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

Laboratory for Analytic Sciences
2024 Research Symposium

Adam **Coscia,** Alex **Endert**
Georgia Institute of Technology, Atlanta, GA

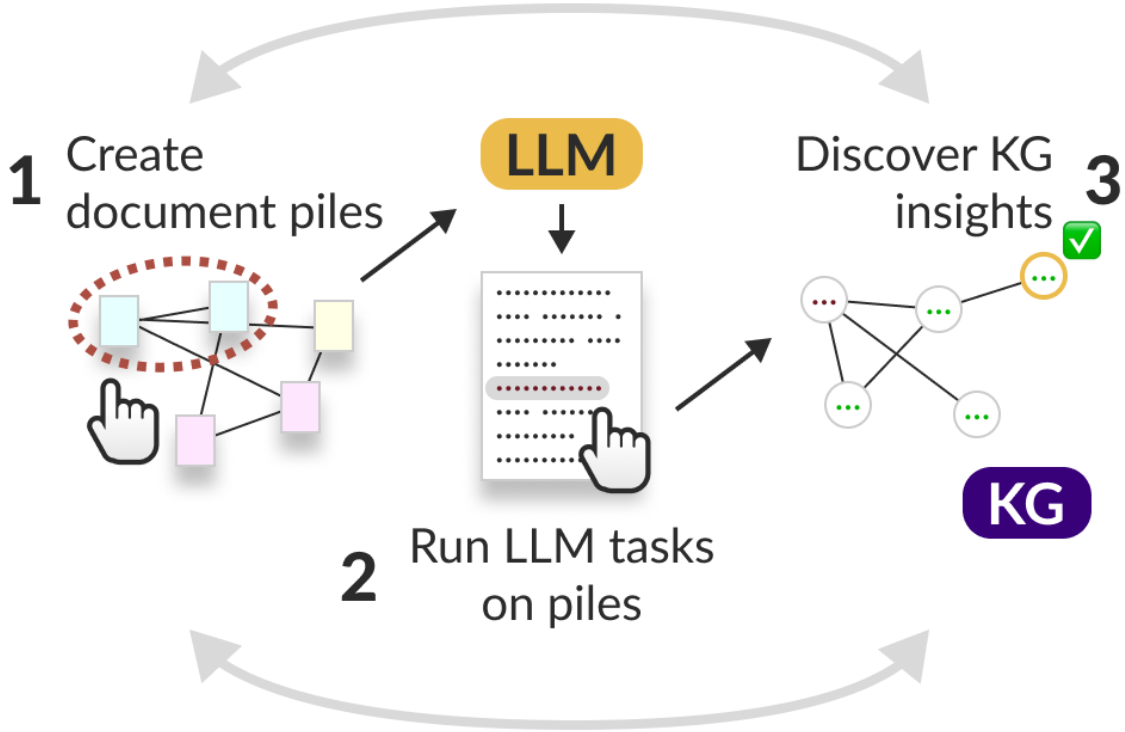
Liz **Richerson,** Sue Mi **K.,** Stephen **S.,** Tim **S.**
Laboratory of Analytic Sciences, Raleigh, NC

Knowledge graphs & large language models are great for **analyzing text data…**

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.

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

Laboratory for Analytic Sciences
2024 Research Symposium



Adam **Coscia,** Alex **Endert**
Georgia Institute of Technology, Atlanta, GA

Liz **Richerson,** Sue Mi **K.,** Stephen **S.,** Tim **S.**
Laboratory of Analytic Sciences, Raleigh, NC

Open questions

1. Where to put LLMs & KGs in **analysis tools**?
2. How will LLMs & KGs affect **sensemaking**?

Intro | Motivation

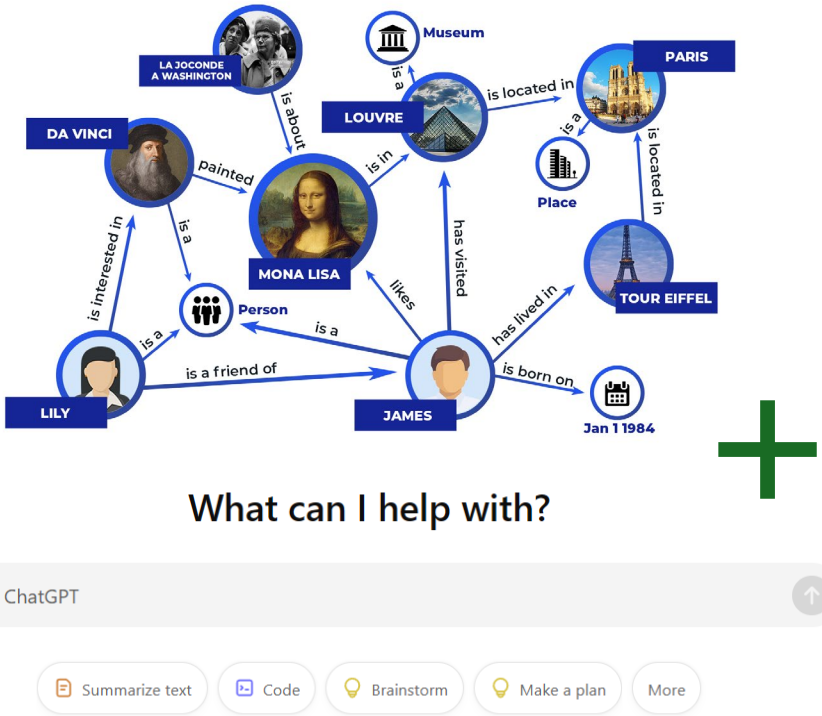
New tech can help with analyzing hundreds of documents:

