DISSERTATION PROJECT HANDOFF DOCUMENT

Compliance Aware, Agentic Pitch Book Generation For Bankers

IOT635 Final Report - Queen Mary University of London

Student: Andrei Rizea (220300153)

Supervisor: Manesha Peiris

This document contains the complete context and generated files for migration to another coding session.

1. PROJECT CONTEXT

1.1 Project Overview

MSc dissertation project developing an AI-powered, compliance-aware pitch book generation system for investment banking. The system combines LLM-based content planning, agentic data retrieval, and programmatic PowerPoint generation to automate the creation of client-facing presentation materials while preserving institutional template styling.

1.2 Key Technical Decisions Made

- **Primary Focus:** Template-aware PowerPoint generation (novel contribution)
- **Agent Framework:** Use existing frameworks (AgentSDK/function-calling) rather than custom orchestration
- **Template Analysis:** Programmatic extraction via python-pptx (VLMs deferred to future work)
- **Data Sources:** Public APIs only (Yahoo Finance, SEC EDGAR)
- **RAG Search:** Designated as stretch goal, not core requirement

1.3 Architecture Overview

```
USER INPUT (deal context, reference template .pptx, pitch book type)
    |
    v
SIMPLE AGENTIC DATA LAYER (AgentSDK or function-calling)
    Tools: Yahoo Finance API, SEC EDGAR API, web search
    |
    v
NOVEL CONTRIBUTION: PPT GENERATOR
    1. Template Analyzer (python-pptx): Extract layouts, styles, placeholders
    2. Content Planner (LLM): Decide slide structure, content per slide
    3. Slide Builder (custom library): Assemble slides using template + content
    |
    v
OUTPUT: .pptx file matching corporate template styling
```

2. WORK COMPLETED IN THIS SESSION

2.1 Introduction, Scope and Context Section (~1,100 words)

Created a completely fresh Chapter 1 for the dissertation, entirely reframed from the interim report to avoid self-plagiarism. The section uses knowledge management theory (Markus 2001, Dalkir 2005) as its theoretical foundation.

2.2 Key Differences from Interim Report

- **Interim:** 'Reduced productivity' → **New:** 'Cognitive misallocation'
- **Interim:** 'Scattered precedence material' → **New:** 'Institutional memory erosion' + tacit/explicit knowledge theory
- **Interim:** 'Tough learning-curves' → **New:** 'Knowledge transfer friction'
- **Interim:** Basic problem list → **New:** Unified Markus (2001) knowledge reuse framework
- **Interim:** Brief PPTAgent mention → **New:** Detailed text-to-slide vs edit-based paradigm evolution

2.3 New Academic Sources Added

- Radhakrishna et al. (2024) - 'Future of Knowledge Management in Investment Banking' - SAGE journal
- Dalkir (2005) - Knowledge Management in Theory and Practice - textbook on tacit/explicit knowledge
- Zheng et al. (2025) - PPTAgent paper with deeper analysis of edit-based paradigm
- McKinsey - 20-25% productivity gains from robust KM systems
- Various BIS, FSB regulatory sources on AI compliance in financial services

2.4 Section Structure

1. The Cognitive Economics of Investment Banking
2. Problem Analysis: Knowledge Reuse Barriers in Document Production
3. Gap Analysis: The Evolution of Presentation Automation
4. Aims and Objectives
5. Scope, Constraints and Success Criteria

3. NEXT STEPS

3.1 Remaining Dissertation Sections

- **Chapter 2:** Literature Review (expand on RAG, agentic AI, PPTAgent research)
- **Chapter 3:** Methodology and Project Plan (Agile SDLC, milestones, risk assessment)
- **Chapter 4:** System Design (architecture diagrams, data models, API specifications)
- **Chapter 5:** Implementation (code samples, technical challenges, solutions)
- **Chapter 6:** Testing and Evaluation (test cases, productivity metrics)
- **Chapter 7:** Results and Discussion
- **Chapter 8:** Conclusion and Future Work

3.2 Git Commands

```
cd path/to/dissertation
git add diss.tex references.bib
git commit -m "Add fresh Introduction, Scope and Context section (~1100 words)"
git push origin main
```

3.3 Compliance Notes

- No Goldman Sachs references in report body (only confidentiality footer if needed)
- Abstract to 'investment banking', 'corporate banking', 'financial services'
- Cannot reference internal data or named stakeholders

4. GENERATED FILE CONTENTS

The following pages contain the complete contents of the two generated files: diss.tex and references.bib. These can be copied directly into your repository.

