Harnessing Generative AI to Support Exploration and Discovery in Library and Archival Collections

- **Richard MARCIANO:**
  - Prof @ U. Maryland iSchool & **Director of AIC** (formerly @ UNC Chapel Hill & @ UCSD Supercomputer Center)
- **Rajesh Kumar GNANASEKARAN:**
  - Assistant Director of **AI Solutions**, Division of IT @ U. Maryland
- **Lori PERINE:**
  - Former **Researcher in Trustworthy AI** at the National Institute of Standards and Technology (**NIST**)
  - See: Identifying & Managing Bias in AI & AI Risk Management Framework Playbook
- **Mark CONRAD:**
  - Former **Digital Archives Specialist** with the U.S. National Archives & Director for Technology Initiatives @ NHPRC

AI Advanced Information Collaboratory

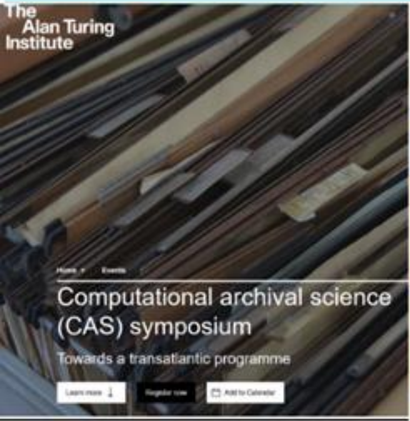
COLLEGE OF INFORMATION STUDIES

LAUNCH OF THE AIC COLLABORATORY

Advanced
Information
Collaboratory

The AIC was launched at The Alan Turing Institute in London, UK on Jan. 20, 2020. It brings together partners from leading academic and cultural institutions from six continents. Its goals are to:

1. **EXPLORE** the opportunities and challenges of "disruptive technologies" for archives and records management (digital curation, machine learning, AI, etc.).
2. **PURSUE** multidisciplinary collaborations to share relevant knowledge across domains.
3. **LEVERAGE** the latest technologies to unlock the hidden information in massive stores of records.
4. **TRAIN** current and future generations of information professionals to think computationally and rapidly adapt new technologies to meet their increasingly large and complex workloads.
5. **PROMOTE** ethical information access and use.



The Alan Turing Institute

Home > Events

Computational archival science (CAS) symposium

Towards a transatlantic programme

AIC Founding Partners:

Dr. Richard Marciano
Professor
UMD iSchool (US)

Mark Conrad
Archives Specialist
NARA - former (US)

Dr. Eirini Goudourali
Head of Dig. Research Progs.
TNA (UK)

Dr. Jane Greenberg
Professor
Dir. Metadata Research Center
Drexel U. (US)

Dr. Mark Hedges
Dep. of Dig. Hum.
King’s College London (UK)

Greg Jansen
Senior Res. Soft. Architect
UMD iSchool (US)

Dr. Michael Kurtz
Asst. Archivist for Rec. Services
NARA - former (US)

Dr. Victoria Lemieux
Assoc. Professor
Blockchain@UBC Cluster
lead (Canada)

Dr. Bill Underwood
Res. Scientist
GTRI Res. Sci (former)
UMD iSchool (US)

Dr. Lyneise Williams
Associate Prof. Art History
Founder VERA Collaborative
UNC Chapel Hill (US)

Advanced Information Collaboratory

North America:

- **MEDiAL Lab @U. Maryland:** Dr. Phil Piety
- **NARA (former):** Bruce Ambacher
- **UCLA:** Dr. Anne Gilliland
- **Kent State U.:** Dr. Karen Gracy
- **U. Missouri:** Dr. Sarah Buchanan
- **Clayton State U.:** Dr. Joshua Kitchens
- **The Smithsonian Institutions (NMAH):** Bob Horton
- **Harvard Library:** Ceilyn Boyd
- **UC Santa Barbara:** Marisol Ramos
- **UC San Diego:** Dr. Andrea Chiba
- **US Holocaust Memorial Museum:** Michael Levy
- **Densho.org:** Geoff Froh
- **Maryland State Archives:** Chris Haley & Maya Davis
- **Spelman College:** Holly Smith
- **Puerto Rican Spring Project:** Marison Ramos, Irmarie Fraticelli, Joel Blanco

South America:

- **U. de Brasilia:** Dr. Cláudio Gottschalg-Duque

UK:

- **Loughborough U.:** Lise Jaillant
- **The Alan Turing Institute:** David Beavan
- **UK TNA:** Pip Wilcox, Mark Bell, Paul Young, Jenny Bunn & Sonia Ranade
- **Oxford U.:** Dr. David De Roure
- **European Holocaust Research Infrastructure:** Dr. Reto Speck

Europe:

- **Hamburg U. Archives:** Francesco Gelati
- **University of Amsterdam:** Dr. Tobias Blanke
- **INESC-ID Portugal:** Dr. Diogo Proença

Africa:

- **U. South Africa:** Dr. Shadrack Katuu

Asia:

- **Central U. of Gujarat (India):** Dr. Bhakti Gala
- **Centre for Dev. of Advanced Computation (India):** Dr. Dinesh Katre
- **Indian Inst. of Management:** Dr. H. Anil Kumar
- **Kyushu U. (Japan):** Dr. Yoichi Tomiura & Dr. Emi Ishita

Australia:

- **U. Canberra:** Dr. Tim Sherratt

CURRENT / RECENT PROJECTS

INFRASTRUCTURE:

