SECTION 4 • DIGITAL WORLDS AND INTELLIGENT TECHNOLOGIES: FROM DATA TO DECISIONS

Application of Multimodal Generative Models and Script Automation for Creating Reproducible English Lectures

Practical Experience of Building an Automated Pipeline

Speaker: Olena Venhrina

ABOUT THE AUTHOR



Olena Venhrina

PhD, Associate Professor at the Department of Cybersecurity and IT

- **Institution:** S. Kuznets KhNUE
- **Experience:** Over 6 years of teaching
- **Interests:** Applied LMMs, Generative AI in education, Human-in-the-Loop pipelines, project management
- **Work email:** olena.venhrina@hneu.net

Generative AI | Human-in-the-Loop | AI-Assisted Scripting (Python / VBA) | Quarto / Reveal.js | TTS Voice Synthesis | Project Management in Construction

Challenges in Creating English-Taught Courses

RELEVANCE

Three Main Obstacles for the Teacher

High Workload

- Manual translation of domain-specific terminology
- Significant time spent recording audio narration
- Preparing a high-quality lecture takes days of work

Language Barrier

- Psychological discomfort and fear of making mistakes
- Uncertainty about pronunciation and speech rhythm
- Difficulty maintaining a steady pace for 30 minutes

Risks of Direct Generation

- Blind AI generation leads to hallucinations
- Distortion of didactic logic and formulas
- Need for mandatory teacher verification

The Human-in-the-Loop Concept

HITL CONCEPT

Separation of Expert Work and Routine

Teacher (Expert)

Control at Every Stage

- Creating source notes with code and formulas
- Expert terminology review after translation
- Aligning speech style and lecturer pace
- Full quality control and final material sign-off

HitL

AI and Scripts

Technical Routine

- Machine translation without breaking Quarto syntax
- Automatic structure generation for draft slides
- Generating draft speaker notes for each slide
- Voice synthesis and presentation synchronization

Why Quarto Instead of Word

TECHNOLOGICAL FOUNDATION

Advantages of Plain Text Over a Word Processor

Traditional Word (.docx)

AI Limitations

- Closed binary format makes script processing difficult
- Embedded math formulas get corrupted during translation
- Code blocks lose indentation and formatting
- High risk of accidental markup changes during copying

VS

Quarto Markdown (.qmd)

Docs-as-Code

- Plain text file, ideal for Git version control
- Native support for mathematical formulas in LaTeX
- Clear separation of code blocks from lecture content
- Guaranteed syntax preservation during API processing

Overcoming Initial Friction

Role of the AI Agent

- Transition requires adaptation: IDE workflow and compiler options
- AI assistant in the IDE handles technical markup routine
- Long-term automation benefits outweigh initial inconvenience

quarto.org

Official documentation

From Lecture Notes to Speaker Script

STAGE 1: CONTENT AND STRUCTURE

First Stage of the Automated Pipeline: Notes, Slide Structure, and Script

STEP 01

Source Notes (QMD)

Author's document with formulas, code, and pedagogical structure.

`Quarto .qmd`

STEP 02

Semantic Translation

Python script via Gemini API translating while protecting code and syntax.

`Gemini 2.5 Flash`

STEP 03

Slide Creation

Loading lecture notes into Gemini Notebook for automated PPTX slide generation.

`NotebookLM (PPTX)`

STEP 04

Speaker Notes

Custom Gem-bot synthesizes voiceover notes for each slide under author review.

`Custom Gem-Bot`

`quarto.org`
Quarto Markdown

`aistudio.google.com`
Google AI Studio API

`notebook.google`
Gemini Notebook

`gemini.google.com`
Google Gemini Gems

From Notes to Ready-to-Use Video

STAGE 2: AUTOMATION

Second Stage of the Automated Pipeline: Audio Synthesis, VBA Script, and Final Video

STEP 05

Export Notes

Script extracts reviewed speaker notes and saves them per slide.

`extract_notes.py`

STEP 06

Speech Synthesis

Model synthesizes natural voiceover with clear technical pronunciation.

`generate_audio.py`

STEP 07

Audio Embedding

Macro attaches audio files to matching slides and reads duration.

`macros.vba`

STEP 08

Slide Sync

Slides transition automatically right after the voiceover ends.

`Autonomous PPTX`

Practical Result: Lecture Fragment

PRACTICAL CASE

Autonomous Slide Playback with Synchronized Audio Narration



VOICE SYNTHESIS

Gemini Neural TTS (clear pronunciation, 24 kHz)

SYNCHRONIZATION

Slide transitions based on audio duration (VBA)

AUTHOR INVOLVEMENT

Fully autonomous during playback

APPLICATION

Live Zoom broadcast, LMS hosting, or YouTube

Practical Benefits for the Educational Process

EFFICIENCY

Time Savings, Academic Accuracy, and Overcoming the Language Barrier

Time Savings

Automating draft structuring, speaker notes, and audio generation drastically reduces lecture preparation compared to manual work.

Academic Accuracy

Preserving author content and correct terminology without fabricated facts due to step-by-step verification.

Language Confidence

Natural English pronunciation of terms and steady speaking pace without the need for multiple voice rerecordings.

COURSE MATERIALS

Open Access Lecture Videos on YouTube

Practical validation of the method in real semester classes

Conclusions and Open Repository

CONCLUSIONS

Research Results, Reproducibility Principles, and Open Materials

1. **Key Role of Human-in-the-Loop**
   Artificial intelligence handles technical routine, while academic responsibility and content verification remain with the teacher.

2. **Combining AI with Accessible Scripting**
   Linking generative models with Python and PowerPoint macros transforms notes into self-contained learning materials.

3. **Reproducibility and Open Access**
   All Quarto templates, prompts, scripts, and macros are open for colleagues in universities to use.



Open-source Project Code:

https://github.com/venhrina-olena/auto_lecture