

Reflection on Design Choices and Limitations

In order to balance functionality, flexibility, and simplicity, the AI-Powered PPTX pipeline was created to automate the entire process of creating presentation decks from raw text. Dividing the workflow into three phases—text-to-slide planning, slide rendering, and evaluation—was a fundamental design decision. This modular approach offers a flexible backup plan in the event that cloud AI access is not possible. It enables both local slide planning using NLTK tokenization and an LLM-powered plan via Gemini.

Different levels of detail were produced for "expert," "graduate," and "executive" users through deliberate audience-variant awareness in slide planning. This illustrates the design principle of adjusting content to the audience's cognitive requirements. Sentence tokenization and heuristic chunking, which strike a balance between simplicity and reasonably coherent slide content, are the foundation of text segmentation.

Compatibility with pre-existing PowerPoint templates was given top priority when making rendering decisions. To create visually consistent slides, placeholder mapping, style extraction (fonts, colors, title/body sizes), and automated insertion of charts, images, and LaTeX math were used. In order to ensure that slides are never empty, the system also supports fallback mechanisms, such as creating demo charts or images when none are supplied.

The evaluation framework, which is supplemented by a CSV template for human review, prioritizes semantic fidelity and design quality through keyword F1 scoring and simple slide layout checks. This offers both automated and human-in-the-loop quality assurance, demonstrating a practical approach.

Despite these advantages, there are still a few drawbacks. The availability of PowerPoint placeholders limits automated layout, and two-column or chart placements might not always be the most aesthetically pleasing. The system does not yet extract or display actual numerical datasets; instead, charts are currently filled with arbitrary demo data. Math inclusion is restricted to demo formulas and is static. Although promising, Gemini-based planning requires robust error handling because it can fail due to malformed outputs or missing API keys. Last but not least, human review is required for final polish in semantic coherence, narrative flow, and audience impact.

In conclusion, the pipeline effectively demonstrates automated text-to-PPTX conversion with modular, variant-aware design, but achieving fully polished, data-driven, and contextually coherent slides requires further integration of intelligent layout algorithms, real-world data handling, and advanced evaluation metrics. With obvious room for improvement in terms of visual appeal, content accuracy, and AI-driven slide intelligence, the design strikes a balance between useful functionality and future extensibility.