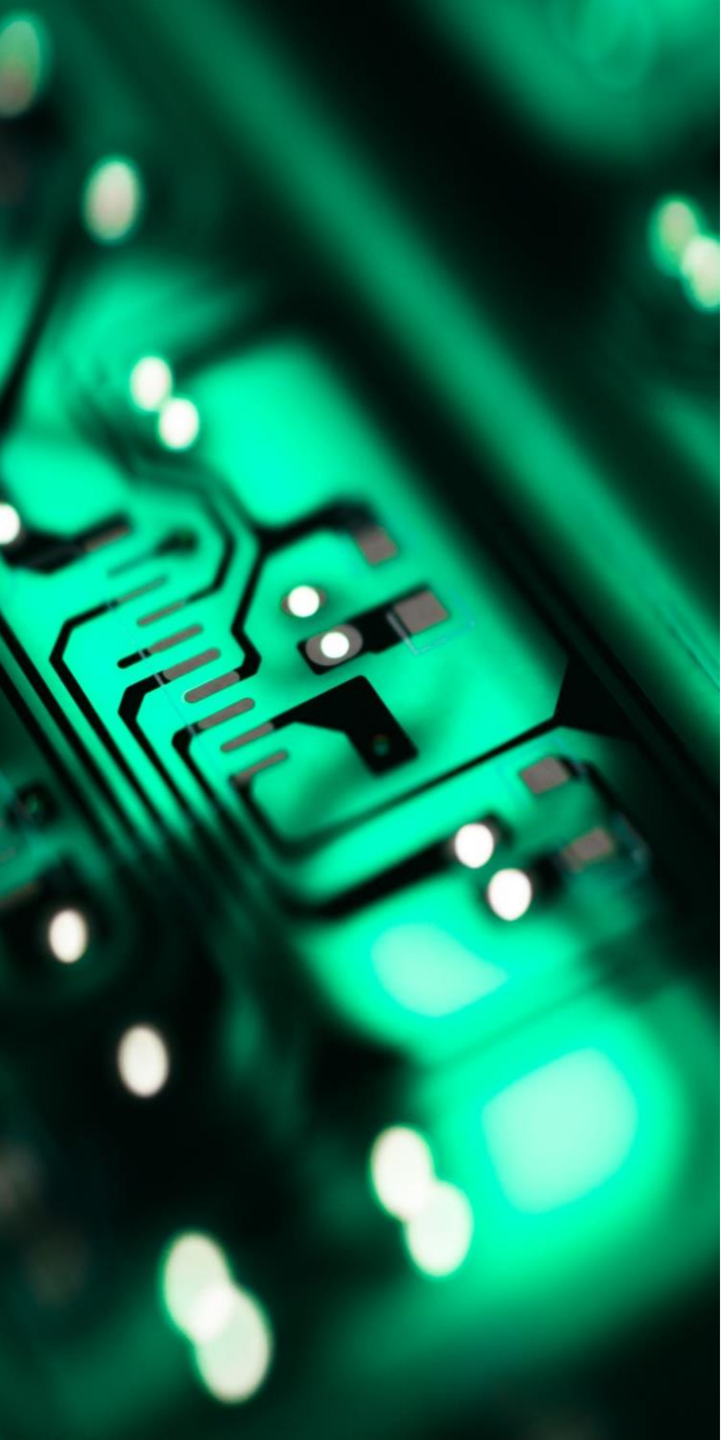
- WIN: a Window Into Neuroregulation *[NSF Convergence: 2019-2024]*
- Developing a Digital Asset Management System for Additive Manufacturing *[ARL: 2020-2025]*
- Developing a Digital Asset Management System for the the Mary McLeod Bethune Historic Site *[NPS: 2019-2022]*
- Improving Fedora 4 to Work with Web-Scale Storage and Services, DRASTIC *[IMLS: 2017-2020]*
- Brown Dog: "Making Sense of Billion-Record Archives" with the NCSA) *[NSF: 2013-2018]*

SOCIAL JUSTICE, HUMAN RIGHTS, CULTURAL HERITAGE:

- Using AI and ML to Optimize Information Discovery in Under-utilized, Holocaust-related Records *[Kurtz Foundation]*
- Harnessing Generative AI to Support Exploration and Discovery in Library and Archival Collections *[IMLS proposal]*
- International Research Portal for Holocaust-Era Cultural Property *[Kurtz Foundation]*
- Measuring the Impact of Urban Renewal *[NSF]*
- Computational Thinking to Unlock the Japanese American WWII Camp Experience *[UMD-FIA]*
- Computational Treatments to re-member the Legacy of Slavery (CT-LOS) *[Kurtz Foundation]*
- Testbed for the Redlining Archives of California's Exclusionary Spaces (T-RACES) *[IMLS]*
- Mapping Inequality – Redlining in New Deal America *[U. Richmond Mellon]*

EDUCATION:

- LIS Education And Data Science Integrated Network Group, LEADING *[IMLS Drexel: 2020-2024]*
- Developing a Computational Framework for Library and Archival Education *[IMLS: 2018-2021]*
- Piloting a Collaborative Network for Integrating CT into Library and Archival Ed. and Practice *[IMLS: 2020-2024]*
- Training of Archival & Library Educators w. iNnovative Technologies *[IMLS: 2022-2023]*



Computational Archival Science (CAS)

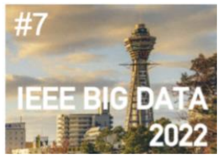
Working Definition:

- A transdisciplinary field grounded in archival, information, and computational science that is concerned with:
  - **the application of** computational methods and resources, design patterns, sociotechnical constructs, and human-technology interaction
  - **to** large-scale (big data) records/archives processing, analysis, storage, long-term preservation, and access problems
  - **with the aim** of improving and optimizing efficiency, authenticity, truthfulness, provenance, productivity, computation, information structure and design, precision, and human technology interaction
  - **in support of** acquisition, appraisal, arrangement and description, preservation, communication, transmission, analysis, and access decisions

https://ai-collaboratory.net/cas/

CAS PORTAL:

https://ai-collaboratory.net/cas/



* Workshops:

- 50+ workshops since 2016
- 9 CAS @ IEEE Big Data Conf. w. 150+ papers

Lessons learned from:

- **CAS#1**: 2016 in Washington, DC
- **CAS#2**: 2017 in Boston
- **CAS#3**: 2018 in Seattle
- **CAS#4**: 2019 in LA
- **CAS#5**: 2020 in Atlanta
- **CAS#6**: 2021 in Orlando
- **CAS#7**: 2022 in Osaka, Japan
- **CAS#8**: 2023 in Sorrento, Italy
- **CAS#9**: 2024 in Washington, DC

* Presentations

* Publications

* Infrastructure

https://ai-collaboratory.net/cas/cas-workshops/2024-9th-cas-workshop/

Welcome!

2024 IEEE International Conference on Big Data (IEEE BigData 2024)

Dec 15-18, 2024 @ Washington DC, USA

IEEE Big Data 2024: CAS #9

https://ai-collaboratory.net/cas/cas-workshops/2024-9th-cas-workshop/

- **Monday, Nov. 4, 2024 (final):** Due date for full workshop papers submission
- **Friday, Nov 15, 2024:** Notification of paper acceptance to authors
- **Wednesday, Nov 20, 2024 (hard deadline):** Camera-ready of accepted papers
- **Dec 15 to 18, 2024:** Day-long CAS workshop (in person) in Washington DC, USA (exact day TBD)

RESEARCH TOPICS COVERED:

- **Application of analytics to archival material**, including AI, ML, text-mining, data-mining, sentiment analysis, network analysis.
- **Analytics in support of archival processing**, including e-discovery, identification of personal information, appraisal, arrangement and description.
- **Scalable services for archives**, including identification, preservation, metadata generation, integrity checking, normalization, reconciliation, linked data, entity extraction, anonymization and reduction.
- **New forms of archives**, including Web, social media, audiovisual archives, and blockchain.
- **Cyber-infrastructures for archive-based research** and for development and hosting of collections
- **Big data and archival theory and practice**
- **Digital curation and preservation**
- **Crowd-sourcing** and archives
- **Big data and the construction of memory and identity**
- **Specific big data technologies** (e.g. NoSQL databases) and their applications
- **Corpora and reference collections** of big archival data
- **Linked data** and archives
- **Big data and provenance**
- **Constructing big data research objects** from archives
- **Legal and ethical issues** in big data archives

SOCIAL JUSTICE, HUMAN RIGHTS, CULTURAL HERITAGE:

