Capstone Project Concept Note and Implementation Plan

Project Title: Lecture Companion: AI-Powered Translation and Summarization tool for Burmese Learners

Team Members
1. Lwin Naing Kyaw
2. Nang Hom Paung
3. Sai Aike Sam
4. Hom Nan Thawe Htun
5. MangShang SauYing
6. Ingyin Khin

1. Project Overview

The Lecture Companion project aims to develop an AI-powered tool that automatically translates, simplifies, and summarizes English lecture transcripts into Burmese, improving accessibility for Myanmar students in higher education. This initiative directly aligns with Sustainable Development Goal 4 (SDG 4): Quality Education, which emphasizes inclusive and equitable access to education for all learners. In Myanmar, a significant language barrier exists in higher education, where most instructional materials are available exclusively in English, making comprehension difficult for students from rural and non-English-speaking backgrounds (UNESCO, 2023). The proposed system addresses this challenge by leveraging state-of-the-art multilingual and multimodal large language models (LLMs) to bridge the educational gap through automated lecture translation and summarization.

The project integrates several critical components in its pipeline: (1) automatic speech recognition (ASR) for transcribing lectures, (2) neural machine translation (NMT) for English-to-Burmese conversion, (3) text simplification and summarization for accessibility, and (4) a retrieval-augmented interface for interactive Q&A. By employing models such as Whisper (Radford et al., 2023) for ASR and Gemini 2.5 Flash (Google DeepMind, 2025) for translation and summarization, the system delivers fast, accurate, and context-aware outputs even for low-resource languages. This approach is particularly impactful for Burmese learners, as Burmese remains an underrepresented language in mainstream AI development, and existing translation tools often fail to preserve domain-specific terminology or scientific accuracy (Nyein et al., 2022).

The Lecture Companion thus contributes to the broader goal of digital inclusion by enabling Myanmar's students to access global educational content in their native language. Beyond immediate pedagogical benefits, the system also has long-term implications for national educational equity, teacher training, and the development of localized AI technologies that support multilingual learning environments. By grounding its approach in responsible AI principles such as fairness, interpretability, and cultural sensitivity, the project positions itself as a prototype for accessible educational technology in low-resource contexts.

2. Objectives

The primary objective of the Lecture Companion project is to develop a scalable and reliable AI-assisted platform that enhances educational accessibility for Burmese learners through automatic translation, simplification, and summarization of lecture content. By integrating advanced speech recognition, neural machine translation, and text simplification models, the project aims to reduce the linguistic barriers that hinder comprehension and engagement in English-based tertiary education across Myanmar.

Specifically, the project is designed to achieve the following objectives:

Automated Speech-to-Text Transcription:
To implement a robust Automatic Speech Recognition (ASR) module capable of accurately transcribing English lecture audio into text, even under conditions of varied speaker accents and background noise. The model of choice, Whisper by OpenAI (Radford et al., 2023), provides high multilingual coverage and strong zero-shot capabilities suitable for academic audio content.

Neural Machine Translation (NMT) to Burmese:
To design a translation pipeline that effectively converts English lecture transcripts into fluent and semantically faithful Burmese text. The translation process will utilize Gemini 2.5 Flash (Google DeepMind, 2025) for context-aware, low-latency generation, complemented by fine-tuned bilingual sentence alignments drawn from datasets like OPUS and NLLB (Costa-jussà et al., 2022).

Text Simplification and Summarization:
To integrate a summarization layer that reduces the cognitive load for learners by producing B1-level simplified Burmese summaries with highlighted key points and formulas. This step enhances accessibility for students with varying English and academic proficiency levels, consistent with pedagogical guidelines for low-literacy education (Crossley et al., 2021).

Retrieval-Augmented Q&A System:
To develop a retrieval component that allows users to query lecture transcripts in Burmese and obtain context-grounded answers, leveraging vector search (e.g., FAISS or Chroma) to facilitate content-based retrieval and interactive learning.

Performance Evaluation and Optimization:
To benchmark model performance using both quantitative metrics (BLEU, ROUGE, and F1 for summarization) and qualitative human evaluation, ensuring fairness and reliability across topics and dialects within the Burmese language.

