

Design First, Code Later: Aesthetically Pleasing Template-Free Slides Generation

Zhiyao Cui^{1,2,3†‡}, Chenxu Wang^{2,4†‡}, Shuyue Hu², Yiqun Zhang^{2‡}
Wenqi Shao^{2,3}, Qiaosheng Zhang^{2,3*}, Zhen Wang^{1*}

¹School of Cybersecurity, Northwestern Polytechnical University,
²Shanghai Artificial Intelligence Laboratory,
³Shanghai Innovation Institution,
⁴Fudan University

{cuizhiyao,wangchenxu,hushuyue,zhangyiqun,shaowenqi,zhangqiaosheng}@pjlab.org.cn,w-zhen@nwpu.edu.cn

Abstract

Producing presentation slides automatically entails coordinating narrative structure with page-level graphic design under strict spatial constraints. For such structured multimodal tasks, a well-organized design process is essential to ensure the final quality of slides. Existing approaches rely on fixed templates or directly emit executable code, thereby both limiting the creative layout-design capabilities of LLMs and bypassing the essential slide-page design step. To address these limitations, this paper: (1) proposes a hierarchical slides generation workflow **DeepSlides** that systematically organizes slide design tasks without any predefined template or style, decoupling slide-page design from implementation; (2) introduces **SlideDesign**, a dataset tailored specifically for slides generation tasks; (3) presents a multi-agent reinforcement learning training paradigm and trains a couple of models **SlideQwens** for slide design and implementation. Experimental results demonstrate that our proposed framework outperforms baseline methods on evaluated metrics and achieves superior performance in human preference evaluations. The dataset and code are available at: <https://github.com/sxswz213/DeepSlides>

1 Introduction

With the rapid advancement of large language models (LLMs), an increasing number of LLM-based agents are capable of automating various tasks, significantly reducing human workloads (Hong et al., 2024; Wang et al., 2025). In particular, multimodal LLMs possess robust creative abilities and multimodal perception (Meng et al., 2025), prompting researchers to explore their application in automated slide generation tasks. As shown in

[†]Co-first, equal contributions.

^{*}Corresponding authors.

[‡]This work was done during the authors' internship at Shanghai Artificial Intelligence Laboratory.

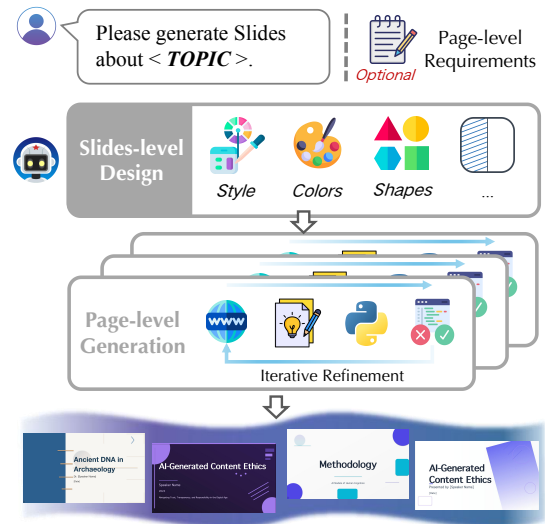


Figure 1: Overview of the DeepSlides workflow. Given a topic and optional user-specified page-level requirements, the system first determines the high-level slide style. Subsequently, each slide is generated through iterative refinement.

Table 1, Google’s NoteBookLM (Google, 2024) utilizes text-to-image models to produce visually impressive slides; some other platforms, such as Kimi (Moonshot-AI, 2025) and Manus (Manus, 2025), employ HTML-based approaches or predefined templates to ensure slide stability. Further research has investigated methods for transforming image-rich documents or research papers into slide presentations (Fu et al., 2022; Zheng et al., 2025; Liu et al., 2026), enabling incremental edits to existing slides under human supervision (Jung et al., 2025; Guo et al., 2024), or generating individual slides directly through executable code (Ge et al., 2025; Tang et al., 2025).

Despite the impressive reasoning and content generation abilities of recent LLMs, the visual quality of automatically generated presentations remains far from satisfactory. Existing systems often produce slide decks that are *visually dull*, *compositionally weak*, and *lack expressive power*,

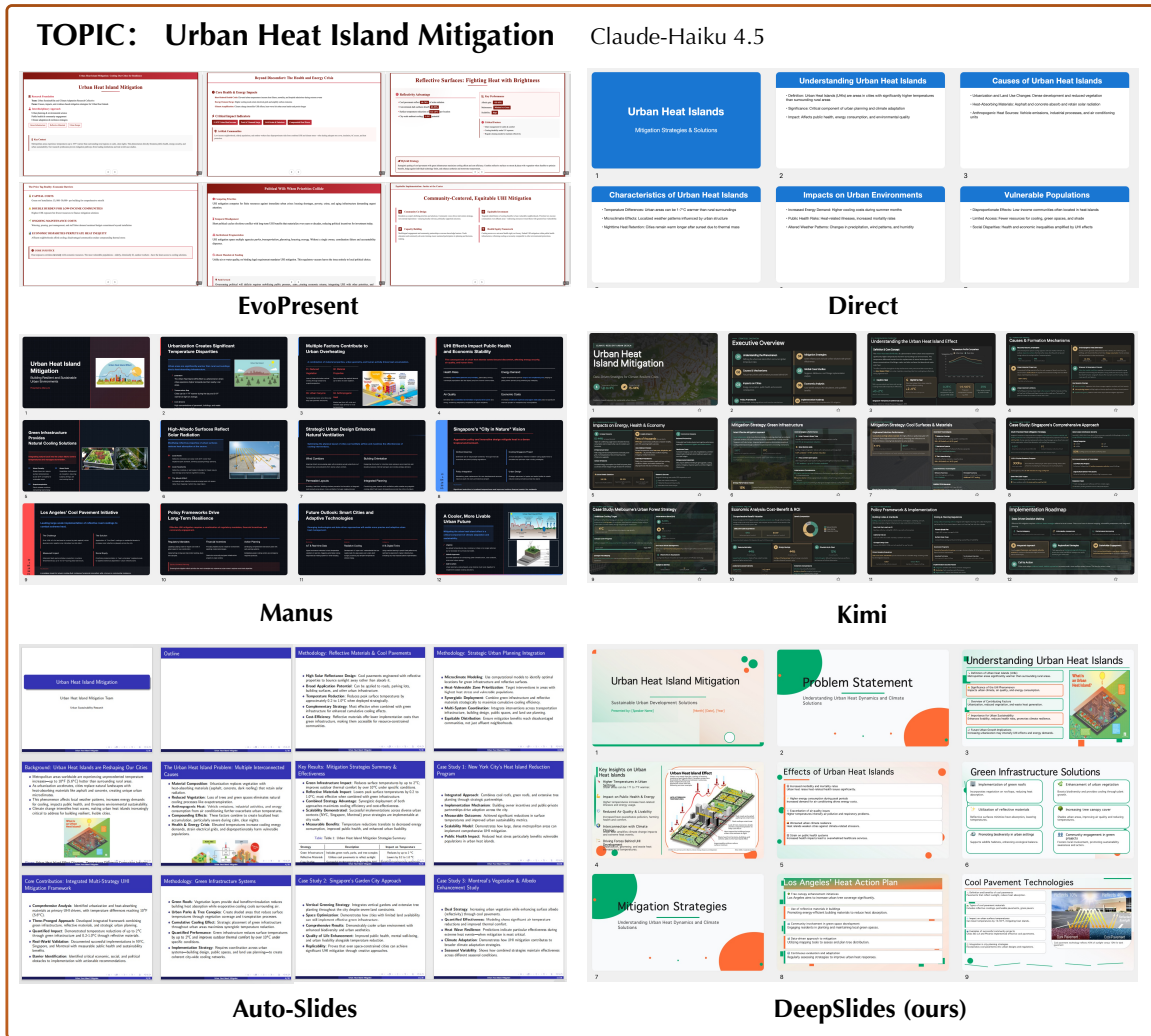


Figure 2: Comparative visualization of slide decks generated for the topic "Urban Heat Island Mitigation", contrasting slides produced by direct model generation, Auto-Slides, EvoPresent, DeepSlides (ours), and two commercial solutions, Manus and Kimi. Additional cases can be found in Appendix F.

failing to capture audience attention in real-world professional settings. A high-quality presentation should function as a sequence of visually engaging scenes, in which layout choices, visual elements, and graphic compositions contribute to the clarity and effectiveness of the presentation. As illustrated in Figure 2, Auto-Slides (Yang et al., 2025) rely heavily on rigid, template-driven layouts; Manus (Manus, 2025) prioritizes decorative aesthetics while neglecting semantic grounding; Kimi (Moonshot-AI, 2025) generates abundant content but suffers from limited visual and structural coherence; and EvoPresent (Liu et al., 2026), although academically rigorous, adopts an overly dense and heavy presentation style. Collectively, these approaches either overemphasize surface-level decoration or focus primarily on content generation, yet still fail to deliver presentations that are visually coherent, creatively composed, and

engaging for real-world use.

To address these limitations, we introduce **DeepSlides**, a hierarchical, design-first framework that conceptualizes slide generation as the synthesis of a sequence of visually coherent scenes, where layout structure, visual elements, and graphic composition collaboratively convey semantic content and direct audience attention. At the core of DeepSlides is a principled decoupling between design reasoning and code realization. The designer module operates within a high-level semantic design space, responsible for orchestrating visual hierarchy, spatial rhythm, typographic flow, and chromatic harmony across scenes without being constrained by low-level PPTX implementation details. In contrast, the coder module is dedicated solely to compiling these finalized design specifications into stable and executable slide code. This modular separation mitigates the distortion of aesthetic intent

	Output Format	Complete Presentation	Template Free
EvoPresent	HTML	✓	×
AutoSlides	LaTeX/PDF	✓	×
SlideGen	PPTX	✓	×
SlideCoder	PPTX	×	×
DOC2PPT	PPTX/PDF	✓	×
PPTAgent	PPTX/HTML	✓	×
PASS	PPTX	✓	×
AutoPresent	PPTX	×	✓
NotebookLM	Image	✓	✓
Kimi	PPTX/Image	✓	✓
Manus	PPTX/PDF	✓	✓
DeepSlides	PPTX/Image	✓	✓

Table 1: Comparison of slide-generation workflows. light blue rows indicate open-source methods, while light orange rows denote closed-source systems.

by implementation artifacts, limits error propagation across stages, and supports the independent optimization of visual expressiveness and execution robustness, thereby enabling DeepSlides to produce coherent, professionally composed slide decks that go beyond prevailing content-centric, template-constrained, or decoration-driven generation paradigms.