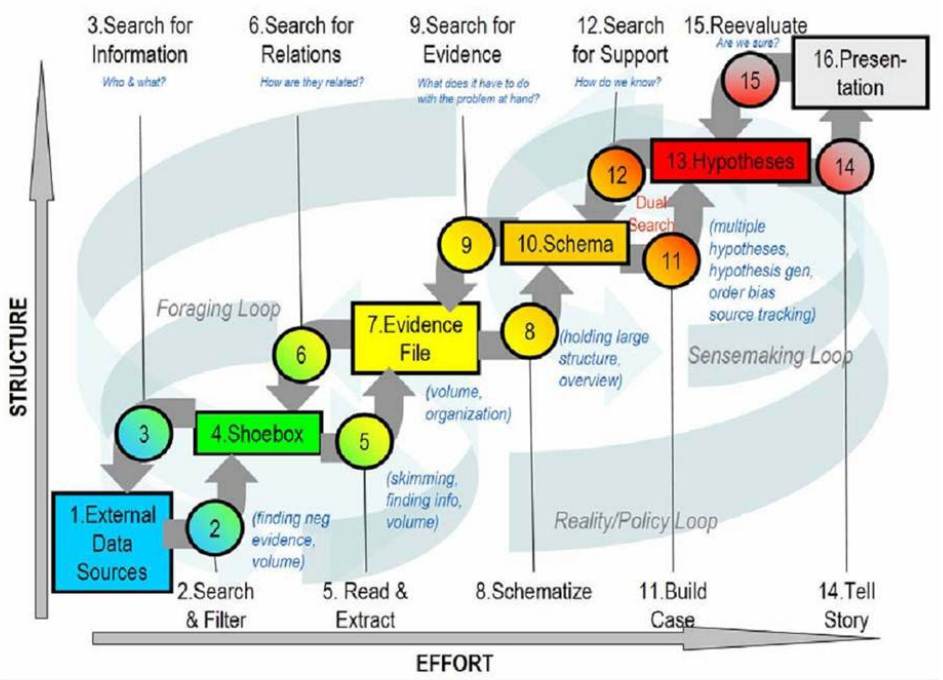
- **Large language models (LLMs)** are good at summarizing text, open-ended question-answering, extracting entities, etc.
- **Knowledge graphs (KG)** provide facts & relationships between entities in text + suggest new sources of information to explore
- **Visual analytics** encourages using interactive interfaces to perform exploration, sensemaking, and decision-making tasks

What are the opportunities for combining LLMs, KGs, and visual analytics to support intelligence analysis?

Intro | Research questions

1. Where to put LLMs & KGs in **analysis tools?**
2. How will LLMs & KGs affect **sensemaking?**

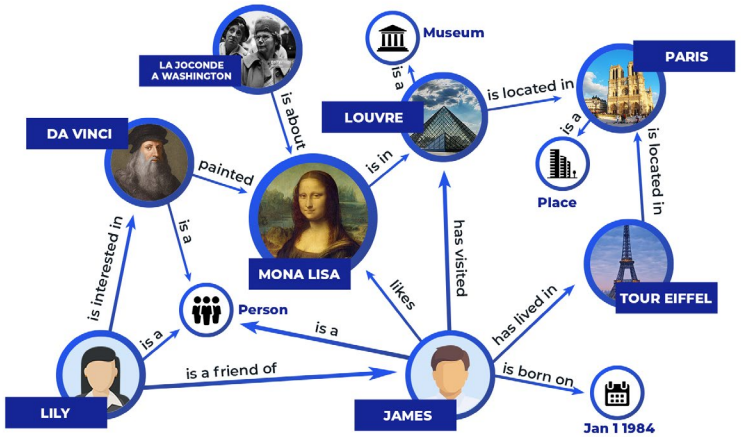


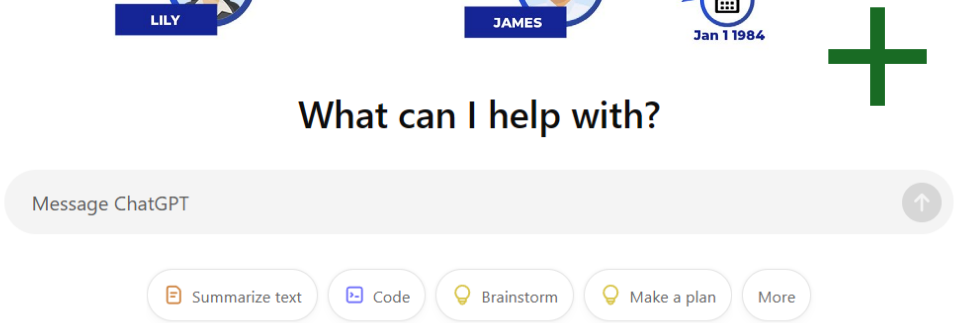


Pirolli and Card, 2005. *The sensemaking process and leverage points for analyst technology as identified through cognitive task analysis.*

Intro | Research questions

1. Where to put LLMs & KGs in **analysis tools?**
2. How will LLMs & KGs affect **sensemaking?**

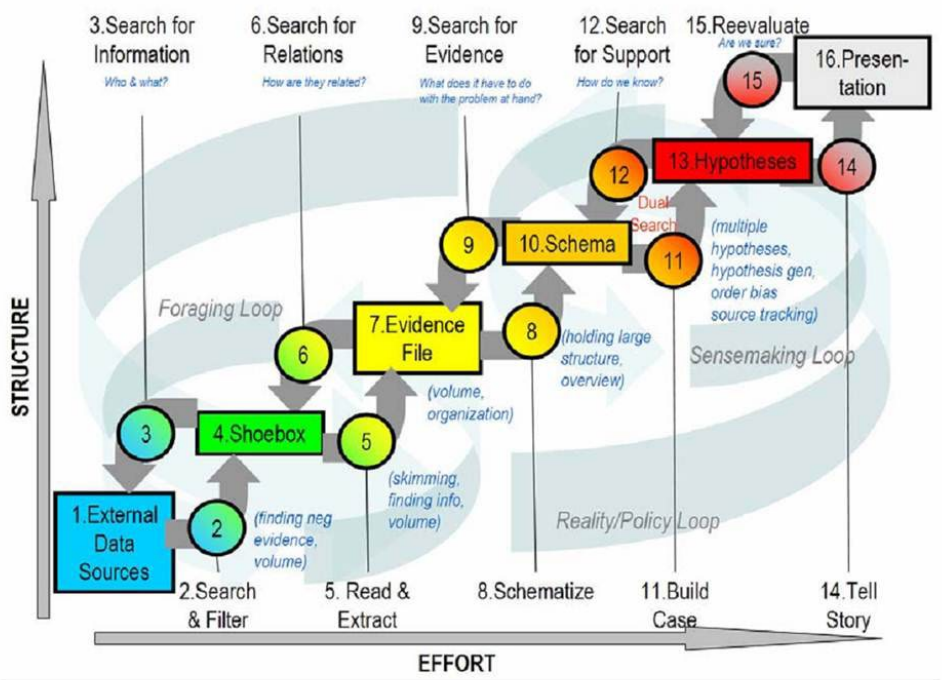




Pirolli and Card, 2005. *The sensemaking process and leverage points for analyst technology as identified through cognitive task analysis.*

Intro | Research questions

1. Where to put LLMs & KGs in **analysis tools?**
2. How will LLMs & KGs affect **sensemaking?**



Pirolli and Card, 2005. *The sensemaking process and leverage points for analyst technology as identified through cognitive task analysis.*

Formative work | Co-design with analysts

We synthesized four user challenges and tasks:

C1 – **Finding relevant documents**

- **Enable** semantic search over documents using LLMs, **Suggest** related documents

C2 – **Analyzing groups of documents**

- **Map** 8 different analysis tasks to LLM prompts + allow custom prompts!

1. Analyze documents
2. Summarize events
3. Extract entities
4. Classify topics
5. Generate questions
6. List tasks
7. Explain concepts
8. Answer questions

C3 – **Validating LLM & KG data**

- **Extract** KG entities from text, **Link** LLM response to related sources

C4 – **Tracking facts & insights**

- **Bookmark** LLM responses + KG facts, **Trace** KG facts to source docs

Formative work | Co-design with analysts

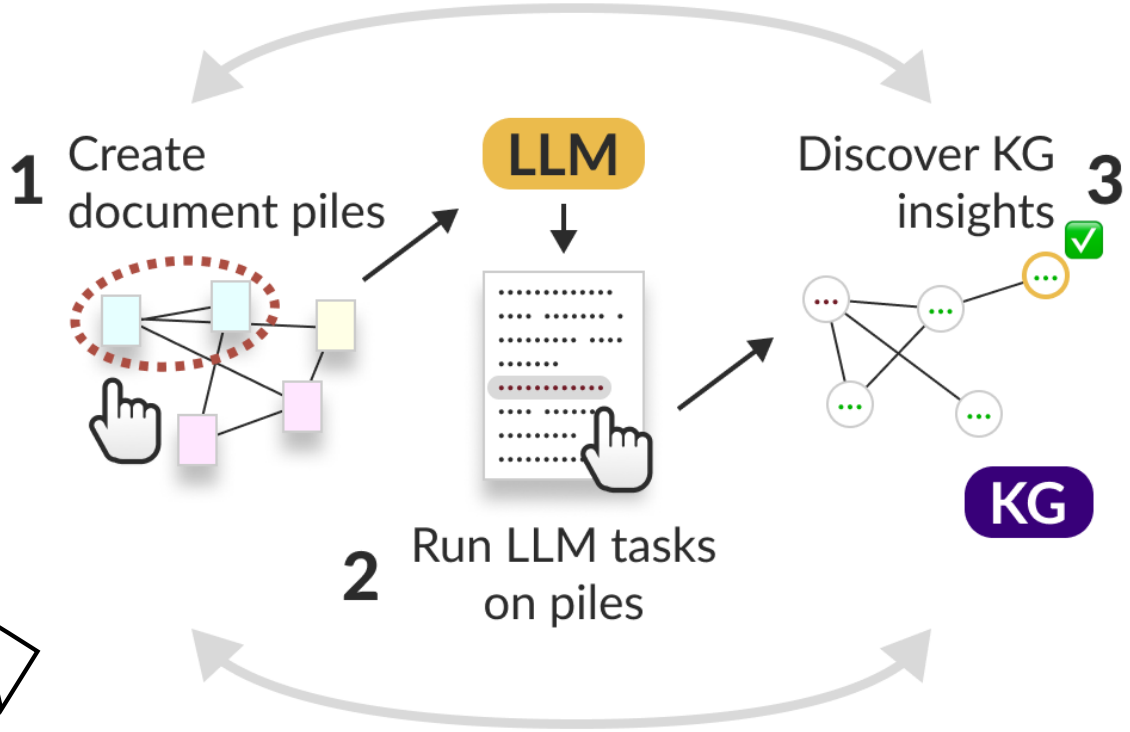
A new document piling workflow:

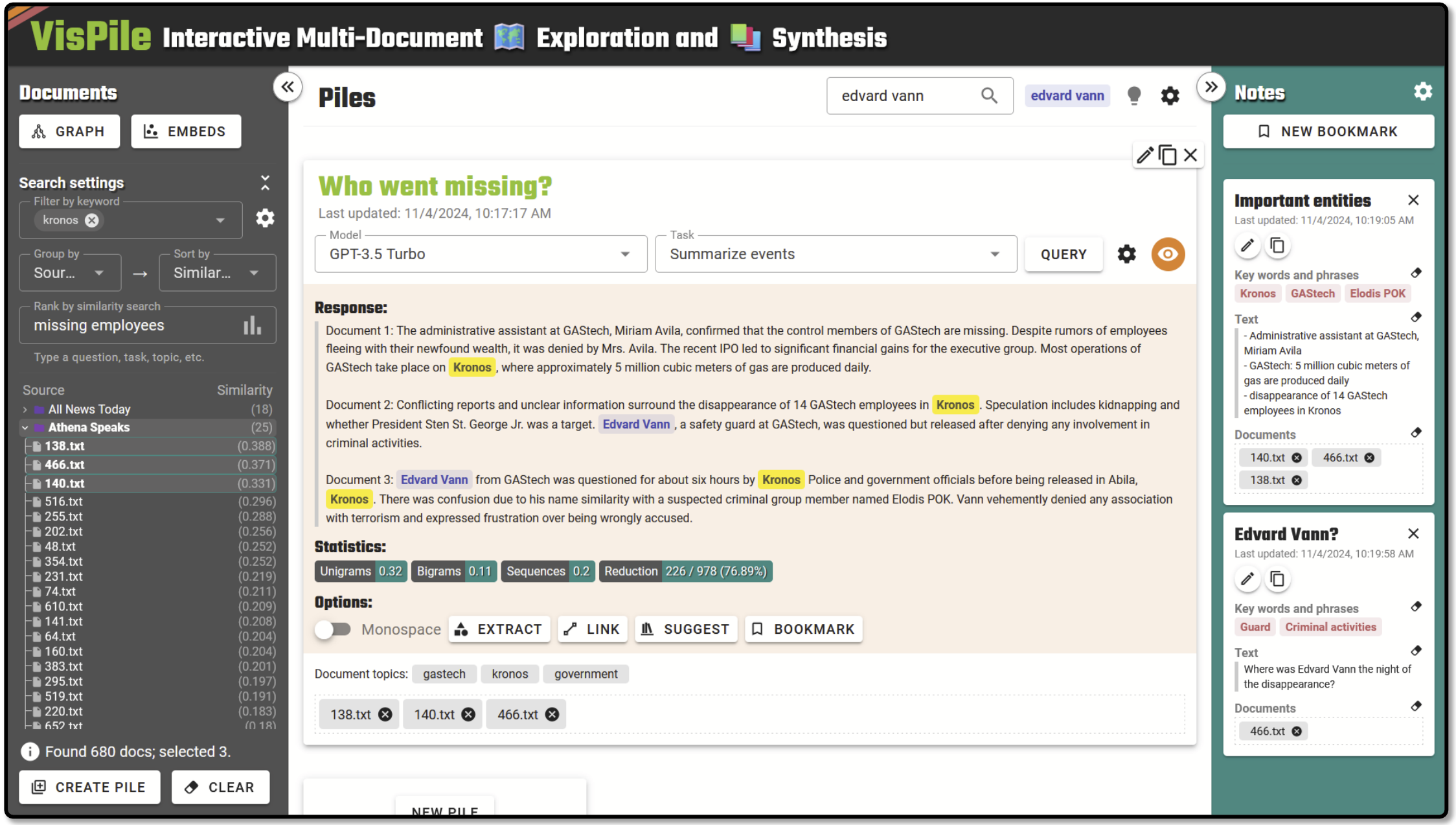
C1 – **Finding relevant documents**

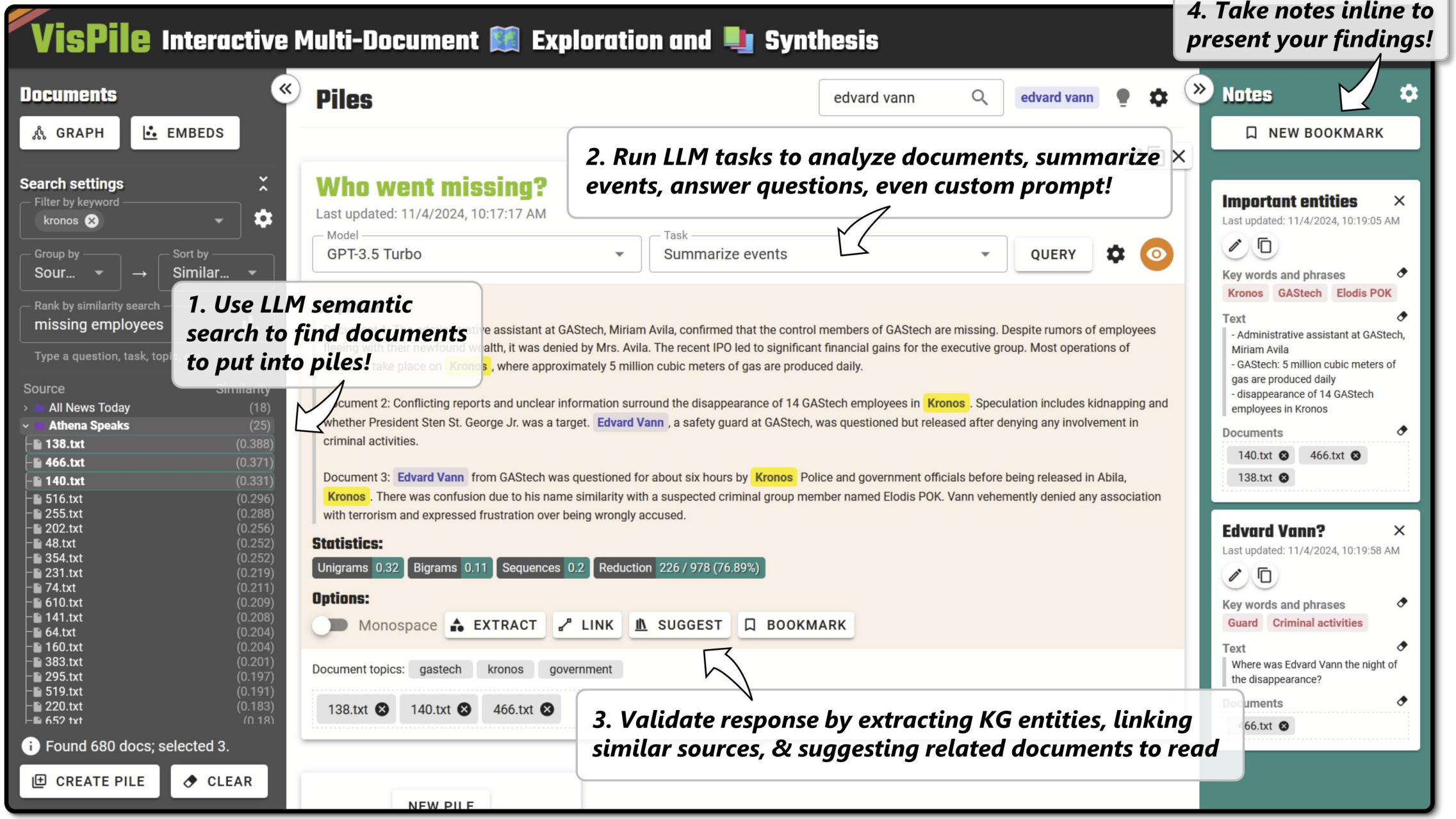
C2 – **Analyzing groups of documents**

C3 – **Validating LLM & KG data**

C4 – **Tracking facts & insights**




VisPile Interactive Multi-Document Exploration and Synthesis
Documents
GRAPH
EMBEDS
Search settings
Filter by keyword
kronos
Group by
Sour...
Sort by
Similar...
Rank by similarity search
missing employees
Type a question, task, topic, etc.
Source
Similarity
All News Today (18)
Athena Speaks (25)
138.txt (0.388)
466.txt (0.371)
140.txt (0.331)
516.txt (0.296)
255.txt (0.288)
202.txt (0.256)
48.txt (0.252)
354.txt (0.252)
231.txt (0.219)
74.txt (0.211)
610.txt (0.209)
141.txt (0.208)
64.txt (0.204)
160.txt (0.204)
383.txt (0.201)
295.txt (0.197)
519.txt (0.191)
220.txt (0.183)
Found 680 docs; selected 3.
CREATE PILE
CLEAR
Piles
edvard vann
edvard vann
Who went missing?
Last updated: 11/4/2024, 10:17:17 AM
Model
GPT-3.5 Turbo
Task
Summarize events
QUERY
Response:
Document 1: The administrative assistant at GAStech, Miriam Avila, confirmed that the control members of GAStech are missing. Despite rumors of employees fleeing with their newfound wealth, it was denied by Mrs. Avila. The recent IPO led to significant financial gains for the executive group. Most operations of GAStech take place on Kronos, where approximately 5 million cubic meters of gas are produced daily.