4.1 FILE: diss.tex

```
\documentclass[12pt,a4paper]{report}

% ============================================
% PACKAGES
% ============================================

% Encoding and fonts
\usepackage[utf8]{inputenc}
\usepackage[T1]{fontenc}
\usepackage{lmodern}

% Page layout
\usepackage[margin=2.5cm]{geometry}
\usepackage{setspace}
\onehalfspacing

% Graphics and figures
\usepackage{graphicx}
\usepackage{float}
\usepackage{subcaption}
\graphicspath{{figures/}}

% Tables
\usepackage{booktabs}
\usepackage{longtable}
\usepackage{tabularx}
\usepackage{multirow}

% Math
\usepackage{amsmath}
\usepackage{amssymb}

% Code listings
\usepackage{listings}
\usepackage{xcolor}

\definecolor{codegreen}{rgb}{0,0.6,0}
\definecolor{codegray}{rgb}{0.5,0.5,0.5}
\definecolor{codepurple}{rgb}{0.58,0,0.82}
\definecolor{backcolour}{rgb}{0.95,0.95,0.92}

\lstdefinestyle{mystyle}{
    backgroundcolor=\color{backcolour},
    commentstyle=\color{codegreen},
    keywordstyle=\color{magenta},
    numberstyle=\tiny\color{codegray},
    stringstyle=\color{codepurple},
    basicstyle=\ttfamily\footnotesize,
    breakatwhitespace=false,
    breaklines=true,
    captionpos=b,
    keepspaces=true,
    numbers=left,
    numbersep=5pt,
    showspaces=false,
    showstringspaces=false,
    showtabs=false,
    tabsize=2
}
\lstset{style=mystyle}

% Hyperlinks
\usepackage[hidelinks]{hyperref}
\usepackage{url}

% Bibliography
\usepackage[
    backend=biber,
    style=authoryear,
    sorting=nyt,
    maxcitenames=2
]{biblatex}
\addbibresource{references.bib}

% Glossaries and acronyms
```

```
\usepackage[acronym,toc]{glossaries}
\makeglossaries

% Headers and footers
\usepackage{fancyhdr}
\pagestyle{fancy}
\fancyhf{}
\fancyhead[L]{\leftmark}
\fancyhead[R]{\thepage}
\renewcommand{\headrulewidth}{0.4pt}

% Appendices
\usepackage[toc,page]{appendix}

% ============================================
% ACRONYMS
% ============================================

\newacronym{pb}{PB}{Pitch Book}
\newacronym{ib}{IB}{Investment Banking}
\newacronym{rag}{RAG}{Retrieval-Augmented Generation}
\newacronym{llm}{LLM}{Large Language Model}
\newacronym{vlm}{VLM}{Vision Language Model}
\newacronym{pm}{PM}{Precedent Material}
\newacronym{api}{API}{Application Programming Interface}
\newacronym{sdlc}{SDLC}{Software Development Lifecycle}
\newacronym{poc}{POC}{Proof of Concept}
\newacronym{hitl}{HITL}{Human-in-the-Loop}

% ============================================
% DOCUMENT INFO
% ============================================

\title{Compliance Aware, Agentic Pitch Book Generation For Bankers}
\author{Andrei Rizea}
\date{\today}

% ============================================
% DOCUMENT
% ============================================

\begin{document}

% --------------------------------------------
% FRONT MATTER
% --------------------------------------------

% Title Page
\begin{titlepage}
    \centering
    \vspace*{2cm}

    {\Huge\bfseries Compliance Aware, Agentic Pitch Book Generation For Bankers\par}

    \vspace{1cm}

    {\Large An all-in-one content generation tool providing agentic workflows with RAG search capabilities.\par}

    \vfill

    {\Large A Queen Mary Final Year Dissertation\par}

    \vspace{1cm}

    {\large\bfseries Andrei Rizea\par}
    {\large Student ID: 220300153\par}

    \vspace{1cm}

    {\large Supervised by: Manesha Peiris\par}

    \vspace{1cm}

    {\large Programme of Study:\par}
    {\large Digital \& Technology Solutions (Software Engineering)\par}

    \vspace{1cm}

    {\large Module: IOT635W\par}
    \vfill
```

```
    \vspace{1cm}

    {\large Queen Mary University of London\par}
    {\large \today\par}

\end{titlepage}

% Abstract
\begin{abstract}
\addcontentsline{toc}{chapter}{Abstract}
% TODO: Write abstract
\end{abstract}

% Acknowledgements
\chapter*{Acknowledgements}
\addcontentsline{toc}{chapter}{Acknowledgements}
% TODO: Write acknowledgements

% Table of Contents
\tableofcontents
\listoffigures
\listoftables
\printglossary[type=\acronymtype,title=List of Acronyms]

% ------------------------------------------
% MAIN MATTER
% ------------------------------------------

\chapter{Introduction, Scope and Context}
\label{ch:introduction}

Knowledge workers in contemporary organisations spend a disproportionate fraction of their time not generating knowledge but

\section{The Cognitive Economics of Investment Banking}
\label{sec:cognitive-economics}

Investment banking represents a paradigmatic knowledge industry, with professional value derived primarily from analytical in

The work of investment banking analysts divides into categories of markedly different cognitive demand. High-value activities

Radhakrishna and colleagues, investigating knowledge management processes in investment banking through structural equation m