Our empirical results demonstrate that DeepSlides consistently outperforms existing baseline methods across key evaluation metrics, with clear advantages in **visual layout, hierarchy construction, and color scheme design**. We conducted a human preference and style evaluation with 20 volunteers, where all methods were presented anonymously in random order to avoid bias. Compared with open-source baselines (Yang et al., 2025; Liu et al., 2026), DeepSlides achieves a win rate of **76.5%**, exceeding the second-best method (9.5%) by a margin of **68.0%**. Even when compared with commercial systems (Manus, 2025; Moonshot-AI, 2025), DeepSlides attains a win rate of **52.0%**, outperforming the strongest commercial baseline (19.5%) by **32.5%**. These results indicate that DeepSlides is consistently perceived as more visually attractive and creative by human evaluators. In the human style evaluation, DeepSlides obtains strong scores across all four dimensions, including clarity, creativity, visualization and aesthetics, further confirming its overall presentation quality.

To further verify the quality of data produced by our workflow, we construct SlideDesign, a multi-disciplinary multimodal dataset under the Field of Science (FOS) taxonomy, which provides paired textual and visual resources for training and evaluation. Based on this dataset, we train a suite of lightweight models on the automatically generated

data. As a result, even with a 0.6B-parameter backbone, the resulting SlideQwens models achieve good performance in slide design and code generation, thereby demonstrating the effectiveness of our workflow.

2 Related Work

2.1 Slides-level generation workflows

While earlier studies had already explored automatic slide generation (Sun et al., 2021), the rapid progress of LLMs and VLMs has dramatically advanced the field, enabling increasingly effective agentic workflows for slide generation (Mondal et al., 2024). PPTAgent (Zheng et al., 2025) decomposes the task into reference analysis and edit actions, while PASS generalizes automated deck creation from Word documents (Aggarwal and Bhand, 2025). SlideGen coordinates collaborative multimodal agents with iterative refinement and layout planning (Liang et al., 2025), and Auto-Slides (Yang et al., 2025) employs a multi-agent paper-to-deck pipeline with interactive refinement and content verification. In parallel, EvoPresent (Liu et al., 2026) focuses on iterative self-improvement via an aesthetic-aware feedback model for multi-round refinement.

Despite strong progress, existing workflows often either (i) rely on template-driven generation, restricting the design space, or (ii) output LaTeX/HTML-style slides that ease rendering but sacrifice native PowerPoint editability and fine-grained interactive, drag-and-drop visual editing.

2.2 Page-level generation, editing and training

A growing body of work targets *single-slide* generation and editing via executable representations to preserve editability. AutoPresent (Ge et al., 2025) introduces SlidesBench and shows that programmatic slide synthesis improves controllability and quality over end-to-end image generation. SlideCoder (Tang et al., 2025) addresses reference-image-to-slide generation with a Slide Complexity Metric and a layout-aware, RAG-enhanced code synthesis framework.

Complementary efforts study *in-place editing* within real presentation software. Talk-to-YourSlides enables instruction-driven PowerPoint editing with a planning–execution architecture and releases TSbench (Jung et al., 2025). PPTC and PPTC-R benchmark multi-turn PowerPoint task completion and robustness to instruction and soft-

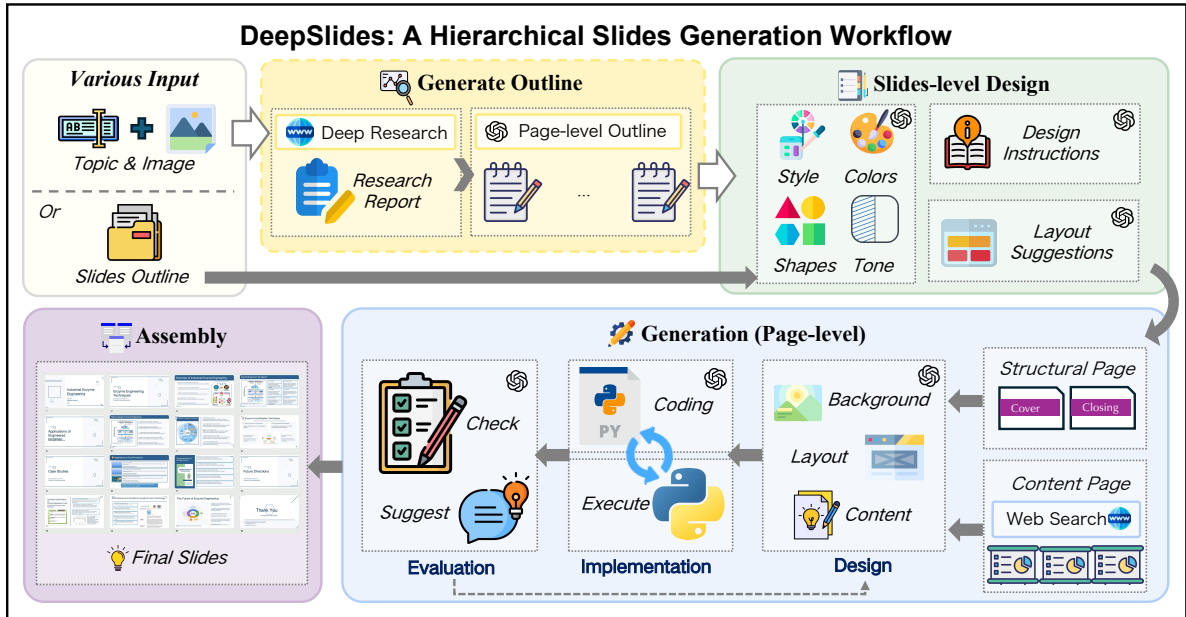


Figure 3: Illustration of the DeepSlides hierarchical slides generation workflow. The framework begins by generating a detailed outline through multimodal deep research or directly from user-provided outlines. Subsequently, slide-level style guidelines are defined, including colors, shapes, and backgrounds. At the page-level, the system performs content expansion via web search and executes detailed design across background, layout, and content layers. Python code is generated to implement each slide, followed by iterative refinement based on evaluation and feedback. Finally, completed slides are assembled into a cohesive presentation.

ware shifts (Guo et al., 2024; Zhang et al., 2024), while PPTArena and PPTBench further evaluate realistic, layout-sensitive edits and generations (Ofengenden et al., 2025; Huang et al., 2025).

From a training perspective, AutoPresent and SlideCoder primarily target code generation. However, design is typically optimized only implicitly through program correctness.

3 DeepSlides

In this section, we first introduce the overall DeepSlides workflow. Given a topic and optional user requirements (e.g., preferred style, color schemes), the system first produces high-level guidance for slides style and layout. It then iteratively refines the content, page design, implementation, and feedback-based revision for each individual slide, and finally aggregates the generated pages into a complete PPTX file. The prompts used in the workflow are provided in Appendix A.

3.1 User Requirement Simulation

In real-world presentation authoring, the process is typically initiated by a domain expert with substantial background knowledge. The expert initiates the process by constructing a high-level outline, which involves determining the necessary sections,

allocating an appropriate number of slides to each section, and specifying the core content of each slide. Only after this process is completed does the expert start authoring individual slides.

To endow the agent with expert-level background knowledge, we introduce a *DeepResearch* module that guides structured knowledge acquisition and integration, enabling the agent to synthesize web search results and retrieved images into multimodal research reports. Based on this comprehensive report, the agent subsequently outlines the slide deck through three steps: (1) determining the appropriate sections; (2) allocating slide budgets to each section; and (3) generating titles and concise bullet points for individual slides. Collectively, these steps yield a structured, multi-level slide outline for the entire presentation.

3.2 Slides-Level Design

At the Slides level, we focus on two aspects: (i) *inter-page coherence* and (ii) *inter-page diversity*. Inter-page coherence requires all pages in a deck to share a consistent style, ensuring they are clearly perceived as part of the same presentation. Inter-page diversity requires that individual pages adopt layouts that are tailored to their specific content, rather than reusing a single template throughout,

which would cause visual fatigue.

Slides features. First, the agent determines the overall style of the slides according to the topic and user requirements, including the global tone, the main color palette, and the font colors for titles and body text. The agent then refines this design by describing the overall visual style in more detail, such as what shapes or patterns should be used as decorative elements, together with their colors, sizes, and approximate positions.

Style Instructions. We categorize pages into two types: *content pages* and *functional pages*. Functional pages, including the cover, end page, and section divider pages, primarily organize the logical structure of the presentation and do not contain substantive content. In contrast, content pages carry the core information and detailed material. Based on the previously generated style description, the agent formulates specific design guidelines tailored for both functional and content pages. It further refines the visual design by providing detailed descriptions of decorative elements, such as recommended shapes or patterns, and specifying their colors, sizes, and approximate placements.

Layout Instructions. Finally, conditioned on the global outline, the agent produces a concise, slide-specific layout specification (e.g., “horizontal blocks”). We perform this step at the slides-level to explicitly promote cross-slide structural diversity while preserving overall stylistic coherence.

3.3 Page-Level Generation

For content pages, we additionally introduce a content expansion stage, whereas this step is unnecessary for functional pages. Specifically, given a slide’s bullet points, the agent first generates targeted search queries and conducts web retrieval to gather supplementary textual and visual materials. The retrieved information is then distilled into candidate text snippets and image descriptions to enrich the slide content. Subsequently, we perform per-slide design, implementation, evaluation and feedback.

Design. We decompose the design of a slide into three layers: *background layer*, *layout layer*, and *content layer*. The agent designs the page by reasoning over these three layers:

- **Background layer.** This layer includes background images, decorative elements, textures,

and other elements that are largely independent of the page’s specific content and layout.

- **Layout layer.** This layer specifies the page layout: the arrangement of structural blocks (e.g., title block, bullet block, image block), their functional roles, spatial positions, and any background panels, borders, or separators associated with these blocks.
- **Content layer.** This layer specifies the concrete textual and visual content to be placed into the structural blocks, including the exact text snippets and images.

Implementation. Given the page-level design, the agent translates the design specification into Python code that programmatically constructs the slide. The resulting slide is then rendered and converted into an image, which will be used for quality assessment. We adopt Python to generate fully editable PPTX files, which better aligns with typical human workflows, rather than HTML- or LaTeX-based outputs (as used in EvoPresent (Liu et al., 2026) and Auto-Slides (Yang et al., 2025)) that are comparatively difficult to edit.

Evaluation and Feedback. The agent evaluates each slide along three dimensions: *completeness*, *compliance*, and *aesthetics*.

- **Completeness** measures whether all input requirements are reflected in the design and whether the textual and visual content matches the outline and retrieved material.
- **Compliance** measures whether the designed text and image blocks overlap improperly, whether the occupied area of each block is appropriate, and whether any content spills outside the page.
- **Aesthetics** measures the visual appeal of the slide, including the harmony of color and typography, the clarity of visual hierarchy, and the overall balance and alignment of elements.

Based on the resulting score and auxiliary feedback signals, the agent decides whether to accept the current design or to iteratively refine the page, feeding targeted revision suggestions provided by the checker into the next design cycle.

4 Multi-Agent Training of DeepSlides

We employed a multi-agent reinforcement learning (MARL) approach to train a set of **QwenSlides** models, consisting of two specialized models: Designer and Coder. We also construct a dataset, **SlideDesign**, for slides generation tasks, as shown in

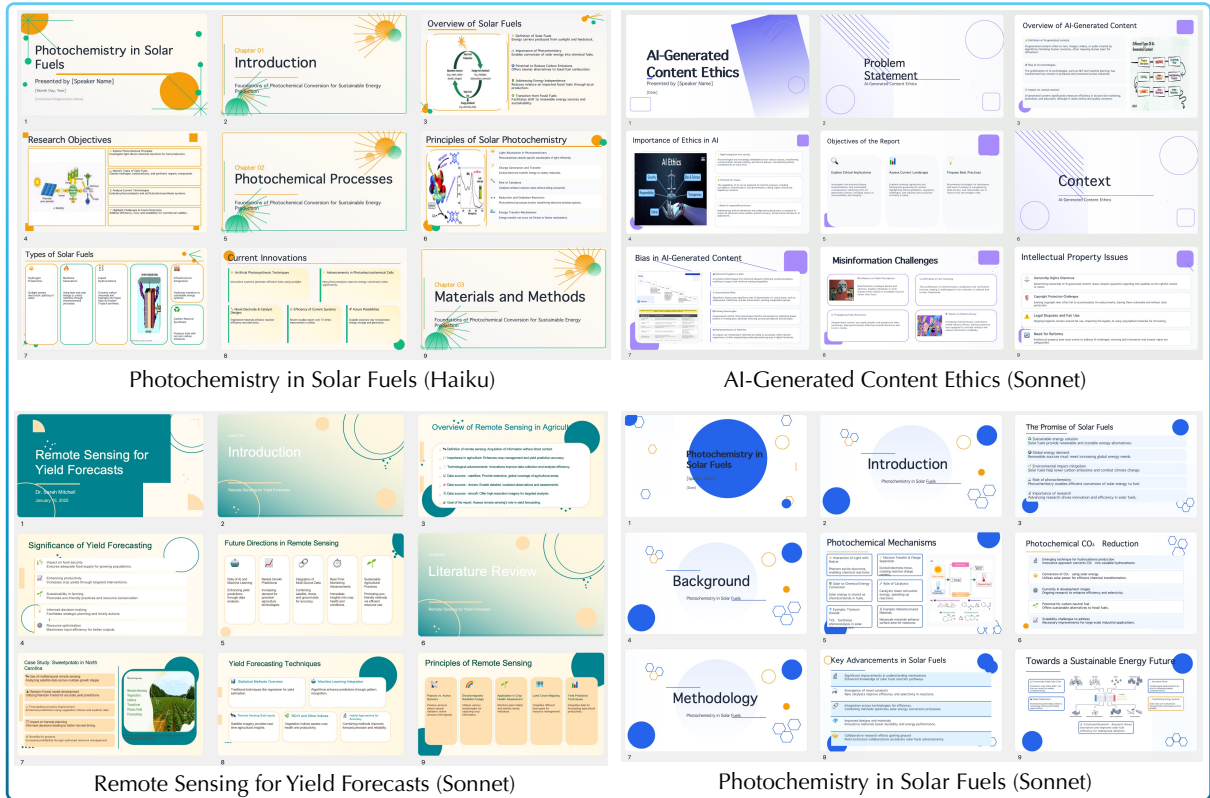


Figure 4: Examples of data in SlideDesign.

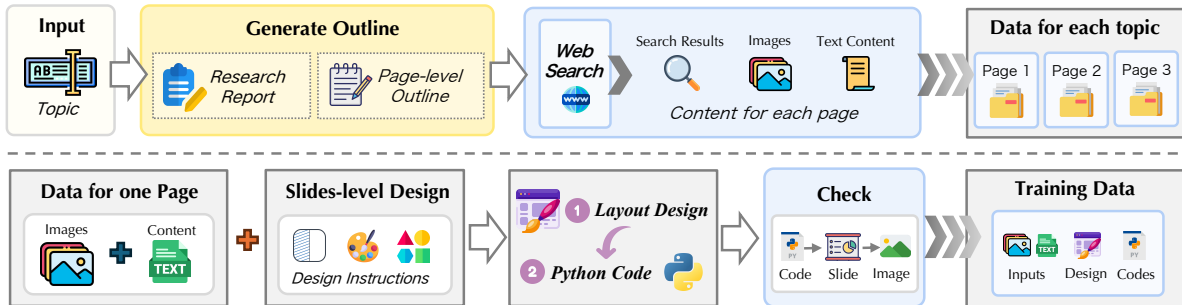


Figure 5: Data generation pipeline

Figure 4.

4.1 Data Collection

As shown in Figure 5, each topic first goes through outline generation to produce structured slide plans (titles and bullet points). We then run outline-guided web searches to collect supporting text and images, and augment each slide with randomly sampled design specifications (e.g., style preferences and color constraints). Based on these inputs, the SFT pipeline generates paired layout designs and executable code, and retains only samples whose code runs successfully, forming the final training set. We construct the **SlideDesign** dataset from 420 topics derived from the Field of Science (FOS) classification as shown in Figure 6. More

details are reported in Appendix B.

4.2 Training process

We first perform SFT as a cold start for both the designer and coder, as prior evidence suggests that SFT-based initialization can rapidly adapt models to new tasks and substantially reduce the computational burden of subsequent RL optimization. Building on these SFT checkpoints, we adopt a multi-agent reinforcement learning framework inspired by ReMA (Wan et al., 2025), where two agents, $Designer_{sft}$ and $Coder_{sft}$, are optimized via interactive rollouts that jointly perform slide layout design and executable code generation. Feedback is provided by a VLM-based reward model that evaluates the outputs along multiple quality



Figure 6: Subject taxonomy and dataset composition. Inner ring: 6 primary areas with counts; outer ring: 42 secondary disciplines.

dimensions. More training details are reported in Appendix C.1.

Specifically, we first initialize two models ($Designer_{sft}$ and $Coder_{sft}$) with SFT. As illustrated in Figure 7, at each timestep t , the designer observes an input prompt s_t and outputs a slide layout design action a_t^D following its policy $\pi_{\theta_D}(a_t^D | s_t)$. Subsequently, the coder receives this design as part of its state s'_t and produces the corresponding Python code action a_t^C according to its policy $\pi_{\theta_C}(a_t^C | s'_t)$. The environment then transitions to the next state, and both agents receive feedback.

We define the coder reward R_C as the product of (i) executability and (ii) implementation completeness. Let $\mathbb{I}_{exec} \in \{0, 1\}$ indicate whether the generated Python code executes successfully, and let $r_{imp} \in [0, 1]$ denote the implementation-completeness score predicted by the reward model (measuring fidelity to the given layout specification). The coder reward is

$$R_C = \mathbb{I}_{exec} \times r_{imp}. \quad (1)$$

The designer reward R_D combines completeness, compliance, and aesthetics. Let $r_{cmp} \in [0, 1]$ and $r_{cpl} \in [0, 1]$ be the completeness and compliance scores produced by the reward model (cf. Section 3). To assess aesthetics, we render the generated slide into an image and obtain a raw aesthetic rating $r_{aes} \in [0, 1]$. To avoid rewarding visually

appealing but incorrectly implemented slides, we gate aesthetics by implementation completeness:

$$r_{sty} = r_{imp} \times r_{aes}. \quad (2)$$

Accordingly, the designer reward is defined as

$$R_D = \alpha r_{cmp} + \beta r_{cpl} + \gamma r_{sty}, \quad (3)$$

where $\alpha, \beta, \gamma \geq 0$ are scalar weights.

This alternating optimization scheme ensures coordinated learning between the designer and coder models.

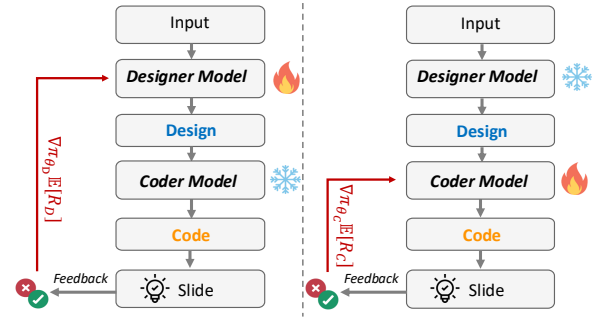


Figure 7: RL train

5 Experiment

Our experiments aim to address the following three questions: (1) Does DeepSlides effectively accomplish slide design tasks while maintaining high-quality content? (2) Does decoupling the design and implementation processes enhance the aesthetic quality of the slides? (3) How is the performance of DeepSlides perceived by human evaluators?

5.1 Baselines

To evaluate DeepSlides, we conducted tests using claude-haiku-4.5 (Anthropic, 2025a) as foundational model. We compared our results against slides directly generated by claude-haiku-4.5 and claude-sonnet-4.5 (Anthropic, 2025b). Additionally, we benchmarked our approach against two prominent slide generation workflows: EvoPresent (Liu et al., 2026) and Auto-Slides (Yang et al., 2025), with claude-haiku-4.5 as the base model. During the user requirement simulation phase, we utilized gpt-4o-mini (OpenAI, 2024) as the foundational model, and the generated research reports served as inputs for the other two workflows.

5.2 VLM as a Judge

We employ two categories of metrics to comprehensively assess slide design quality: objective metrics (Success Rate, Balance) and Vision-Language

Method	Objective Metrics		Content		Visual			Avg.
	Success Rate (%)	Balance (\uparrow)	Clarity	Coherence	Layout	Hierarchy	Color	
Direct _{Claude Haiku}	99	0.53	3.99	3.64	2.64	3.87	3.14	3.46
Direct _{Claude Sonnet}	97	0.67	<u>3.79</u>	3.47	2.67	3.64	3.13	3.34
Auto-Slides _{Claude Haiku}	87	0.81	3.53	3.64	2.49	3.50	3.08	3.25
EvoPresent _{Claude Haiku}	100	0.77	3.53	3.84	<u>3.46</u>	<u>3.95</u>	<u>3.69</u>	<u>3.69</u>
DeepSlides _{Claude Haiku}	95	<u>0.78</u>	<u>3.79</u>	<u>3.81</u>	3.55	4.04	3.71	3.78

Table 2: We compare baseline systems and DeepSlides variants using objective metrics and VLM-as-judge scores for content quality and visual design. Best and second-best results are highlighted in **bold** and underline, respectively.

Model (VLM)-derived metrics. *Success Rate* measures the fraction of generations that can be successfully rendered without errors, while *Balance* quantifies the spatial evenness of the layout distribution. The VLM-derived metrics include content dimensions (Clarity, Coherence) and visual dimensions (Layout, Hierarchy, Color Scheme), assessed using GPT-5.2 (OpenAI, 2025). Detailed definitions are provided in Appendix C.2.

As summarized in Table 2, DeepSlides exhibits superior overall performance relative to baseline approaches, obtaining the highest average VLM-based evaluation score (3.78). While the Success Rate of DeepSlides (95%) slightly trails EvoPresent (100%), DeepSlides notably surpasses EvoPresent and other baselines (Direct, Auto-Slides) in crucial content metrics such as Layout (3.55 vs. 3.46), Hierarchy (4.04 vs. 3.95), and Color Scheme (3.71 vs. 3.69). Auto-Slides attains a higher Balance score than DeepSlides primarily because it generates slides via \LaTeX , where the layout is largely fixed and thus exhibits less spatial variation across outputs, leading to more stable balance measurements. These findings confirm that DeepSlides effectively accomplishes slide design tasks without compromising content quality.

5.3 Human Evaluation

To assess DeepSlides’ alignment with human preferences, we recruited 20 volunteers to rank slides generated on identical topics across frameworks, with the slide order randomly shuffled to prevent presentation bias. We randomly sample 40 topics for human evaluation and use Claude Haiku 4.5 and Claude Sonnet 4.5 as base models (20 topics per model). The evaluators are split into two groups of ten, with one group completing the ranking task and the other completing the scoring task.

We report the frequency of first-place selections and the pairwise win rates of our framework against each competing method. Additionally, evaluators rated DeepSlides’ slides across four dimensions.

Detailed evaluation criteria and questionnaires are presented in Appendix D.

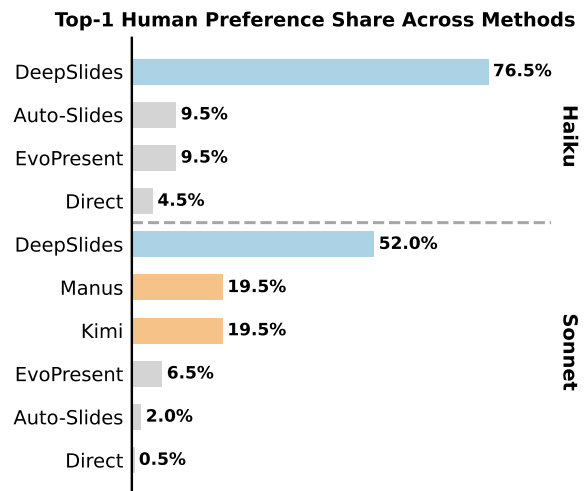


Figure 8: We report the fraction of topics for which each framework is ranked #1 by human evaluators.

DeepSlides is most frequently ranked as the top method in human evaluation. As shown in Figure 8, it achieves the highest Top-1 preference rate under both Claude Haiku 4.5 and Claude Sonnet 4.5 settings, with its Top-1 cumulative preference consistently exceeding all competing approaches. This indicates that human raters most often select DeepSlides as their first-choice system, reflecting stronger overall preference relative to the baselines.

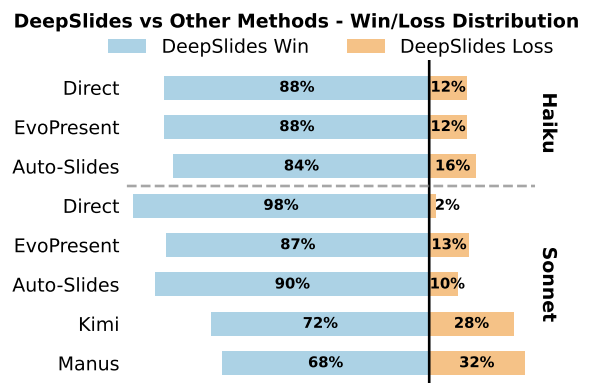


Figure 9: Pairwise win rate.

DeepSlides is consistently preferred by human evaluators over all baselines. As shown in Figure 9, DeepSlides achieves a dominant win rate against all baselines under both Claude Haiku 4.5 and Claude Sonnet 4.5 settings, with only a small fraction of ties or losses. Notably, it remains strongly preferred even when compared to commercial proprietary solutions (e.g., Kimi and Manus), indicating that the proposed approach delivers superior perceived quality beyond open-source and template-based alternatives.

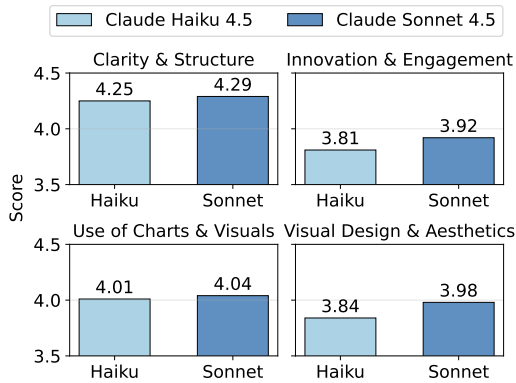


Figure 10: Human evaluation scores.

DeepSlides Achieves High Scores in Human Evaluation. As shown in Figure 10, human evaluators provided detailed ratings across various dimensions. DeepSlides received strong ratings for *Clarity & Structure*, achieving scores of 4.25 for Haiku and 4.29 for Sonnet, indicating excellent content organization. Ratings for *Innovation & Engagement* and *Use of Charts & Visuals* are also notable, demonstrating DeepSlides’ capability to produce engaging and visually informative slides. Furthermore, the high scores in *Visual Design & Aesthetics* underline the effectiveness of the proposed decoupling strategy in enhancing visual appeal.

Model	Success Rate	Balance	Visual
Qwen3-0.6B	0.0	-	-
SlideQwens _{sft}	81.4	0.77	3.23
SlideQwens _{full}	86.0	0.77	3.38

Table 3: Model Training Performance Comparison.

5.4 Training Results

To validate our training strategy, we evaluate the base, SFT, and SFT+RL models on single-slide generation, where each model must jointly produce a layout and executable code given page content and randomly sampled design guidelines. As shown in

Table 3, the base model (*Qwen3-0.6B*) fails to complete the task, whereas SFT enables *SlideQwens_{sft}* to reach an 81.4% success rate (Balance 0.77, Visual 3.23). Adding MARL further improves reliability and visual quality: *SlideQwens_{full}* achieves 86.0% success rate and higher Visual score (3.38) with comparable Balance (0.77), indicating that SFT is necessary for executability while MARL yields additional quality gains.

6 Conclusion

In this paper, we propose **DeepSlides**, a hierarchical workflow for automated slide generation that explicitly decouples slide design from implementation, addressing the limitations of prior work that either rely on fixed templates or bypass the essential design stage. We curate **SlideDesign**, a dataset tailored for slide-page design and implementation, and introduce a multi-agent reinforcement learning training paradigm to train the **SlideQwens** model family. Extensive empirical evaluations, including objective metrics, vision-language model judgments, and human preference studies, demonstrate that DeepSlides achieves superior completeness, compliance, and aesthetic quality compared with competitive baselines. Our ablation studies further confirm that staged supervised fine-tuning followed by reinforcement learning is critical for balancing content quality and visual attractiveness in structured multimodal slide tasks. Collectively, these findings validate that decoupling design and implementation within a structured workflow substantially improves automatic slide generation, producing outputs that better align with human evaluators in both content and visual design.

Acknowledgement

This research was supported by the Fundamental and Interdisciplinary Disciplines Breakthrough Plan of the Ministry of Education of China (JYB2025XDXM116), the National Natural Science Foundation of China (No. U22B2036, No. 62506186), the Technological Innovation Team of Shaanxi Province (No. 2025RS-CXTD009), and the International Cooperation Project of Shaanxi Province (No. 2025GH-YBXM-017). This work was supported by Shanghai Artificial Intelligence Laboratory.

Limitations

Despite improved controllability via decoupling design from implementation and feedback-driven specialization, our method has several limitations. The multi-stage pipeline increases inference latency and operational complexity compared to single-pass generation. While we aim for template-free design, achieving consistent brand-level styling (e.g., strict typography and color constraints) may still require user-provided style references or stronger conditioning, which we do not systematically study. Moreover, our automatic evaluators provide only a proxy for human judgments of aesthetics and communicative effectiveness, leading to failure modes such as over-decoration, suboptimal whitespace, and occasional divergences between design intent and rendered output. Finally, experiments are conducted primarily in academic slide settings with a fixed set of rendering backends; broader generalization to other domains and authoring environments remains an open question.

From a deployment perspective, the system entails additional risks. It may inadvertently violate intellectual property or licensing constraints in text and visual assets without explicit provenance tracking and license-aware filtering. Because the pipeline emits executable slide code and relies on external rendering backends, it also introduces security considerations (e.g., prompt injection and unsafe HTML/SVG/code paths), necessitating sandboxing and rigorous sanitization. Additionally, learned design priors may encode cultural bias or undermine accessibility, and the ability to generate high-quality persuasive decks increases misuse potential, motivating policy enforcement and auditing in practical deployments.

Ethics Statement

Human experts participated voluntarily in providing evaluation ratings, and no personally identifiable or sensitive data were collected; all procedures complied with institutional ethical guidelines.

References

Tushar Aggarwal and Aarohi Bhand. 2025. Pass: Presentation automation for slide generation and speech. *arXiv preprint arXiv:2501.06497*.

Anthropic. 2025a. Introducing claude haiku 4.5. <https://www.anthropic.com/news/claude-haiku-4-5>. Anthropic Claude Haiku 4.5 model release and system card details.

Anthropic. 2025b. Introducing claude sonnet 4.5. <https://www.anthropic.com/news/claude-sonnet-4-5>. Anthropic Claude Sonnet 4.5 model release and capabilities.

Tsu-Jui Fu, William Yang Wang, Daniel McDuff, and Yale Song. 2022. Doc2ppt: Automatic presentation slides generation from scientific documents. In *Proceedings of the AAAI Conference on Artificial Intelligence (AAAI)*, volume 36, pages 634–642.

Jiaxin Ge, Zora Zhiruo Wang, Xuhui Zhou, Yi-Hao Peng, Sanjay Subramanian, Qinyue Tan, Maarten Sap, Alane Suhr, Daniel Fried, Graham Neubig, and Trevor Darrell. 2025. Autopresent: Designing structured visuals from scratch. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*.

Google. 2024. Notebooklm. Accessed: 2025-12-02.

Yiduo Guo, Zekai Zhang, Yaobo Liang, Dongyan Zhao, and Nan Duan. 2024. Pptc benchmark: Evaluating large language models for powerpoint task completion. In *Findings of the Association for Computational Linguistics: ACL 2024*, pages 8682–8701.

Sirui Hong, Mingchen Zhuge, Jonathan Chen, Xiawu Zheng, Yuheng Cheng, Jinlin Wang, Ceyao Zhang, Zili Wang, Steven Ka Shing Yau, Zijuan Lin, and 1 others. 2024. Metagpt: Meta programming for a multi-agent collaborative framework. In *The Twelfth International Conference on Learning Representations*.

Zheng Huang, Xukai Liu, Tianyu Hu, Kai Zhang, and Ye Liu. 2025. Pptbench: Towards holistic evaluation of large language models for powerpoint layout and design understanding. *arXiv preprint arXiv:2512.02624*.

Kyudan Jung, Hojun Cho, Jooyeol Yun, Soyoun Yang, Jaehyeok Jang, and Jaegul Choo. 2025. Talk to your slides: High-efficiency slide editing via language-driven structured data manipulation. *arXiv preprint arXiv:2505.11604*.

Xin Liang, Xiang Zhang, Yiwei Xu, Siqi Sun, and Chenyu You. 2025. Slidegen: Collaborative multimodal agents for scientific slide generation. *arXiv preprint arXiv:2512.04529*.

Chengzhi Liu, Yuzhe Yang, Kaiwen Zhou, Zhen Zhang, Yue Fan, Yanan Xie, Peng Qi, and Xin Eric Wang. 2026. Presenting a paper is an art: Self-improvement aesthetic agents for academic presentations. In *International Conference on Learning Representations*. Poster.

Manus. 2025. Manus. Accessed: 2025-12-02.

Fanqing Meng, Lingxiao Du, Zongkai Liu, Zhixiang Zhou, Quanfeng Lu, Daocheng Fu, Tiancheng Han, Botian Shi, Wenhai Wang, Junjun He, Kaipeng Zhang, Ping Luo, Yu Qiao, Qiaosheng Zhang,

- and Wenqi Shao. 2025. Mm-eureka: Exploring the frontiers of multimodal reasoning with rule-based reinforcement learning. *arXiv preprint arXiv:2503.07365*.
- Ishani Mondal, Shwetha S, Anandhavelu Natarajan, Aparna Garimella, Sambaran Bandyopadhyay, and Jordan Boyd-Graber. 2024. Presentations by the humans and for the humans: Harnessing LLMs for generating persona-aware slides from documents. In *Proceedings of the 18th Conference of the European Chapter of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 2664–2684, St. Julian’s, Malta. Association for Computational Linguistics.
- Moonshot-AI. 2025. Kimi. Accessed: 2025-12-02.
- Michael Ofengenden, Yunze Man, Ziqi Pang, and Yu-Xiong Wang. 2025. Pptarena: A benchmark for agentic powerpoint editing. *arXiv preprint arXiv:2512.03042*.
- OpenAI. 2024. Gpt-4o mini model documentation. <https://platform.openai.com/docs/models/gpt-4o-mini>. OpenAI official model page for GPT-4o-mini.
- OpenAI. 2025. Update to GPT-5 system card: GPT-5.2. https://cdn.openai.com/pdf/3a4153c8-c748-4b71-8e31-aecbde944f8d/oai_5_2_system-card.pdf. Accessed: 2025-12-30.
- Edward Sun, Yufang Hou, Dakuo Wang, Yunfeng Zhang, and Nancy XR Wang. 2021. D2s: Document-to-slide generation via query-based text summarization. In *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 1405–1418.
- Wenxin Tang, Jingyu Xiao, Wenxuan Jiang, Xi Xiao, Yuhang Wang, Xuxin Tang, Qing Li, Yuehe Ma, Junliang Liu, Shisong Tang, and Michael R. Lyu. 2025. Slidecoder: Layout-aware rag-enhanced hierarchical slide generation from design. *Preprint*, arXiv:2506.07964.
- Ziyu Wan, Yunxiang Li, Xiaoyu Wen, Yan Song, Hanjing Wang, Linyi Yang, Mark Schmidt, Jun Wang, Weinan Zhang, Shuyue Hu, and 1 others. 2025. Rema: Learning to meta-think for llms with multi-agent reinforcement learning. *arXiv preprint arXiv:2503.09501*.
- Xin Wang, Zhiyao Cui, Hao Li, Ya Zeng, Chenxu Wang, Ruiqi Song, Yihang Chen, Kun Shao, Qiaosheng Zhang, Jinzhuo Liu, Siyue Ren, Shuyue Hu, and Zhen Wang. 2025. Perpilot: Personalizing vlm-based mobile agents via memory and exploration. *Preprint*, arXiv:2508.18040.
- Yuheng Yang, Wenjia Jiang, Yang Wang, Yiwei Wang, and Chi Zhang. 2025. Auto-slides: An interactive multi-agent system for creating and customizing research presentations. *Preprint*, arXiv:2509.11062.
- Zekai Zhang, Yiduo Guo, Yaobo Liang, Dongyan Zhao, and Nan Duan. 2024. Pptc-r benchmark: Towards evaluating the robustness of large language models for powerpoint task completion. In *Findings of the Association for Computational Linguistics: EMNLP 2024*, pages 12387–12402.
- Hao Zheng, Xinyan Guan, Hao Kong, Wenkai Zhang, Jia Zheng, Weixiang Zhou, Hongyu Lin, Yaojie Lu, Xianpei Han, and Le Sun. 2025. Pptagent: Generating and evaluating presentations beyond text-to-slides. In *Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing*, pages 14402–14418, Suzhou, China. Association for Computational Linguistics.

A Prompts

Prompt for Designing Slide

You are a seasoned slide designer responsible for designing slide layouts. Generate a slide layout description in JSON format based on the following details:

Title: <slide title>

Detailed points: <Main points>

<Slides features>

<Style Instructions>

<Layout Instructions>

Here are some images you may optionally use in the PPT: <Image list>

<Suggestions for improving design>

OUTPUT: <Design format>

Prompt for Coding

Generate Python code that creates slides using the python-pptx library based on the following detailed slide description:

Title: <slide title>

Detailed points: <Main points>

Design Details: <Agent’s Design for the Slide>

<Image list>

<Tool list>

Code requirements:

1. Import the necessary libraries.
2. Create the slides and ensure the widescreen standard aspect ratio: 16:9 (13.33 inches × 7.5 inches).
3. According to the detailed description, add the title, bullet points, and images at specified positions; set fonts and styles; explicitly set the size of each element to prevent overlap/occlusion; ensure text wraps automatically.
4. Only the provided image URLs can be used. Do not reserve any positions for any images that are not provided, and do not use text descriptions to fill the gaps. Or you can also manually create some flowcharts using various graphics, but don’t just leave an empty space or just provide a textual description.
5. All the text should be placed on the top layer.
6. Save the file as: <Save target>

Previous code and errors (if any): <previous code> <error message>

Please provide complete, executable Python code based on this information.

Note: output Python code only, do not output any other text. Code will be save in utf-8 encoding.

B Dataset

We allocate 10 topics per secondary field (420 topics total). Our released dataset only contains (i) source webpage URLs and (ii) concise webpage summaries, and does not redistribute raw webpage content. To mitigate potential privacy and safety risks, we apply automatic screening to remove or mask personally identifying information (e.g., email addresses, phone numbers, and person names when present) from summaries, and filter samples using keyword- and model-based checks for potentially offensive content. We additionally perform manual spot-checking on a subset of the collected entries and discard flagged items. We randomly select 50 topics for supervised fine-tuning (SFT) cold-start data, 200 topics for reinforcement learning (RL) training data and 100 topics for the main experiments. We randomly select 500 pages of content for the ablation study. For generating SFT data, the pipeline employs layout design and Python code generation steps utilizing the gpt-5 model (designer) and *claude-haiku-4.5* model (coder). Outputs undergo a rigorous check against quality thresholds, resulting in a final dataset of 1,286 training pairs.

C Experiment Details

C.1 Training Details

SFT Cold Start. We conduct supervised fine-tuning (SFT) of a Qwen3-0.6B model using full-parameter training. We use AdamW with a learning rate of 1×10^{-4} , cosine scheduling, weight decay of 0.01, and gradient clipping at 1.0. The per-device batch size is 1 with 8-step gradient accumulation (effective batch size 8). We train for 10 epochs with a warm-up ratio of 0.1, using bf16 and gradient checkpointing.

MARL. We train the MARL policy with a GRPO-based pipeline using a Qwen3-0.6B backbone. We set the train/validation batch size to 8 and cap both prompt and response lengths at 8,000

tokens. Rollouts use single-turn generation with $n = 10$ samples per prompt. GRPO uses a learning rate of 1×10^{-6} with mini-batch size 2 and micro-batch size 4 per GPU. We enable a switch-agent mechanism every 10 steps between two specialized agents. Training runs for 10 epochs (500 steps), with evaluation every 100 steps and checkpointing every 50 steps.

All training was conducted on NVIDIA H200 GPUs; supervised fine-tuning used 2 GPU-hours, and RL training used 400 GPU-hours (402 GPU-hours in total).

C.2 Metrics

Balance. We define a visual-weight map $w(x, y)$ on a grayscale image I as a mixture of background-relative deviation and local contrast:

$$w(x, y) = (1 - \lambda) |I(x, y) - b| + \lambda |I(x, y) - (\mathcal{G}_\sigma * I)(x, y)|.$$

where $b = \text{median}(I)$ and \mathcal{G}_σ is a Gaussian blur. Let the total weight be $M = \sum_{x,y} w(x, y)$.

The weight center of mass (c_x, c_y) yields a centeredness score

$$s_{\text{com}} = 1 - \frac{1}{\sqrt{2}} \sqrt{\left(\frac{|c_x - W/2|}{W/2}\right)^2 + \left(\frac{|c_y - H/2|}{H/2}\right)^2}.$$

Define left/right and top/bottom weight sums M_L, M_R, M_T, M_B . Then

$$s_{\text{lr}} = 1 - \frac{|M_L - M_R|}{M}, \quad s_{\text{tb}} = 1 - \frac{|M_T - M_B|}{M}.$$

The final balance score is the average of these three components,

$$\text{Balance} = \text{clip}_{[0,1]} \left(\frac{s_{\text{com}} + s_{\text{lr}} + s_{\text{tb}}}{3} \right).$$

Clarity. Assess how easily the information can be understood at a glance, focusing solely on the content’s clarity and logical structure. This includes the clarity of language, how well the information is explained, and whether the key points are immediately apparent. A high score means the content is direct, concise, and easy to understand without unnecessary complexity.

Coherence. Evaluate how logically the content is organized. A clear narrative flow is critical, with well-structured sections and seamless transitions between ideas. We encourage the use of both text

and visuals to enhance clarity and support the narrative. A high score means the content is organized in a way that guides the viewer through the slide’s message and is complemented effectively by visual elements.

Layout. Analyze the alignment, spacing, and overall balance of the slide’s elements. Creative and thoughtful typography and element placement enhance the visual appeal and engagement. Single-block content layouts are discouraged, as they tend to be overly monotonous and can fail to capture the audience’s interest. Additionally, the slide should incorporate tasteful decorative elements that enhance the overall design without overwhelming the content. A high score reflects a well-structured, balanced slide with sufficient white space, and ideally some creative flair in the layout design.

Hierarchy. This metric assesses whether there is a clear visual distinction between primary and secondary information. It evaluates if the audience can immediately identify titles, main points, and supporting details. This is typically achieved through variations in font size, weight (boldness), color, and placement.

Color Scheme. Evaluate how rich and aesthetically pleasing the color palette is. This includes assessing the use of a variety of colors in a harmonious way, with good contrast that enhances readability and supports the slide’s overall theme. The slide should also feature some decorative elements with colors used thoughtfully and reasonably, contributing to both visual appeal and a balanced design. A high score means the colors are vibrant, well-coordinated, and support the slide’s aesthetic and thematic cohesion. The color of the slides should correspond to the content.

C.3 Other Details

For all LLM calls, we set temperature=0 and max_tokens=8192; all other parameters follow the default settings of the corresponding API/package.

D Human Evaluation

We recruit 20 graduate students from diverse disciplines as human evaluators; participation is unpaid. All human evaluators participated voluntarily with informed consent, and were informed that their annotations would be used solely for research purposes. Our human expert evaluators were aged 20–50, all had at least a university-level education,

demonstrated strong English proficiency, and were based in Asia. We store and report the collected ratings in anonymized and aggregated form. Our questionnaire is shown in Figure 11 and Figure 12.

E License and Availability

All released artifacts (including code, evaluation scripts, and accompanying resources) are distributed under the MIT License. We use all third-party artifacts (datasets, models, and toolchains) in accordance with their stated intended use and access conditions. Our released artifacts are intended for research and evaluation only; we do not redistribute any third-party data beyond what their original terms permit, and any derived resources remain subject to the same restrictions.

F Case Study

As shown in Figures 13-17, we present several comparative examples.

Questionnaire for Slides Ranking

Task: The slide decks below were generated based on an identical topic. Please conduct a comparative assessment of their overall quality (e.g., content clarity, structure, coherence, and visual presentation) and provide a complete ranking from the best-performing deck to the weakest.

GROUP: 1

TOPIC: Dispute Resolution Innovation

1

2

3

4

Please enter your ranking (from best to worst): _____

Figure 11: Detailed information on the human evaluation questionnaire.

Questionnaire for Slides Evaluation

TOPIC: Dispute Resolution Innovation



Task: Please rate the slide deck quality across the following four dimensions.

1. Clarity & Structure

- 5: The content is highly clear and logically rigorous, with a well-organized structure.
- 4: The content is clear and the structure is generally sound, though certain parts may appear slightly cluttered or somewhat redundant.
- 3: The content is reasonably clear, but the structure is insufficiently explicit.
- 2: The content lacks clarity; information is overly dispersed or difficult to comprehend.
- 1: The content is disorganized with poor structure, making it hard to follow.

2. Use of Charts & Visuals

- 5: Charts and visual elements are used appropriately and effectively, substantially aiding explanation and comprehension.
- 4: Charts and visuals are used well overall, but may occasionally be redundant or not sufficiently precise.
- 3: Charts and visuals are limited, or in some places do not provide meaningful support for understanding.
- 2: Charts and visuals are largely absent or potentially misleading.
- 1: No charts or visual elements are provided, or the visuals are entirely irrelevant to the content.

3. Innovation & Engagement

- 5: The design and content are highly engaging and demonstrate strong creativity, effectively capturing audience attention.
- 4: The presentation shows a degree of innovation and generally sustains audience interest.
- 3: The presentation is relatively conventional, with only moderate innovation; audience engagement may not be sustained.
- 2: The presentation lacks novelty and may lead to reduced audience interest.
- 1: The presentation is entirely unoriginal; both content and design are monotonous and unengaging.

4. Visual Design & Aesthetics

- 5: The visual design is outstanding, with excellent color harmony, well-varied layout, and effective integration of text and images.
- 4: The design is aesthetically pleasing overall, though some aspects could be further refined.
- 3: The design quality is average; color choices and/or layout require improvement, and certain elements feel inconsistent.
- 2: The design is cluttered, with incoherent color schemes and a confusing mix of text and images.
- 1: The design is poor, with weak visual quality; color and layout are confusing and impede readability.

Figure 12: Detailed information on the human evaluation questionnaire.

TOPIC: Phytoremediation Field Trials

Claude-Haiku 4.5

| | | | |
|--|--|--|---|
| Phytoremediation
The use of plants to clean up environmental pollution. | Overview of Phytoremediation
Phytoremediation is a cost-effective and environmentally friendly technology for cleaning up contaminated sites. It involves the use of plants to absorb, degrade, or stabilize pollutants in the soil and water. | Objectives of the Report
The primary objective of this report is to provide a comprehensive overview of phytoremediation field trials, including their objectives, methodologies, and results. The secondary objective is to highlight the challenges and opportunities associated with this technology. | Significance of Research
Research in phytoremediation is significant because it offers a sustainable and eco-friendly solution for addressing environmental pollution. It also contributes to the development of new plant species and technologies for remediation. |
| Mechanisms of Phytoremediation
Plants can remove pollutants from the environment through various mechanisms, including phytoextraction, phytodegradation, phytostabilization, and phytovolatilization. | Types of Contaminants
Phytoremediation is effective for a wide range of contaminants, including heavy metals, organic pollutants, and nutrients. The choice of plant species depends on the specific contaminants to be remediated. | Historical Context
The concept of phytoremediation has been known for centuries, but it was not until the late 20th century that it gained widespread attention as a viable remediation technology. | Field Trial Design
Field trials are essential for evaluating the effectiveness of phytoremediation in real-world conditions. They involve the selection of plant species, site preparation, and data collection over time. |
| Case Study: Water Research
A case study focusing on the use of plants to clean up contaminated water bodies, such as rivers and lakes. | Case Study: Wastewater Treatment
A case study focusing on the use of plants to treat wastewater, reducing the need for chemical treatments and sludge production. | Comparative Methods Approach
This approach involves comparing different phytoremediation techniques to determine the most effective one for a given site and contaminant. | Successful Applications
Phytoremediation has been successfully applied in various settings, including industrial sites, agricultural lands, and urban areas. It has shown promising results in cleaning up contaminated soil and water. |
| Challenges Faced
Several challenges can hinder the success of phytoremediation field trials, including slow remediation rates, limited plant species diversity, and potential secondary pollution. | Overcoming Barriers
Researchers are working to overcome these barriers by developing more efficient plant species, improving site preparation techniques, and exploring innovative remediation strategies. | Future Field Trials
Future field trials should focus on optimizing plant species selection, improving site preparation, and exploring innovative remediation strategies to enhance the effectiveness of phytoremediation. | Key Findings from Field Trials
Key findings from field trials include the effectiveness of different plant species, the impact of site preparation, and the potential for secondary pollution. These findings provide valuable insights into the practical application of phytoremediation. |