Document 2: Conflicting reports and unclear information surround the disappearance of 14 GAStech employees in Kronos. Speculation includes kidnapping and whether President Sten St. George Jr. was a target. Edvard Vann, a safety guard at GAStech, was questioned but released after denying any involvement in criminal activities.
Document 3: Edvard Vann from GAStech was questioned for about six hours by Kronos Police and government officials before being released in Abila, Kronos. There was confusion due to his name similarity with a suspected criminal group member named Elodis POK. Vann vehemently denied any association with terrorism and expressed frustration over being wrongly accused.
Statistics:
Unigrams 0.32
Bigrams 0.11
Sequences 0.2
Reduction 226 / 978 (76.89%)
Options:
Monospace
EXTRACT
LINK
SUGGEST
BOOKMARK
Document topics:
gastech
kronos
government
138.txt
140.txt
466.txt
NEW PILE
Notes
NEW BOOKMARK
Important entities
Last updated: 11/4/2024, 10:19:05 AM
Key words and phrases
Kronos
GAStech
Elodis POK
Text
- Administrative assistant at GAStech, Miriam Avila
- GAStech: 5 million cubic meters of gas are produced daily
- disappearance of 14 GAStech employees in Kronos
Documents
140.txt
466.txt
138.txt
Edvard Vann?
Last updated: 11/4/2024, 10:19:58 AM
Key words and phrases
Guard
Criminal activities
Text
Where was Edvard Vann the night of the disappearance?
Documents
466.txt