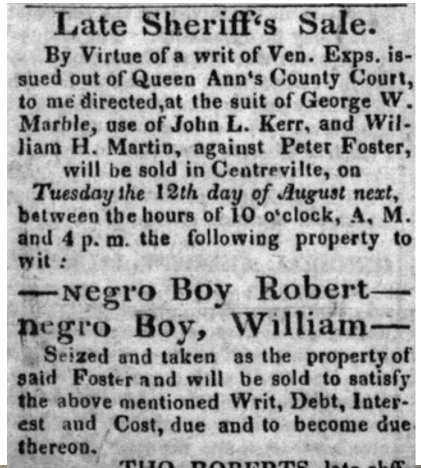
Upcoming Notebooks showcasing the use of AI / ML / GenAI / LLM:

- **Large Language Models (LLM) / ChatGPT: MSA Legacy of Slavery Project**
  - **Conversing with the Past: Re-examining the Legacy of Slavery in Domestic Traffic Newspaper Advertisements with OpenAI's GPT3 LLM,**
    - Rajesh Kumar Gnanasekaran, C. E. Haley, R. Marciano, UCL PRESS: *Artificial Intelligence for Cultural Heritage Organizations* (2024)
- **Transfer Learning / BERT: MSA Legacy of Slavery Project**
  - **Using Transfer Learning to contextually Optimize Optical Character Recognition (OCR) output and perform new Feature Extraction on a digitized cultural and historical dataset,**
    - Aravind Inbasekaran, Rajesh Kumar Gnanasekaran, R. Marciano, 2021 IEEE International Conference on Big Data, Dec. 15, 2021, Orlando, FL.
    - See: https://ai-collaboratory.net/wp-content/uploads/2021/11/3_Inbasekaran.pdf
- **Digital Curation / ML: FDR Presidential Library Morgenthau Holocaust Collections Project**
  - **Digital Curation to Support Machine Learning,**
    - Teddy Randby, Richard Marciano, 2020 IEEE International Conference on Big Data, Dec. 11, 2020, Atlanta, GA.
    - See: https://ai-collaboratory.net/wp-content/uploads/2020/11/Randby.pdf.
- **Computer Vision / ML: FDR Presidential Library Morgenthau Holocaust Collections Project**
  - **Using AI and ML to Optimize Information Discovery in Under-utilized, Holocaust-related Records**,
    - Kirsten Strigel Carter, Abby Gondek, William Underwod, Teddy Randby, Richard Marciano, AI & Society [Journal of Knowledge, Culture and Communication], Special issue on “Born Digital” – Shedding Light into the Darkness of Digital Culture, Fall 2021.
    - See: https://ai-collaboratory.net/wp-content/uploads/2021/10/Carter-et-al-Using-AI-and-ML-to-Optimize-Information-Discovery_MHCP_AIC.pdf
- **Computer Vision / ML: Spelman College Archives Photograph Collection**
  - **“An AI-Assisted Framework for Rapid Conversion of Descriptive Photo Metadata into Linked Data”**,
    - Jennifer Proctor, Richard Marciano, MTSR 2021 Virtual Conference, 15th International Conference on Metadata and Semantics Research, Nov. 29 – Dec. 3, 2021.
    - See: https://ai-collaboratory.net/wp-content/uploads/2021/10/2021-Proctor_paper.pdf
- **NLP / NER: NARA War Relocation Authority (WRA)**
  - **“Computational Curation of a Digitized Series of WWII Japanese-American Internment,”**
    - Underwood, B., Marciano, et al., IEEE Big Data 2017’s 2nd Computational Archival Science (CAS) Workshop, Boston, MA, Dec. 13, 2017.
    - See: https://ai-collaboratory.net/wp-content/uploads/2020/04/Underwood-CAS-2017.pdf

Data: Domestic Traffic Ads

MSA Legacy of Slavery Project (1824 to 1864)

"[D]omestic traffic is defined as the interstate and intrastate trade of enslaved men, women, women and children. Similar to runaway ads and committal notices, domestic traffic ads were a means of communicating to the general public the subscriber's desire to buy or sell a slave(s). Ads could be placed by private slave dealers and agents, gentry in need of domestic help, yeomen in need of extra field hands, or a public sale of an estate by the orphan's court." MSA



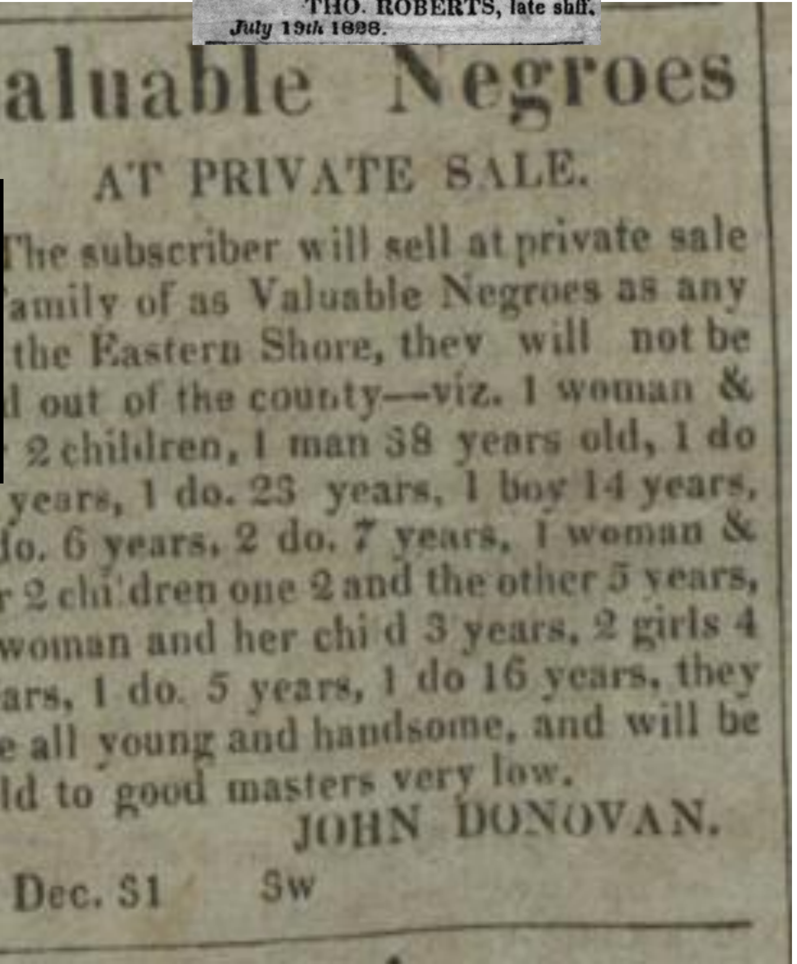
Late Sheriff's Sale.

By Virtue of a writ of Ven. Exps. issued out of Queen Ann's County Court, to me directed,at the suit of George W. Marble, use of John L. Kerr, and William H. Martin, against Peter Foster, will be sold in Centrevilte, on *Tuesday the 12th day of August next,* between the hours of 10 o'clock, A. M. and 4 p. m. the following property to wit:

—Negro Boy Robert—
negro Boy, William—

Seized and taken as the property of said Foster and will be sold to satisfy the above mentioned Writ, Debt, Interest and Cost, due and to become due thereon.

THO. ROBERTS, late shff.
July 19th 1828.



