

# Preventing the Next Enron:

Using Graph-Based RAG on Enterprise  
Data for Fraud Detection, Compliance,  
and Operational Efficiency

BUSN 30135 AI and Financial Information  
Winter 2025 - Final Project

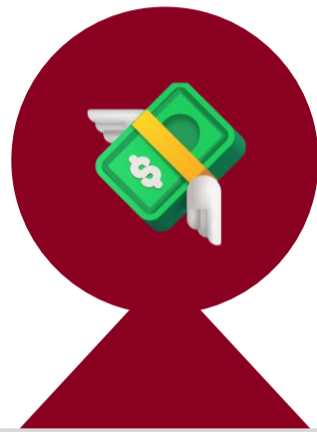


The University of Chicago Booth School of Business

Michael Mott  
Danny Elfman  
Diego Sinay  
Santiago Etchepare  
Matias Zamorano

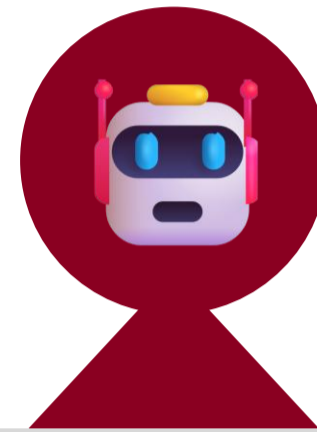
# Motivation & Solution: Our Method

## Why does this matter?



### Massive Economic Impact

Corporate fraud and compliance failures cost businesses over \$3.7 trillion annually<sup>1</sup>



### Traditional RAG is Inadequate

Conventional NLP models miss relational fraud patterns and fail to connect cross-document insights



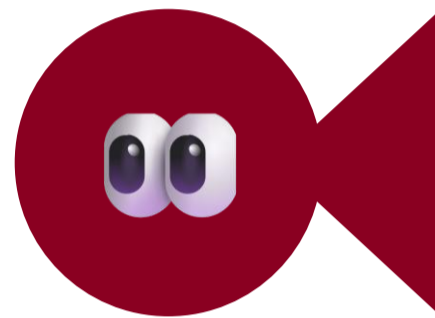
### Workflow based on GraphRAG Bridges the Gap:

By integrating LLMs with graph-based reasoning, we enhance detection, compliance monitoring, and operational efficiency



# Data Sample: Enron's Emails from '98 to '00

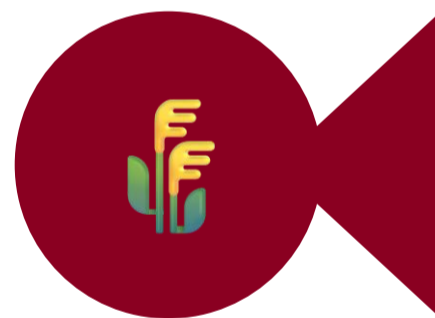
## Sample Emails



From: [dcarter@allmort.com](mailto:dcarter@allmort.com)  
To: [bill.williams@enron.com](mailto:bill.williams@enron.com)  
Subject:  
Cc:  
Bcc:  
Date: Mon, 26 Nov 2001 16:23:24 -0800 (PST)

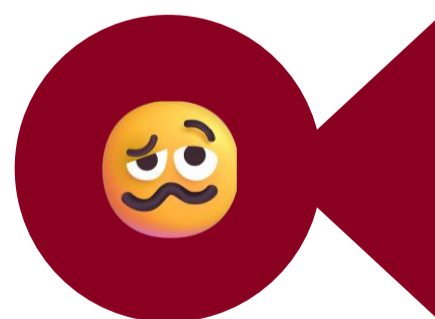
So . . . you were looking for a one night stand afterall . . . ??

DC



From: [richard.sanders@enron.com](mailto:richard.sanders@enron.com)  
To: [barton.clark@enron.com](mailto:barton.clark@enron.com)  
Subject: Re: Dark Star  
Cc:  
Bcc:  
Date: Thu, 10 Jun 1999 02:28:00 -0700 (PDT)

After further thought, it seems to me that in light of our fear of litigation w/ American Coal, we should keep the documents. To further insulate the Coal Group and you from any claim that Enron misused the information,I suggest that you transfer the information to me and I will hold it for safekeeping.

