AXAM: When There Is No Internet, AI-Powered Learning For the Last Mile

AUTHOR

Emmanuel Olim Kasigazi
M.S. in Data Analytics and Visualization

FACULTY MENTOR

Andrew Catlin, Industry Professor

ACKNOWLEDGEMENTS

The author thanks Professor Andrew Catlin for faculty mentorship and guidance; MIT OpenCourseWare for making 7,617+ lecture transcripts freely available under Creative Commons; the National Science Foundation I-Corps program for supporting early-stage customer discovery; and the Katz School of Science and Health.

Katz
Katz School of Science and Health

Introduction

- Across the world, millions of learners sit in classrooms rich in curiosity but poor in connectivity. Over 1.3 billion children globally lack internet access at home, and 40% of the world's population cannot access education in a language they understand (UNESCO, 2023). Despite rapid advances in AI-powered education, existing tools overwhelmingly require stable internet connectivity — where connectivity is absent, so too is access. This creates structural exclusion of the learners who need support most.

- The integration of AI into education has accelerated since 2022, yet attention remains concentrated on well-resourced, high-connectivity environments. Retrieval-Augmented Generation (RAG) has emerged as a promising framework for grounding AI responses in curated knowledge bases (Lewis et al., 2020), but existing RAG implementations assume cloud infrastructure. Their application in fully offline, low-resource educational settings remains largely absent from the literature.

- This project aims to design, build, and evaluate AXAM, an offline AI educational platform for low-connectivity environments, built on 7,617 MIT OpenCourseWare lecture transcripts (MIT OpenCourseWare, 2024) and a multilingual embedding model supporting 100+ languages (Chen et al., 2024). Objectives include retrieval accuracy evaluation (Precision@3, MRR), Large Language Model (LLM) as-judge answer quality scoring, and cross-lingual retrieval benchmarking across five languages.

- It is hypothesized that a fully offline RAG pipeline can deliver retrieval accuracy and answer quality sufficient for meaningful educational use — proving that world-class AI-powered learning does not require a world-class internet connection.

Approach

AXAM was built to work like an offline AI tutor — one that can answer student questions without needing the internet. To make this possible, three pieces had to come together:

1. **A library of knowledge.** I started with 7,617 free MIT OpenCourseWare lectures and turned them into a searchable digital library. Imagine taking thousands of class lectures and breaking them into small, easy-to-find pieces — like creating an index for every important idea inside every lecture. This means when a student asks a question, AXAM can quickly find the most relevant pieces of those lectures.
2. **A search engine that understands meaning**. Instead of matching keywords like a regular search bar, AXAM uses an AI model that understands what a student is actually asking — even if they phrase it differently or ask in another language. It works in over 100 languages, so a student in rural Texas asking in English and a student in another country asking in Spanish can both find the same answer.
3. **A small but capable AI tutor**. Once the right lecture content is found, a compact AI model uses that content to write a clear, accurate answer back to the student — entirely on the device, with no internet required.

To check how well AXAM works, the platform was tested in three ways: (1) how often it finds the right lecture content, (2) how well it answers questions across five languages, and (3) how high-quality the answers are when reviewed by an independent AI evaluator.

Approach, Cont'd

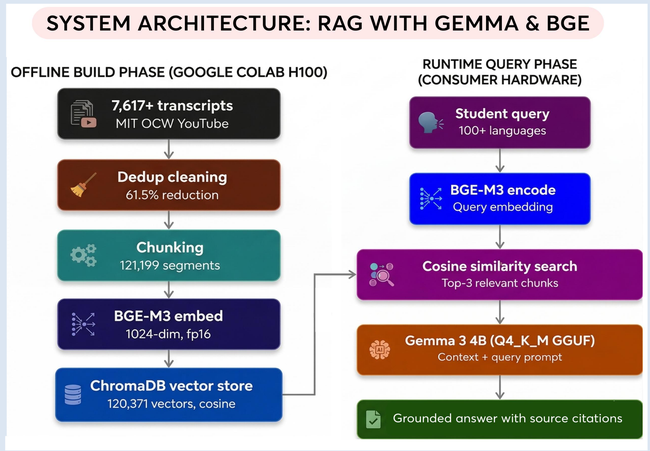

Figure 1: System Architecture: RAG with Gemma & BGE

Outcomes

AXAM finds the right lecture content nearly every time. Tested with 50 questions across 16 subjects in English, the system delivered correct results 99% of the time — and kept performing across five different languages.

Table 1. Cross-Lingual Retrieval Performance

Language	MRR	P@3	Avg Top-1 Similarity
English	1.000	1.000	68.0%
French	0.904	0.800	63.2%
Spanish	0.883	0.733	62.9%
Swahili	0.617	0.433	59.6%
Hindi	0.556	0.383	59.8%

How accurately does AXAM find the right lecture content when students ask questions in different languages? French and Spanish stayed close to English in accuracy — around 90% as good. Swahili and Hindi worked too, though less reliably, pointing to clear areas for future improvement.

Table 2. LLM-as-Judge Answer Quality (N=15)

Criterion	Mean	Min	Max
Relevance	4.60	1	5
Faithfulness	4.40	1	5
Completeness	3.93	1	5
Clarity	4.53	2	5
Overall	**4.37**		

Quality scores from an independent AI evaluator (1 = poor, 5 = excellent). The evaluator graded AXAM's answers on relevance, accuracy, completeness, and clarity. Overall score: 4.37 out of 5.0 — answers students can trust.

Conclusions

AXAM proves that AI-powered learning can work without the internet. The system found the right educational content nearly every time it was asked, and an independent AI evaluator rated its answers 4.37 out of 5.0 — accurate, clear, and trustworthy. It also worked across five languages, opening the door to learners far beyond English-only environments. The entire system fits in about 5.5 GB of storage, small enough to run on a school laptop or a USB stick — which means students can get high-quality learning support anywhere, regardless of whether they have internet access.

Limitations:

1. Answer generation defaults to English due to English-only corpus, despite multilingual retrieval success;
2. Reduced retrieval accuracy for lower-resource languages such as Swahili and Hindi;
3. Response times were benchmarked on cloud GPUs (Colab H100/T4), not representative of target consumer hardware;
4. Evaluation relied on keyword-based relevance heuristics bootstrapped from the system's own retrieval.

Future work: will focus on adding more languages to the content library, building automatic translation for harder questions, adding text-to-speech for students with disabilities, testing speed on real classroom hardware, and packaging AXAM as a simple desktop app for schools worldwide.

References: