no moderator is

available, posts are held as pending and re-screened every ten minutes. Every decision is logged for review. There are posting limits of 10 per hour and 40 per day. **Gaps:** no user reporting, editing or deleting of one's own posts yet, and automated moderation will make mistakes in both directions.

10.5 Academic integrity

The Terms prohibit using Unravix to cheat in live examinations or graded assessments and prohibit sharing or requesting leaked papers, and the forum moderator rejects such requests. The guess-first protocol is itself a partial safeguard, since Unravix is a poor tool for copying answers. It cannot stop a determined student from using it during an unsupervised assessment.

10.6 Curricular and cultural bias

The generator is told to prefer Indian syllabus context (NCERT first) where relevant, which suits the launch wedge but may produce India-centric examples for a student in Lagos or Leeds. Most curated knots are physical-science misconceptions from the JEE and NEET syllabi. Misconceptions in the humanities, languages and professional domains are less well documented in the research literature and harder to verify automatically.

10.7 Dependence on free infrastructure

Every component runs on a free tier. Quotas change without notice, providers can add card requirements (one already has), and terms of service can change. The rotation absorbs individual failures, but a broad tightening of free tiers would slow fresh answers to the queue path. The project accepts this risk in exchange for being free to students.

10.8 Other limitations

- **No outcome evidence yet.** The pedagogical claims rest on prior research, not on Unravix data.
- **Generated library quality.** 856 AI-generated knots are unreviewed and labelled as such.
- **Retention promises need automation.** The Privacy Policy states retention periods (for example, 12 months for the moderation log and 24 months for feedback), and the audit flagged that scheduled purge jobs must enforce them.
- **Accessibility.** No formal WCAG audit has been done.
- **Single maintainer.** Unravix is built and run by one developer, which limits response time and creates key-person risk.

11

Conclusion

Generative AI has made answers abundant. The research reviewed here explains why abundance alone does not produce understanding: learners hold specific, often coherent, wrong ideas that survive correct explanations unless those ideas are first made visible, then shown to fail, then replaced and tested in a new situation. Unravix turns that body of work into a product. Its unit of learning is the knot rather than the question; its answers are structured diagnoses rather than paragraphs; and its trust comes from an independent check rather than from fluency.

As of October 2026 the system exists and works on free infrastructure. It has a curated core of 42 hand-written misconceptions, an AI-generated library across a 329-level catalogue, a diagnose-first Ask agent with cross-family verification and a verified-only cache, a 32-entry provider rotation with circuit breakers, library and queue fallbacks, moderated forums, and privacy and terms written for a global audience of students that includes minors. Its weaknesses are just as clear: fresh answers slow down under concurrent load, the generated library is unreviewed, age is self-declared, and there is no outcome evidence yet.

The next step is therefore measurement, not more features: run the A/B and concept-inventory studies, audit the *Verified* badge, and publish the results whatever they show. If diagnose-first answers do produce more durable understanding, the long-term opportunity is the one the project started with: a shared, growing map of where humans get stuck and what reliably gets them unstuck, built partly by the students who were stuck there.

Find the knot. Watch it click.