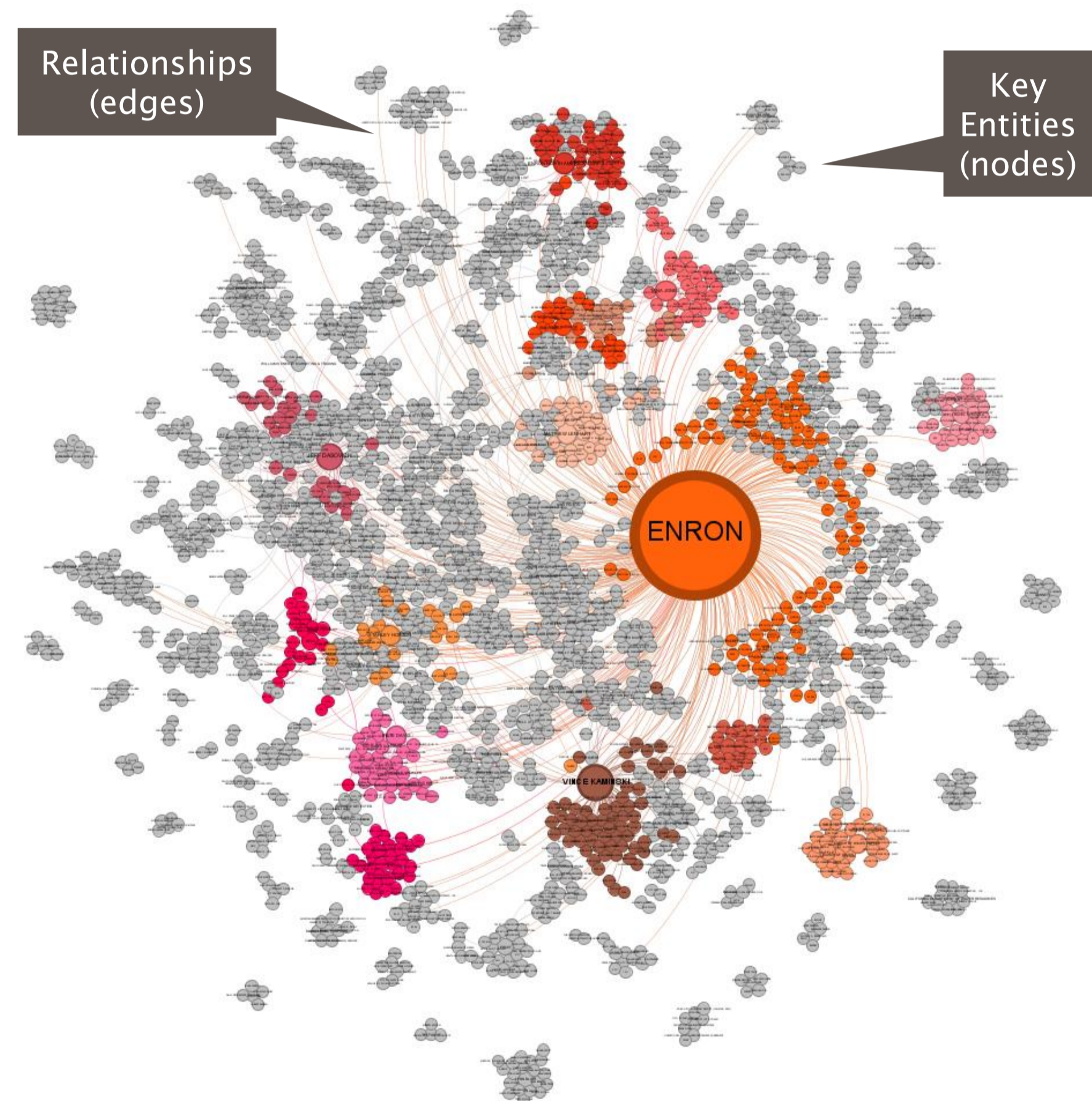


From: [chris.dorland@enron.com](mailto:chris.dorland@enron.com)  
To: [mmolloy@oebi.com](mailto:mmolloy@oebi.com)  
Subject: Re:  
Cc:  
Bcc:  
Date: Mon, 14 Aug 2000 07:16:00 -0700 (PDT)

The wedding was great! Too much fun. Once again I'm hung over on a monday. No great stories except that I nearly wiped out while walking down the aisle. I tripped on a cord to a video camera. Mildly embarrassing. My friend was a rock. I think I was more nervous than him. Ft. Mac is a total shit hole. On my tombstone it wil read " He went to Ft. MacMurray, once." No Houston for me. Thank god. I have to go to a party at my boss's on Saturday. I think I might try to play golf as well. I will be taking it very easy. No more booze. My liver hurts.

Chris

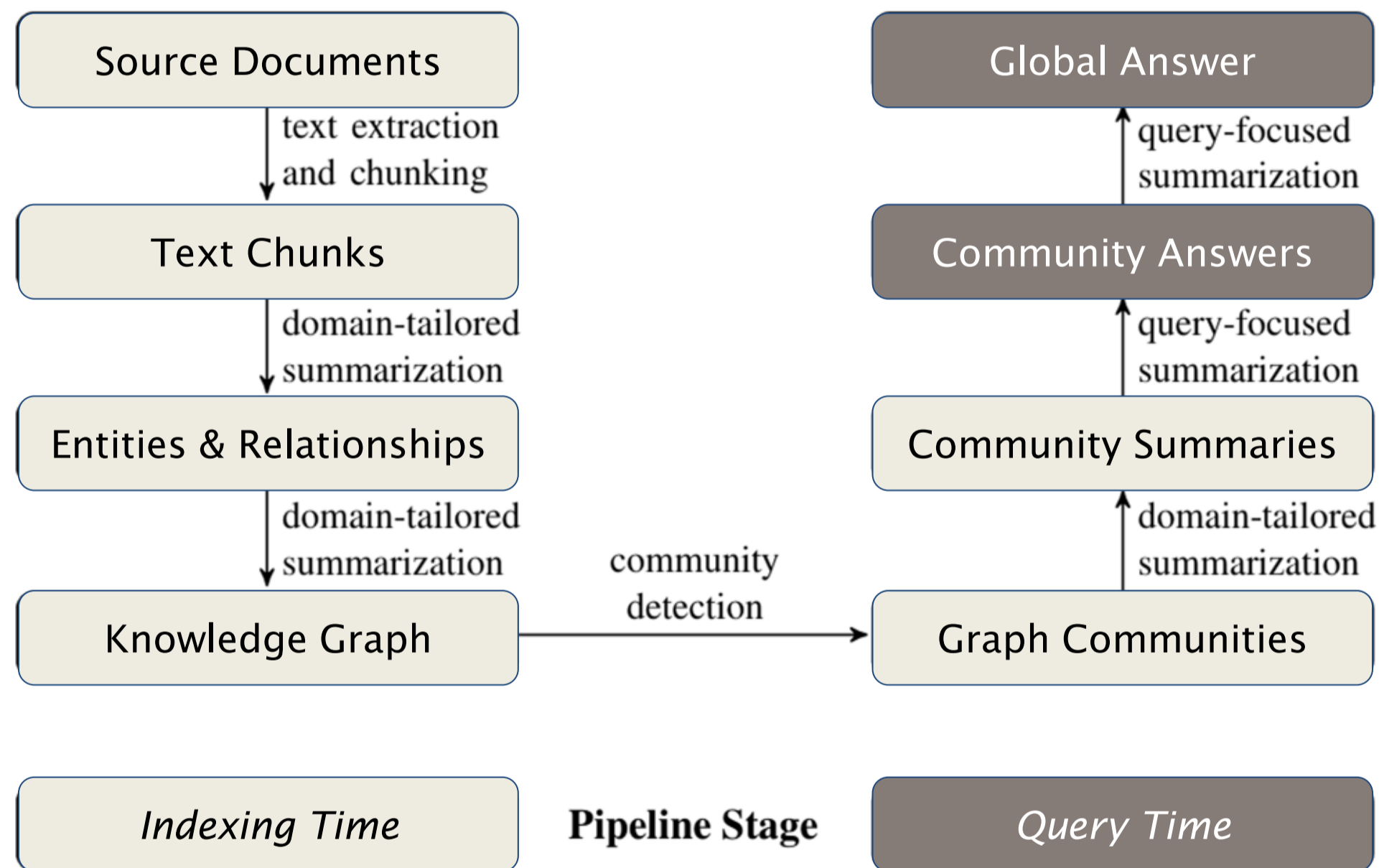
## Knowledge Graph



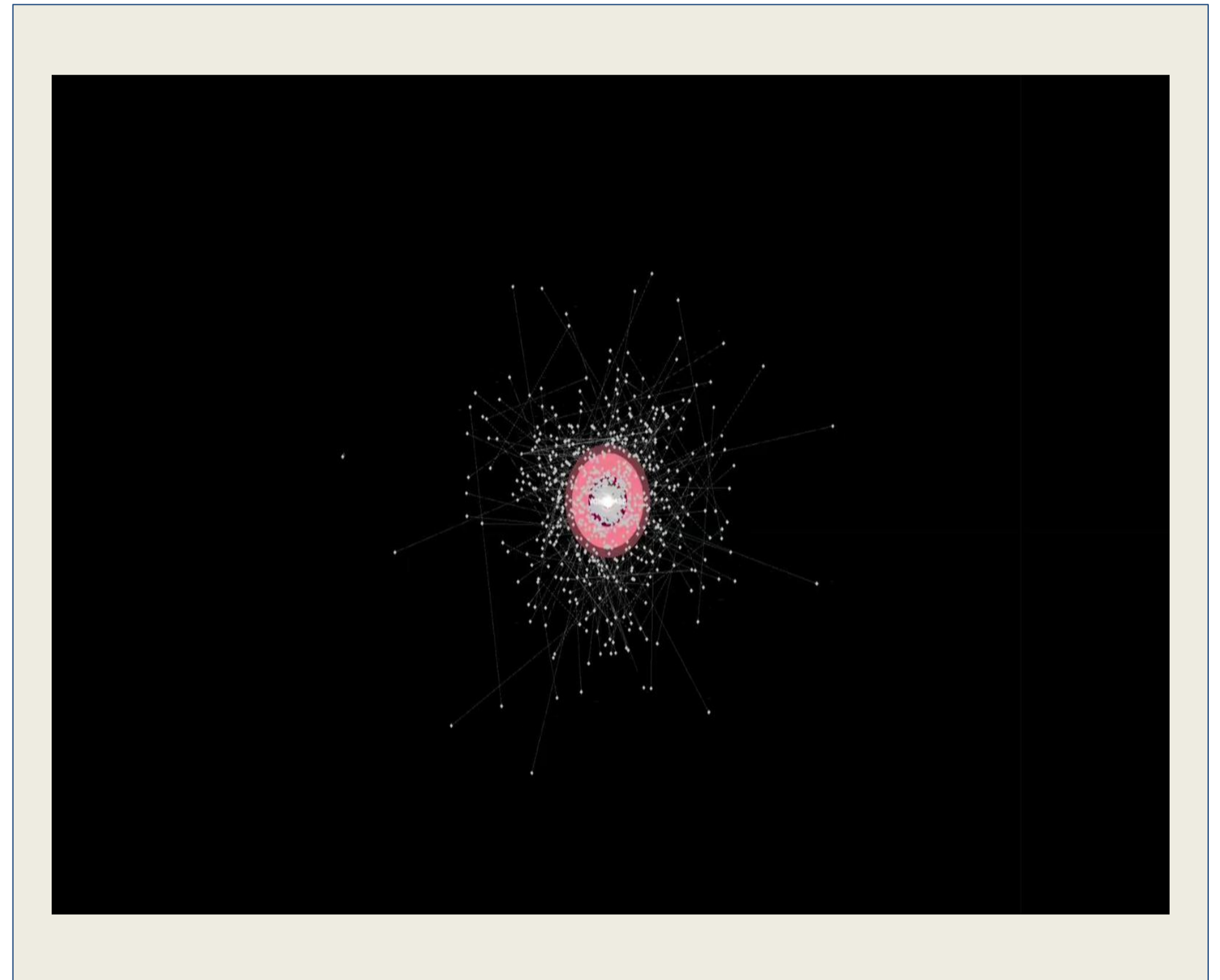


# Workflow: Tuning Prompts to Ensure Domain-Specific Answer

## Workflow Visualization

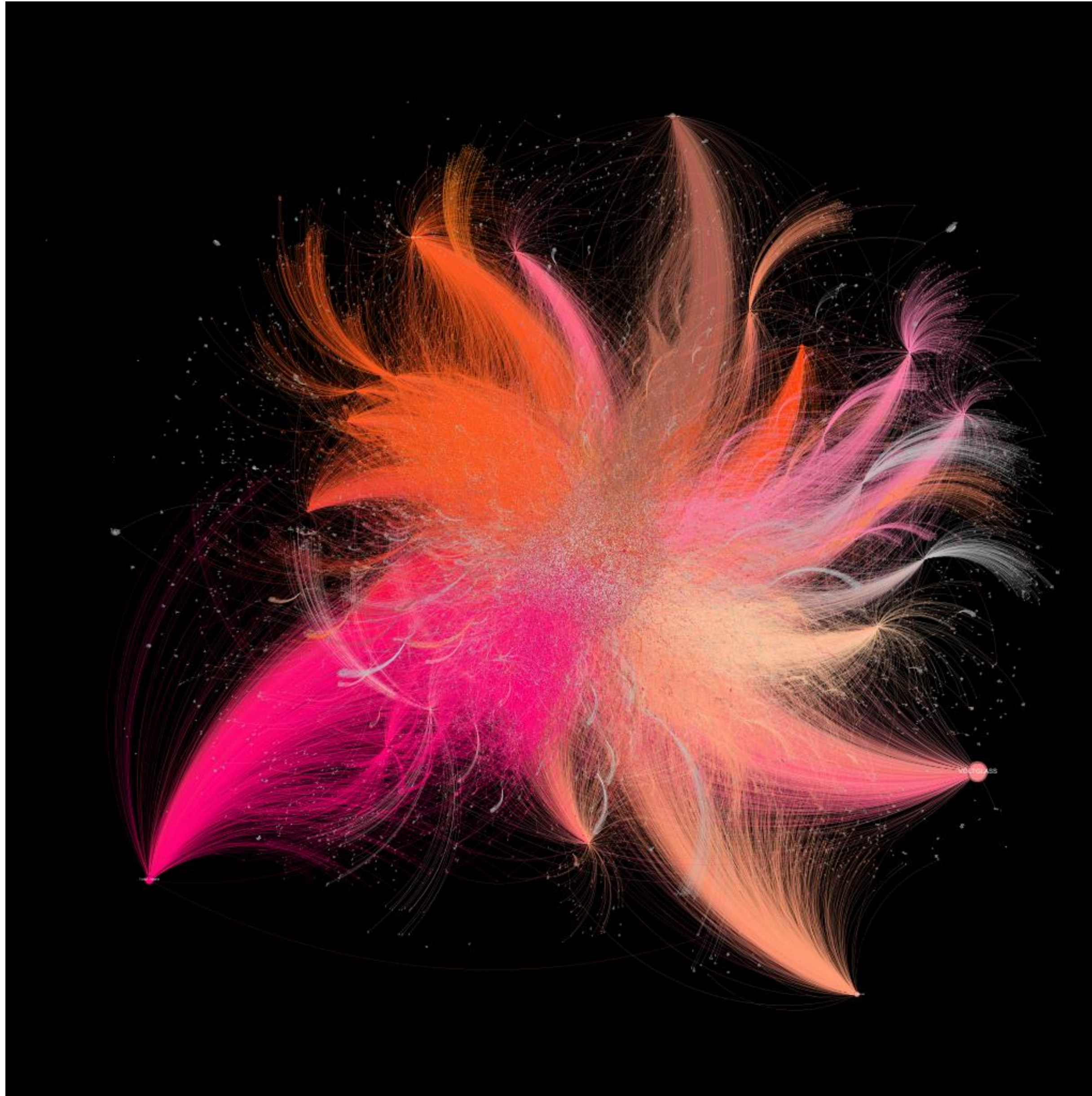


## Knowledge Graph Visualization





# Graphical Representation of Extracted Entities and Relationships



62,000 nodes and 125,000 edges



# Comparative Advantages: Example 1

**Query: Cite specific instances of employee discussions regarding the manipulation of financial indicators to present a more favorable view of the organization's financial health**

Traditional  
RAG

Based on the provided email content, **it is not possible to cite specific instances of employee discussions regarding the manipulation of financial indicators.** The emails appear to be related to a legal issue (gouging) and a discussion about market power law, but there is no mention of employee discussions or manipulation of financial indicators.

GraphRAG

There are several documented instances of employee discussions regarding the manipulation of financial indicators at Voltglass, particularly focusing on presenting a more favorable view of the organization's financial health. These discussions highlight concerns about financial practices and the integrity of financial reporting.

**\*\*Accounting Considerations in Deal Structures\*\***

An email from Bob Chandler discusses **potential deal structures aimed at moving income from one fiscal year to another to take advantage of favorable pricing.** This indicates a strategic approach to financial reporting that could manipulate the appearance of financial health by **shifting income across reporting periods** [Data: Sources (62759)].

...

These instances collectively illustrate a culture of awareness and concern regarding the manipulation of financial indicators within Voltglass, reflecting the complexities and ethical considerations involved in financial reporting practices.

## Comparative Advantages: Example 2

**Query: Are there any instances of employees potentially engaging in manipulation of financial metrics, specifically VAR manipulation?**

Traditional  
RAG

Based on the provided email content, there is **no direct evidence of employees engaging in manipulation of financial metrics**, specifically VAR (Value-at-Risk) manipulation. However, several red flags and indicators suggest that there may be potential issues with how companies are handling risk management and disclosing information to investors.

GraphRAG

The data provided indicates several instances where employees at Voltglass have been involved in discussions and actions that suggest potential manipulation or concerns regarding the financial metric known as Value at Risk (VAR). Below, I summarize the relevant findings from the reports, entities, and relationships that highlight these instances.

**\*\*Concerns Over VAR Calculations\*\***: There are multiple communications indicating that **employees are aware of the potential discrepancies** in VAR calculations. For example, discussions have taken place regarding the **need for a "band-aid" to the term VAR curve** to address issues before they lead to significant discrepancies in reported metrics. [Data: Sources (51020, 90147)]

...

The data suggests that there are multiple instances where employees at Voltglass have engaged in discussions that raise concerns about the manipulation of VAR.

# Real-World Applications

## Example use cases:



**Financial Sector (AML & Fraud):** Detects collusion by linking suspicious communication threads to specific transactions. Flags entities involved in anomalous multi-hop relationships (e.g., **account X → shell company → external partner**)

**Corporate Compliance:** Monitors communications for policy breaches (e.g., insider trading hints). Simplifies audits with transparent, explainable insights backed by a traceable knowledge graph



**Operations and Workflow Optimization:** Identifies communication bottlenecks between teams or departments. Recommends process improvements by analyzing cross-functional email patterns



# Preventing the Next Enron:

Using Graph-Based RAG on Enterprise  
Data for Fraud Detection, Compliance,  
and Operational Efficiency

