

Incorporating Knowledge Graphs and Large Language Models into Visual Text Analysis Tools

Adam Coscia, Alex Endert

Liz Richerson, Sue Mi K., Stephen S., Tim S.

Introduction

- Analyzing 100+ documents for insights is **time-consuming**; can we **speed it up**?
- Large language models (LLMs) and knowledge graphs (KGs) can help **summarize, extract, explain, and validate** text in documents!
- How can LLMs & KGs be put in **visual text analysis tools**? How will they affect **sensemaking**?

Methods

- We **co-designed** a prototype **visualization tool** that lets analysts create **document piles** and...
 - Run **LLM tasks** like **summarize, extract, and answer questions** directly on documents in piles
 - Search/extract **KG facts** + **trace the source**, to find **new docs** & put them in piles for **follow-up analysis**
- We ran a study w/ **845 docs**, to observe how **LLMs** and **KGs** affect **sensemaking** during **analysis**

Watch our video!



acoscia6@gatech.edu
<https://adamcoscia.com/papers/vispile>

LAS Laboratory for Analytic Sciences
 2024 Research Symposium

This material is based upon work done, in whole or in part, in coordination with the Department of Defense (DoD). Any opinions, findings, conclusions, or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the DoD and/or any agency or entity of the United States Government.

1

We built a tool, **VisPile**, to explore how LLMs & KGs can help analysts find relevant sources, analyze document, & validate LLM responses!

1. Use LLM semantic search to find documents to put into piles!

2. Run LLM tasks to analyze documents, summarize events, answer questions, even custom prompt!

3. Validate response by extracting KG entities, linking similar sources, & suggesting related documents to read

4. Take notes inline to present your findings!

2

We then ran a study w/ students. **VisPile** enabled new LLM- & KG-assisted workflows, yet raised issues of trust and understanding!

Results

- Some students **blindly trusted** LLM responses, **assumed** the system “**knew what it was doing...**”
- LLMs/KGs saved **reading time** ↓ but increased **time spent validating** ↑ **Follow-up prompts** helped!