\section{Problem Analysis: Knowledge Reuse Barriers in Document Production}
\label{sec:problem-analysis}

The inefficiencies observable in pitch book production reflect broader challenges in organisational knowledge management that

Three interconnected barriers impede the realisation of this reuse potential. The first concerns the misallocation of cogniti

The second barrier involves the erosion of institutional memory through inadequate knowledge codification. Knowledge manageme

The third barrier relates to the friction inherent in transferring tacit knowledge to new practitioners. Each institution mai

\section{Gap Analysis: The Evolution of Presentation Automation}
\label{sec:gap-analysis}

Research in automated presentation generation has progressed through distinct paradigms, each addressing different dimensions

The recent PPTAgent framework advances beyond this paradigm by adopting an edit-based approach that analyses reference presen

However, PPTAgent and comparable systems target general presentation scenarios, leaving domain-specific requirements unaddres

Commercial platforms offering AI-assisted presentation creation similarly address generic business contexts. Furthermore, ins

Retrieval-Augmented Generation provides an architectural foundation for grounding generative outputs in authoritative knowled

Agentic AI architectures extend these capabilities by enabling autonomous planning and execution of multi-step processes. Rat

The synthesis of these capabilities---template-aware generation, domain-specific data retrieval, and agentic orchestration---

\section{Aims and Objectives}
\label{sec:aims-objectives}

This project aims to design, implement, and evaluate a proof-of-concept system demonstrating how AI-powered document generati

From an organisational perspective, the aim is to demonstrate a workflow enabling analysts to generate substantive first draf

From an educational perspective, the aim is to investigate the technical integration of large language model orchestration wi

These aims translate into five concrete objectives:
```

```
\begin{enumerate}
    \item Design and implement a template analysis component that programmatically extracts layout structures, colour palette
    \item Develop an agentic data retrieval layer that autonomously gathers company financials, market data, and regulatory f
    \item Construct an LLM-powered content planning module that generates structured specifications for slide content, adapti
    \item Build a slide assembly component that maps planned content to extracted template structures, producing PowerPoint f
    \item Evaluate system outputs against metrics encompassing generation efficiency, template compliance, and content approp
\end{enumerate}

\section{Scope, Constraints and Success Criteria}
\label{sec:scope}

The project scope encompasses a functional proof of concept demonstrating the complete workflow from user input specification

Several constraints define implementation boundaries. The agentic orchestration layer employs established frameworks rather t

The project explicitly excludes real-time market data integration, production deployment infrastructure, and comprehensive re

Success evaluation applies four criteria. First, the system should generate complete pitch book drafts within five minutes of

Throughout, the system is designed for human-in-the-loop operation, producing drafts that support rather than supplant profes

% ---------------------------------------------
% SUBSEQUENT CHAPTERS (placeholders)
% ---------------------------------------------

\chapter{Literature Review}
\label{ch:literature-review}

% TODO: Expand literature review

\chapter{Methodology}
\label{ch:methodology}

% TODO: Methodology content

\chapter{System Design}
\label{ch:system-design}

% TODO: System design content

\chapter{Implementation}
\label{ch:implementation}

% TODO: Implementation details

\chapter{Testing and Evaluation}
\label{ch:testing}

% TODO: Testing methodology and results

\chapter{Results and Discussion}
\label{ch:results}

% TODO: Present and discuss results

\chapter{Conclusion}
\label{ch:conclusion}

% TODO: Conclusions and future work

% ---------------------------------------------
% BACK MATTER
% ---------------------------------------------

% Bibliography
\printbibliography[heading=bibintoc]

% Appendices
\begin{appendices}

\chapter{Risk Assessment}
\label{app:risk-assessment}

% TODO: Include risk assessment table

\chapter{Project Plan}
\label{app:project-plan}

% TODO: Include Gantt chart or project timeline
\chapter{KSB Mapping}
```

```
\label{app:ksb-mapping}

% TODO: Map project to apprenticeship Knowledge, Skills, and Behaviours

\end{appendices}

\end{document}
```