BUSN 30135 AI and Financial Information  
Winter 2025 - Final Project



The University of Chicago Booth School of Business

Michael Mott  
Danny Elfman  
Diego Sinay  
Santiago Etchepare  
Matias Zamorano

# GraphRAG: AI-Driven Insights for Fraud Detection, Compliance, and Operational Efficiency

University of Chicago Booth School of Business

<sup>1</sup>March 2025 - Prototype Trained on Enron Emails Dataset

## Motivation & Solution: Our Product

### Why does this matter?

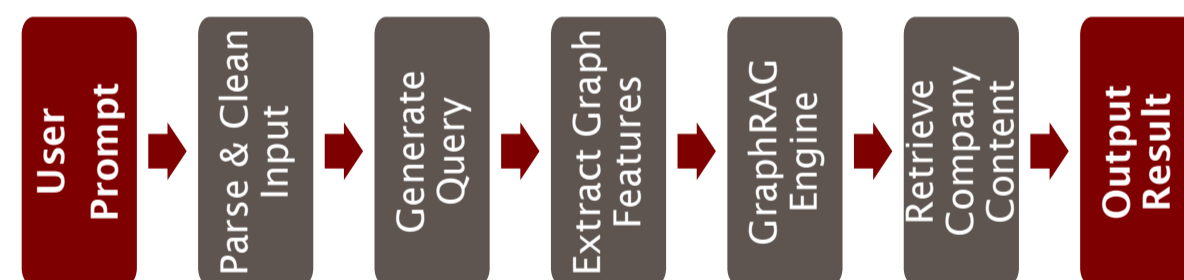
- Massive Economic Impact:** Corporate fraud and compliance failures cost businesses over \$3.7 trillion annually<sup>1</sup>
- Traditional AI is Inadequate:** Conventional NLP models miss relational fraud patterns and fail to connect cross-document insights
- GraphRAG Bridges the Gap:** By integrating LLMs with graph-based reasoning, we enhance detection, compliance monitoring, and operational efficiency

## Example

### User UI

- User:** "Who are the workers prone to illicit behavior?"
- GraphRAG:** "Bradford, Alec, Joseph as they have XYZ"