3. Background

Myanmar's higher education system faces a persistent language divide, where most tertiary instruction and learning resources are available primarily in English. While English proficiency remains low across the country, especially in rural regions, university students are expected to learn from English-language textbooks, research papers, and lectures (UNESCO, 2023). This mismatch significantly hinders comprehension, classroom participation, and academic performance. According to the British Council (2022), less than 20% of Myanmar university students can effectively engage with English academic materials without assistance. As a result, many learners rely on rote memorization rather than conceptual understanding, thereby widening the educational gap between English-proficient and non-proficient students.

Existing solutions for lecture translation and accessibility are predominantly designed for high-resource languages such as English, Spanish, or Mandarin. Tools like Google Translate or Microsoft Translator exhibit low accuracy when translating complex academic content into Burmese due to limited parallel corpora and poor handling of domain-specific terminology (Nyein et al., 2022). Additionally, Burmese presents unique linguistic challenges such as agglutinative morphology, flexible word order, and script-based segmentation, making traditional rule-based or statistical approaches ineffective (Myint et al., 2021). This linguistic complexity underscores the need for a neural machine translation (NMT) approach specifically adapted to Burmese.

Recent advances in Large Language Models (LLMs) and multimodal AI systems offer promising solutions to overcome these limitations. End-to-end architectures like Whisper (Radford et al., 2023) can perform robust multilingual speech recognition without extensive fine-tuning, while Gemini 2.5 Flash (Google DeepMind, 2025) provides ultra-fast multimodal translation and summarization capabilities optimized for low-resource adaptation. These advances enable the development of a pipeline capable of processing real-world educational audio and producing contextually faithful Burmese summaries.

Moreover, previous initiatives in low-resource language processing, such as Meta AI's No Language Left Behind (NLLB) (Costa-jussà et al., 2022), have demonstrated that multilingual pretraining and transfer learning can significantly enhance translation performance for underrepresented languages. By leveraging such foundation models and adapting them for educational contexts, Lecture Companion situates itself at the intersection of language technology, accessibility, and educational equity.

In addition to technical innovation, the project addresses the broader socio-economic goal of inclusive education under Sustainable Development Goal 4 (Quality Education). By enabling Burmese learners to understand lectures in their native language, the tool promotes equal learning opportunities, supports teacher-led instruction, and enhances lifelong learning. Furthermore, it contributes to Myanmar's digital transformation efforts by introducing localized AI solutions that empower learners in resource-constrained environments.

4. Methodology

The Lecture Companion system is designed as a modular pipeline that integrates speech recognition, neural machine translation, text simplification, summarization, and retrieval-based interaction. Each component is built using modern deep learning architectures optimized for low-resource adaptation and real-time inference. The overarching goal is to transform raw lecture audio into accessible Burmese summaries while maintaining contextual accuracy, fluency, and educational relevance.

4.1 System Overview

The system pipeline comprises five sequential stages:

1. Input and Preprocessing
2. Speech Recognition (ASR)
3. Machine Translation (NMT)
4. Text Simplification and Summarization
5. Retrieval-Augmented Q&A

Each stage is supported by specific algorithms and models chosen for their efficiency, multilingual capabilities, and compatibility with Burmese text structures.

4.2 Automatic Speech Recognition (ASR)

The ASR module uses Whisper (Radford et al., 2023), a Transformer-based model trained on 680,000 hours of multilingual audio. Whisper operates via an encoder–decoder Transformer architecture where: The encoder converts MFCC features into latent representations using self-attention. The decoder autoregressively predicts text tokens using a cross-attention mechanism between the encoder outputs and past token embeddings.

Mathematically, for an input sequence $x = (x1, ..., xT)$ and token sequence $y = (y1, ..., yN)$, the conditional probability is modeled as:

$$P(y|x) = \prod_{t=1}^{n} P(y_t | y_{<t}, h)$$

Where h= encoder outputs

4.3 Neural Machine Translation (NMT)

The translated output is generated using **Gemini 2.5 Flash (Google DeepMind, 2025)**, a multimodal large language model optimized for translation and summarization. It integrates:

- **Multilingual embedding layer** pretrained on diverse text pairs
- **Transformer encoder-decoder backbone** with **sparse mixture-of-experts (MoE)** layers for efficient scaling
- **Contrastive decoding** to retain semantic fidelity between English and Burmese

4.4 Text simplification and summarization

To ensure accessibility, the translated Burmese text is simplified using a two-step sequence-to-sequence model:

1. **Simplification**: Uses fine-tuned **Gemini Flash's summarization head** with a rule-based complexity controller.