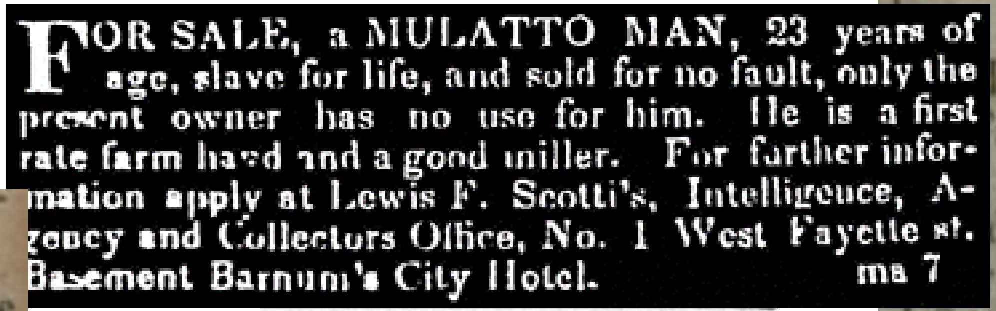
Valuable Negroes

AT PRIVATE SALE.

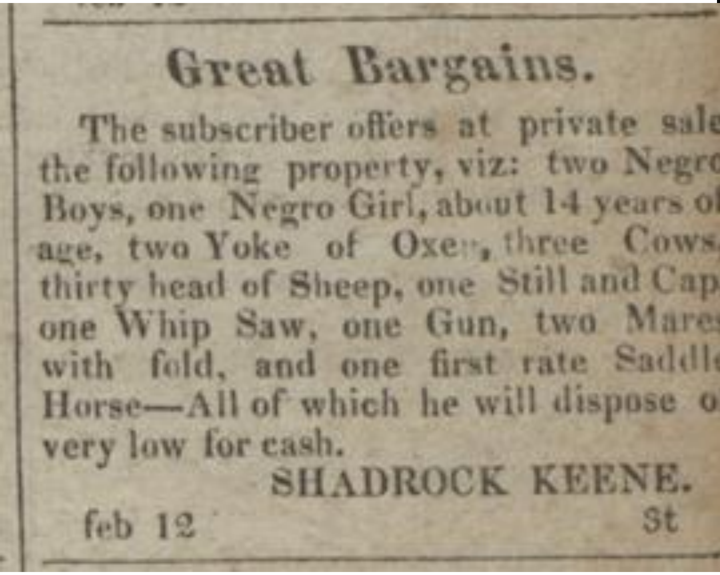
he subscriber will sell at private sale amily of as Valuable Negroes as any the Eastern Shore, they will not be l out of the county—viz. 1 woman & 2 children, 1 man 38 years old, 1 do 20 years, 1 do. 23 years, 1 boy 14 years, 1 do. 6 years, 2 do. 7 years, 1 woman & her 2 chi'dren one 2 and the other 5 years, 1 woman and her chi d 3 years, 2 girls 4 years, 1 do. 5 years, 1 do 16 years, they are all young and handsome, and will be sold to good masters very low.

JOHN DONOVAN.

Dec. 31 3w



FOR SALE, a MULATTO MAN, 23 years of age, slave for life, and sold for no fault, only the present owner has no use for him. He is a first rate farm hand and a good miller. For further information apply at Lewis F. Scotti's, Intelligeuce, Agency and Collectors Office, No. 1 West Fayette st. Basement Barnum's City Hotel. ma 7

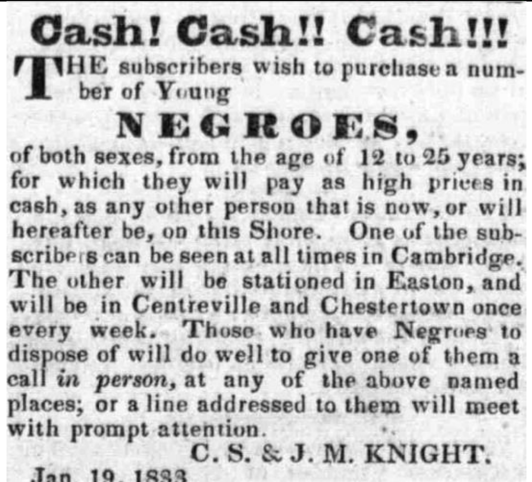


Great Bargains.

The subscriber offers at private sale the following property, viz: two Negro Boys, one Negro Girl, about 14 years of age, two Yoke of Oxen, three Cows, thirty head of Sheep, one Still and Cap, one Whip Saw, one Gun, two Mares with fold, and one first rate Saddle Horse—All of which he will dispose of very low for cash.

SHADROCK KEENE.

feb 12 3t



Cash! Cash!! Cash!!!

THE subscribers wish to purchase a number of Young

NEGROES,

of both sexes, from the age of 12 to 25 years; for which they will pay as high prices in cash, as any other person that is now, or will hereafter be, on this Shore. One of the subscribers can be seen at all times in Cambridge. The other will be stationed in Easton, and will be in Centreville and Chestertown once every week. Those who have Negroes to dispose of will do well to give one of them a call *in person,* at any of the above named places; or a line addressed to them will meet with prompt attention.

C. S. & J. M. KNIGHT.

Jan 19, 1833

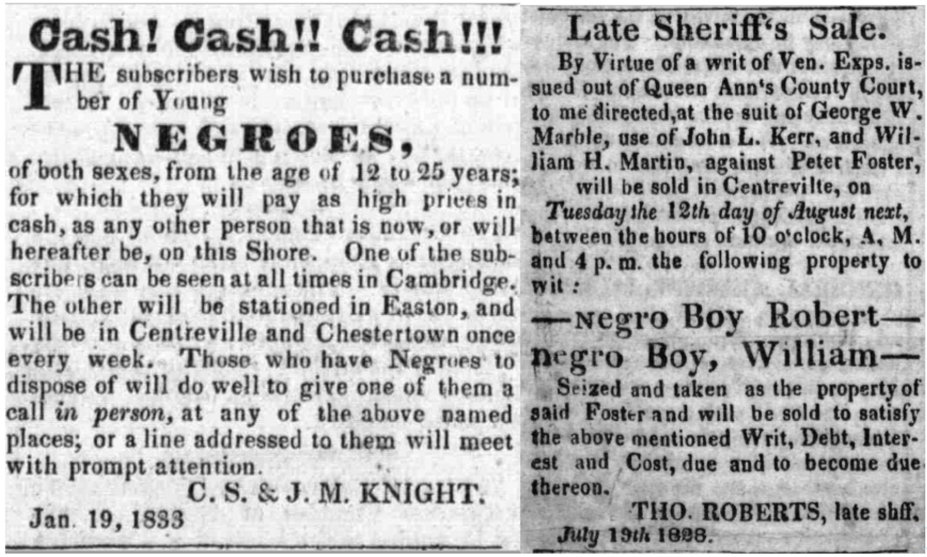
2023 Pilot:

on incorporating language-based generative AI technologies into future library and archival services

- UCL Press/AEOLIAN 2024 edited collection on

Navigating Artificial Intelligence for Cultural Heritage Organisations:

- **“Conversing with the Past: Re-examining the Legacy of Slavery in Domestic Traffic Newspaper Advertisements with OpenAI's GPT3 LLM”**.
  - Rajesh Kumar Gnanasekaran, Christopher E. Haley, and Richard Marciano.

DTA advertisement samples for a purchase ad and a sale ad	**GPT-powered AI bot queries** with a subset of 764 ads	**Lessons learned from the pilot**
 **Cash! Cash!! Cash!!!** THE subscribers wish to purchase a number of Young **NEGROES,** of both sexes, from the age of 12 to 25 years; for which they will pay as high prices in cash, as any other person that is now, or will hereafter be, on this Shore. One of the subscribers can be seen at all times in Cambridge. The other will be stationed in Easton, and will be in Centreville and Chestertown once every week. Those who have Negroes to dispose of will do well to give one of them a call *in person*, at any of the above named places; or a line addressed to them will meet with prompt attention. C. S. & J. M. KNIGHT. Jan. 19, 1833 **Late Sheriff‘s Sale.** By Virtue of a writ of Ven. Exps. issued out of Queen Ann‘s County Court, to me directed,at the suit of George W. Marble, use of John L. Kerr, and William H. Martin, against Peter Foster, will be sold in Centrevilte, on *Tuesday the 12th day of August next,* between the hours of 10 o‘clock, A. M. and 4 p. m. the following property to wit : **—Negro Boy Robert—** **negro Boy, William—** Seized and taken as the property of said Foster and will be sold to satisfy the above mentioned Writ, Debt, Interest and Cost, due and to become due thereon. THO. ROBERTS, late shff. *July 19th 1828.*	1. How many ads are in the dataset? 2. How many ads were placed by County? 3. Can you report ad counts for both public auctions and public sales? 4. Were any ads placed on Christmas Day (how many and which years)? 5. How many cooks were on sale? 6. Can you show the number of ads placed by each ad source and county? 7. How many ads were placed for individuals under 10 years old?	● DTA content was not AI-ready and needed significant pre-processing to lend itself to generative AI. ● GPT models can understand natural language statements to automatically choose the columns needed for aggregate data analysis. ● GPT models need to understand the context of the DTA collection and be further trained. ● GPT models seem to be able to duplicate some of the level of analysis seen in non-AI exploratory case studies.

Project Overview

- Proposal to the IMLS National Leadership Grant for Libraries (NLG-L) program (Aug. 2024 ?)
- NLG-L Goal 3 & Objective 3.2: ‘***Improve** the ability of libraries and archives to provide broad **access** to and use of information and collections*’ & ‘***Support** innovative approaches to **digital collection management**.*’
- Community Partners/Collaborators:

Maryland State Archives
Legacy of Slavery Program

Kennard African American Cultural Heritage Center
(Queen Anne’s County, Maryland)

Research Methodology: **Exploratory Case-Study with Stakeholder Participation**

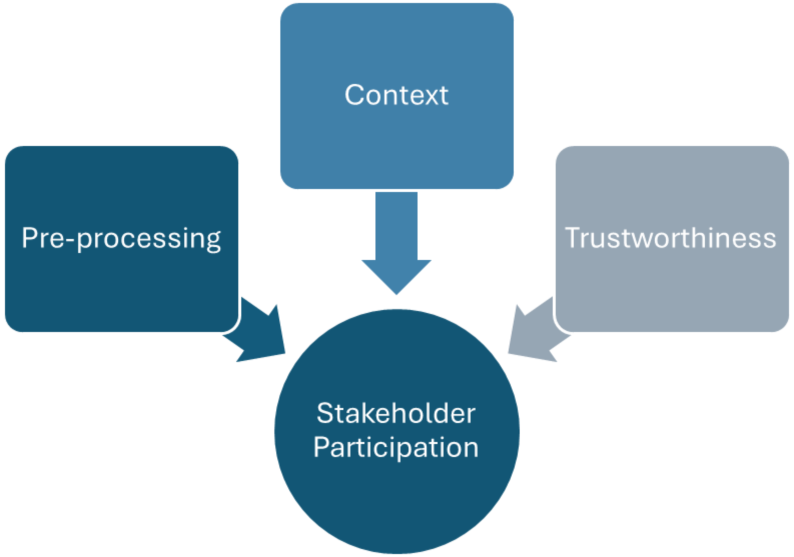
Research Questions	Research Methods
(RQ1) How should we further **curate** library and archival **collections** to make them **AI-ready**?	Research collection curation techniques and AI-readiness to support generative AI work and refine GPT models to understand a collection’s context.
(RQ2) How can we **groundtruth GPT models** with traditional, non-AI exploratory data analysis models?	Interrogate the DTA dataset using traditional, non-AI approaches for comparison with GPT model results.
(RQ3) How can we **incorporate socio-technical considerations to promote trustworthiness and mitigate potential bias** arising from the use of GPT models with library and archival collections?	Engage domain experts and community members in GPT model design and evaluation via surveys and focus groups; and incorporate their contextual input for data preparation, prompt engineering, and model training.

(RQ1) *How should we further **curate** library and archival **collections** to make them **AI-ready**?*

NSF workshop -- ID4: Institute of Data Drive Dynamical Design, AI-Ready Data: Navigating the Dynamic Frontier of Metadata and Ontologies, Apr. 15-16, 2024. Drexel U.

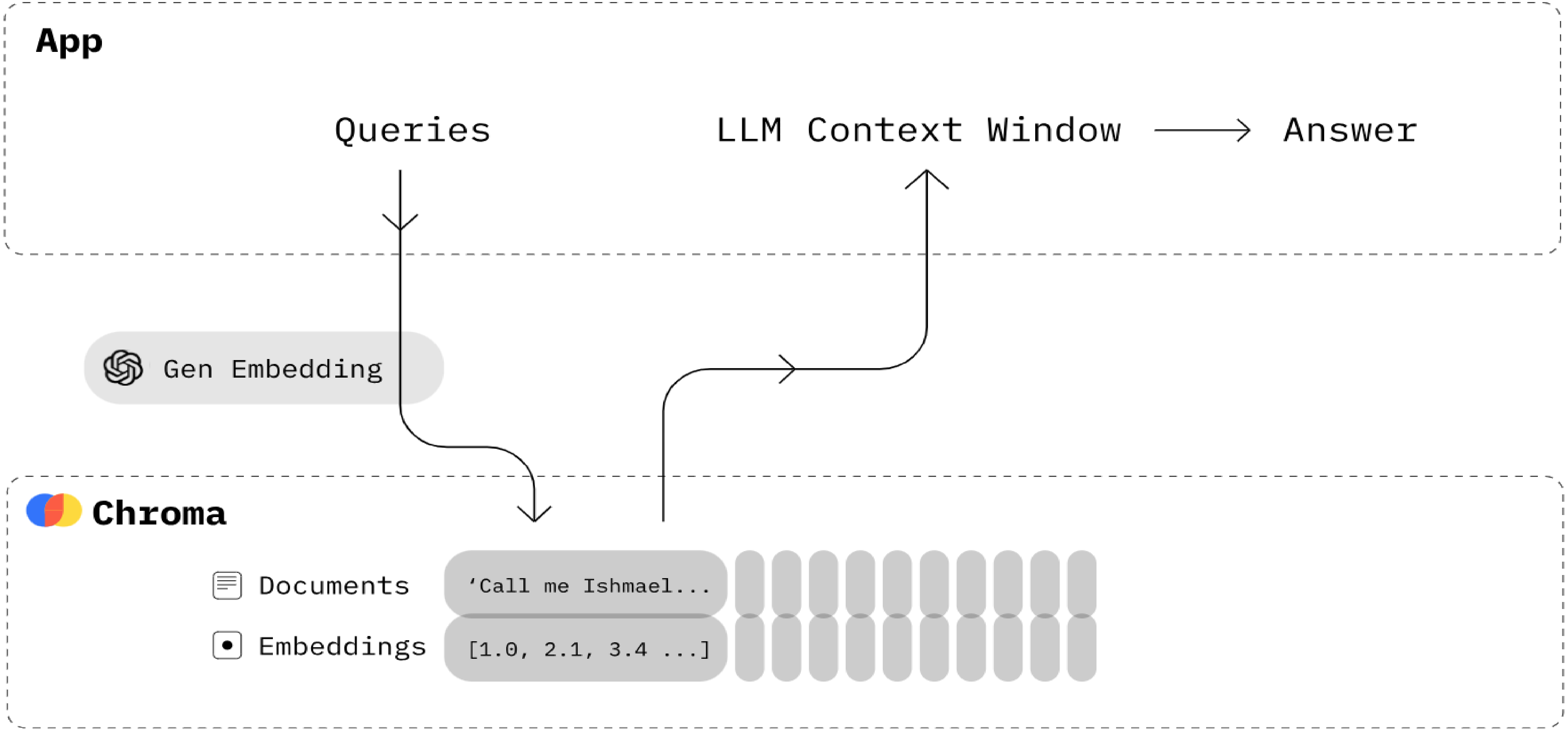
Data Preparation

- Explore operationalization of focus group expectations and concerns as metadata, features, or functionality to be used in the initial data preparation and/or model training. →**METADATA SPECIFICATION & USE IN PRE-PROCESSING FOR AI-READINESS**



- Metadata added during pilot (***record***)
  (1) Terms of Service,
  (2) Trade Reason
  (3) Features
  (4) Terms of Sale
  (5) Owner
  (6) Optical Character Recognition (OCR) text of each DTA ad image.
- Examples of potential metadata developed through stakeholder participation (***context/use/trustworthiness focus***)
  - Metadata for specific use cases (e.g. historical research, K-12 education, higher education, look-up services in libraries or archives, interactive cultural heritage displays in museums, personal and professional genealogy, etc)
  - Historical contextual information
  - Labels reflecting evolution of knowledge Culturally aware labels
  - Labels reflecting user sensitivities or risk tolerances

RQ1 - Using DTA with LLM to create a Contextualized AI Chatbot



Retrieval Augmented Generation

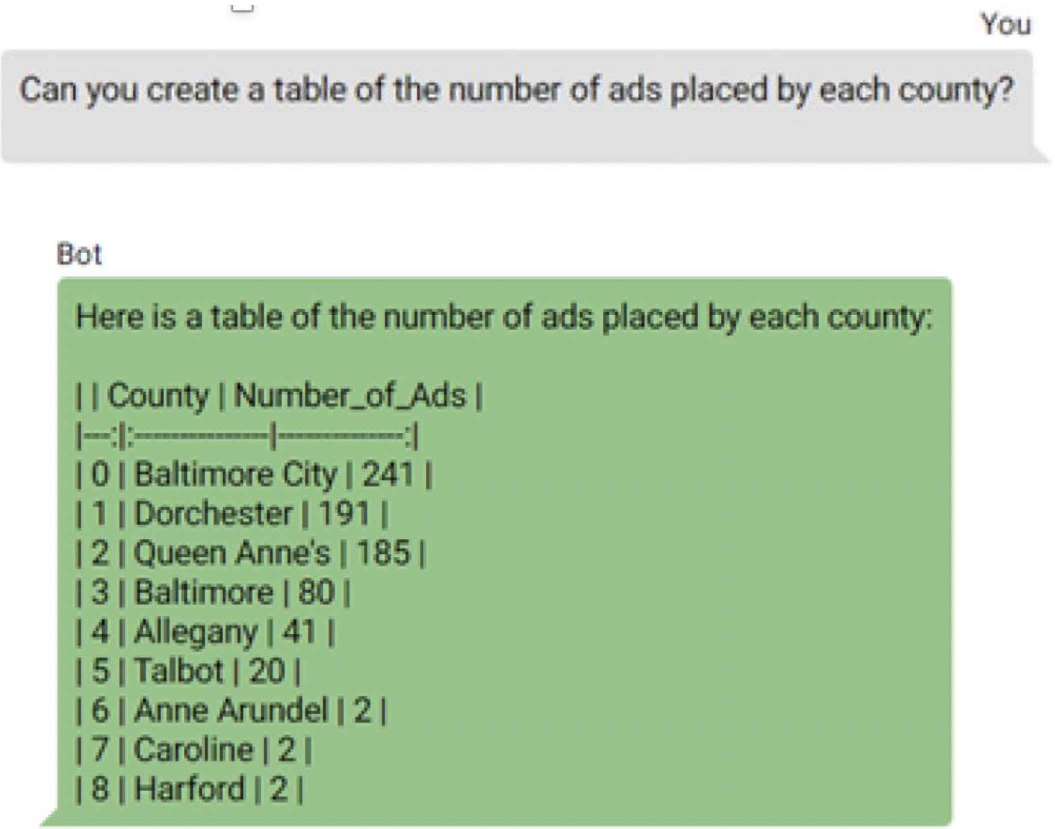
Source - https://heidloff.net/article/retrieval-augmented-generation-chroma-langchain/

(RQ2) *How can we* ***compare GPT models with traditional, non-AI*** *exploratory data analysis models?*

International Symposium on Grids and Clouds (ISGC March 2024): Leveraging OpenAI's LLMs to Create Learning-as-a-Service (LaaS) Solutions for Culturally Rich Conversational AI: Chatbot: ChatLoS - A Study Using the Legacy of Slavery Dataset, Rajesh Gnanasekaran

LLM Chatbot vs. Tableau Comparative Data Analysis – Test Results:

https://docs.google.com/document/d/1KFMMo1yDm2ygtcXvzhZJ2E7HdV9o6FAWcokfX8c--MA/edit?usp=sharing



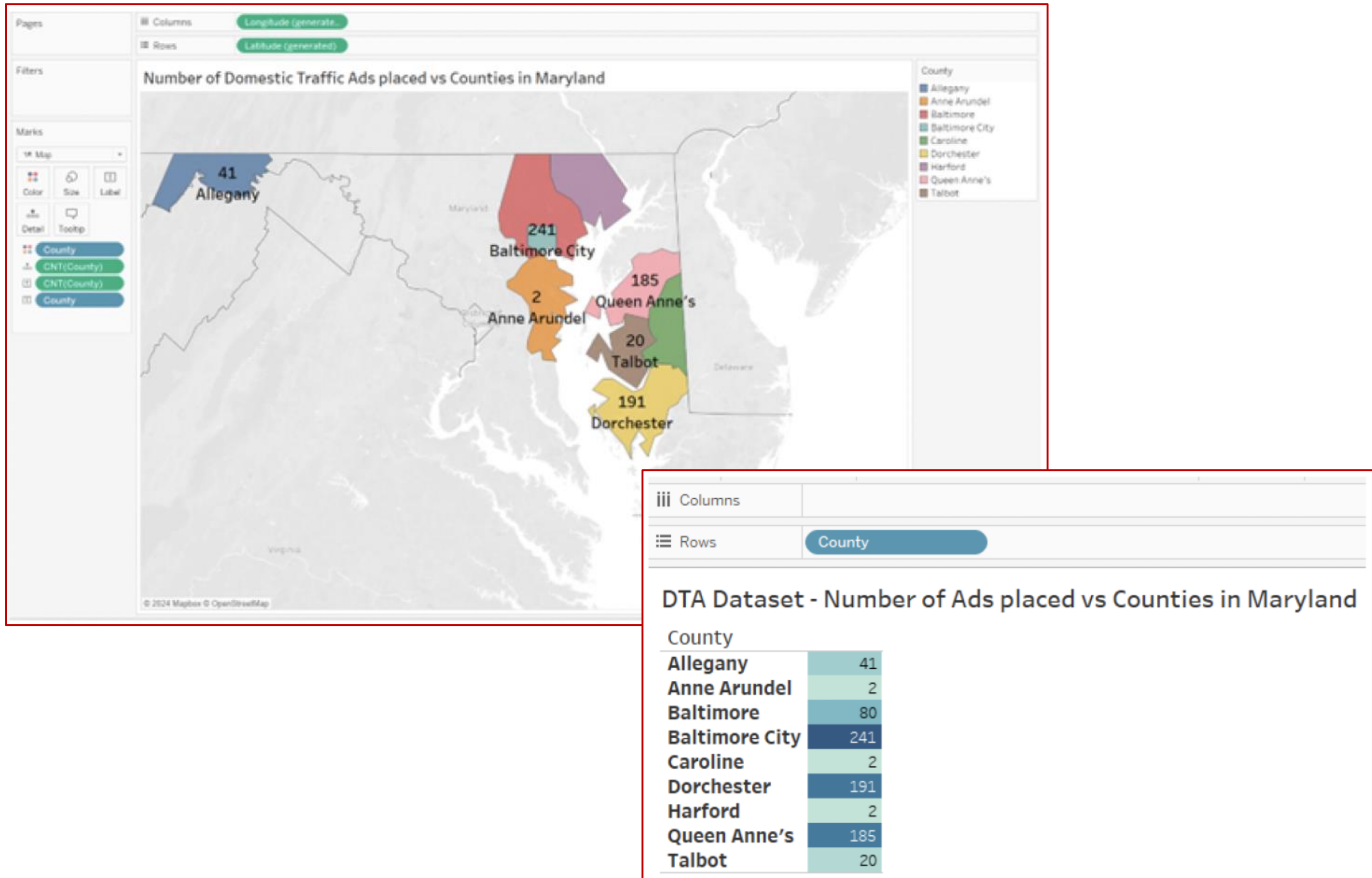
You

Can you create a table of the number of ads placed by each county?

Bot

Here is a table of the number of ads placed by each county:

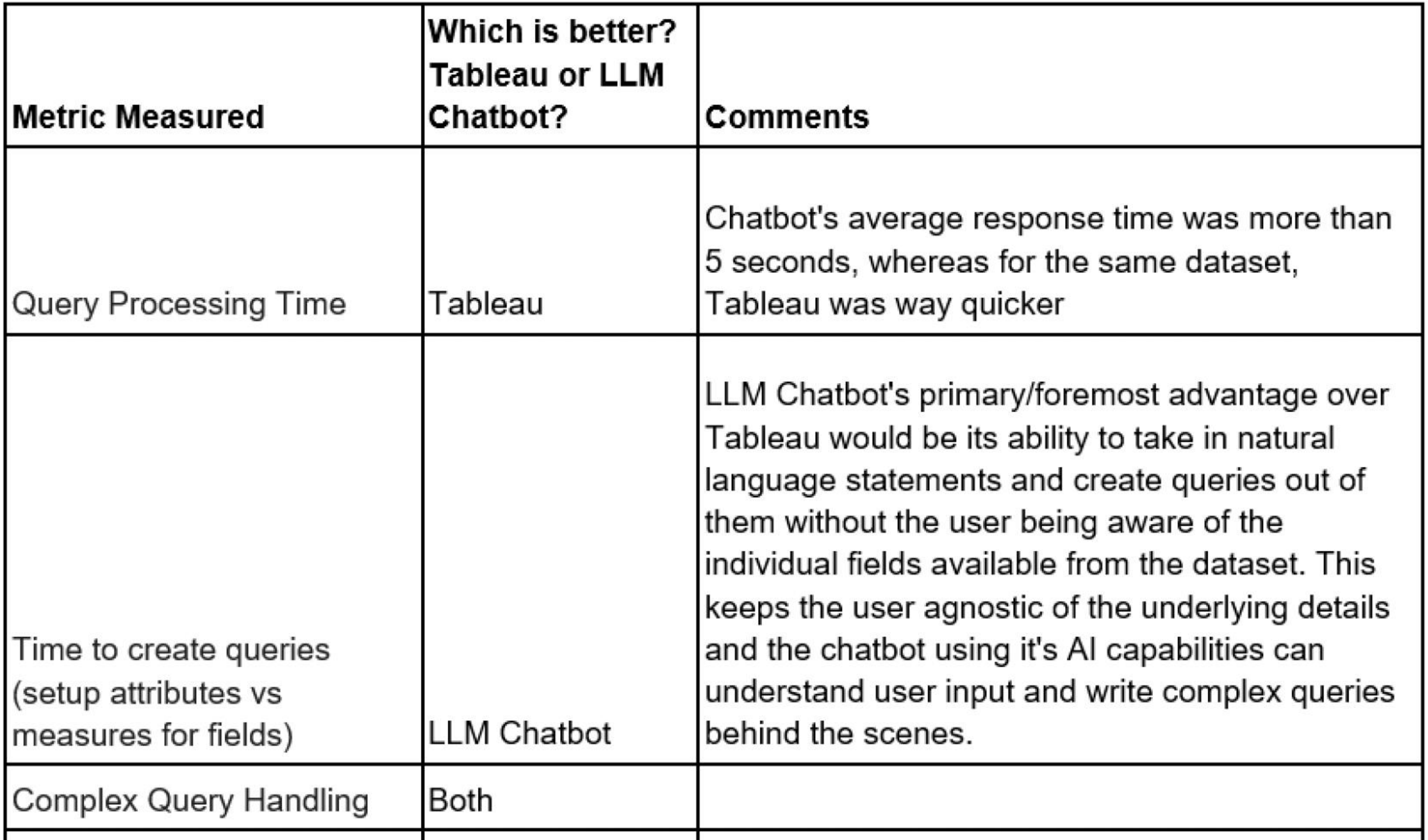
```
| | County | Number_of_Ads |
|---:|:-------------|-------------:|
| 0 | Baltimore City | 241 |
| 1 | Dorchester | 191 |
| 2 | Queen Anne's | 185 |
| 3 | Baltimore | 80 |
| 4 | Allegany | 41 |
| 5 | Talbot | 20 |
| 6 | Anne Arundel | 2 |
| 7 | Caroline | 2 |
| 8 | Harford | 2 |
```