References

- Bastani, H., Bastani, O., Sungu, A., Ge, H., Kabakcı, Ö., & Mariman, R. (2025). Generative AI without guardrails can harm learning: Evidence from high school mathematics. *Proceedings of the National Academy of Sciences*, 122(26), e2422633122. <https://doi.org/10.1073/pnas.2422633122>
- Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice*, 5(1), 7–74. <https://doi.org/10.1080/0969595980050102>
- Bloom, B. S. (1984). The 2 sigma problem: The search for methods of group instruction as effective as one-to-one tutoring. *Educational Researcher*, 13(6), 4–16. <https://doi.org/10.3102/0013189X013006004>
- Brainly. (2023a, April 19). *Brainly's new AI-powered capabilities enable a personalized learning experience for every student* [Press release]. Business Wire. <https://www.businesswire.com/news/home/20230419006023/en/>
- Brainly. (2023b, September 7). *Brainly enrolls new AI-powered tools for more personalized and accessible learning* [Press release]. Business Wire. <https://www.businesswire.com/news/home/20230907148279/en/>
- Business Standard. (2023, December 4). *Exams trainer Allen acquires edtech platform Doubtnut to improve technology*. https://www.business-standard.com/companies/news/allen-career-institute-acquires-ai-enabled-edtech-platform-doubtnut-123120400620_1.html
- Chi, M. T. H. (2005). Commonsense conceptions of emergent processes: Why some misconceptions are robust. *Journal of the Learning Sciences*, 14(2), 161–199. https://doi.org/10.1207/s15327809jls1402_1
- Chi, M. T. H., Bassok, M., Lewis, M. W., Reimann, P., & Glaser, R. (1989). Self-explanations: How students study and use examples in learning to solve problems. *Cognitive Science*, 13(2), 145–182. https://doi.org/10.1207/s15516709cog1302_1
- Chi, M. T. H., Slotta, J. D., & de Leeuw, N. (1994). From things to processes: A theory of conceptual change for learning science concepts. *Learning and Instruction*, 4(1), 27–43. [https://doi.org/10.1016/0959-4752\(94\)90017-5](https://doi.org/10.1016/0959-4752(94)90017-5)
- Crouch, C. H., & Mazur, E. (2001). Peer Instruction: Ten years of experience and results. *American Journal of Physics*, 69(9), 970–977. <https://doi.org/10.1119/1.1374249>
- diSessa, A. A. (1993). Toward an epistemology of physics. *Cognition and Instruction*, 10(2–3), 105–225. <https://doi.org/10.1080/07370008.1985.9649008>
- Dunlosky, J., Rawson, K. A., Marsh, E. J., Nathan, M. J., & Willingham, D. T. (2013). Improving students' learning with effective learning techniques: Promising directions from cognitive and educational psychology. *Psychological Science in the Public Interest*, 14(1), 4–58. <https://doi.org/10.1177/1529100612453266>
- Eedi. (n.d.-a). *Diagnostic engine*. Retrieved October 18, 2026, from <https://www.eedi.com/diagnostic-engine>
- Eedi. (n.d.-b). *Eedi impact evaluation: FAQ*. Retrieved October 18, 2026, from <https://rct.eedi.com/faq>
- Eedi. (n.d.-c). *How do quizzes work?* Eedi Help Centre. Retrieved October 18, 2026, from <https://help.eedi.co.uk/en/articles/6368626-how-do-quizzes-work>
- European Commission. (2023, March 28). *Commission clears acquisition of Photomath by Google* [Press release IP/23/1927]. https://ec.europa.eu/commission/presscorner/detail/en/ip_23_1927
- Google. (2025, August 6). *Guided Learning in Gemini: From answers to understanding*. The Keyword. <https://blog.google/outreach-initiatives/education/guided-learning/>
- Hake, R. R. (1998). Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses. *American Journal of Physics*, 66(1), 64–74. <https://doi.org/10.1119/1.18809>