2. **Summarization**: Applies extractive + abstractive hybrid summarization, combining TextRank (Mihalcea & Tarau, 2004) for key sentence extraction with transformer-based summarization for natural phrasing.

4.5 Retrieval Augmented Q&A

A retrieval-augmented generation (RAG) module supports user interaction. The simplified transcript and summaries are embedded into a FAISS vector index using Sentence-BERT embeddings (Reimers & Gurevych, 2019). Queries in Burmese are vectorized and matched to the most semantically similar transcript chunks for contextual answers generated by Gemini Flash.The retrieval score is computed as cosine similarity:

$$sim(q, d) = \frac{q \cdot d}{||q|| \ ||d||}$$

This architecture allows learners to ask concept-level questions (e.g., “Explain Ohm’s Law”) and receive grounded responses from the lecture context.

5. Architecture Design Diagram

The overall system architecture of the Lecture Companion project integrates multiple AI components into a modular, end-to-end pipeline that converts English lecture content into simplified Burmese transcripts and interactive learning materials. The design emphasizes accessibility, scalability, and modularity, separating automatic speech recognition (ASR), translation and summarization, and retrieval-augmented question-answering (RAG) into distinct yet connected layers. Figure 1 provides a high-level overview of this dataflow, illustrating how lecture inputs are processed through each stage and visualized in the bilingual learning interface.

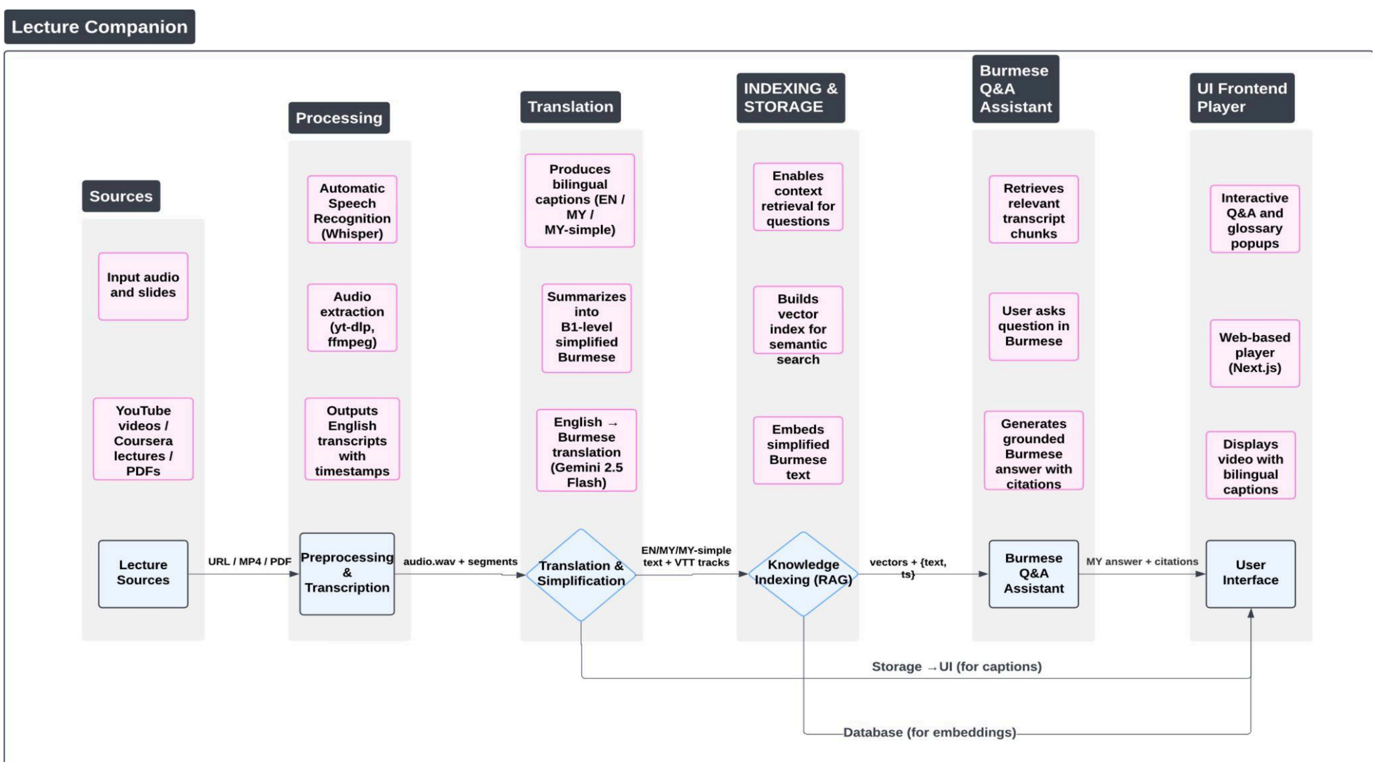

Figure1 - Overview of the main components in the Lecture Companion system, showing the flow from lecture sources through preprocessing, translation, and indexing to the bilingual user interface with Q&A support.