4.2 FILE: references.bib

```
% Bibliography file for the dissertation
% Compliance Aware, Agentic Pitch Book Generation For Bankers
% Updated with fresh sources for Introduction, Scope and Context

% ============================================
% AI AND DOCUMENT GENERATION
% ============================================

@article{zheng2025pptagent,
  author    = {Zheng, Hao and Guan, Xinyan and Kong, Hao and Zheng, Jiangjie and Zhou, Wei and Lin, Hongye and Lu, Yutao and
  title     = {{PPTAgent}: Generating and Evaluating Presentations Beyond Text-to-Slides},
  journal   = {arXiv preprint},
  year      = {2025},
  eprint    = {2501.03936},
  archiveprefix = {arXiv},
  url       = {https://arxiv.org/abs/2501.03936}
}

@article{liu2024structured,
  author    = {Liu, Michael Xieyang and Liu, Frederick and Fiannaca, Alexander J. and Koo, Terry and Dixon, Lucas and Terry,
  title     = {{``We Need Structured Output''}: Towards User-centered Constraints on Large Language Model Output},
  booktitle = {Extended Abstracts of the CHI Conference on Human Factors in Computing Systems},
  year      = {2024},
  publisher = {ACM},
  address   = {Honolulu, HI, USA},
  doi       = {10.1145/3613905.3650756}
}

@article{tang2024strucbench,
  author    = {Tang, Xiangru and Zong, Yiming and Phang, Jason and Zhao, Yilun and Zhou, Wangchunshu and Cohan, Arman and Ger
  title     = {{Struc-Bench}: Are Large Language Models Good at Generating Complex Structured Tabular Data?},
  booktitle = {Proceedings of the 2024 Conference of the North American Chapter of the Association for Computational Linguist
  year      = {2024},
  pages     = {12--34},
  publisher = {Association for Computational Linguistics}
}

@article{digenio2024semistructured,
  author    = {Di Genio, Giuseppe and others},
  title     = {{LLM} Based Multi-Agent Generation of Semi-structured Documents from Semantic Templates in the Public Administ
  journal   = {arXiv preprint},
  year      = {2024},
  eprint    = {2402.14871},
  archiveprefix = {arXiv}
}

% ============================================
% RAG AND KNOWLEDGE RETRIEVAL
% ============================================

@article{lewis2020rag,
  author    = {Lewis, Patrick and Perez, Ethan and Piktus, Aleksandra and Petroni, Fabio and Karpukhin, Vladimir and Goyal, N
  title     = {Retrieval-Augmented Generation for Knowledge-Intensive {NLP} Tasks},
  journal   = {Advances in Neural Information Processing Systems},
  volume    = {33},
  pages     = {9459--9474},
  year      = {2020}
}

@article{gao2023rag,
  author    = {Gao, Yunfan and Xiong, Yun and Gao, Xinyu and Jia, Kangxiang and Pan, Jinliu and Bi, Yuxi and Dai, Yi and Sun,
  title     = {Retrieval-Augmented Generation for Large Language Models: A Survey},
  journal   = {arXiv preprint},
  year      = {2023},
  eprint    = {2312.10997},
  archiveprefix = {arXiv}
}

@article{schneider2025ragbise,
  author    = {Schneider, Johannes and Feuerriegel, Stefan and others},
  title     = {Retrieval-Augmented Generation ({RAG})},
  journal   = {Business \& Information Systems Engineering},
  year      = {2025},
  publisher = {Springer},
  doi       = {10.1007/s12599-025-00945-3}
```

```
}

@article{chen2025ragsystematic,
  author    = {Chen, Jiaying and others},
  title     = {Retrieval-Augmented Generation ({RAG}) and Large Language Models ({LLMs}) for Enterprise Knowledge Management
  journal   = {Applied Sciences},
  volume    = {16},
  number    = {1},
  pages     = {368},
  year      = {2025},
  publisher = {MDPI}