Direct

| | | | |
|--|---|---|---|
| Phytoremediation Field Trials
Phytoremediation: Research Team
November 15, 2023 | Objectives
• Research and Review
• Case Studies
• Methodology
• Results
• Discussion and Conclusion | Phytoremediation Process: Overview
Phytoremediation is a cost-effective and environmentally friendly technology for cleaning up contaminated sites. It involves the use of plants to absorb, degrade, or stabilize pollutants in the soil and water. | Methodology: Plant Species Selection and Site Characterization
• Site Assessment: Conduct a thorough site assessment to identify the type and extent of contamination.
• Plant Selection: Choose plant species based on their ability to tolerate and remediate the specific contaminants.
• Site Preparation: Prepare the site for planting, including soil testing and site cleanup.
• Planting and Maintenance: Plant the selected species and provide ongoing care and maintenance.
• Data Collection: Monitor the progress of the remediation process and collect data on plant growth and pollutant levels. |
| Background: Environmental Remediation's Evolving Landscape
The landscape of environmental remediation is rapidly changing, driven by increasing regulatory pressure, public awareness, and the need for more sustainable and cost-effective solutions. | The Challenge: Limitations of Conventional Remediation
Conventional remediation methods, such as excavation and incineration, are often expensive, disruptive, and generate secondary waste. They also have a high carbon footprint and can be time-consuming. | Advanced Remediation Techniques
Advanced remediation techniques, such as phytoremediation, offer a more sustainable and cost-effective alternative. They are also more aesthetically pleasing and can be integrated with other remediation methods. | Field Trial 1: Water Quality & Nutrient Treatment
• Study Focus: Control nutrient runoff from agricultural fields to improve water quality in nearby streams.
• Plant Species: Select native plants known for their ability to absorb and store nutrients.
• Site Preparation: Prepare the site for planting, including soil testing and site cleanup.
• Planting and Maintenance: Plant the selected species and provide ongoing care and maintenance.
• Data Collection: Monitor the progress of the remediation process and collect data on plant growth and nutrient levels. |
| Key Contribution: Phytoremediation in Sustainable Remediation
Phytoremediation offers a sustainable and cost-effective solution for addressing environmental pollution. It is a key component of a holistic remediation strategy that prioritizes environmental protection and community well-being. | Phytoremediation Mechanisms: Four Pathways to Success
Phytoremediation operates through four main pathways: phytoextraction, phytodegradation, phytostabilization, and phytovolatilization. Each pathway has its own strengths and limitations, and the choice of pathway depends on the specific contaminants and site conditions. | Field Trial 2: Heavy Metal Extraction from Agricultural Soil
• Study Focus: Extract heavy metals from agricultural soil to improve soil health and crop safety.
• Plant Species: Select plants known for their ability to absorb and store heavy metals.
• Site Preparation: Prepare the site for planting, including soil testing and site cleanup.
• Planting and Maintenance: Plant the selected species and provide ongoing care and maintenance.
• Data Collection: Monitor the progress of the remediation process and collect data on plant growth and heavy metal levels. | Field Trial 3: Contaminant Removal in Industrial Wastewater
• Study Focus: Remove contaminants from industrial wastewater to protect the environment and public health.
• Plant Species: Select plants known for their ability to absorb and degrade contaminants.
• Site Preparation: Prepare the site for planting, including soil testing and site cleanup.
• Planting and Maintenance: Plant the selected species and provide ongoing care and maintenance.
• Data Collection: Monitor the progress of the remediation process and collect data on plant growth and contaminant levels. |

Auto-Slides

| | | |
|--|---|---|
| PHYTOREMEDIATION FIELD TRIALS: BRIDGING THEORY'S GAPS TO REAL-World ENVIRONMENTS
• Introduction: Phytoremediation is a cost-effective and environmentally friendly technology for cleaning up contaminated sites. It involves the use of plants to absorb, degrade, or stabilize pollutants in the soil and water. | THE POLLUTION CRISIS: WHY PLANTS MATTER
• The Problem: Pollution is a global crisis that threatens human health and the environment. It is caused by a wide range of sources, including industrial activities, agriculture, and urban development. | CONVENTIONAL REMEDIATION: FROM SLOWLY TO MORE EFFICIENT
• Traditional Methods: Conventional remediation methods, such as excavation and incineration, are often expensive, disruptive, and generate secondary waste. They also have a high carbon footprint and can be time-consuming. |
| WATER QUALITY: AESTHETIC PHYTOREMEDIATION
• Study Focus: Improve water quality in rivers and lakes by using plants to absorb and degrade pollutants. | FROM CONCEPTS TO FIELD: METHODS MATTERS: THEORETICAL FIELD TRIALS
• Methodology: Field trials are essential for evaluating the effectiveness of phytoremediation in real-world conditions. They involve the selection of plant species, site preparation, and data collection over time. | ADVANTAGES OF PHYTOREMEDIATION: ECONOMIC, ENVIRONMENTAL, AND SOCIAL BENEFITS
• Cost-Effectiveness: Phytoremediation is a cost-effective alternative to conventional remediation methods. It is also more aesthetically pleasing and can be integrated with other remediation methods. |
| ECOSYSTEM CO-BENEFITS: HEALING BEYOND CONTAMINATION CONTROL
• Ecosystem Health: Phytoremediation can also provide co-benefits for the environment, such as improving soil health, increasing biodiversity, and providing habitat for wildlife. | THE FUTURE PERSPECTIVE: OPTIMIZING PHYTOREMEDIATION'S CAPABILITIES
• Research and Development: Researchers are working to optimize the capabilities of phytoremediation by developing more efficient plant species, improving site preparation techniques, and exploring innovative remediation strategies. | IMPLEMENTATION FRAMEWORK: FROM CONCEPT TO FIELD DEPLOYMENT
• Site Assessment: Conduct a thorough site assessment to identify the type and extent of contamination.
• Plant Selection: Choose plant species based on their ability to tolerate and remediate the specific contaminants.
• Site Preparation: Prepare the site for planting, including soil testing and site cleanup.
• Planting and Maintenance: Plant the selected species and provide ongoing care and maintenance.
• Data Collection: Monitor the progress of the remediation process and collect data on plant growth and pollutant levels. |

EvoPresent

| | | | |
|--|---|---|---|
| Phytoremediation Field Trials
Introduction | Methodology
Methodology | What is Phytoremediation?
What is Phytoremediation? | Importance of Research
Importance of Research |
| Results
Results | Mechanisms of Phytoremediation
Mechanisms of Phytoremediation | Field Trials Case Studies
Field Trials Case Studies | Summary of Findings
Summary of Findings |
| Discussion
Discussion | Advantages of Phytoremediation
Advantages of Phytoremediation | Limitations and Challenges
Limitations and Challenges | Thank You
Thank You |
| Discussion
Discussion | Future Perspectives
Future Perspectives | Conclusion
Conclusion | |

DeepSlides (ours)

| | | | |
|---------------------------------------|---|---|---|
| Introduction
Introduction | Background
Background | Objectives
Objectives | Methodology
Methodology |
| Results
Results | Discussion
Discussion | Conclusion
Conclusion | References
References |
| Appendix
Appendix | Field Trial Design
Field Trial Design | Plant Species Selection
Plant Species Selection | Site Characterization
Site Characterization |
| Water Quality
Water Quality | Heavy Metal Extraction
Heavy Metal Extraction | Industrial Wastewater
Industrial Wastewater | Future Perspectives
Future Perspectives |

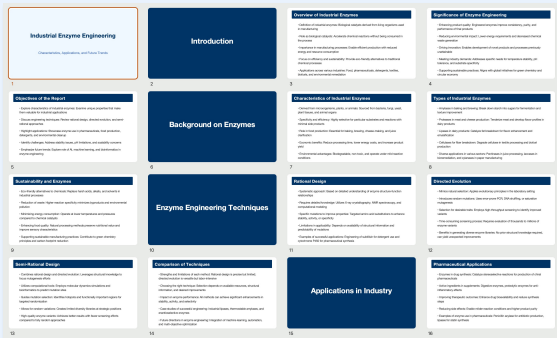
Manus

| | | | |
|---|---|---|---------------------------------------|
| Phytoremediation Field Trials
Introduction | Objectives
Objectives | Methodology
Methodology | Results
Results |
| Discussion
Discussion | Conclusion
Conclusion | References
References | Appendix
Appendix |
| Field Trial Design
Field Trial Design | Plant Species Selection
Plant Species Selection | Site Characterization
Site Characterization | Water Quality
Water Quality |
| Heavy Metal Extraction
Heavy Metal Extraction | Industrial Wastewater
Industrial Wastewater | Future Perspectives
Future Perspectives | |

Kimi

Figure 13: Case Study.

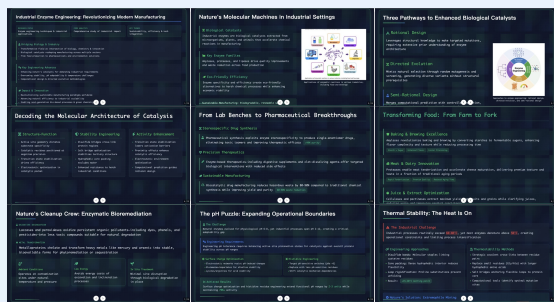
TOPIC: Industrial Enzyme Engineering Claude-Sonnet 4.5



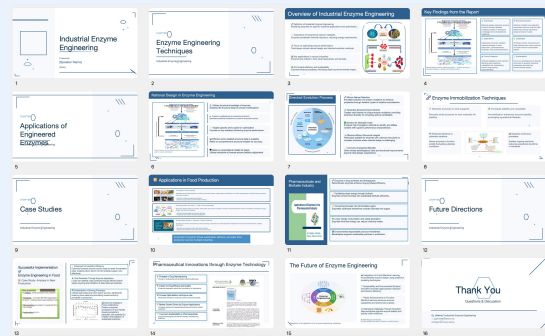
Direct



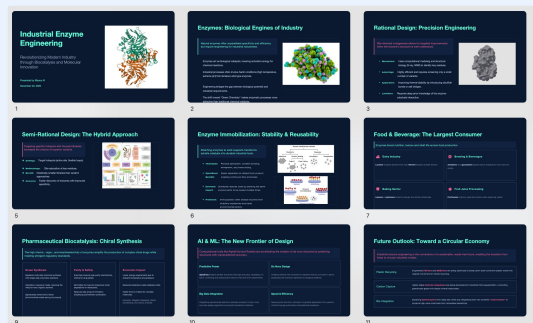
Auto-Slides



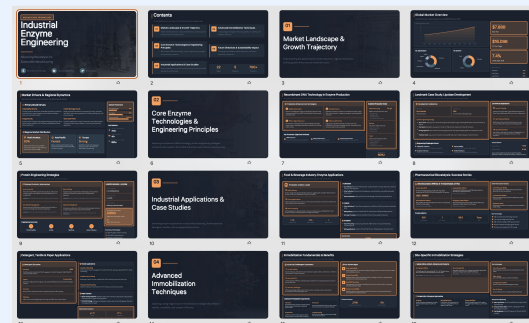
EvoPresent



DeepSlides (ours)



Manus



Kimi

Figure 14: Case Study.