5.1 High-level dataflow (end-to-end)

As shown in Figure 2, the Lecture Companion follows a linear yet modular data pipeline. Lecture media are first downloaded and converted to audio using yt-dlp and ffmpeg, then passed to the ASR module (Whisper) for transcription. The resulting English transcripts are translated and simplified using Gemini 2.5 Flash, generating bilingual captions and simplified Burmese summaries. These texts are segmented, embedded, and stored in a vector database to support retrieval-augmented generation (RAG). Finally, the frontend application presents the processed results as synchronized captions and an interactive Burmese Q&A feature, completing the transformation from raw lecture input to accessible learning output.

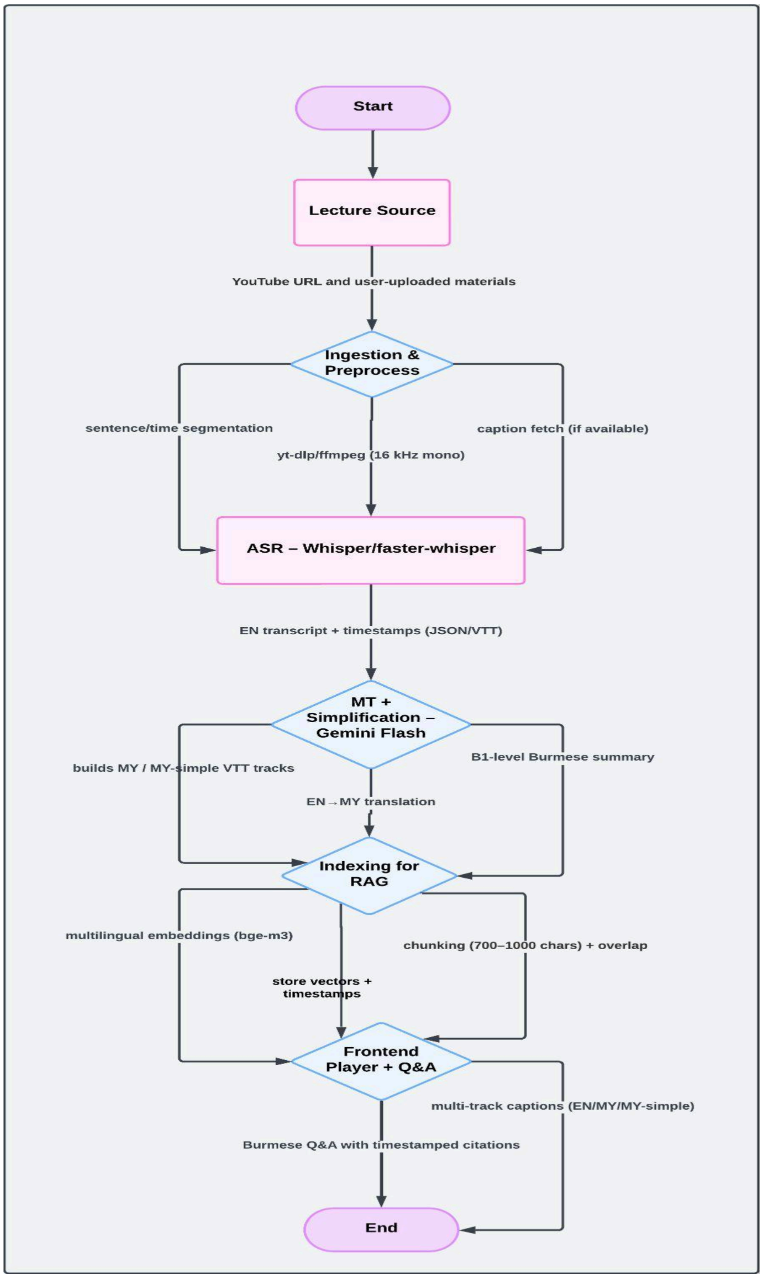

FIgure 2 - End-to-end dataflow illustrating how lecture inputs are transcribed, translated, simplified, and indexed before being presented as bilingual captions and interactive Burmese Q&A.

