Writing Revision of Sustainability Reports Using Generative Al

Sustainability reporting is becoming more complex and mandatory under new global regulations like CSRD. This project explores using generative AI – specifically a multi-agent system with Retrieval-Augmented Generation (RAG) – to automate the annual revision of ESG reports. Our prototype shows AI can enhance efficiency in updating key sections while maintaining factual accuracy.

Authors

Professor Sean Shin Mr. Max Wang Team: Meng Hongyu, Pan Haochen, Thunder Tan, Wang Shiyue Advisor: Mira Wang

Affiliations

National University of Singapore, Business School (SGF) Singapore Exchange Limited (SGX)

1 Introduction

As sustainability standards evolve, companies face growing pressure to update ESG reports annually. These updates are labor-intensive, involving not just data changes but also narrative revisions to align with new frameworks and expectations. Manual revisions consume significant time and often risk inconsistency. This project explores how generative AI can assist in this process by automating routine updates while preserving tone and compliance.

2 Objectives

Board Message Revision

Develop and test a multi-agent AI system to revise CEO/Board statements across years. The goal is to preserve tone and narrative structure while integrating updated achievements and priorities.

Tone Identification and Adjustment

Evaluate whether AI agents can detect and modify report tone from different stakeholder perspectives (e.g., regulatory, investor), improving narrative consistency and clarity.

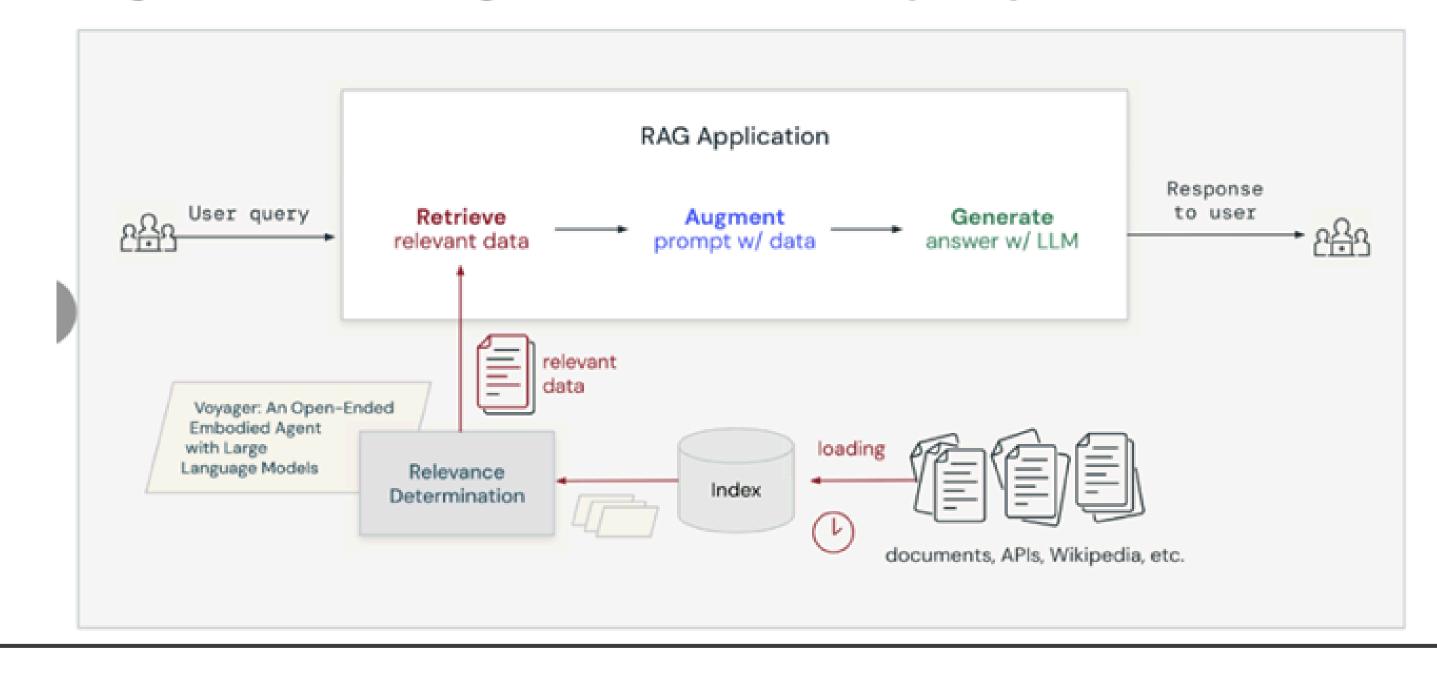
Framework Transition

Assess the Al's ability to align disclosures with evolving frameworks by identifying reporting gaps and adapting content to meet new standards.

3 Methodology

We built a local prototype using AutoGen, an open-source multi-agent framework, to simulate writing team collaboration through iterative dialogue among agents with distinct roles (research, drafting, critique). RAG integration ensured accuracy by retrieving ESG knowledge base content (annual reports, board statements, etc.) via semantic similarity for generation. Key to contextual relevance were prompt engineering, dialogue control, and RAG integration. The prototype was locally deployed for data privacy and tested with HSBC and Singtel examples.

Augmentation: Adding context to the user's prompt



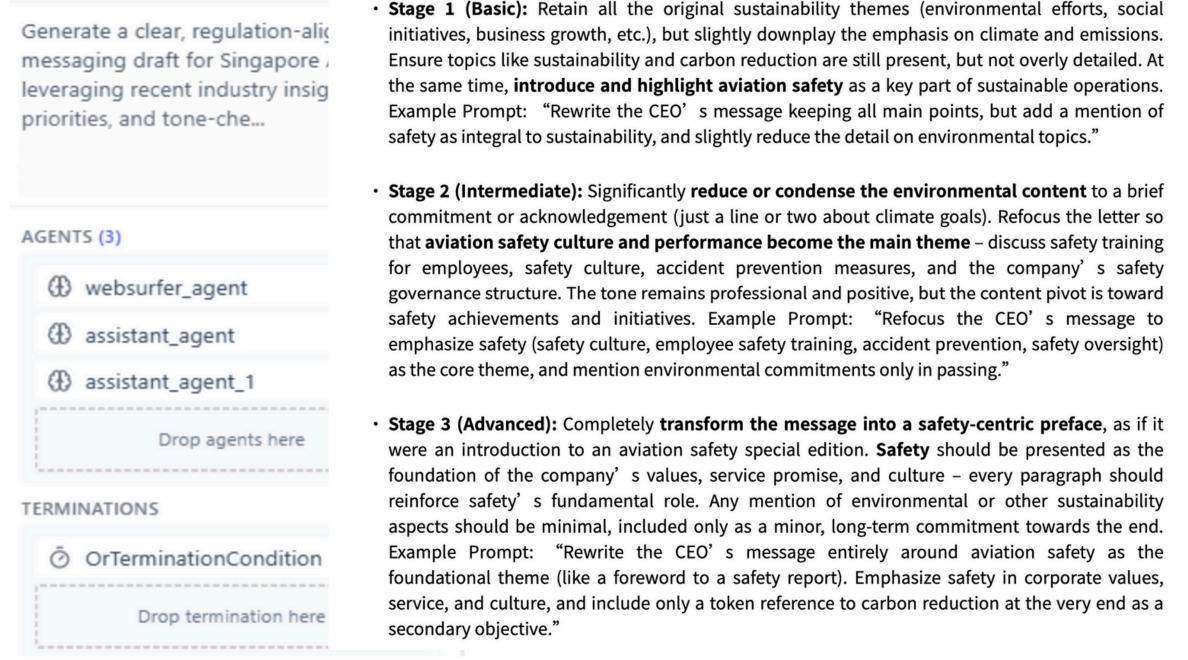
4 Analysis

Board Message Revision

Key Findings and Challenges

ESG Draft Generator Test team

Al-led revision of ESG CEO messages using a multi-agent framework offers promising efficiency gains and structural alignment with existing board message formats. Specialised agents (research, strategy, critic) enhance strategic focus, thematic responsiveness, and regulatory compliance. Prompt testing confirms sensitivity to tone shifts and iterative refinement. However, the system faces significant issues: fragmented messaging due to agent specialisation, high setup complexity, and time-intensive debugging. These limitations highlight the need for streamlined coordination or alternative approaches like prompt-engineered single-LLM systems.

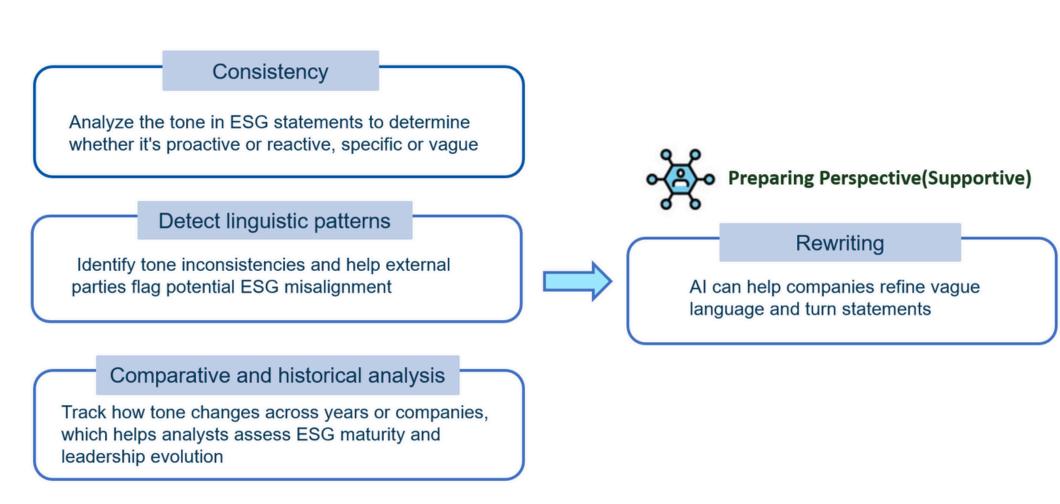


Prompt Design Suggestions for Each Stage

Tone Change

Multi-agent LLMs improve ESG tone analysis by assigning distinct tasks—extraction, sentiment scoring, benchmarking, and validation—to specialized agents. This modular design enhances interpretability, factual accuracy, and longitudinal tracking, while internal critique loops support auditability and regulatory alignment.

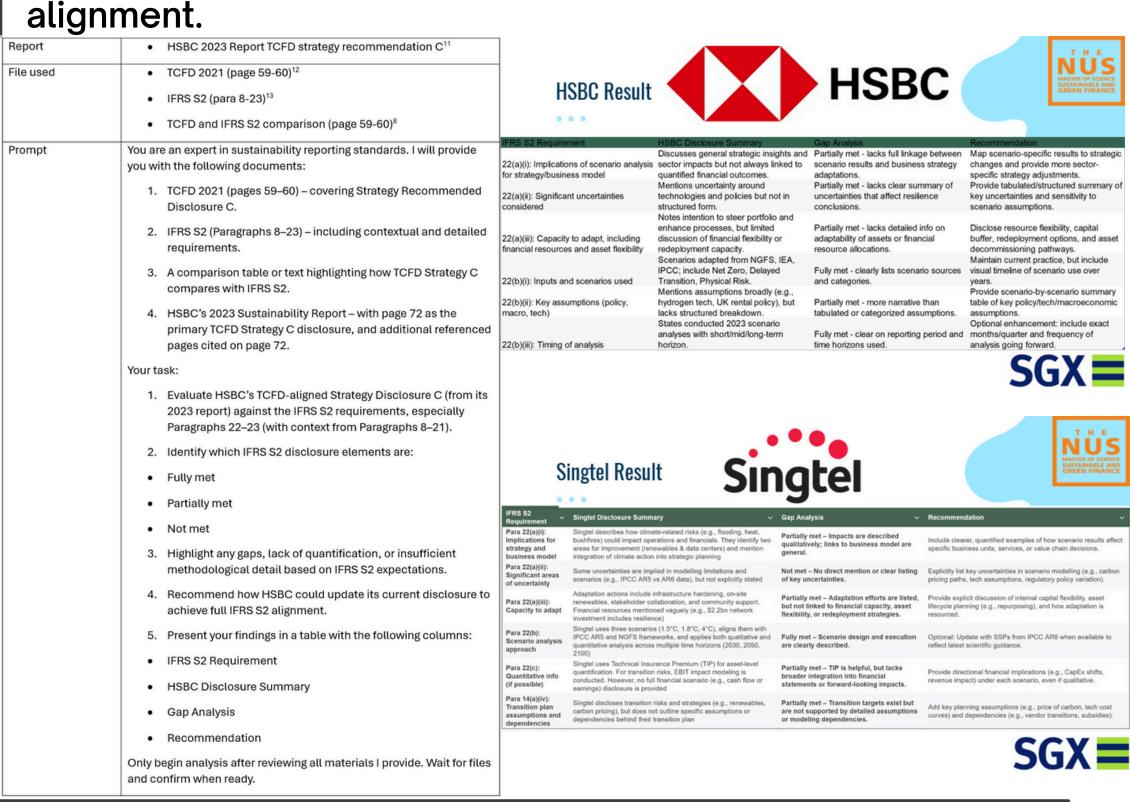
However, the system faces key limitations: fragmented outputs due to agent specialization, high coordination complexity, and the absence of a synthesis mechanism to ensure narrative cohesion. These challenges may limit its effectiveness compared to streamlined single-model approaches for end-to-end ESG disclosure drafting.



Framework Transition

Key Findings and Challenges

Our study compared HSBC and Singtel's TCFD-aligned disclosures against IFRS S2 (Paras. 22–23) using ChatGPT-4o. While the AI accurately mapped HSBC's preprocessed report, it struggled with Singtel's longer, less-structured version—misinterpreting sections and adding nonexistent references. Key challenges identified include file size limitations, inconsistent outputs due to vague prompts, and interpretative bias in unstructured data. To improve AI mapping, reports should be standardised, prompts made more precise, and outputs reviewed by experts to ensure accuracy and reliability in disclosure



5 Conclusion

As sustainability standards evolve, companies face growing pressure to update ESG reports annually. These updates are labor-intensive, involving not just data changes but also narrative revisions to align with new frameworks and expectations. Manual revisions consume significant time and often risk inconsistency. This project explores how generative AI can assist in this process by automating routine updates while preserving tone and compliance.