- Halloun, I. A., & Hestenes, D. (1985). The initial knowledge state of college physics students. *American Journal of Physics*, 53(11), 1043–1055. <https://doi.org/10.1119/1.14030>
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81–112. <https://doi.org/10.3102/003465430298487>
- Hestenes, D., Wells, M., & Swackhamer, G. (1992). Force Concept Inventory. *The Physics Teacher*, 30(3), 141–158. <https://doi.org/10.1119/1.2343497>
- Ji, Z., Lee, N., Frieske, R., Yu, T., Su, D., Xu, Y., Ishii, E., Bang, Y. J., Madotto, A., & Fung, P. (2023). Survey of hallucination in natural language generation. *ACM Computing Surveys*, 55(12), 1–38. <https://doi.org/10.1145/3571730>
- Kapur, M. (2008). Productive failure. *Cognition and Instruction*, 26(3), 379–424. <https://doi.org/10.1080/07370000802212669>
- Khan Academy. (n.d.). *Khanmigo pricing*. Retrieved October 18, 2026, from <https://www.khanmigo.ai/pricing>
- Kluger, A. N., & DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin*, 119(2), 254–284. <https://doi.org/10.1037/0033-2909.119.2.254>
- Kornell, N., Hays, M. J., & Bjork, R. A. (2009). Unsuccessful retrieval attempts enhance subsequent learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 35(4), 989–998. <https://doi.org/10.1037/a0015729>
- OpenAI. (2025, July 29). *Introducing study mode*. <https://openai.com/index/chatgpt-study-mode/>
- Photomath. (n.d.). *About Photomath*. Retrieved October 18, 2026, from <https://photomath.com/about/>
- Posner, G. J., Strike, K. A., Hewson, P. W., & Gertzog, W. A. (1982). Accommodation of a scientific conception: Toward a theory of conceptual change. *Science Education*, 66(2), 211–227. <https://doi.org/10.1002/sce.3730660207>
- Roediger, H. L., III, & Karpicke, J. D. (2006). Test-enhanced learning: Taking memory tests improves long-term retention. *Psychological Science*, 17(3), 249–255. <https://doi.org/10.1111/j.1467-9280.2006.01693.x>
- Sadler, P. M. (1998). Psychometric models of student conceptions in science: Reconciling qualitative studies and distractor-driven assessment instruments. *Journal of Research in Science Teaching*, 35(3), 265–296. [https://doi.org/10.1002/\(SICI\)1098-2736\(199803\)35:3<265::AID-TEA3>3.0.CO;2-P](https://doi.org/10.1002/(SICI)1098-2736(199803)35:3<265::AID-TEA3>3.0.CO;2-P)
- Smith, J. P., III, diSessa, A. A., & Roschelle, J. (1994). Misconceptions reconceived: A constructivist analysis of knowledge in transition. *Journal of the Learning Sciences*, 3(2), 115–163. https://doi.org/10.1207/s15327809jls0302_1
- Treagust, D. F. (1988). Development and use of diagnostic tests to evaluate students' misconceptions in science. *International Journal of Science Education*, 10(2), 159–169. <https://doi.org/10.1080/0950069880100204>
- VanLehn, K. (2011). The relative effectiveness of human tutoring, intelligent tutoring systems, and other tutoring systems. *Educational Psychologist*, 46(4), 197–221. <https://doi.org/10.1080/00461520.2011.611369>
- Vosniadou, S. (1994). Capturing and modeling the process of conceptual change. *Learning and Instruction*, 4(1), 45–69. [https://doi.org/10.1016/0959-4752\(94\)90018-3](https://doi.org/10.1016/0959-4752(94)90018-3)
- Vosniadou, S., & Brewer, W. F. (1992). Mental models of the earth: A study of conceptual change in childhood. *Cognitive Psychology*, 24(4), 535–585. [https://doi.org/10.1016/0010-0285\(92\)90018-W](https://doi.org/10.1016/0010-0285(92)90018-W)

APPENDIX A

Sample Curated Knots

A sample of the 42 hand-written seed knots in knots.json (two per physics chapter, one per chapter elsewhere), exactly as shipped. Each knot is stored with `curated = true`, so library fallbacks can label it *Verified*. The correct option is in bold; the other options are the misconception distractors.

ID	Subject / chapter	Knot	Probe (correct option in bold)	Exams
kin1	Physics Kinematics	v = 0 means a = 0 If velocity is zero at an instant, acceleration must be zero too.	A ball thrown up reaches its highest point. Its acceleration there is: Zero · g downward · g upward	JEE·NEET
kin2	Physics Kinematics	Speed vs velocity average Average speed is always equal to the magnitude of average velocity.	You run one full lap of a 400 m track in 80 s. Your average velocity is: 5 m/s · 0 m/s · 2.5 m/s	JEE·NEET
n1m1	Physics Laws of Motion	Motion needs force A moving object needs a constant force to keep moving.	A puck slides on perfectly frictionless ice. To keep it moving at constant velocity you need: A constant push · No force at all · An increasing push	JEE·NEET
n1m2	Physics Laws of Motion	Action-reaction cancel Action and reaction forces cancel each other, so nothing can move.	A horse pulls a cart. Why can they move if the forces are equal and opposite? The horse's pull is slightly bigger · The pair acts on different bodies · Reaction appears a bit later	JEE·NEET
wpe1	Physics Work, Energy & Power	Holding = doing work Holding a heavy bag still means you are doing work on it.	You hold a 10 kg bag still for 5 minutes. Work done on the bag by you is: Positive · Zero · Negative	JEE·NEET
wpe2	Physics Work, Energy & Power	Work always positive Work done by a force is always positive.	Work done by friction on a block sliding to a stop is: Positive · Zero · Negative	JEE
e1e1	Physics Current Electricity	Current gets used up Current is used up as it passes through a bulb.	In a series circuit with one bulb, current before the bulb vs after it is: More before · Same · More after	JEE·NEET
e1e2	Physics Current Electricity	More resistors, more R Adding a resistor always increases the total resistance.	Adding a second identical resistor in PARALLEL to one resistor R makes the total: 2R · R/2 · R	JEE·NEET
e1e3	Physics Electrostatics	E = 0 means V = 0 Where the electric field is zero, the potential must also be zero.	Inside a charged hollow conducting sphere, the potential is: Zero · Constant, equal to surface value · Increasing inward	JEE·NEET
e1e4	Physics Electrostatics	Charges follow field lines A charge released in a field always moves along the field line.	A charge starts from rest on a CURVED field line. Its path: Follows the line exactly · Can leave the line · Moves perpendicular to it	JEE·NEET
mol1	Chemistry Mole Concept	Equal mass, equal moles Equal masses of two gases contain the same number of molecules.	Which has more molecules: 2 g of H ₂ or 2 g of O ₂ ? H ₂ · O ₂ · Same	JEE·NEET
eq1	Chemistry Chemical Equilibrium	Equilibrium = reaction stops At equilibrium, the reaction has stopped.	At chemical equilibrium: Both reactions stop · Forward and backward rates are equal · Only the forward reaction runs	JEE·NEET
goc1	Chemistry GOC / Organic Basics	Resonance = flipping Resonance structures flip back and forth rapidly between forms.	Benzene's C-C bond lengths are: Alternating single and double · All equal, in between · Constantly switching	JEE·NEET