6. Data Sources

The Lecture Companion project integrates a combination of open academic corpora and user-generated data to support multilingual translation, summarization, and pedagogical text simplification tasks. The design of our data strategy emphasizes representativeness, accessibility, and adaptability to low-resource language contexts, specifically, Burmese. By combining curated English lecture transcripts, standardized language proficiency corpora, and authentic user-uploaded lecture data, we aim to establish a realistic and scalable foundation for both model development and real-world deployment.

For prototyping and benchmarking, we employ publicly available English lecture datasets, notably the **Andrew Ng Machine Learning Course Transcripts from Kaggle (Ng, 2021)**. This corpus provides domain-rich, technical lecture text that is ideal for evaluating translation fidelity and summarization accuracy in academic content. Each transcript contains clear speaker structure and contextual segments, allowing for structured alignment during the ASR and translation phases. Such data are crucial for testing domain-specific translation quality, particularly for scientific and engineering vocabulary that Burmese learners often find challenging to comprehend in English.

In parallel, the system will process **YouTube lecture videos and user-uploaded materials**, enabling dynamic collection of authentic, diverse speech patterns and accents. This approach not only improves model generalization across educational contexts but also aligns with the system's deployment goal, allowing learners to upload or stream any English lecture and receive Burmese translations and summaries in real time. All uploaded data will undergo preprocessing steps, including audio normalization, silence trimming, and automatic sentence segmentation, to ensure consistent input quality for ASR and translation modules.

To guide the simplification and readability calibration, we utilize reference corpora based on Common European Framework of Reference (CEFR)-aligned datasets. Specifically, we adopt the **CEFR-English dataset available on Hugging Face** (Azpiazu & Pera, 2019), which includes sentence-level CEFR proficiency annotations. This dataset supports the generation of English language vocabulary classifier and helper. Although the CEFR data are English-based, they serve as a linguistic reference for defining simplification thresholds and fine-tuning summary generation prompts within Gemini 2.5 Flash.

7. Literature Review

Work on Burmese (Myanmar) NLP consistently shows that segmentation and orthography are first-order constraints on downstream ASR/MT/summarization quality. Classic and recent studies argue that Myanmar text lacks reliable whitespace tokenization and mixes legacy encodings with Unicode; consequently, effective pipelines must incorporate syllable- or dictionary-aware segmentation prior to modeling (or learn to be robust without explicit words). Thu et al. (2014) demonstrate that injecting small dictionary resources into an unsupervised segmenter "improved the F-score from 0.48 to 0.66," highlighting how even light linguistic priors can materially lift Myanmar tokenization in low-resource conditions. Parallel findings from broader low-resource Myanmar NLU emphasize infrastructure challenges e.g., non-standard encodings and sparse corpora, stating that Myanmar content “varies in quality significantly, and very few good datasets…are available,” which directly motivates normalization and quality-controlled data curation in new systems.

For machine translation, Myanmar-specific neural MT studies report that preprocessing with rule-based syllable breaking and normalization is beneficial; Lwin & Wai explicitly note that a “rule-based syllable breaking approach is used” before attention-based NMT, yielding their best BLEU under modest data (≤100k pairs), an empirical pattern our pipeline mirrors via Unicode normalization and syllable-aware tokenization before LLM translation and simplification. On the speech side, end-to-end Myanmar ASR under data scarcity has achieved strong baselines with CTC/BLSTM or Transformer variants; Chit & Lin report a best CER 4.72% / SER 12.38% on a ~26-hour corpus, which we take as evidence that compact, well-regularized ASR backbones can be effective when paired with careful segmentation and domain tuning. Finally, Myanmar-focused evaluation work in cross-lingual NLU (myXNLI) further codifies low-resource pain points and shows that judicious data augmentation and transfer can yield small but consistent gains, reinforcing our choice to combine Myanmar-aware preprocessing with multilingual LLMs for translation, simplification, and retrieval, inspired by multilingual transfer paradigms (e.g., NLLB) without directly adopting their training recipe in this capstone.