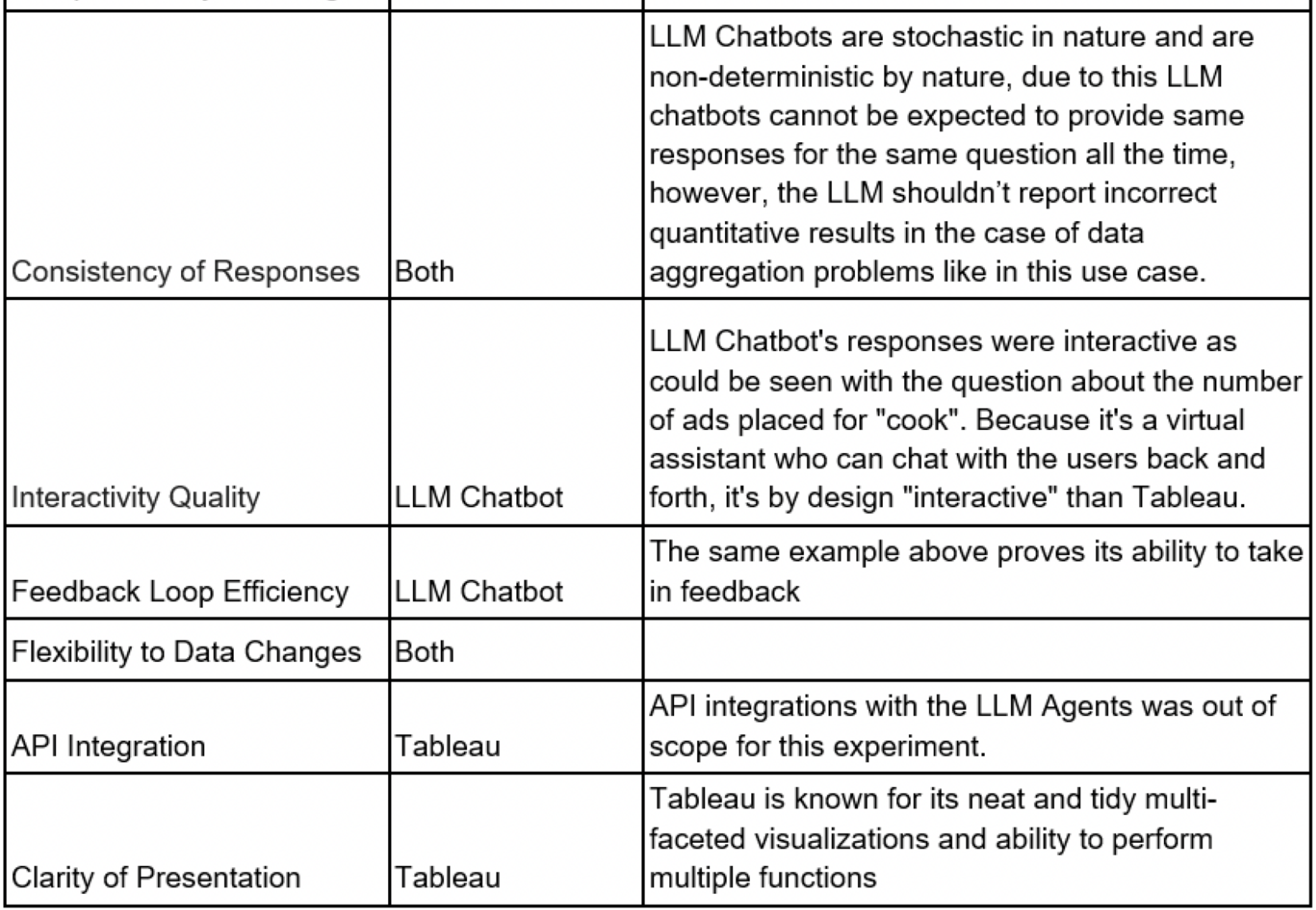
DTA Dataset - Number of Ads placed vs Counties in Maryland

County	
Allegany	41
Anne Arundel	2
Baltimore	80
Baltimore City	241
Caroline	2
Dorchester	191
Harford	2
Queen Anne's	185
Talbot	20

Metric Comparison between LLM Chatbot and Tableau:



Metric Measured	Which is better? Tableau or LLM Chatbot?	Comments
Query Processing Time	Tableau	Chatbot's average response time was more than 5 seconds, whereas for the same dataset, Tableau was way quicker
Time to create queries (setup attributes vs measures for fields)	LLM Chatbot	LLM Chatbot's primary/foremost advantage over Tableau would be its ability to take in natural language statements and create queries out of them without the user being aware of the individual fields available from the dataset. This keeps the user agnostic of the underlying details and the chatbot using it's AI capabilities can understand user input and write complex queries behind the scenes.
Complex Query Handling	Both	



Consistency of Responses	Both	LLM Chatbots are stochastic in nature and are non-deterministic by nature, due to this LLM chatbots cannot be expected to provide same responses for the same question all the time, however, the LLM shouldn’t report incorrect quantitative results in the case of data aggregation problems like in this use case.
Interactivity Quality	LLM Chatbot	LLM Chatbot's responses were interactive as could be seen with the question about the number of ads placed for "cook". Because it's a virtual assistant who can chat with the users back and forth, it's by design "interactive" than Tableau.
Feedback Loop Efficiency	LLM Chatbot	The same example above proves its ability to take in feedback
Flexibility to Data Changes	Both	
API Integration	Tableau	API integrations with the LLM Agents was out of scope for this experiment.
Clarity of Presentation	Tableau	Tableau is known for its neat and tidy multi-faceted visualizations and ability to perform multiple functions

Status/Opportunities/Challenges

- Opportunity
  - “Strengthen the ability of libraries and archives to serve the public and research communities by unlocking new dimensions of access, discovery, and understanding within their collections.”
  - Support FAIR (Findable, Accessible, Interoperable, and Reusable) principles and reproducible computational research (RCR)
  - Prototype participatory methods that may have more general application (the broader AI community can learn from library/archive practices)
  - Contribute to generalizing use of ontologies that incorporate stakeholder participation
- Challenges
  - Adopting/adapting methods for human-human along with human-tech feedback
  - Operationalizing stakeholder feedback in data processing and/or model functionality – limitations? conceptual integrity?
  - Potential complexity and contradiction in stakeholder input
  - Misuse of GenAI to obscure the cultural record

Thank you

Contact Information:

Richard Marciano marciano@umd.edu

Advanced Information Collaboratory

COLLEGE OF INFORMATION STUDIES