

01 The bottleneck

Oral examination is the most reliable way to detect genuine understanding. It is also the least scalable assessment we have.

The gap. Not "an AI that asks questions" — those drift off-syllabus and score in ways no examination board would accept. What a department can adopt is an examiner **grounded in the lecturer's own material**, with a human approving every question first.

24 hours
of examining for one course, one semester — one examiner, 120 students, twelve minutes each. Students leave with a grade rather than a diagnosis.

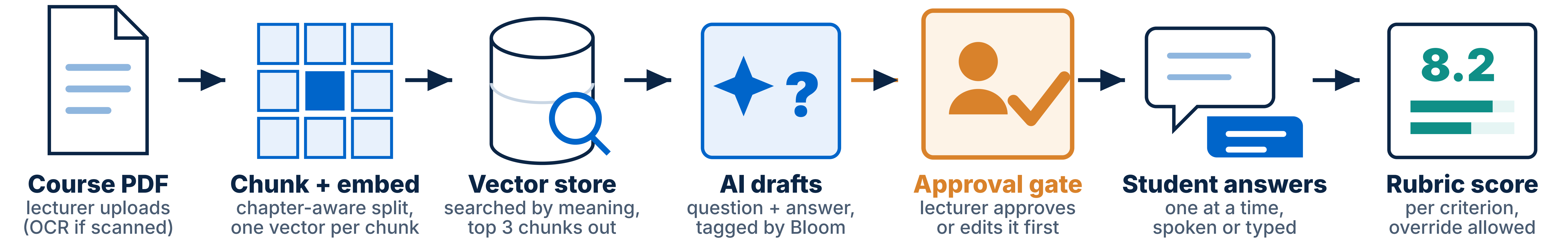
3
session modes — exam, practice, job interview

6
Bloom levels, selected hierarchically

100% human-approved
no question and no reference answer reaches a student without an explicit lecturer approval

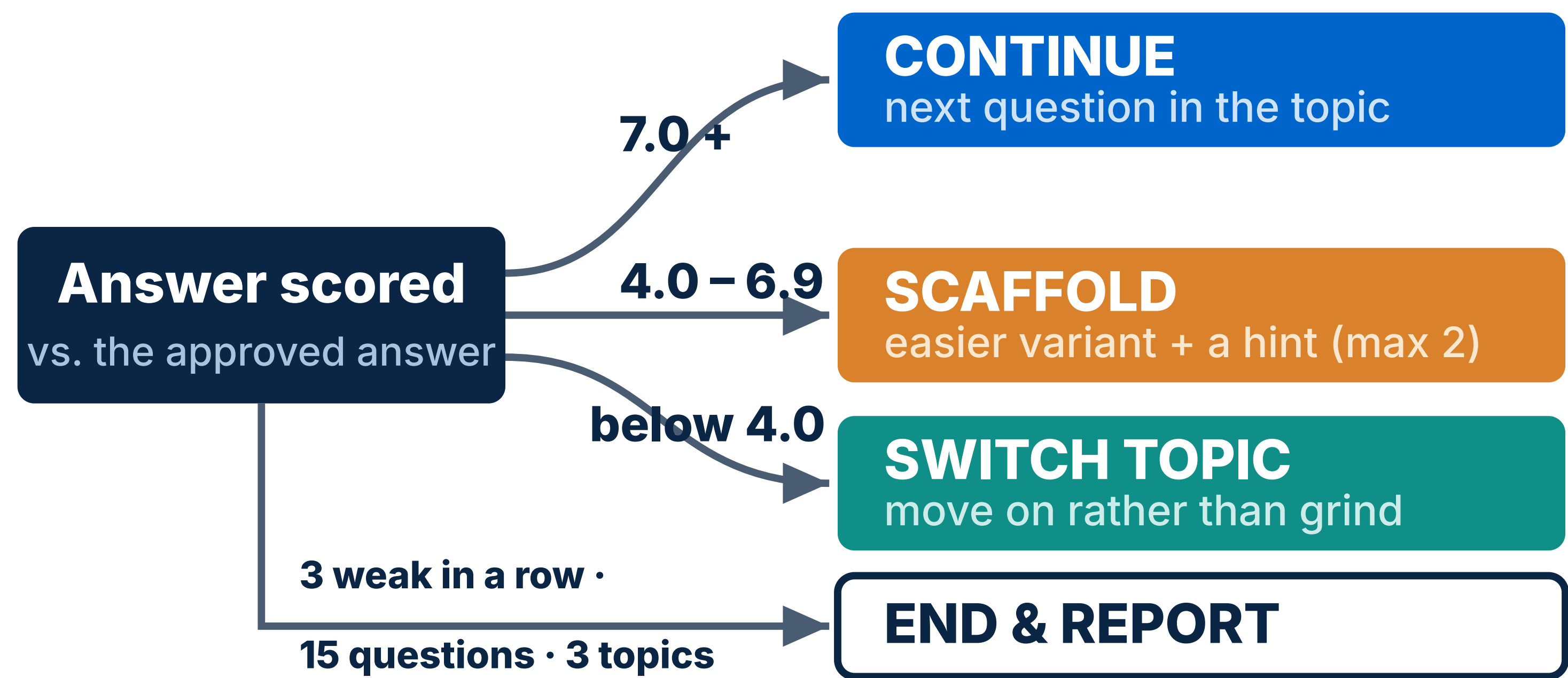
02 How it works

Retrieval-grounded generation · human approval gate · deterministic adaptation



PDF + OCR > chapter-aware chunking > text-embedding-004 > pgvector top-3 > Gemini 2.5 Flash > lecturer approval > Vietnamese or English > per-criterion rubric

03 Adaptation without hand-waving



The next move comes from **ordinary code with published thresholds**, not from the model. The model supplies wording; the policy supplies the decision.