VisPile Interactive Multi-Document Exploration and Synthesis
4. Take notes inline to present your findings!
Documents
GRAPH
EMBEDS
Search settings
Filter by keyword
kronos
Group by
Sour...
Sort by
Similar...
Rank by similarity search
missing employees
Type a question, task, topic
Source
All News Today (18)
Athena Speaks (25)
138.txt (0.388)
466.txt (0.371)
140.txt (0.331)
516.txt (0.296)
255.txt (0.288)
202.txt (0.256)
48.txt (0.252)
354.txt (0.252)
231.txt (0.219)
74.txt (0.211)
610.txt (0.209)
141.txt (0.208)
64.txt (0.204)
160.txt (0.204)
383.txt (0.201)
295.txt (0.197)
519.txt (0.191)
220.txt (0.183)
Found 680 docs; selected 3.
CREATE PILE
CLEAR
1. Use LLM semantic search to find documents to put into piles!
Piles
edvard vann
edvard vann
Who went missing?
Last updated: 11/4/2024, 10:17:17 AM
2. Run LLM tasks to analyze documents, summarize events, answer questions, even custom prompt!
Model
GPT-3.5 Turbo
Task
Summarize events
QUERY
e assistant at GAStech, Miriam Avila, confirmed that the control members of GAStech are missing. Despite rumors of employees
alth, it was denied by Mrs. Avila. The recent IPO led to significant financial gains for the executive group. Most operations of
take place on Kronos, where approximately 5 million cubic meters of gas are produced daily.
cument 2: Conflicting reports and unclear information surround the disappearance of 14 GAStech employees in Kronos. Speculation includes kidnapping and whether President Sten St. George Jr. was a target. Edvard Vann, a safety guard at GAStech, was questioned but released after denying any involvement in criminal activities.
Document 3: Edvard Vann from GAStech was questioned for about six hours by Kronos Police and government officials before being released in Abila, Kronos. There was confusion due to his name similarity with a suspected criminal group member named Elodis POK. Vann vehemently denied any association with terrorism and expressed frustration over being wrongly accused.
Statistics:
Unigrams 0.32
Bigrams 0.11
Sequences 0.2
Reduction 226 / 978 (76.89%)
Options:
Monospace
EXTRACT
LINK
SUGGEST
BOOKMARK
Document topics: gastech kronos government
138.txt
140.txt
466.txt
3. Validate response by extracting KG entities, linking similar sources, & suggesting related documents to read
NEW PILE
Notes
NEW BOOKMARK
Important entities
Last updated: 11/4/2024, 10:19:05 AM
Key words and phrases
Kronos GAStech Elodis POK
Text
- Administrative assistant at GAStech, Miriam Avila
- GAStech: 5 million cubic meters of gas are produced daily
- disappearance of 14 GAStech employees in Kronos
Documents
140.txt
466.txt
138.txt
Edvard Vann?
Last updated: 11/4/2024, 10:19:58 AM
Key words and phrases
Guard
Criminal activities
Text
Where was Edvard Vann the night of the disappearance?
466.txt

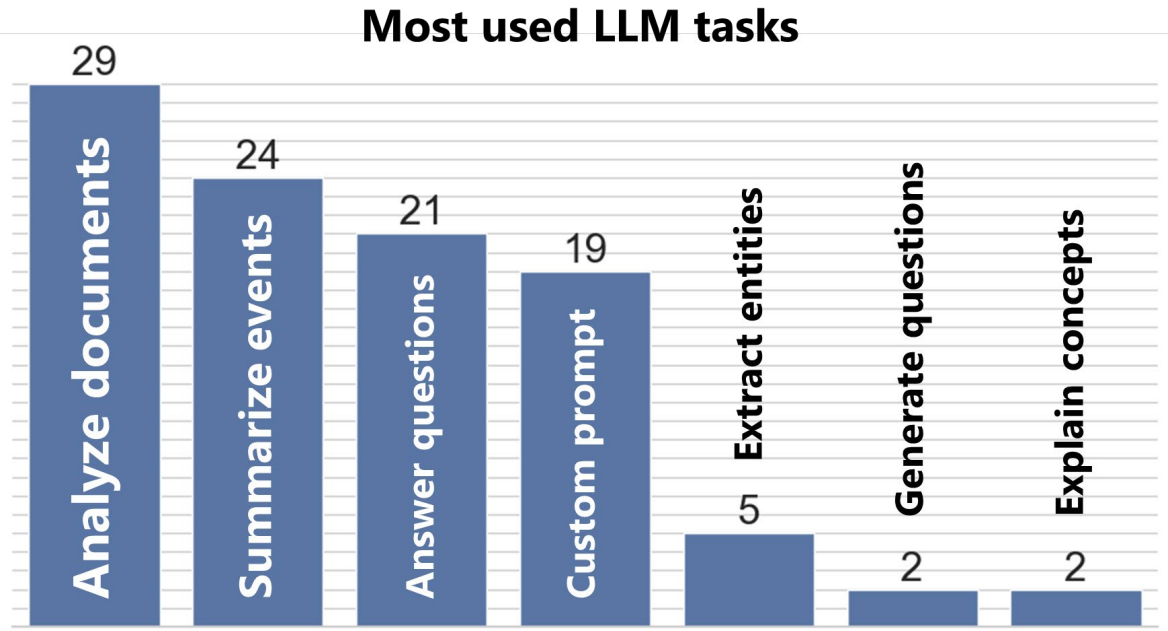
Evaluation | User study

Study design

- N: 17 participants | Task: identify relationships across 845 text documents
- Measures: clicks + time spent + think-aloud feedback | Analysis: mixed-methods

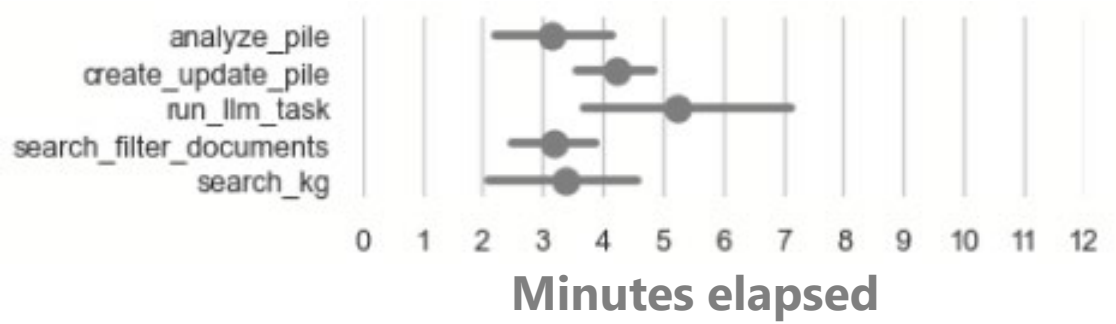
Interactions

- 1 – 7 piles created (**2-3 piles on average**), ranging in size from 1 to 64 documents, with most b/w **10 & 16 docs in a pile**
- Many participants stuck with **one task throughout**, a few combined them
- Many cited the LLM results "**verbatim**" without fact-checking!