Building on these findings, our system: (i) adopts Unicode normalization + syllable-aware segmentation (following Thu et al.; Lwin & Wai) before any modeling; (ii) uses end-to-end ASR proven viable for Myanmar to feed clean transcripts to downstream modules (following CTC-style evidence); and (iii) layers LLM-based translation + simplification with retrieval augmentation to address domain terminology and pedagogy, targeting not only literal adequacy but educational usability (readability-controlled Burmese summaries and Q&A grounded in the lecture transcript). This hybrid, Myanmar-first design operationalizes the core lesson from the literature: lightweight linguistic priors + robust end-to-end models + careful data hygiene outperform naïve, word-level pipelines in low-resource Burmese settings.

Implementation Plan

1. Technology Stack

The *Lecture Companion* system is developed using a combination of **Python-based machine learning frameworks**, **large language model APIs**, and **modern deployment tools**. The selected stack prioritizes efficiency, scalability, and accessibility for low-resource language contexts like Burmese, ensuring smooth integration across transcription, translation, and retrieval modules.

Programming Language

- **Python (v3.10+)** — chosen for its extensive ecosystem of libraries in natural language processing (NLP), speech recognition, and data handling.

Core Machine Learning and Data Libraries

Component	Library / Tool	Purpose
Speech Recognition (ASR)	`faster-whisper`	Efficient implementation of OpenAI's Whisper model for converting lecture audio into English text. Optimized for multilingual and accent-robust transcription.
Translation and Simplification	`google-generativeai` (Gemini 2.5 Flash API)	Performs English-to-Burmese neural machine translation and B1-level simplification, generating learner-friendly summaries.
Sentence Embeddings	`sentence-transformers` (multilingual model)	Converts text into dense vector representations for semantic retrieval in the RAG system.
Vector Search / Retrieval	`faiss-cpu` (for prototype) / `faiss-gpu` (for production)	High-speed similarity search over vectorized transcript segments, enabling efficient question-answer retrieval.
Data Ingestion	`PyPDF2`, `youtube-transcript-api`, `yt-dlp`, `ffmpeg-python`	Extracts and preprocesses text or audio data from online lecture sources such as YouTube and Coursera.

Evaluation and Analysis	`sacrebleu`, `scikit-learn`, `pandas`, `matplotlib`	Computes BLEU, ROUGE, and other performance metrics; supports data visualization and quantitative evaluation.

Application Framework and Deployment

Layer	Technology	Function
Backend API	**FastAPI**	Manages the machine learning pipeline endpoints (ASR → MT → Simplify → RAG) with async processing.
Frontend/UI	**Next.js 14 + TailwindCSS** *(planned)*	Provides an interactive interface for bilingual transcript viewing and Burmese Q&A.
Database and Storage	**PostgreSQL (+ pgvector)**, **MinIO/S3**, **Redis**	Stores processed transcripts, vector embeddings, and cached model outputs.
Task Queue	**Celery + Redis**	Handles long-running jobs such as transcription and translation asynchronously.
Deployment Environment	**Docker / Docker Compose**	Containerized setup ensuring reproducibility and smooth cross-platform deployment; GPU-enabled for ASR/LLM acceleration.

Hardware and Cloud Environment

- **Development:** Local GPU workstation or Google Colab for prototyping.
- **Production:** Cloud VM (e.g., GCP A2 / AWS g4dn) with ≥ 16 GB GPU memory for real-time inference.

2. Timeline

The overall draft project timeline is summarized in **the excel sheet** (link here). The tasks for the project have been delegated and the assigning of date and role are still in progress.
The gantt chart covers **10 weeks (from October to December 2025)** and illustrates the sequential development of the three main pipelines:

1. **Pipeline 1 – Core Lecture Processing (ASR → Translation → Summarisation)**
    - *Weeks 1–5:* Data ingestion, ASR setup (Faster-Whisper), and Gemini translation-simplification integration.
    - *Deliverable:* End-to-end transcript generation producing English and simplified Burmese VTT files.
2. **Pipeline 2 – RAG Q&A System (Embedding → Retrieval → LLM Answer)**
    - *Weeks 5–8:* Vectorisation of processed transcripts, FAISS/pgvector index construction, and integration of Gemini for context-grounded Q&A in Burmese.
    - *Deliverable:* Interactive query API returning time-stamped grounded responses.
3. **Pipeline 3 – CEFR Vocabulary Classifier (Difficulty → Glossary → Frontend Highlight)**
    - *Weeks 7–9:* Development of lexical-difficulty classifier, CEFR annotation, and front-end vocabulary visualization.
    - *Deliverable:* Color-coded transcript viewer showing CEFR-based word highlights and glossary pop-ups.
4. **Final Integration and Evaluation (Weeks 9–10)**
    - Combine all modules into the unified *Lecture Companion* interface.
    - Conduct user evaluation, collect feedback, and finalize documentation and presentation.