05 What exists today

Implementation status

Flask + Docker

Next.js 16 / React 19

Supabase Postgres + pgvector

Gemini 2.5 Flash

Tesseract OCR

47 REST endpoints

8 service modules

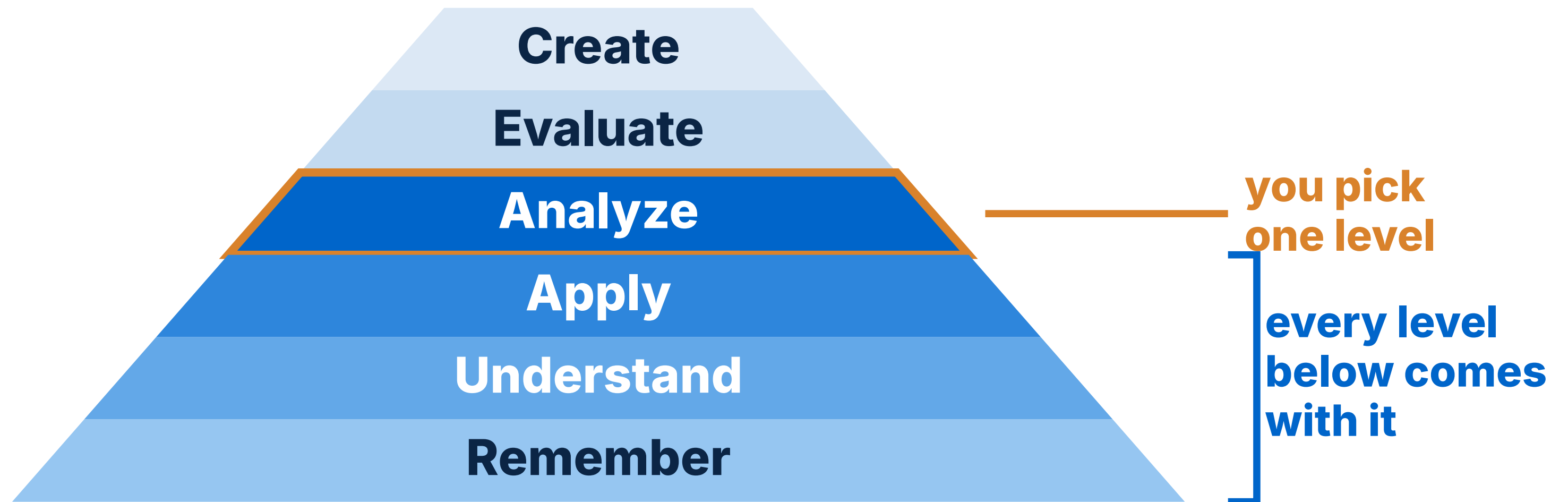
GitHub Actions CD

Honest status. Running, but no controlled trial yet — a NEU cohort pilot is next.

04 One pipeline, three settings

☒ Examination	LECTURER-OWNED
A whole cohort, one approved question set, one password-protected window.	
☒ Adaptive practice	STUDENT-INITIATED
Unlimited rehearsal, paced by the policy on the left.	
☒ Job interview	CAREER PREPARATION
Interviewed against a real CV and a real job description.	

Bloom levels are cumulative



A session climbs the ladder rather than sitting on one rung.

Left: an oral exam. Right: the lecturer's review. Below: a job interview, end to end

06 What it looks like in use

iView · Oral exam

Data Structures

ANALYZE Question 4 of 12

06:12 left

"Compare inserting into a hash table with inserting into a balanced BST. When would you choose each, and why?"

Speak or type your answer...

iView · Review

42 submissions

Nguyen Van A. 8.2

Tran Thi B 6.5

Accuracy 8.5, Depth 7.0, Clarity 9.0, Bloom hit L4

Override score & feedback

A job interview, end to end — the same pipeline, pointed at a CV instead of a syllabus

1 Bring your CV and the posting

CV — nguyenvana.pdf

JD — Data Analyst, fintech

Role: Data Analyst, Level: Fresher, Language: English, Length: 8 questions

2 Be interviewed

Question 3 of 8 Fresher · English

12:40 left

"Your CV lists a churn-prediction project. The model did well in training and badly in production. What did you do about it?"

"I checked the training and serving data first, and found the churn window had been computed differently..."

3 Get the report

7.6 overall Fresher level

Fit to the posting: 8.4, Structure of the answer: 6.8, Technical depth: 7.9, Clarity: 7.2

Every score stays overridable by a lecturer.

07 What we take from building it

- **Grounding plus a human gate** is what makes an LLM adoptable by a department.
- **A deterministic policy** makes the adaptation explainable.
- **One pipeline** serves the exam hall and the graduate interview.

Next: server-side speech · IRT calibration · reliability study · LMS integration.

Come and try it

Bring a PDF of your own teaching material to the stand — we will build you a viva from it while you wait.

github.com/Lampx83/iView

Source code & demo