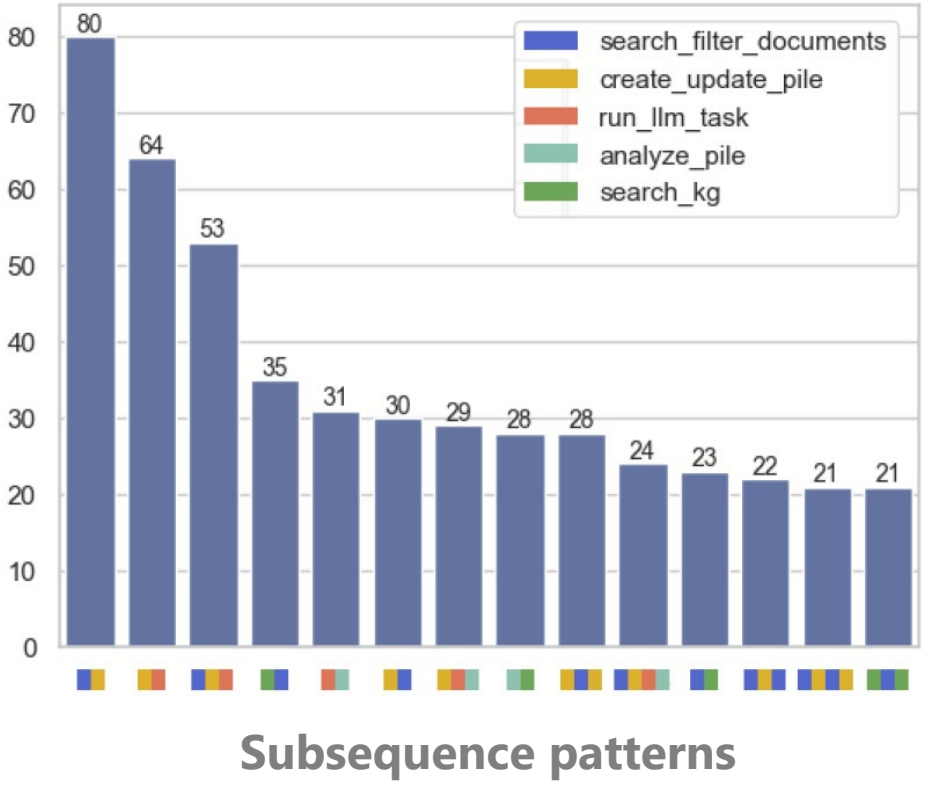


Evaluation | User study

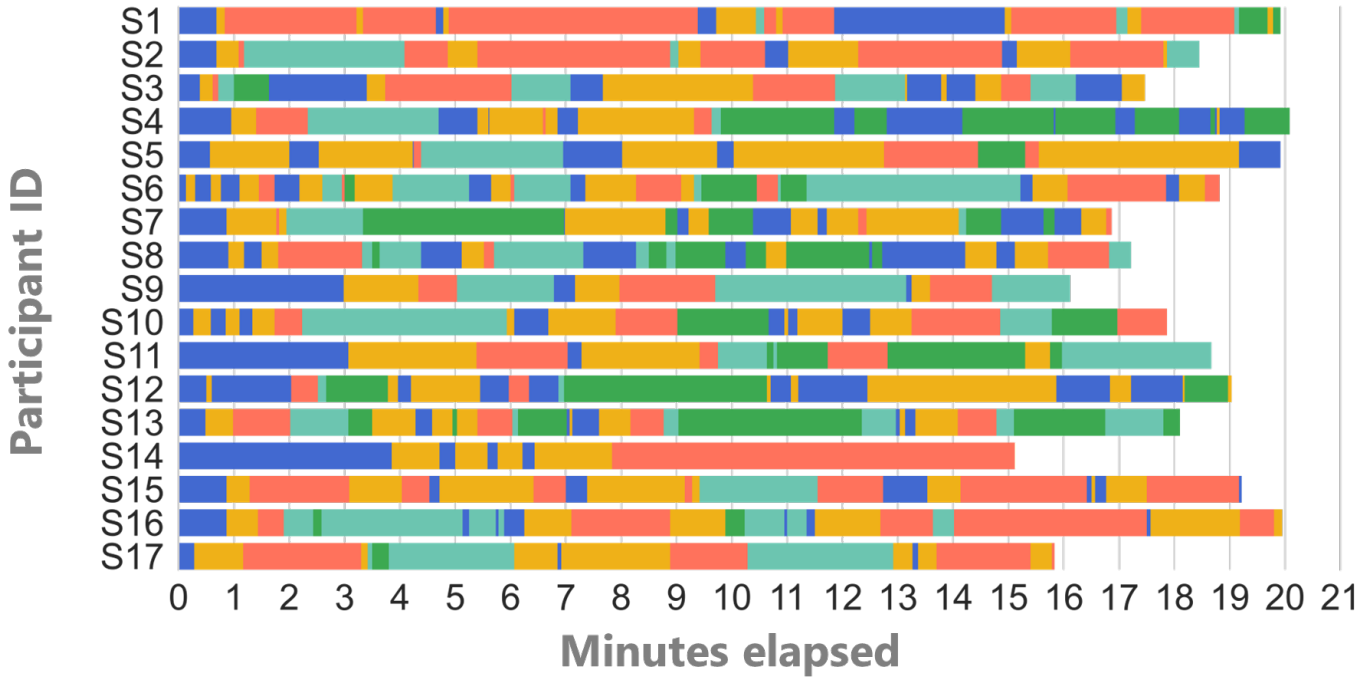
95% confidence interval of time spent



Common activity sequences



Time spent searching docs / piling docs / querying LLM tasks / validating LLM response / searching KG



