Enhancing Access to Legacy of Slavery Records using Generative Al

Rajesh Kumar Gnanasekaran

Lori Perine

Richard Marciano

Mark Conrad





Background on MSA's LoS

Title: Why focus on Maryland State Archives' Legacy of Slavery (LoS)?

- The Maryland State Archives (MSA) undertook a multi-year effort with about 100 staff since 2001 to digitize historical records from the LoS project. Project to discover stories of unknown 'heroes' of slave flight and resistance.
- This effort resulted in a large, digital database containing over 400,000 records, including Certificates of Freedom (CoF), Domestic Traffic Ads (DTA), and Manumissions.
- Despite this significant achievement, the actual usage and engagement with the dataset remain low.





Leveraging Generative AI & Knowledge Graphs

"ChatLoS" using Generative AI & Knowledge Graph-RAG Framework

- Chatbot ("ChatLoS") that pulls context from the *Domestic Traffic Ads (DTA)*dataset and provides *context-aware* responses
- Demonstrates interactive natural language conversational capabilities
- Knowledge Graphs (KGs) serve the purpose of getting the data "Al-Ready" by providing structured source of truth context (through nodes, relationships)
- Study included three datasets from LoS together (CoF, DTA, Manumissions) for cross-dataset insights.
- Also includes a systematic user evaluation grounded on Activity Theory by involving two separate IRB approved user studies with well experienced MSA's experts
- MARY AND tions only looked at using one dataset in a context-aware marinely and with non-Al ready data

Leveraging Generative AI & Knowledge Graphs

"ChatLoS" using Generative AI & Knowledge Graph-RAG Framework

- Also includes a systematic user evaluation grounded on Activity Theory by involving two separate IRB approved user studies with well experienced MSA's experts
- Prior solutions only looked at using one dataset in a context-aware manner and with non-Al ready data
- User study findings were positive with room for improvement.





How KG-RAG works with ChatLoS?

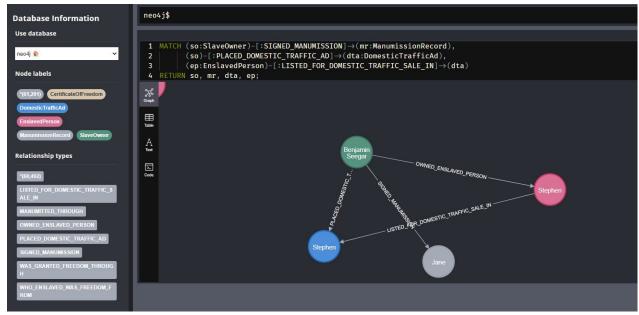
- Knowledge Graphs (KGs) provide structured source of truth context (through nodes, relationships)
 - Nodes and relationships built by connecting multiple datasets together by integrating CoF, DTA, Manumissions for cross-dataset insights.
- KG-RAG connects the KG and GPT to create GraphQL or Cypher Queries dynamically and provide that to the GPT to generate natural language responses, similar to simple-RAG but more powerful
- Used Neo4J desktop to import datasets and create nodes and relationships





Node Label		
EnslavedPerson		
SlaveOwner		
ManumissionRecord		
CertificateOfFreedom		
DomesticTrafficAd		

Relationship Type	Source Node	Target Node
OWNED_ENSLAVED_PERSON	SlaveOwner	EnslavedPerson
PLACED_DOMESTIC_TRAFFIC_AD	SlaveOwner	DomesticTrafficAd
LISTED_FOR_DOMESTIC_TRAFFIC_SALE_IN	EnslavedPerson	DomesticTrafficAd
SIGNED_MANUMISSION	SlaveOwner	ManumissionRecord
MANUMITTED_THROUGH	EnslavedPerson	ManumissionRecord
WAS_GRANTED_FREEDOM_THROUGH	EnslavedPerson	CertificateOfFreedom
WHO_ENSLAVED_WAS_FREEDOM_FROM	CertificateOfFreedom	SlaveOwner