}

% ============================================
% AGENTIC AI AND WORKFLOWS
% ============================================

@online{ibm2024agentic,
  author    = {{IBM}},
  title     = {What are Agentic Workflows?},
  year      = {2024},
  url       = {https://www.ibm.com/think/topics/agentic-workflows},
  urldate   = {2025-11-18}
}

@article{mckinsey2025agentic,
  author    = {{McKinsey \& Company}},
  title     = {Seizing the Agentic {AI} Advantage},
  journal   = {McKinsey Insights},
  year      = {2025},
  url       = {https://www.mckinsey.com/capabilities/quantumblack/our-insights/seizing-the-agentic-ai-advantage}
}

@article{bcg2025agentic,
  author    = {{Boston Consulting Group}},
  title     = {How Agentic {AI} is Transforming Enterprise Platforms},
  journal   = {BCG Publications},
  year      = {2025},
  url       = {https://www.bcg.com/publications/2025/how-agentic-ai-is-transforming-enterprise-platforms}
}

@article{deloitte2025agentic,
  author    = {{Deloitte}},
  title     = {Agentic {AI} Strategy},
  journal   = {Deloitte Insights: Tech Trends 2026},
  year      = {2025},
  url       = {https://www.deloitte.com/us/en/insights/topics/technology-management/tech-trends/2026/agentic-ai-strategy.html
}

@article{mitsloan2025agentic,
  author    = {{MIT Sloan Management Review} and {Boston Consulting Group}},
  title     = {The Emerging Agentic Enterprise: How Leaders Must Navigate a New Age of {AI}},
  journal   = {MIT Sloan Management Review},
  year      = {2025},
  url       = {https://sloanreview.mit.edu/projects/the-emerging-agentic-enterprise-how-leaders-must-navigate-a-new-age-of-ai
}

% ============================================
% AI IN FINANCIAL SERVICES
% ============================================

@article{mirishli2025regulating,
  author    = {Mirishli, Shahmar},
  title     = {Regulating {AI} in Financial Services: Legal Frameworks and Compliance Challenges},
  journal   = {arXiv preprint},
  year      = {2025},
  eprint    = {2503.14541},
  archiveprefix = {arXiv}
}

@article{nature2025aifinance,
  author    = {Various},
  title     = {{AI} Integration in Financial Services: A Systematic Review of Trends and Regulatory Challenges},
  journal   = {Humanities and Social Sciences Communications},
  year      = {2025},
  publisher = {Nature Portfolio},
  doi       = {10.1038/s41599-025-04850-8}
}

@report{fsb2024aistability,
```

```
  author    = {{Financial Stability Board}},
  title     = {The Financial Stability Implications of Artificial Intelligence},
  year      = {2024},
  url       = {https://www.fsb.org/uploads/P14112024.pdf}
}

@report{bis2024regulating,
  author    = {{Bank for International Settlements}},
  title     = {Regulating {AI} in the Financial Sector: Recent Developments and Main Challenges},
  series    = {FSI Insights},
  number    = {63},
  year      = {2024},
  url       = {https://www.bis.org/fsi/publ/insights63.pdf}
}

@article{mckinsey2023genai,
  author    = {{McKinsey \& Company}},
  title     = {Been There, Doing That: How Corporate and Investment Banks Are Tackling Gen {AI}},
  journal   = {McKinsey Financial Services Insights},
  year      = {2023},
  month     = sep,
  url       = {https://www.mckinsey.com/industries/financial-services/our-insights/been-there-doing-that-how-corporate-and-in
}

@online{alphasense2025genai,
  author    = {{AlphaSense}},
  title     = {Generative {AI} in Investment Banking: Use Cases and Future Outlook},
  year      = {2025},
  url       = {https://www.alpha-sense.com/blog/trends/generative-ai-in-investment-banking/}