TOPIC: Spacecraft Systems Engineering Claude-Haiku 4.5

| | | | |
|--|---|--|---|
| 1
Spacecraft Systems Engineering
Introduction and Scope | 2
Overview of Spacecraft Systems Engineering | 3
Objectives of the Report | 4
Fundamental Principles |
| 5
Risk of Systems Engineer | 6
Lifecycle of Spacecraft Systems Engineering | 7
Design and Development Process | 8
Verification and Validation (V&V) |
| 9
Testing Methods | 10
Subsystem Integration | 11
Phases of Systems Engineering | 12
Requirements Driven Design |
| 13
Trade-Off Analysis | 14
Risk Management Strategies | 15
Current Challenges | 16
Case Studies of Challenges |

Direct

| | | | |
|---|---|--|--|
| 1
Spacecraft Systems Engineering
Introduction | 2
Guides | 3
The Systems Engineering Lifecycle: From Concept to Operations | 4
The V&V Maturity Framework: Verification & Validation Challenges |
| 5
Requirements Engineering: Demands, Uncertainties, Evolving Challenges | 6
The Challenge: Managing Extreme Complexity in Spacecraft Design | 7
Requirements Definition Phase: Establishing the Foundation | 8
Design Trade-Off Analysis: Balancing Competing Objectives |
| 9
Core Contribution: Comprehensive Systems Engineering Framework | 10
Foundational Principles of Spacecraft Systems Engineering | 11
Advanced Integration: Achieving Seamless System-Level Performance | 12
Verification Activities: Ensuring Design Compliance |

Auto-Slides

| | | |
|--|---|---|
| 1
Advanced System Engineering: Interdisciplinary, Iterative & Systemic | 2
The Systems Engineering Maturity Model: Beyond Compliance | 3
From Mission Concept to Reality: The Design Journey |
| 4
Validation: Proving It Works in the Real World | 5
SYSTEMS ENGINEERING: A HOLISTIC APPROACH | 6
The Mission System: Complex, Interconnected, Evolving |
| 7
Complexity, Uncertainty & Risk: Engineering the Unpredictable | 8
Requirements Engineering: From Mission to Architecture | 9
Advanced Integration: Achieving Seamless System-Level Performance |

EvoPresent

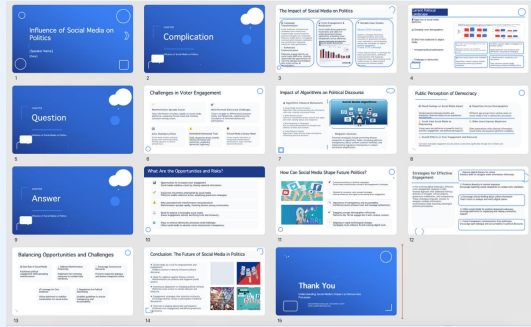
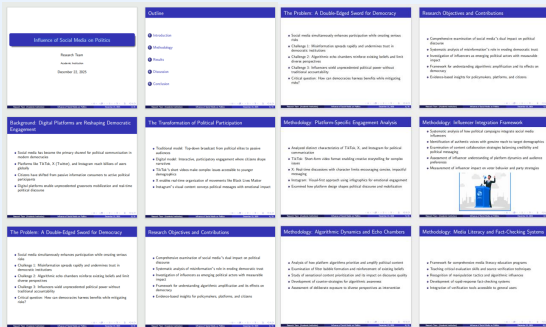
| | | | |
|--|--|---|--|
| 1
Spacecraft Systems Engineering
Introduction | 2
Introduction | 3
Overview of Spacecraft Systems Engineering | 4
Systems of Spacecraft Engineering |
| 5
Spacecraft Systems Overview | 6
Key Components of Spacecraft Systems | 7
Integration of Subsystems | 8
Engineering Processes |
| 9
Attitude Control Systems Innovations | 10
Key Technologies | 11
Challenges and Future Directions | 12
Advanced Propulsion Technologies |
| 13
Challenges and Future Directions | 14
Challenges and Future Directions | 15
Key Challenges in Spacecraft Systems Engineering | 16
Future Directions in Spacecraft Engineering |

DeepSlides (ours)

Figure 15: Case Study.

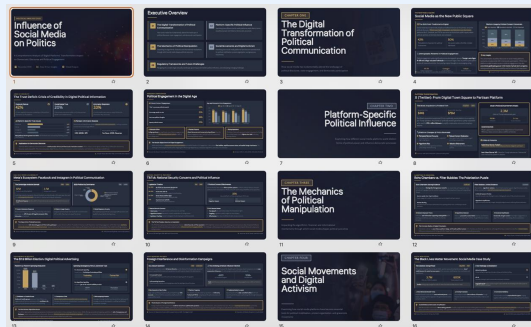
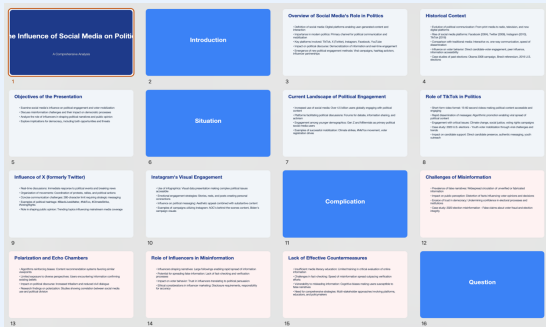
TOPIC: Influence of Social Media on Politics

Claude-Sonnet 4.5



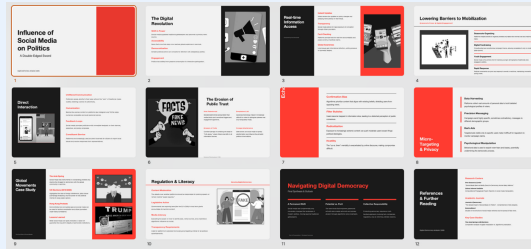
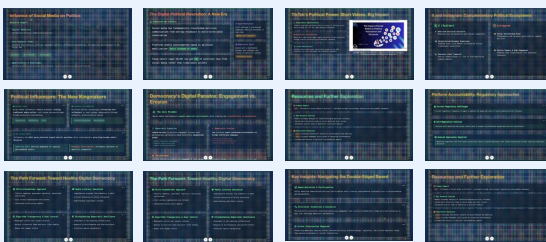
Auto-Slides

DeepSlides (ours)



Direct

Kimi



EvoPresent

Manus

Figure 16: Case Study.

TOPIC: Immunology of Vaccines

Claude-Sonnet 4.5

This grid shows 12 slide thumbnails from a presentation on immunology. The slides are arranged in a 3x4 grid. The first slide is the title slide 'Immunology of Vaccines'. The second slide is 'Introduction to Immunology'. The third slide is 'Importance of Vaccination'. The fourth slide is 'Overview of Immune System'. The fifth slide is 'History of Vaccination'. The sixth slide is 'Types of Vaccines'. The seventh slide is 'Live Attenuated Vaccines'. The eighth slide is 'Inactivated Vaccines'. The ninth slide is 'Subunit and mRNA Vaccines'. The tenth slide is 'Mechanism of Action'. The eleventh slide is 'Introduction to Immune Response'. The twelfth slide is 'Cells and Antibody Production'. The thirteenth slide is 'T Cells in Immune Response'. The fourteenth slide is 'Establishment of Immunological Memory'. The fifteenth slide is 'Immune Response to Vaccination'.

This grid shows another set of 12 slide thumbnails. The first slide is 'History of Vaccines'. The second slide is 'Vaccine Classification: Live Attenuated Vaccines'. The third slide is 'Live Attenuated Vaccines: Potential Pathogens'. The fourth slide is 'Live Attenuated Vaccines: Potential Pathogens'. The fifth slide is 'Subunit Vaccines: Protein Pathogens'. The sixth slide is 'Subunit Vaccines: Protein Pathogens'. The seventh slide is 'Subunit Vaccines: Protein Pathogens'. The eighth slide is 'Subunit Vaccines: Protein Pathogens'. The ninth slide is 'Subunit Vaccines: Protein Pathogens'. The tenth slide is 'Subunit Vaccines: Protein Pathogens'. The eleventh slide is 'Subunit Vaccines: Protein Pathogens'. The twelfth slide is 'Subunit Vaccines: Protein Pathogens'.

Direct

Auto-Slides

This grid shows 12 slide thumbnails. The first slide is 'Immunology of Vaccines'. The second slide is 'The Cellular Orchestra: How Your Body Learns to Fight'. The third slide is 'Immunological Memory: Your Body's Permanent Defense Archive'. The fourth slide is 'Live Attenuated Vaccines: Training Pathogens for Protection'. The fifth slide is 'Inactivated Vaccines: Maximum Safety, Scheduled Protection'. The sixth slide is 'Subunit Vaccines: Precision Targeting with Minimal Payload'. The seventh slide is 'mRNA Vaccines: Genetic Instructions for DIF Immunity'. The eighth slide is 'Vaccine Hesitancy: When Misinformation Becomes a Pathogen'. The ninth slide is 'Next-Gen Vaccines: New Technologies for Emerging Threats'.

This grid shows 12 slide thumbnails. The first slide is 'Immunology of Vaccines'. The second slide is 'Complication'. The third slide is 'Understanding Immunology'. The fourth slide is 'Vaccine Mechanisms'. The fifth slide is 'Question'. The sixth slide is 'Challenges with Immunization'. The seventh slide is 'Impact of Reduced Vaccination'. The eighth slide is 'Answer'. The ninth slide is 'What Types of Vaccines Exist?'. The tenth slide is 'How Do Vaccines Affect Public Health?'. The eleventh slide is 'Conclusion: The Role of Vaccines'. The twelfth slide is 'Thank You'.

EvoPresent

DeepSlides (ours)

This grid shows 12 slide thumbnails. The first slide is 'Immunology of Vaccines'. The second slide is 'Introduction to Vaccination'. The third slide is 'The Anatomy of a Vaccine'. The fourth slide is 'Classification of Vaccines'. The fifth slide is 'The Immune System'. The sixth slide is 'Adaptive Immune Response'. The seventh slide is 'Adaptive Immune Response: T Cells'. The eighth slide is 'Immunological Memory'. The ninth slide is 'Next-Gen Pathways: mRNA & DNA Vaccines'. The tenth slide is 'Breakthroughs in 2024-2025'. The eleventh slide is 'Challenges and Future Directions'. The twelfth slide is 'Conclusion'.

This grid shows 12 slide thumbnails. The first slide is 'Immunology of Vaccines'. The second slide is 'Contents'. The third slide is 'Introduction to Immunology'. The fourth slide is 'Immune Response'. The fifth slide is 'The Immune System'. The sixth slide is 'Adaptive Immune Response'. The seventh slide is 'Immunological Memory'. The eighth slide is 'Next-Gen Pathways: mRNA & DNA Vaccines'. The ninth slide is 'Breakthroughs in 2024-2025'. The tenth slide is 'Challenges and Future Directions'. The eleventh slide is 'Conclusion'. The twelfth slide is 'References'.

Manus

Kimi

Figure 17: Case Study.