Week / Dates	Major Focus	Pipeline(s)	Key Tasks	Deliverables / Checkpoints
Week 1 (Oct 20 – Oct 26)	Setup & Data Collection	P1	• Team setup & role assignment • Collect lecture videos (YouTube/Coursera) • Extract audio (yt-dlp + ffmpeg) • Initialize Git & shared Drive structure	Data folder + project repo ready
Week 2 (Oct 27 – Nov 2)	ASR Implementation	P1	• Run Faster-Whisper medium model on sample lectures • Transcript cleaning & segmentation • Preliminary WER test	English transcripts (ASR) v1
Week 3 (Nov 3 – Nov 9)	Translation & Simplification Prototype	P1	• Integrate Gemini 2.5 Flash API • Prompt engineering for EN→MY translation and B1-level simplification • Evaluate BLEU/ROUGE	Bilingual VTT captions v1
Week 4 (Nov 10 – Nov 16)	End-to-End Pipeline Testing	P1 + P2 (start)	• Automate ASR→Translate→Summarize flow • Build evaluation scripts • Start embedding experiments (sentence-transformers)	Core pipeline demo + metrics
Week 5 (Nov 17 – Nov 23)	RAG Q&A System Development	P2	• Create vector index (FAISS/pgvector) • Implement query API • Integrate Gemini for context-grounded answers • Begin UI prototype (Q&A panel)	Working RAG endpoint
Week 6 (Nov 24 – Nov 30)	Testing & Evaluation	P1 + P2	• Human evaluation of translation quality • Measure retrieval similarity scores • Refine Gemini prompt templates • Collect user feedback (round 1)	Evaluation report v1

Week 7 (Dec 1 – Dec 7)	CEFR Classifier & Glossary and Translation Integration	P3	• Train CEFR difficulty model • Highlight vocab in transcript • Connect free dictionary/glossary and google translate API • Front-end highlight tooltips	CEFR annotated transcript
Week 8 (Dec 8 – Dec 14)	Final Integration & UI Polish	P1 + P2 + P3	• Combine all pipelines into FastAPI + Next.js app • User interface testing • Prepare demo video • Freeze code for final submission	Unified Lecture Companion demo
Week 9 (Dec 15 – Dec 21)	Evaluation & Presentation	All	• Finalize report and slides • Collect final metrics • Internal presentation rehearsal • FTL Capstone Day presentation (Dec 21)	Final submission & presentation

3. Milestones

Milestone No.	Description	Expected Output / Deliverable	Target Week
M1	*Core Pipeline Functional* – Lecture video can be transcribed, translated, and summarised into Burmese.	ASR → EN text → MY translation → B1 summary (VTT files).	Week 3
M2	*End-to-End Integration (Phase 1)* – Pipeline 1 fully automated and reproducible from video input.	Scripted pipeline; evaluation scores (WER, BLEU).	Week 5
M3	*RAG Q&A Module Ready* – Transcript embeddings + vector database built; Gemini answers grounded in context.	Working FastAPI endpoint for Burmese Q&A.	Week 7
M4	*CEFR Vocabulary Classifier Prototype* – Predicts difficulty and highlights vocabulary in UI.	Annotated transcripts with CEFR color coding & tooltips.	Week 8
M5	*Unified Web Interface* – All modules deployed in Next.js frontend with FastAPI backend.	Fully functional “Lecture Companion” demo site.	Week 9
M6	*Final Evaluation & Presentation* – Quantitative metrics + user feedback collected; report submission.	Capstone report + poster + presentation slides.	Week 10

4. Challenges and Mitigations

Challenge: API quotas and costs for Gemini.
Mitigation: Secure a research/educational grant or paid plan. Implement robust caching of all API results (transcripts, translations, summaries) to avoid redundant calls.
Challenge: Poor translation/simplification quality for highly niche technical terms.
Mitigation: Continuously iterate on the Gemini prompts. Integrate the planned "Glossary" feature to inject correct Burmese terms. Use human evaluation to identify and fix common simplification errors.
Challenge: ASR inaccuracy for strong non-native accents.
Mitigation: faster-whisper is generally robust, but we will test different model sizes (e.g., medium) if the small model proves insufficient.
Challenge: RAG system returns irrelevant context.

