



AI-Powered Slides Generation and Chat for Financial Reports

Yongwon Cho

Liping Yin

Rongxiang Zhi

Yaohai Zhou



Overview

- A software application for finance professionals
- Transform 10-K financial report into PowerPoint presentation
- Automation and AI (GPT 3.5)
- Visually engaging components (charts and summaries)
- Chat: efficiently locate and extract information



Significance

- Make users more productive preparing presentations
- Explain document clearly and concisely
- Help audience understand and stay engaged
- Chat feature: extract info about personalized topics



Assumption

- Support one company (Apple) instead of multiple companies
- Internal tool instead of commercial product
- Ensure delivery of MVP (Minimum Viable Product)
- No objective way to control or evaluate AI content

Demo



Methodology

“Easy” Parts

- First 10 Slides
- Topics Manually Chosen
- Chart Types Manually Chosen
- Predictable

“Difficult” Parts

- Last 2 Slides
- Topics AI Chosen
- No Chart - Only Summaries
- Experimental



“Easy” - Section Extraction

- Use SEC Extractor API to extract text from 10-K
- No longer used when AI chooses topics
- More flexible method required



“Easy” - Chart Generation

- Pandas + Matplotlib
- Matplotlib: graphics too basic and inflexible
- Switch to native charts in python-pptx
- More attractive
- Can be edited afterwards



“Easy” - Text Summarization

- Use AI to summarize certain sections
- Prompt: format response in bullet points
- Output: Four bullet points
- Format adjustments for later processing



“Easy” - PowerPoint Assembly

- Python-pptx: make and change PowerPoint in Python
- Easy to generate native PowerPoint charts
- Customization: name, title and templates



“Difficult” - Chat

- Ask anything about the document
- Not as simple as sending the whole document!
- Only send information related to question
- Use llama_index data framework
- Indexing + querying



Chat - Indexing

- Create database index
- Quicker to find right information
- Break down document into chunks
- Transform each chunk to embedding
- Capture semantic meaning
- Result: a database of all embeddings



Chat - Querying

- Transform question to embedding
- Same transformation method applied
- Pick closest embeddings from database >> “context info”
- Context info + original question >> AI
- Support response “streaming”



“Difficult” - AI Topic Selection

- Experimental feature based on chat
- Use AI to suggest 2 topic titles >> last 2 slides in PowerPoint
- Related to company and frequently appear
- Demonstrate clear difference from previous year
- Use suggested topics to pull out details from database



Performance

- Test **100+** times
- Different years and templates
- Automatic retry for all AI parts
- **100%** no crash or error



“Easy” - Chart Correctness

- **3** people worked on three different methods
- Manual calculation + matplotlib + python-pptx
- Avoid personal mistakes
- **3** results match
- Implementation **100%** accurate



“Easy” - Summarization Quality

- Text summaries we **manually** pick topics for
- Test **100+** times
- Factually correct and no unrelated info
- Success rate **100%**
- Work well because topics are predetermined
- Limited room for variation or mistakes



“Difficult” - AI Topic Selection Quality

- Test **40+** times
- **80%** success: fitting and important
- Failures include:
 - Use of placeholders (e.g. X%, \$X)
 - Content too general, unrelated to document
 - Abnormal Topic Titles

Topic 1: Expansion into Wearables Market

- ▶ In 2022, the company faced intense competition in the wearables market, with competitors offering lower-priced products and larger customer bases. This resulted in a decrease in product margins and potential revenue losses
- ▶ The company's ability to maintain its functional and design advantages in wearables is crucial for its financial success. However, competitors imitating the company's product features and applications pose a significant challenge, potentially impacting revenue growth
- ▶ While the company is focused on expanding its market opportunities in wearables, it must be cautious about competitors collaborating to offer integrated solutions that are more competitive than its current offerings. This could potentially lead to a loss of market share and revenue

Figure: Good Topic Selection

Topic 2: Risk Factors and Legal Proceedings

- ▶ The Company's revenue is at risk due to the potential impact of the COVID-19 pandemic. The forward-looking statements in the 10-K document highlight the uncertainty surrounding future business operations and results
- ▶ The Company's financial condition and stock price can be materially and adversely affected by various factors, including those described in the risk factors section of the 10-K document. Past financial performance may not be a reliable indicator of future performance
- ▶ Investors should not rely on historical trends to anticipate future results or trends. The risk factors section emphasizes that other factors, in addition to those discussed, can affect the Company's results of operations and financial condition

Figure: Unrelated or Too General



AI Topic Selection Quality

General/unrelated content - WHY?

- Document divided into chunks
- AI only sees specific parts
- No understanding of whole picture
- Bad topic selection: not often appear in document
- Bad content - relevant info is not there!



“Difficult” - Chat Quality

- **70+** test questions
- Cover different question types:
 - Numbers or text
 - Different scales
 - Creative



Chat Quality - Test Questions

Question Type	Sample Question
Numerical	What is the value of iPhone sales in 2022?
Text, Small Scale	Why does sales in Greater China increase this year?
Text, Large Scale	How has COVID-19 affected the company?
Creative	Assume I am the CFO assistant. Write a 200-word internal email for me to the board of the company to inform them that the latest 10-K is released and what major updates they should expect. Use detailed examples.



Chat Quality - Test Results

- Match the question
- Accurate info
- Relevant info from right parts of document
- **90%** success
- **Main reason for failure:** unable to locate (especially numbers)
- Small details lost in embedding transformation



Finding

- **Limited Applicability and Usefulness:**
 - Only support one company
 - Need smarter way to extract information
- **Challenges in AI Decision Making**
 - Challenging for AI to determine good topics
 - Problematic for AI to locate and output structured info
- **Dilemma of AI Creativity**



Finding

“For professional tasks like finance, it's probably a good idea to mix AI skills with human expertise. At least for now, that seems to work best.”

Our chat feature is a good example!



Recommendation

- Support for 10-K filings from diverse companies
- Improving quality of AI topic selection
- Using AI to determine useful chart types
- Ethical considerations in AI-generated content