ID	Subject / chapter	Knot	Probe (correct option in bold)	Exams
fun1	Maths Functions & Limits	Limit = value at the point The limit at a point is just the function's value there.	$f(x) = (x^2 - 1)/(x - 1)$. As $x \rightarrow 1$, the limit is: Undefined · 2 · 0	JEE
prb1	Maths Probability	Gambler's fallacy After many heads, tails becomes more likely.	A fair coin shows heads 5 times in a row. P(tails next) is: More than $\frac{1}{2}$ · Exactly $\frac{1}{2}$ · Less than $\frac{1}{2}$	JEE
bio1	Biology Photosynthesis & Respiration	Plants don't respire Plants photosynthesise instead of respiring.	Plant cells carry out respiration: Only at night · Day and night · Never	NEET
bio3	Biology Genetics & Inheritance	Dominant = common Dominant alleles are always the most common in a population.	Polydactyly (extra fingers) is dominant in humans. It is: Very common · Rare · Present in everyone	NEET
bio5	Biology Cell Biology	Only animals have mitochondria Plant cells have chloroplasts instead of mitochondria.	A leaf cell has: Only chloroplasts · Chloroplasts and mitochondria · Only mitochondria	NEET

Glossary

Knot

A specific, testable misconception, written as what the student wrongly believes.

Confusion Graph

The data structure that indexes learning by knots, linking them to subjects, chapters, levels, prerequisite roots, the next likely knot and, in future, peer explanations.

Probe

A three-option question shown before any explanation; its distractors are the misconceptions.

Click moment

The short explanation that resolves a knot, starting from the wrong belief. In the planned peer layer, a 15–30 second student explanation.

Transfer check

A new-situation question testing the same idea after the explanation.

Retry

A second explanation from a different angle, followed by another check.

Roots

Prerequisite ideas a knot depends on.

Next knot

The misconception students typically encounter next.

Knot Scan

An eight-question diagnostic for a chosen level that ends in a Trap Report.

Trap Report

The scan result: which misconceptions the student showed, not a score.

Verified

Badge meaning an automated check by a model from a different family passed. It is not human review.

Needs review

Label for answers that were not verified, or where the verifier found issues.

Cross-family verification

Checking a model's output with a model from a different provider family.

Circuit breaker

Logic that pauses a failing provider for a period that depends on how it failed.

Library fallback

Full-text match against verified cache rows and curated knots, used when no provider is available.

Ask queue

Stores questions that could not be answered immediately; a cron job answers them and the student polls a ticket.

Normalised gain (g)

The fraction of possible improvement achieved between pre-test and post-test (Hake, 1998).

RLS

Row-level security: database policies that limit which rows each role can read or write.

Wedge

The narrow first market (JEE, NEET and GATE students) used to reach the global one.