Mitigation: Fine-tune the text segmentation (chunking) strategy. Experiment with different embedding models if the initial one performs poorly on Burmese.

5. Ethical Considerations

Data Privacy: Our primary data sources are publicly available educational lectures. We will not be processing any private or personal user data. If a "file upload" feature is implemented, the UI will clearly state that content is processed by external APIs and not stored long-term.
Bias: AI models can carry inherent biases. The ASR model may have higher error rates for certain non-native English accents. The LLM may carry cultural biases. We will be mindful of these risks, but a full debiasing is outside our project's scope. We will document any significant biases observed during evaluation.
Accuracy and Impact: The primary ethical risk is providing an incorrect translation or simplification that leads to a student's misunderstanding.
Mitigation: We will not claim 100% accuracy. The tool will be explicitly branded as a "Lecture Companion" or "Study Aid," not a replacement for the original lecture. The UI will always provide the original English text alongside the Burmese translation and summary, encouraging users to compare and use their own judgment.

6. References

Azpiazu, I. M., & Pera, M. S. (2019). CEFR-based lexical and reading comprehension datasets for assessing English language proficiency. In Proceedings of the 14th Workshop on Innovative Use of NLP for Building Educational Applications (pp. 274–284). Association for Computational Linguistics.

British Council. (2022). *English proficiency and higher education in Myanmar: Barriers and opportunities.* British Council Myanmar.

Chit, K. M., & Lin, C. (2023). An end-to-end speech recognition system for the Myanmar language. In *2023 7th International Conference on Information Technology (InCIT)* (pp. 1–6). IEEE. https://doi.org/10.48550/arXiv.2105.06253

Costa-jussà, M. R., Cross, J., Ćelebi, O., Elbayad, M., Heafield, K., Heffernan, K., & Fan, A. (2022). *No language left behind: Scaling human-centered machine translation.* arXiv preprint arXiv:2207.04672.

Crossley, S. A., Skalicky, S., & McNamara, D. S. (2021). Text simplification and accessibility: Advances in computational linguistics approaches. *Computational Linguistics Journal, 47*(2), 391–419.

Google DeepMind. (2025). *Gemini 2.5 Flash: Scaling efficient multimodal reasoning.* arXiv preprint arXiv:2507.06261.

Lwin, K. T., & Wai, Y. M. (2021). Myanmar language neural machine translation using rule-based syllable breaking approach. In *2021 IEEE International Conference on Computer Science and Artificial Intelligence (CSAI)* (pp. 52–58). IEEE.

Mihalcea, R., & Tarau, P. (2004). TextRank: Bringing order into texts. In *Proceedings of the 2004 Conference on Empirical Methods in Natural Language Processing (EMNLP 2004)* (pp. 404–411). Association for Computational Linguistics.

Myint, M. T., Aye, P. P., & Tun, Z. M. (2021). Challenges in machine translation for the Burmese language. *Myanmar Language Technology Workshop Proceedings.*

Ng, A. (2021). Machine Learning Course Transcripts [Dataset]. Kaggle. https://www.kaggle.com/datasets/thebrownviking20/andrew-ng-machine-learning-course

Nyein, K., Khin, I., & Aung, M. (2022). Challenges of machine translation for low-resource languages: A case study on Burmese-English. *Myanmar Language Technology Workshop.*

Radford, A., Kim, J. W., Xu, T., Brockman, G., McLeavey, C., & Sutskever, I. (2023). *Robust speech recognition via large-scale weak supervision.* arXiv preprint arXiv:2302.03540.

Reimers, N., & Gurevych, I. (2019). Sentence-BERT: Sentence embeddings using Siamese BERT-networks. *arXiv preprint arXiv:1908.10084.*

Thu, Y. K., Finch, A., Utiyama, M., & Sumita, E. (2014). A joint word segmentation and POS tagging model for Myanmar language. In *Proceedings of the 2014 International Workshop on Asian Language Resources* (pp. 1–8). Association for Computational Linguistics.

UNESCO. (2023). *Education in Myanmar: Barriers to access and equity.* United Nations Educational, Scientific and Cultural Organization.

Win, K. M., Marasinghe, A., & Aye, T. T. (2023). Low-resource natural language understanding and translation for Myanmar (Burmese): Challenges and recent advances. *Language Resources and Evaluation, 57*(2), 483–504