Evaluation | User study

Analysis process

- Several participants "**reused**" piles, performed follow-up prompts to (1) get more information or (2) "try again" to beat **hallucinations**
- A mix of **validation strategies** – (1) **Link** sentences to source documents; (2) **Extract** KG entities to confirm fact exists in other docs; (3) **Suggest** documents to complete search

Trusting and understanding AI

- Tension between analysis and validation – some participants "**blindly trusted**" LLM results, others spent additional time "**vetting**" LLM & KG before even starting analysis
- Interface encouraged validation in some places; in others it **wasn't clear** if it was needed
- Some preferred bigger piles to "**catch everything**"; some preferred smaller to ensure they "**know where the results are coming from**"

Evaluation | Expert analyst feedback

We asked 6 LAS analysts to give formative feedback after a 1-hour session using VisPile to analyze documents

- Feedback was overall positive! ❤️
- **Summarization** was by far most popular task; most analysts "trusted it". Those that knew what semantic search was liked the "**RAG**" approach…
  - Every analyst requested additional **validation** checks on LLM- and KG-generated results
- Some felt the LLM could do a better job at finding information, while others were inherently distrustful. Most did not want LLM to "**synthesize**", only "**search**" & "**extract**"
  - Similar split in preferred pile size; some wanted to use LLMs on **large** piles to narrow search, while others preferred to search **first** before involving LLMs
- More **distrust** of **KG** overall; hard to trace **source** of KG facts + unclear what **relationship** between document piles and KG was…

Discussion | Revisiting our questions

Guidelines for putting LLMs & KGs into visual text analysis

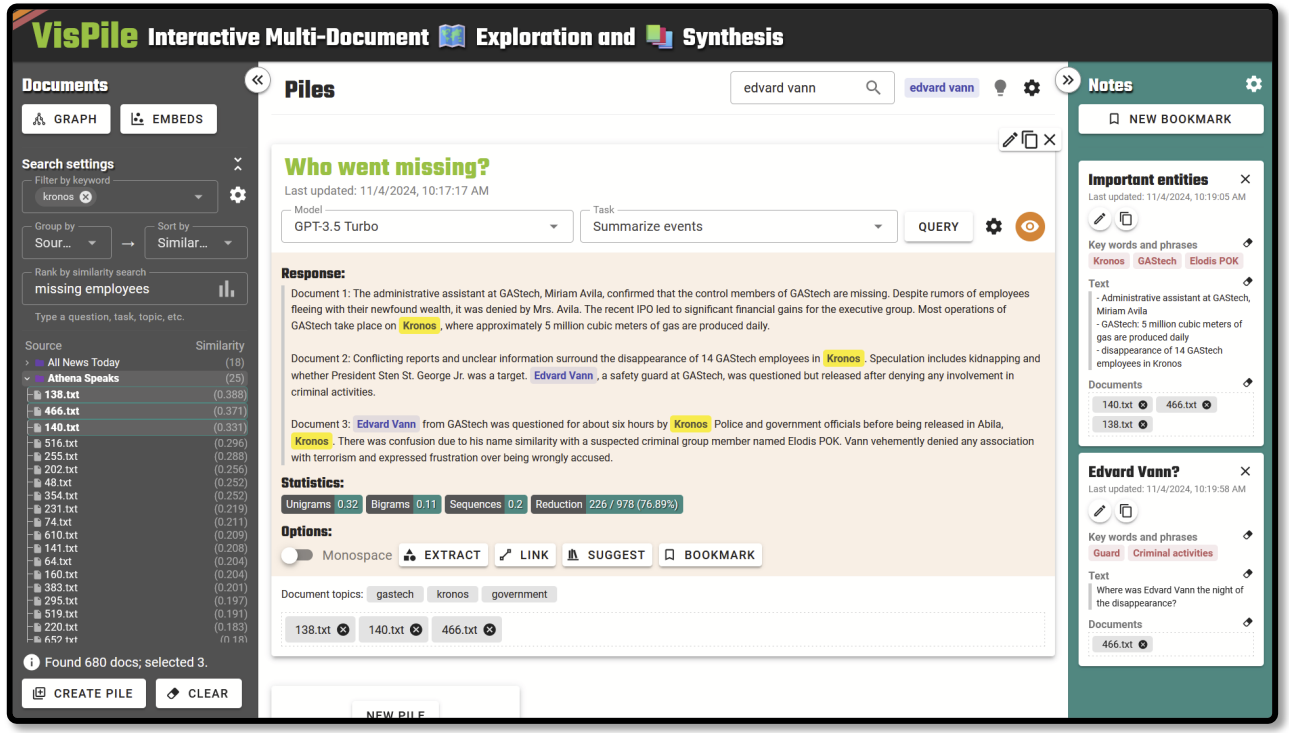
1. Where to put LLMs & KGs in **analysis tools?**
   - Support a few tasks with **tight integration** b/w LLM & KG results
   - Let analysts **combine** tasks and edit prompts to encourage **exploration**
   - Run validation **automatically** without user input to improve **confidence**
2. How will LLMs & KGs affect **sensemaking?**
   - Make it obvious when/where users should **check their results**
   - Prioritize allowing **follow-up operations** to support lines of inquiry

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

Try our demo!

https://adamcoscia.com/papers/vispile/

Laboratory for Analytic Sciences
2024 Research Symposium



Adam **Coscia,** Alex **Endert**
Georgia Institute of Technology, Atlanta, GA

Liz **Richerson,** Sue Mi **K.,** Stephen **S.,** Tim **S.**
Laboratory of Analytic Sciences, Raleigh, NC

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.