6

Cross-dataset querying with DTA, Manumissions, and Certificates of Freedom

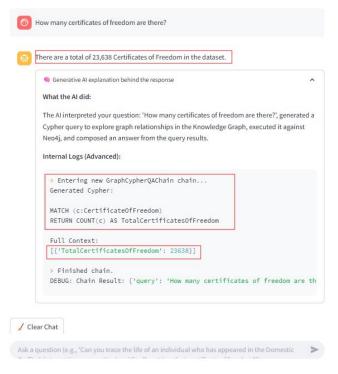




✓ Clear Chat

Figure 37: ChatLoS v3's answer showing "explainable" response for an aggregate multi-attribute filter question

ChatLoS enhanced with KG-RAG using Neo4j and Generative AI— Cross-dataset querying with DTA, Manumissions, and Certificates of Freedom



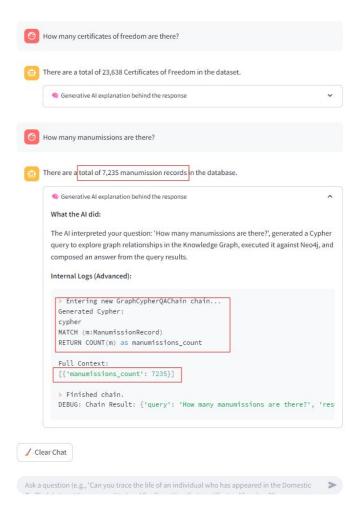
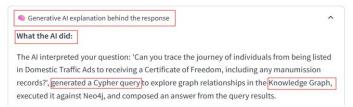


Figure 38, 39: ChatLoS v3's KG RAG-based answers responding for multi-dataset aggregate questions

ChatLoS enhanced with KG-RAG using Neo4j and Generative AI— Cross-dataset querying with DTA, Manumissions, and Certificates of Freedom

- Can you trace the journey of individuals from being listed in Domestic Traffic Ads to receiving a Certificate of Freedom, including any manumission records?
- Yes, based on the available data, we can trace the journey of an individual named John Howard from being listed in Domestic Traffic Ads to receiving his Certificate of Freedom.
 - Domestic Traffic Ads: John Howard was listed for sale in a Domestic Traffic Ad on August 15, 1829. You can view the original ad through the following source: 1829 Domestic Traffic Ad for John Howard.
 - Manumission Records: Approximately a year later, his Manumission Record shows that he was granted freedom on September 28, 1830.
 - Certificate of Freedom: Subsequently, he was officially recognized as a free individual through his Certificate of Freedom dated October 25, 1830.

This progression strongly suggests that John Howard transitioned from being enslaved, being listed for sale, gaining his freedom, and then receiving official recognization of that freedom. This kind of detailed tracing can further our understanding of the process of liberation during this period.



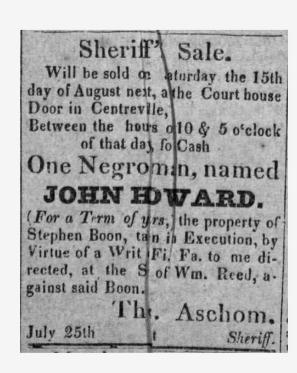
This progression strongly suggests that John Howard transitioned from being enslaved, being listed for sale, gaining his freedom, and then receiving official recognization of that freedom. This kind of detailed tracing can further our understanding of the process of liberation during this period.







Figure 40, 41: ChatLoS v3's KG RAG-based answer for life tracing questions



10

Figure 42: ChatLoS v3's answer source linking to the traceable ad link on MSA's website

Future Work - Advancing KG-RAG with Agentic Al and Community Integration

- Expand KG schema:
 - include skills, occupations, aliases to resolve missing field contradictions
- Implement GPT Actions:
 - enable real-time interactions with MSA databases for live updates
- Use AMCP (Agentic Model Context Protocol):
 - manage multi-step archival workflows across datasets
- Introduce real-time user feedback loop:
 - thumbs up/down, comments, source flagging
- Flipped prompting for clarification:
 - Al prompts users when ambiguity is detected
- Phased rollout with community-led testing: archivists, genealogists, and students





Conclusion