}

% ============================================
% INVESTMENT BANKING PRODUCTIVITY
% ============================================

@online{cfi2025analyst,
  author    = {{Corporate Finance Institute}},
  title     = {Investment Banking Analyst Job Description, Hours, and Salary},
  year      = {2025},
  url       = {https://corporatefinanceinstitute.com/resources/career/investment-banking-analyst-job-description-hours-salary
  urldate   = {2025-02-24}
}

@online{wso2024hours,
  author    = {{Wall Street Oasis}},
  title     = {Investment Banking Hours: Typical Schedule, Work-Life Balance, and Expectations},
  year      = {2024},
  url       = {https://www.wallstreetoasis.com/resources/careers/jobs/investment-banking-hours}
}

@online{streetofwalls2024analyst,
  author    = {{Street of Walls}},
  title     = {Investment Banking Analyst},
  year      = {2024},
  url       = {https://www.streetofwalls.com/finance-training-courses/investment-banking-overview-and-behavioral-training/inv
}

% ============================================
% KNOWLEDGE MANAGEMENT
% ============================================

@article{markus2001knowledge,
  author    = {Markus, M. Lynne},
  title     = {Toward a Theory of Knowledge Reuse: Types of Knowledge Reuse Situations and Factors in Reuse Success},
  journal   = {Journal of Management Information Systems},
  volume    = {18},
  number    = {1},
  pages     = {57--93},
  year      = {2001},
  publisher = {Taylor \& Francis}
}

@article{radhakrishna2024km,
  author    = {Radhakrishna, Vedapradha and Ravi, Hariharan and Elango, Sudha and Divyashree, Venkoba and Jaganathan, Ashok a
  title     = {Future of Knowledge Management in Investment Banking: Role of Personal Intelligent Assistants},
  journal   = {SAGE Open},
  year      = {2024},
  volume    = {14},
  number    = {4},
  publisher = {SAGE Publications},
```

```
  doi       = {10.1177/20597991241287118}
}

@book{dalkir2005km,
  author    = {Dalkir, Kimiz},
  title     = {Knowledge Management in Theory and Practice},
  publisher = {Elsevier Butterworth-Heinemann},
  year      = {2005},
  address   = {Burlington, MA}
}

@article{pmcknowledge2020,
  author    = {Various},
  title     = {The Current Understanding of Knowledge Management Concepts: A Critical Review},
  journal   = {Journal of Knowledge Management},
  year      = {2020},
  publisher = {PMC},
  url       = {https://pmc.ncbi.nlm.nih.gov/articles/PMC7787024/}
}

@article{sciencedirect2013financeKM,
  author    = {Various},
  title     = {The Review of Knowledge Management in Financial Industry},
  journal   = {Procedia - Social and Behavioral Sciences},
  year      = {2013},
  publisher = {Elsevier},
  doi       = {10.1016/j.sbspro.2012.12.183}
}

% ============================================
% HUMAN-IN-THE-LOOP AND AI SAFETY
% ============================================

@online{ibm2024hitl,
  author    = {{IBM}},
  title     = {What is Human in the Loop ({HITL})?},
  year      = {2024},
  url       = {https://www.ibm.com/think/topics/human-in-the-loop},
  urldate   = {2025-11-18}
}

% ============================================
% PRODUCTIVITY MEASUREMENT
% ============================================

@article{sauermann2023performance,
  author    = {Sauermann, Jan},
  title     = {Performance Measures and Worker Productivity},
  journal   = {IZA World of Labor},
  year      = {2023},
  publisher = {Institute of Labor Economics},
  doi       = {10.15185/izawol.260.v2}
}
```