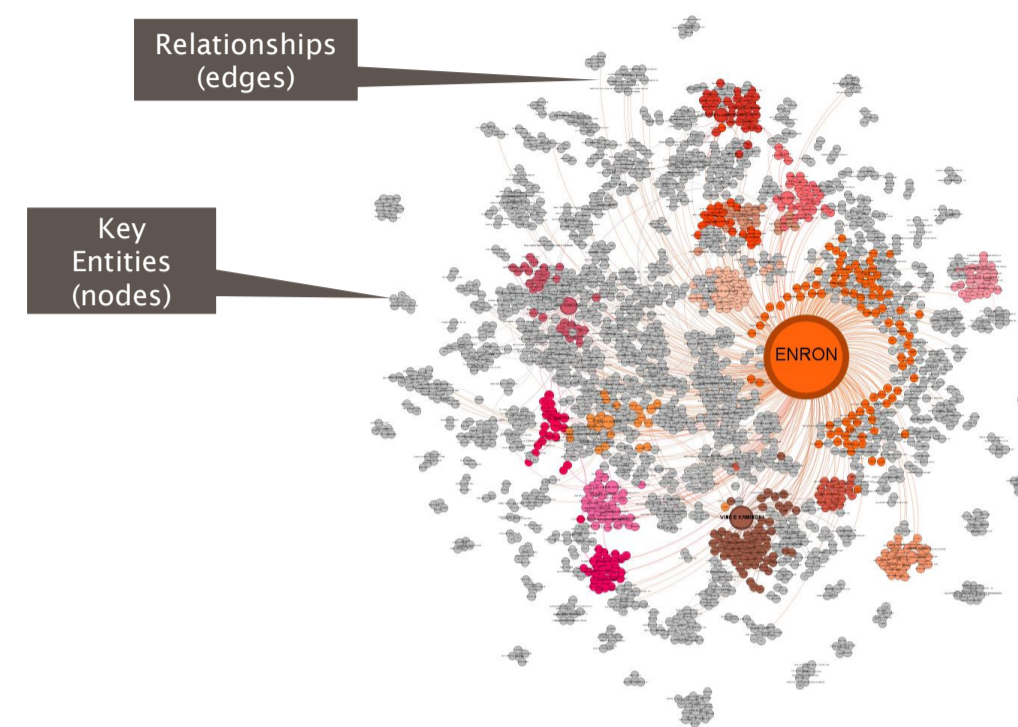
### Internal Process



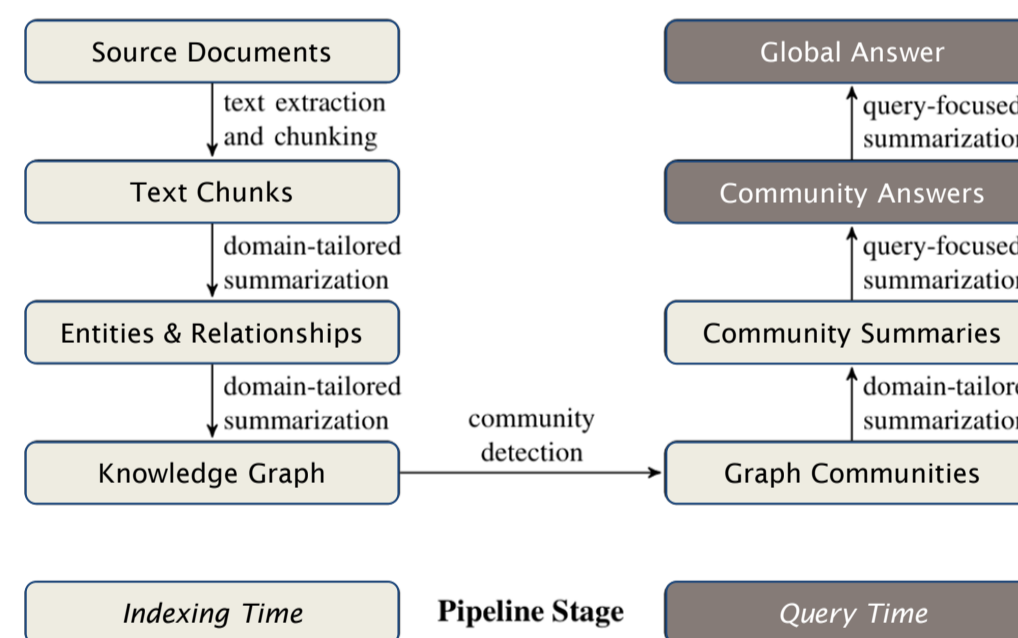
## Real-World Applications

- Financial Sector (AML & Fraud):** Detects collusion by linking suspicious communication threads to specific transactions. Flags entities involved in anomalous multi-hop relationships (e.g., account X → shell company → external partner)
- Corporate Compliance:** Monitors communications for policy breaches (e.g., insider trading hints). Simplifies audits with transparent, explainable insights backed by a traceable knowledge graph
- Operations and Workflow Optimization:** Identifies communication bottlenecks between teams or departments. Recommends process improvements by analyzing cross-functional email patterns

## Data Sample



## Workflow Visualization



## Theoretical Framework

**GraphRAG builds a knowledge graph, identifying key entities and their relationships from the corpus**

- Partitions knowledge graph** into hierarchical communities (sub-clusters) of related entities and generates bottom-up summaries for each community
- Recursively integrates lower-level community summaries into higher-level overviews**, providing comprehensive insights into the corpus
- Queries are answered via map-reduce** — each community summary generates partial answers (map) that are combined into a global response (reduce)

## Comparative Advantages

Criteria	Traditional	Our Approach
Processing	Isolated chunks or keywords, little context	Holistic knowledge graph integrates context and relationships across corpus
Queries	Limited to exact matches or top-K similar snippets	Global insight queries and local deep dives via graph traversal
Patterns	Difficult - requires manual correlation of multiple sources.	Automatic multi-hop reasoning connects entities, revealing hidden patterns
Accuracy	Prone to missing context or hallucinating	High accuracy. LLM is grounded by graph facts, yielding comprehensive, factual answers
Explainability	Low - "why was this flagged?" often unclear.	High - graph provides a transparent chain of reasoning for each result
Adaptability	Rigid and slow- new rules/keywords must be added manually	Flexible - Extraction generalizes to new data
Performance	Often siloed results	Demonstrated ~70-80% win rate in answer quality vs. baseline RAG

## Key Features

- Compliance Tracking:** Automatically monitors conversations for policy adherence, highlights regulatory red flags in emails, chats, and documents
- Fraud Detection:** Identifies suspicious patterns and anomalies by mapping entity relationships across large-scale data (e.g., internal logs, financial records)
- Operational Efficiency:** Analyzes communication bottlenecks, optimizing internal processes and cross-team collaboration
- Enriched Analytics:** Graph-based retrieval combined with LLMs produces contextual insights that pure keyword or vector search often misses

## Additional Information

### Security & Privacy

- Data Protection:** End-to-end encryption ensures sensitive info remains in-house.
- Privacy by Design:** Configurable filtering of PII and compliance with major privacy laws (e.g., GDPR)
- Continuous Monitoring:** Automated re-checks of newly ingested data to maintain up-to-date

### Contacts

Michael Mott, Danny Elfman, Diego Sinay, Santiago Etchepare, Matias Zamorano  
**Phone:** +1-800-GRAG-NOW1  
**Press Release:** <https://tinyurl.com/GRAGNOW>



# Appendix A: Entity Extraction Prompt

Example 1:

Entity\_types: PERSON, ROLE

Text:

"Date: 2000-07-20 03:10:00,

From: john.lavorato@acme.com,

To: rob.milnthorp@acme.com, paul.devries@acme.com, mark.davis@acme.com,

Subject: Mime-Version: 1.0,

Body: It sounds like I need to clarify my plans for trading and risk management of power in Ontario.

1. Garrett is in Toronto and should be doing everything possible to try to understand all the ins and outs of the Ontario Power Market including building forward curves. He has dual reporting to Paul Devries (head of the ACME's Toronto operation) and Dana Davies (head of ENA power trading for the Northeast U.S.). Dana has ultimate risk management responsibility for Ontario.

2. My current opinion is that we should trade Ontario out of the Toronto office with risk management responsibilities rolling into Dana. This could always change in the future if it doesn't make sense.

3. John McKay, although he has been helping Bill out, will ultimately have no power trading responsibilities.

4. Tom May is the cash trader in N.Y. reporting to Dana. I expect us to use Tom's expertise with respect to Ontario to build up our competency there, however, Tom's main role will be N.Y.

5. Garrett should spend some time (probably soon) in Houston getting to know the people and the technical resources available to him. If anyone has any questions concerns please give me a call. John"

#####

Output:

("entity"{tuple\_delimiter}JOHN LAVORATO{tuple\_delimiter}PERSON{tuple\_delimiter}John Lavorato is clarifying trading and risk management plans for Ontario)

{record\_delimiter}

("entity"{tuple\_delimiter}GARRETT{tuple\_delimiter}PERSON{tuple\_delimiter}Garrett works in Toronto, focusing on Ontario Power Market details and forward curves)

{record\_delimiter}

("entity"{tuple\_delimiter}PAUL DEVRIES{tuple\_delimiter}PERSON{tuple\_delimiter}Paul Devries helps oversee Garrett in Toronto)

{record\_delimiter}

("entity"{tuple\_delimiter}HEAD OF ACME'S TORONTO OPERATION{tuple\_delimiter}ROLE{tuple\_delimiter}Role held by Paul Devries)

{record\_delimiter}

{completion\_delimiter}

# Appendix B: Community Report Generation Prompt

# Example Input

-----

SOURCES

id, text

- 1, Text: From: compliance.office@acme.com To: management.team@acme.com Cc: legal.team@acme.com, risk@acme.com Date: Wed, 12 Jul 2000 08:30:00 -0600 (CST) Subject: Quick Update on Compliance & Risk Efforts
- 2, Quick update on what's been cooking in Compliance and Risk Management. Risk Management is stepping up — They've been tightening up on our financial risk assessments and mitigation strategies since early this year.
- 3, Their efforts are key to keeping us on solid ground financially and in compliance with the latest market regulations as of mid-2000. It's crucial for our strategic planning and helps us stay ahead.
- 5, Legal's keeping us in check — The Legal Compliance team is on top of ensuring all our operations are up to scratch with legal standards. They're especially focused on improving our corporate governance and contract management as of the second quarter of 2000. This is critical for keeping our operations smooth and legally sound.
- 9, Working together — Risk Management and Legal Compliance have been syncing up better than ever since the start of Q2 2000. They're making sure our strategies are not just effective but also fully compliant. This coordination is essential for our integrated governance approach.
- 10, Your thoughts? — How do these updates impact your area? Got ideas on how we can do better? Give your department heads a shout.
- 11, Thanks for staying engaged. Let's keep pushing for better and smarter ways to work. Cheers, Jane Doe

Output:

```
{{
  {{
    "summary": "Risk Management Operational Scope",
    "explanation": "The Risk Management Department at AMCE plays a pivotal role in identifying, assessing, and mitigating financial risks. Their proactive approach, highlighted from the beginning of 2000, helps safeguard ACME against potential financial pitfalls and ensures continuous compliance with evolving market regulations. Effective risk management not only prevents financial anomalies but also supports the company's strategic decision-making processes.
```

```
[Data: Sources (2, 3), Date_Range ((2000, 01, 01), (2000, 07, 12))]"
  }},
```



# Appendix C: Extracted Entity and Relationship Example Output

```
{"result": {"id": "chatcpl-BA3i01MWb2WyRodhVnrGjJ7bFdETu", "choices": [{"finish_reason": "stop", "index": 0, "logprobs": null, "message": {"content": "(\n\"entity\"\n<|>JAIME WILLIAMS\n<|>PERSON\n<|>Jaime Williams is sending daily historic consumption data to John Griffith)\n##\n(\n\"entity\"\n<|>JOHN GRIFFITH\n<|>PERSON\n<|>John Griffith is the recipient of the daily historic consumption data from Jaime Williams)\n##\n(\n\"entity\"\n<|>COPAMEX\n<|>ORGANIZATION\n<|>Copamex is the entity for which the daily historic consumption data is being reported)\n##\n(\n\"entity\"\n<|>DAILY HISTORIC CONSUMPTION\n<|>FINANCIALMETRIC\n<|>Refers to the daily data on gas consumption for Copamex)\n##\n(\n\"relationship\"\n<|>JAIME WILLIAMS\n<|>JOHN GRIFFITH\n<|>Jaime Williams is sending consumption data to John Griffith\n<|>7)\n##\n(\n\"relationship\"\n<|>DAILY HISTORIC CONSUMPTION\n<|>COPAMEX\n<|>The daily historic consumption data pertains to Copamex\n<|>6)\n##\n<|COMPLETE|>"},
```

.....

```
n-Real Data-\n#####\nEntity_types:\nPERSON,FINANCIALMETRIC,ROLE,ORGANIZATION,PROJECT\nText: Date: 2000-12-08 01:36:00 From:\njaime.williams@Voltglass.com To: john.griffith@Voltglass.com Subject: CONSUMOS GAS Body:\nJohn, I am sending you the daily historic consumption of Copamex....I think this is the\nsame information as before. Thanks and regards.\n#####\n
```